



Department
for Work &
Pensions

Improving Lives: Helping Workless Families

Supporting methodology document

April 2017

Contents

- Contents** _____ **2**
- Introduction** _____ **3**
- Overview of the analysis** _____ **5**
 - 1) Part one: children in workless families** _____ **5**
 - 2) Indicator two: parental conflict** _____ **28**
 - 3) Indicator three: poor parental mental health** _____ **36**
 - 4) Targeted provision to reduce parental conflict** _____ **38**
- Description of the main data sources used** _____ **40**
- Definitions and terminology within the statistics** _____ **43**
- General limitations of this analysis** _____ **46**

Introduction

Purpose of document

This supporting methodology document is a technical accompaniment to the Improving Lives: Helping Workless Families Analysis and Research Pack, available online: <https://www.gov.uk/government/publications/improving-lives-helping-workless-families-evidence-base>.

This document primarily supports specific sections of the Analysis and Research Pack where DWP analysts have conducted new survey-based analysis (as listed below). It provides more detailed methodological descriptions and definitions as well as detailed outputs from the regression analyses. It provides a detailed description of the data sources used in our analyses, and discusses some general limitations to this analysis given the methodology and data sources used. Finally, this document also includes estimates of how many families we estimate are eligible for new targeted provision to reduce parental conflict.

This document is also accompanied by detailed data tables, which present full results from all of the tables, charts and statistics presented in the Analysis and Research Pack. This is also available online:

<https://www.gov.uk/government/publications/improving-lives-helping-workless-families-evidence-base>.

Structure and contents

This document contains:

- An overview of the analysis conducted in the following sections:
 1. Part one: children in workless families (pages 7 to 39 of the Analysis and Research Pack);
 2. Indicator two: parental conflict (pages 50 to 69);
 3. Indicator three: poor parental mental health (pages 70 to 78);
 4. Estimates of how many families are eligible for new targeted provision to reduce parental conflict (new information not contained in the Analysis and Research pack);
- A detailed description of the data sources used for this analysis;
- The precise definition of characteristics we have derived in our analysis, and;
- A discussion of the general limitations of these statistics.

Potential users of this document

This document is aimed at interested users of these statistics who would like a more detailed understanding of the methodology and data sources underpinning the analysis presented in the Analysis and Research Pack. This may include: policy and analytical teams within DWP, the Devolved Administrations and other government departments, local authorities, parliament, academics, journalists, and the voluntary

Improving Lives: Helping Workless Families - Supporting Methodology Paper

sector. Researchers and analysts outside government could use the statistics and data to further examine topics such as worklessness and disadvantage.

We welcome any enquiries from interested users:

helpingworklessfamilies@dwp.gsi.gov.uk

Overview of the analysis

1) Part one: children in workless families

Overview of section

This section of the Analysis and Research Pack sets out the evidence behind some of the issues associated with persistent worklessness, how these disadvantages are often interrelated, and how they impact on children's outcomes. In particular, we conducted regression analysis (logistic regression) to help further explore the association between worklessness, or moving into or out of worklessness, and a range of selected parent and family characteristics available in the Understanding Society survey. We also built a series of multiple linear regression models to explore the association between worklessness and selected child outcomes using the Millennium Cohort Study.

Data sources

The analysis in this section is predominantly based on the Understanding Society survey and the Millennium Cohort Study. See the Data Sources section for a full description of these data sources.

Underlying sample population

Understanding Society

The underlying sample for this section was children who were present in all waves of the survey. Children could only join or leave the sample in subsequent waves if they were born into the sample, or stopped being a dependent child.¹ For basic descriptive statistics, all children in the sample were included in comparisons. For transitions statistics, children who were present in the study for 2 or more waves were included (unless otherwise specified).

In certain places where sample sizes were small, we 'pooled' the data. This is clearly specified underneath the relevant charts in the Analysis and Research Pack. This means that children can be counted more than once if they participated in a number of waves of data collection. As children may experience multiple family work states, and multiple work transitions, this takes account of the experience of the child at each time point (wave). Standard errors are clustered at the individual and family level to account for individual/family effects.

Millennium Cohort Study

The underlying sample for this section was children who were present in all waves of the survey (up until wave 5 – aged 11).

¹ Dependent children are defined as children aged between birth and 16 years of age, or 17 and 18 year olds who are living at home and in full-time education.

Weights used and attrition

Combined cross sectional and longitudinal weights provided by Understanding Society and Millennium Cohort Study were used to adjust for initial sampling bias and attrition during the study. In this paper, the supporting tables, and the Analysis and Research Pack, all values are weighted. There may be unobservable characteristics which are associated with both non-response and attrition which could not be modelled in the data, and which would likely lead to lower representation of less advantaged groups. However, the data (with weighting) generally aligns well with known statistics on the UK population.

Regression analysis on parental characteristics associated with worklessness

Methodology

We conducted logistic regression analysis to help further understand the relationship between worklessness, or moving into or out of worklessness, and a range of selected parent and family characteristics available in the Understanding Society survey. For this analysis, data was pooled across waves and clustered at the individual child level. Only children who were able to experience a transition (i.e. are present in the survey for 2 or more consecutive waves) were included.

Three separate logistic regression analyses were run:

1. For the first regression, the dependent variable was whether the child was living in a workless family². Characteristics here were taken from the same wave as the measure of work status.
2. For the second regression, the dependent variable was whether the child was living in a family that had made a move from worklessness into work.
3. For the third regression, the dependent variable was whether the child was living in a family that had made a move from in work to worklessness.

Characteristics here were taken from the wave before the transition occurred (i.e. if the family moved from work into worklessness, characteristics are measured when the family were in work). In this way, the longitudinal nature of the data is exploited to get a better understanding of how family characteristics potentially impact on work status. Because the transition into or out of work occurs after the characteristic is measured, it cannot cause a change in that characteristic (for example, poor parental mental health may occur after falling into worklessness, however in this case we measure poor parental mental health whilst the family are still in work. This does not deal with the issue entirely, however, as parents may anticipate entry into worklessness (or into work), and thus their characteristics may change as a response to this.

Robust standard errors were clustered on the individual and family, and longitudinal weights provided by Understanding Society, were used. Data was pooled across the five waves used in the sample, and year (wave) of response is included as a control in all regressions.

² For a definition see the Definitions and Terminology section, on page 42.

Improving Lives: Helping Workless Families - Supporting Methodology Paper

Responses were first constructed at the child level, such that parental characteristics were counted more than once if they had more than one child. The regression was then repeated at family level to ensure results were not biased by this. This gave very similar results. Results from the first regression are presented in the Analysis and Research Pack (see pages 13 to 27).

Missing values on independent variables were included as dummy variables; however regressions were only run for children for whom we knew their parental work status.

Results from regression analysis

Table 1: Relationship between work states and number of potential barriers to work the family faces: analysis at child level (Understanding Society)

Selected potential barriers	Workless household	Move into work	Become workless
Nil	0 (.)	0 (.)	0 (.)
One	0.197*** (0.014)	-0.0614 (0.044)	0.0331*** (0.005)
Two	0.336*** (0.014)	-0.160*** (0.042)	0.0617*** (0.005)
Three or more	0.483*** (0.013)	-0.249*** (0.041)	0.0791*** (0.006)
Parents from an ethnic minority group	0.0343*** (0.008)	0.0184 (0.016)	0.0125*** (0.003)
<i>Observations</i>	50,764	5,797	33,345
<i>R2</i>	0.27	0.04	0.08

Data: Understanding Society survey, 2010-2015. Logistic Regression. Marginal effects are shown with standard errors in parentheses. * p < 0.10, ** p < 0.05, *** p < 0.01 Dummy variables are included for missing values, and binary variables for wave of response are included.

Table 2: Relationship between work states and number of potential barriers to work the family face: analysis at family level (Understanding Society)

Selected potential barriers	Workless household	Move into work	Become workless
Nil	0 (.)	0 (.)	0 (.)
One	0.182*** (0.017)	-0.0127 (0.061)	0.0296*** (0.006)
Two	0.320*** (0.017)	-0.0986* (0.059)	0.0597*** (0.007)
Three or more	0.463*** (0.017)	-0.207*** (0.058)	0.0739*** (0.007)
Parents from an ethnic minority group	0.0267*** (0.009)	0.0171 (0.021)	0.0102** (0.004)
<i>Observations</i>	28,098	3,030	18,416
<i>R2</i>	0.27	0.04	0.08

Data: Understanding Society survey, 2010-2015. Logistic Regression. Marginal effects are shown with standard errors in parentheses. * p < 0.10, ** p < 0.05, *** p < 0.01 Dummy variables are included for missing values, and binary variables for wave of response are included.

Table 3: Relationship between work states, family and parent characteristics: analysis at child level (Understanding Society)

	Workless household	Move into work	Become workless
Couple parents	0 (.)	0 (.)	0 (.)
Lone parent with a child under 5	0.405*** (0.019)	-0.130*** (0.020)	0.0582*** (0.010)
Lone parent with a child 5+	0.201*** (0.010)	-0.0665*** (0.020)	0.0220*** (0.005)
Oldest adult is under 30	0.0872*** (0.009)	-0.0427*** (0.016)	0.0218*** (0.005)
Lives in a large family (3+ children)	0.0589*** (0.005)	-0.0397*** (0.014)	0.0118*** (0.002)
Lives in social housing	0.143*** (0.007)	-0.0696*** (0.013)	0.0443*** (0.004)
At least one parent reports poor mental health	0.0442*** (0.004)	-0.0481*** (0.016)	0.00981*** (0.003)
At least one parent has a longstanding limiting illness and/or disability	0.108*** (0.006)	-0.153*** (0.014)	0.0130*** (0.003)
Parents have qualifications below GCSE (5A*-C) level or equivalent	0.128*** (0.011)	-0.0735*** (0.015)	0.0327*** (0.007)
White	0 (.)	0 (.)	0 (.)
Mixed	-0.0222** (0.011)	0.00784 (0.039)	-0.0000 (0.004)
Indian	-0.0111 (0.015)	-0.000411 (0.042)	0.00451 (0.007)
Pakistani	0.0405*** (0.014)	-0.0449** (0.022)	0.00304 (0.005)
Bangladeshi	0.0326** (0.014)	0.0668** (0.029)	0.00793 (0.007)
Black Caribbean	-0.00476 (0.015)	-0.0654** (0.026)	-0.00882* (0.005)
Black African	0.000317 (0.012)	0.0978*** (0.031)	0.0181** (0.007)
Other	0.0184 (0.014)	0.104*** (0.036)	0.0161* (0.009)
<i>Observations</i>	<i>50,764</i>	<i>5,797</i>	<i>33,345</i>
<i>R2</i>	<i>0.40</i>	<i>0.07</i>	<i>0.13</i>

Data: Understanding Society survey, 2010-2015. Logistic Regression. Marginal effects are shown with standard errors in parentheses. * p < 0.10, ** p < 0.05, *** p < 0.01 Dummy variables are included for missing values, and binary variables for wave of response are included.

Table 4: Relationship between work states, family and parent characteristics: analysis at family level (Understanding Society)

	Workless household	Move into work	Become workless
Couple parents	0 (.)	0 (.)	0 (.)
Lone parent with a child under 5	0.390*** (0.021)	-0.136*** (0.027)	0.0498*** (0.013)
Lone parent with a child 5+	0.184*** (0.011)	-0.0510* (0.027)	0.0295*** (0.006)
Oldest adult is under 30	0.0856*** (0.011)	-0.0384* (0.021)	0.0278*** (0.007)
Lives in a large family (3+ children)	0.0572*** (0.007)	-0.0400* (0.021)	0.0103*** (0.004)
Lives in social housing	0.131*** (0.009)	-0.0636*** (0.018)	0.0371*** (0.005)
At least one parent reports poor mental health	0.0450*** (0.005)	-0.0520** (0.021)	0.0106*** (0.004)
At least one parent has a longstanding limiting illness and/or disability	0.114*** (0.007)	-0.159*** (0.019)	0.0167*** (0.005)
Parents have qualifications below GCSE (5A*-C) level or equivalent	0.127*** (0.014)	-0.0929*** (0.020)	0.0239*** (0.009)
White	0 (.)	0 (.)	0 (.)
Mixed	-0.0271*** (0.010)	0.0364 (0.047)	0.00356 (0.006)
Indian	0.00331 (0.018)	-0.00932 (0.052)	0.00570 (0.009)
Pakistani	0.0530*** (0.018)	-0.0461 (0.034)	0.00400 (0.008)
Bangladeshi	0.0469** (0.019)	0.0413 (0.045)	0.0107 (0.010)
Black Caribbean	0.00593 (0.019)	-0.0491 (0.036)	-0.00263 (0.007)
Black African	0.00238 (0.015)	0.0713 (0.044)	0.0123 (0.010)
Other	0.0111 (0.016)	0.0866* (0.045)	0.00219 (0.009)
<i>Observations</i>	<i>28,098</i>	<i>3,030</i>	<i>18,416</i>
<i>R2</i>	<i>0.39</i>	<i>0.07</i>	<i>0.12</i>

Data: Understanding Society survey, 2010-2015. Logistic Regression. Marginal effects are shown with standard errors in parentheses. *p < 0.10, ** p < 0.05, *** p < 0.01 Dummy variables are included for missing values, and binary variables for wave of response are included.

Regression analysis on association between worklessness and child outcomes

Methodology

To understand the relationship between parental work status and child outcomes, linear regression models were fitted with three model specifications. The dependent variables in the regressions were the child outcomes, which included: cognitive ability, internalising problems, externalising problems, and pro-social behaviours³. All child outcomes were standardised so that coefficients refer to standard deviation differences based on the independent variables in the models. Full model specifications are outlined in the table below.

	Description of the model	Variables used
Model 1	Mean difference in outcomes by family work status in wave 5, adjusting for pre-determined variables (measured at baseline – wave 1).	<ul style="list-style-type: none"> • Country of birth • Sex of child • Age in months at interview (quadratic) • Month-year of birth (of child) • Survey stratum
Model 2	Mean difference in outcomes by worklessness status adjusting for: Model 1 plus measures of family circumstance (measured at baseline), which may confound work patterns and child development:	<ul style="list-style-type: none"> • Low qualification household • Main language spoken in HH is not English • Ethnicity of main carer at baseline • Teenager mother (aged under 19 years at birth) • Number of children in family at birth (quadratic) • Deciles of multiple area deprivation (IMD) • Low birth weight (<2500g) • Born prematurely • Unplanned pregnancy • At least one carer experiences anxiety/depression on Rutter inventory scale • Depressed parent at 9 months • Lone parent family at 9 months • Grandfather in work when main carer was age 14 years • <i>Health and well-being and family</i> • At least one carer has severe mental distress • At least one carer has a longstanding illness which limits them in some way • At least one carer has low life satisfaction (bottom quartile of distribution) • At least one change in family structure, no change (couple), no change (lone parent).
Model 3	Model 3 plus quintile of equivalised disposable income (using the OECD equivalisation scale)	

³ For a definition see page 33 of the Analysis and Research Pack.

Results from regression

Table 5: Association of household work status with child cognitive ability at age 11 years: wave 5 unless otherwise noted. (Millennium Cohort Study)

VARIABLES	(Model 1)	(Model 2)	(Model 3)
Workless household	-0.42*** (0.04)	-0.19*** (0.06)	-0.04 (0.06)
Sex of child (male = 0)	-0.08*** (0.02)	-0.07*** (0.02)	-0.07*** (0.02)
Low qualification household		-0.16** (0.07)	-0.08 (0.07)
English second language		-0.13* (0.07)	-0.15** (0.07)
Teen mother		-0.17** (0.08)	-0.04 (0.08)
Number of children at birth		-0.22*** (0.04)	-0.17*** (0.05)
Number of children at birth, squared		0.02*** (0.01)	0.02** (0.01)
IMD deciles		0.04*** (0.01)	0.03*** (0.01)
Low birth weight (< 2500 g)		-0.12** (0.05)	-0.08 (0.05)
Premature birth		0.01 (0.03)	0.01 (0.03)
Unplanned pregnancy		0.01 (0.03)	0.03 (0.03)
Depressed parent at 9 months		-0.14* (0.07)	-0.11 (0.07)
Lone parent household at wave 1		-0.12* (0.06)	0.02 (0.06)
Grandfather of child not in work when parent aged 14		-0.13** (0.06)	-0.09 (0.06)
At least one carer has severe mental distress (wave 4)		-0.20** (0.09)	-0.16* (0.09)
At least one carer has limiting health condition (wave 4)		-0.01 (0.04)	0.00 (0.04)
At least one carer has low life satisfaction (wave 4)		-0.02 (0.03)	-0.01 (0.03)
Always lone parent (base=always couple)		-0.02 (0.10)	-0.05 (0.10)
At least one transition (base=always couple)		-0.04 (0.04)	-0.00 (0.04)

Continued on the next page...

Improving Lives: Helping Workless Families - Supporting Methodology Paper

	(Model 1)	(Model 2)	(Model 3)
<i>Table 5 continued...</i>			
Equivalised Disposable Income Quintiles (base = bottom quintile)			0.26*** (0.05)
3			0.40*** (0.06)
4			0.52*** (0.06)
Top quintile			0.65*** (0.06)
Constant	-2.09 (11.66)	-19.77 (13.86)	-19.08 (13.31)
<i>Observations</i>	10,399	7,605	7,605
<i>R-squared</i>	0.05	0.08	0.10

Data source: Millennium Cohort Study. Robust standard errors, clustered by survey design cluster, in parentheses. All models also adjust for ethnicity of main carer, survey stratum, month-year of child's birth, quadratic in child's age in months at survey interview, country of birth, and are weighted using the wave 5 survey weights. *** p<0.01, ** p<0.05, * p<0.1

Table 6: Association of household work pattern with child cognitive ability at age 11 years: wave 5 unless otherwise noted. (Millennium Cohort Study)

VARIABLES	(Model 1)	(Model 2)	(Model 3)
Episodic worklessness (base=never workless)	-0.29*** (0.04)	-0.16*** (0.05)	-0.05 (0.05)
Persistent worklessness (base=never workless)	-0.48*** (0.04)	-0.27*** (0.06)	-0.09 (0.06)
Sex of child (male = 0)	-0.08*** (0.02)	-0.07*** (0.02)	-0.07*** (0.02)
Low qualification household		-0.15** (0.07)	-0.07 (0.07)
English second language		-0.12* (0.07)	-0.14** (0.07)
Teen mother		-0.13 (0.08)	-0.03 (0.08)
Number of children at birth		-0.22*** (0.04)	-0.17*** (0.05)
Number of children at birth, squared		0.02*** (0.01)	0.02** (0.01)
IMD		0.04*** (0.01)	0.03*** (0.01)
Low birth weight (< 2500 g)		-0.12** (0.05)	-0.08 (0.05)
Premature birth		0.01 (0.03)	0.01 (0.03)
Unplanned pregnancy		0.01 (0.03)	0.03 (0.03)
Depressed parent at 9 months		-0.12* (0.07)	-0.10 (0.07)
Lone parent household at wave 1		-0.05 (0.06)	0.04 (0.06)
Grandfather of child not in work when parent aged 14		-0.12** (0.06)	-0.08 (0.06)
At least one carer has severe mental distress (wave 4)		-0.20** (0.09)	-0.16* (0.09)
At least one carer has limiting health condition (wave 4)		-0.00 (0.04)	0.01 (0.04)
At least one carer has low life satisfaction (wave 4)		-0.02 (0.03)	-0.01 (0.03)
Always lone parent (base=always couple)		0.01 (0.10)	-0.03 (0.10)
At least one transition (base=always couple)		-0.01 (0.04)	0.01 (0.04)

Continued on the next page...

Improving Lives: Helping Workless Families - Supporting Methodology Paper

	(Model 1)	(Model 2)	(Model 3)
<i>Table 6 continued...</i>			
Equivalised Disposable Income Quintiles (base = bottom quintile)			0.25*** (0.06)
3			0.39*** (0.06)
4			0.50*** (0.06)
Top quintile			0.63*** (0.07)
Constant	-0.94 (11.49)	-19.95 (13.96)	-19.03 (13.46)
<i>Observations</i>	10,399	7,605	7,605
<i>R-squared</i>	0.06	0.08	0.10

Data source: Millennium Cohort Study. Robust standard errors, clustered by survey design cluster, in parentheses. All models also adjust for ethnicity of main carer, survey stratum, month-year of child's birth, quadratic in child's age in months at survey interview, country of birth, and are weighted using the wave 5 survey weights. *** p<0.01, ** p<0.05, * p<0.1

Table 7: Association of household work status with child externalising problems at age 11 years: wave 5 unless otherwise noted. (Millennium Cohort Study)

VARIABLES	(Model 1)	(Model 2)	(Model 3)
Workless household	0.55*** (0.05)	0.17*** (0.06)	0.06 (0.06)
Sex of child (male = 0)	-0.36*** (0.02)	-0.34*** (0.03)	-0.34*** (0.02)
Low qualification household		0.22*** (0.07)	0.16** (0.07)
English second language		-0.14** (0.06)	-0.12* (0.06)
Teen mother		0.04 (0.09)	-0.06 (0.09)
Number of children at birth		-0.06 (0.07)	-0.10 (0.07)
Number of children at birth, squared		0.02 (0.02)	0.02 (0.02)
IMD		-0.03*** (0.01)	-0.02*** (0.01)
Low birth weight (< 2500 g)		0.20*** (0.07)	0.17** (0.07)
Premature birth		0.05** (0.02)	0.05** (0.02)
Unplanned pregnancy		0.06** (0.03)	0.04 (0.03)
Depressed parent at 9 months		0.24** (0.09)	0.22** (0.09)
Lone parent household at wave 1		0.10 (0.07)	-0.00 (0.07)
Grandfather of child not in work when parent aged 14		0.11* (0.06)	0.07 (0.06)
At least one carer has severe mental distress (wave 4)		0.51*** (0.09)	0.48*** (0.09)
At least one carer has limiting health condition (wave 4)		0.14*** (0.04)	0.13*** (0.04)
At least one carer has low life satisfaction (wave 4)		0.20*** (0.03)	0.19*** (0.03)
Always lone parent (base=always couple)		0.14 (0.11)	0.15 (0.11)
At least one transition (base=always couple)		0.15*** (0.04)	0.12*** (0.04)

Continued on the next page...

Improving Lives: Helping Workless Families - Supporting Methodology Paper

	(Model 1)	(Model 2)	(Model 3)
<i>Table 7 continued...</i>			
Equivalised Disposable Income Quintile. 2 (base = bottom quintile)			-0.09 (0.06)
3			-0.26*** (0.06)
4			-0.38*** (0.07)
Top quintile			-0.43*** (0.07)
Constant	8.19 (11.76)	-2.99 (15.90)	-4.94 (15.71)
<i>Observations</i>	10,159	7,633	7,633
<i>R-squared</i>	0.08	0.14	0.16

Data source: Millennium Cohort Study Robust standard errors, clustered by survey design cluster, in parentheses. All models also adjust for ethnicity of main carer, survey stratum, month-year of child's birth, quadratic in child's age in months at survey interview, country of birth, and are weighted using the wave 5 survey weights. *** p<0.01, ** p<0.05, * p<0.1

Table 8: Association of persistence of household work status with child externalising problems at age 11 years: wave 5 unless otherwise noted. (Millennium Cohort Study)

VARIABLES	(Model 1)	(Model 2)	(Model 3)
Episodic worklessness (base=never workless)	0.39*** (0.04)	0.18*** (0.05)	0.08 (0.05)
Persistent worklessness (base=never workless)	0.63*** (0.05)	0.28*** (0.07)	0.15* (0.08)
Sex of child (male = 0)	-0.36*** (0.02)	-0.34*** (0.03)	-0.34*** (0.03)
Low qualification household		0.21*** (0.07)	0.15** (0.07)
English second language		-0.12 (0.08)	-0.11 (0.08)
Teen mother		-0.01 (0.10)	-0.07 (0.10)
Number of children at birth		-0.06 (0.07)	-0.10 (0.08)
Number of children at birth, squared		0.01 (0.02)	0.02 (0.02)
IMD		-0.03*** (0.01)	-0.02** (0.01)
Low birth weight (< 2500 g)		0.15** (0.07)	0.11 (0.07)
Premature birth		0.04 (0.03)	0.04 (0.03)
Unplanned pregnancy		0.04 (0.03)	0.02 (0.03)
Depressed parent at 9 months		0.25** (0.10)	0.23** (0.10)
Lone parent household at wave 1		0.04 (0.08)	-0.03 (0.08)
Grandfather of child not in work when parent aged 14		0.11* (0.06)	0.08 (0.06)
At least one carer has severe mental distress (wave 4)		0.34*** (0.08)	0.31*** (0.08)
At least one carer has limiting health condition (wave 4)		0.06* (0.03)	0.06* (0.03)
At least one carer has low life satisfaction (wave 4)		0.19*** (0.03)	0.18*** (0.03)
Always lone parent (base=always couple)		0.06 (0.11)	0.08 (0.11)
At least one transition (base=always couple)		0.09** (0.04)	0.07* (0.04)

Continued on the next page...

Improving Lives: Helping Workless Families - Supporting Methodology Paper

	(Model 1)	(Model 2)	(Model 3)
<i>Table 8 continued...</i>			
Equivalised Disposable Income 2 (base = bottom quintile)			-0.09 (0.07)
3			-0.24*** (0.06)
4			-0.38*** (0.07)
Top quintile			-0.44*** (0.07)
Constant			-7.94 (12.04)
<i>Observations</i>	10,159	7,558	7,558
<i>R-squared</i>	0.10	0.13	0.14

Data source: Millennium Cohort Study Robust standard errors, clustered by survey design cluster, in parentheses. All models also adjust for ethnicity of main carer, survey stratum, month-year of child's birth, quadratic in child's age in months at survey interview, country of birth, and are weighted using the wave 5 survey weights. *** p<0.01, ** p<0.05, * p<0.1

Table 9: Association of household work status with child internalising problems at age 11 years: wave 5 unless otherwise noted. (Millennium Cohort Study)

VARIABLES	(Model 1)	(Model 2)	(Model 3)
Workless household	0.48*** (0.05)	0.14** (0.06)	0.08 (0.07)
Sex of child (male = 0)	-0.02 (0.03)	-0.01 (0.03)	-0.01 (0.03)
Low qualification household		0.16** (0.07)	0.13* (0.07)
English second language		-0.04 (0.05)	-0.04 (0.05)
Teen mother		-0.14* (0.08)	-0.18** (0.08)
Number of children at birth		-0.28*** (0.09)	-0.29*** (0.09)
Number of children at birth, squared		0.05** (0.02)	0.05** (0.02)
IMD		-0.03*** (0.01)	-0.02*** (0.01)
Low birth weight (< 2500 g)		0.12* (0.06)	0.10 (0.06)
Premature birth		0.04 (0.03)	0.04 (0.03)
Unplanned pregnancy		0.07*** (0.03)	0.06** (0.03)
Depressed parent at 9 months		0.25*** (0.09)	0.24*** (0.09)
Lone parent household at wave 1		0.00 (0.07)	-0.04 (0.07)
Grandfather of child not in work when parent aged 14		0.04 (0.06)	0.03 (0.06)
At least one carer has severe mental distress (wave 4)		0.56*** (0.11)	0.55*** (0.11)
At least one carer has limiting health condition (wave 4)		0.25*** (0.04)	0.24*** (0.04)
At least one carer has low life satisfaction (wave 4)		0.24*** (0.03)	0.24*** (0.03)
Always lone parent (base=always couple)		0.22* (0.13)	0.22* (0.13)
At least one transition (base=always couple)		0.13*** (0.04)	0.11*** (0.04)

Continued on the next page...

Improving Lives: Helping Workless Families - Supporting Methodology Paper

	(Model 1)	(Model 2)	(Model 3)
<i>Table 9 continued...</i>			
Equivalised Disposable Income Quintiles 2 (base = bottom quintile)			-0.01 (0.07)
3			-0.15** (0.07)
4			-0.13* (0.08)
Top quintile			-0.19** (0.08)
Constant	12.28 (11.21)	2.99 (13.36)	1.61 (13.30)
<i>Observations</i>	10,172	7,637	7,637
<i>R-squared</i>	0.04	0.11	0.12

Data source: Millennium Cohort Study Robust standard errors, clustered by survey design cluster, in parentheses. All models also adjust for ethnicity of main carer, survey stratum, month-year of child's birth, quadratic in child's age in months at survey interview, country of birth, and are weighted using the wave 5 survey weights. *** p<0.01, ** p<0.05, * p<0.1

Table 10: Association of persistence of household work status with child internalising problems at age 11 years: wave 5 unless otherwise noted. (Millennium Cohort Study)

VARIABLES	(Model 1)	(Model 2)	(Model 3)
Episodic worklessness (base=never workless)	0.33*** (0.04)	0.14*** (0.05)	0.10* (0.05)
Persistent worklessness (base=never workless)	0.56*** (0.05)	0.21*** (0.07)	0.16** (0.07)
Sex of child (male = 0)	-0.02 (0.02)	-0.01 (0.03)	-0.01 (0.03)
Low qualification household		0.15** (0.07)	0.13* (0.07)
English second language		-0.05 (0.05)	-0.04 (0.05)
Teen mother		-0.18** (0.08)	-0.20** (0.08)
Number of children at birth		-0.28*** (0.09)	-0.29*** (0.09)
Number of children at birth, squared		0.05** (0.02)	0.05** (0.02)
IMD		-0.03*** (0.01)	-0.02*** (0.01)
Low birth weight (< 2500 g)		0.12* (0.06)	0.10 (0.06)
Premature birth		0.04 (0.03)	0.04 (0.03)
Unplanned pregnancy		0.07*** (0.03)	0.06** (0.03)
Depressed parent at 9 months		0.24*** (0.09)	0.24** (0.09)
Lone parent household at wave 1		-0.05 (0.07)	-0.07 (0.07)
Grandfather of child not in work when parent aged 14		0.03 (0.06)	0.02 (0.06)
At least one carer has severe mental distress (wave 4)		0.55*** (0.11)	0.54*** (0.11)
At least one carer has limiting health condition (wave 4)		0.24*** (0.04)	0.24*** (0.04)
At least one carer has low life satisfaction (wave 4)		0.24*** (0.03)	0.24*** (0.03)
Always lone parent (base=always couple)		0.20 (0.13)	0.20 (0.13)
At least one transition (base=always couple)		0.10** (0.04)	0.09** (0.04)

Continued on the next page...

Improving Lives: Helping Workless Families - Supporting Methodology Paper

	(Model 1)	(Model 2)	(Model 3)
<i>Table 10 continued...</i>			
Equivalised Disposable Income Quintiles 2 (base = bottom quintile)			0.01 (0.07)
3			-0.12 (0.07)
4			-0.10 (0.08)
Top quintile			-0.15* (0.08)
Constant	11.54 (11.30)	2.78 (13.27)	1.55 (13.26)
<i>Observations</i>	10,172	7,637	7,637
<i>R-squared</i>	0.05	0.11	0.12

Data source: Millennium Cohort Study Robust standard errors, clustered by survey design cluster, in parentheses. All models also adjust for ethnicity of main carer, survey stratum, month-year of child's birth, quadratic in child's age in months at survey interview, country of birth, and are weighted using the wave 5 survey weights. *** p<0.01, ** p<0.05, * p<0.1

Table 11: Association of household work status with child pro-social behaviours at age 11 years: wave 5 unless otherwise noted. (Millennium Cohort Study)

VARIABLES	(Model 1)	(Model 2)	(Model 3)
Workless household	-0.29*** (0.06)	-0.19** (0.08)	-0.13* (0.07)
Sex of child (male = 0)	0.37*** (0.02)	0.36*** (0.03)	0.36*** (0.03)
Low qualification household		-0.14 (0.10)	-0.11 (0.09)
English second language		-0.13 (0.08)	-0.13* (0.08)
Teen mother		0.11 (0.10)	0.16 (0.10)
Number of children at birth		0.26** (0.12)	0.27** (0.12)
Number of children at birth, squared		-0.05* (0.03)	-0.05* (0.03)
IMD		0.01 (0.01)	0.00 (0.01)
Low birth weight (< 2500 g)		0.01 (0.07)	0.03 (0.07)
Premature birth		-0.04 (0.03)	-0.04 (0.03)
Unplanned pregnancy		0.02 (0.03)	0.03 (0.03)
Depressed parent at 9 months		-0.05 (0.08)	-0.04 (0.08)
Lone parent household at wave 1		-0.08 (0.06)	-0.02 (0.07)
Grandfather of child not in work when parent aged 14		0.04 (0.07)	0.05 (0.07)
At least one carer has severe mental distress (wave 4)		-0.15 (0.10)	-0.13 (0.10)
At least one carer has limiting health condition (wave 4)		0.02 (0.04)	0.03 (0.04)
At least one carer has low life satisfaction (wave 4)		-0.12*** (0.03)	-0.12*** (0.03)
Always lone parent (base=always couple)		-0.20 (0.14)	-0.21 (0.13)
At least one transition (base=always couple)		-0.06 (0.04)	-0.04 (0.04)

Continued on the next page...

Improving Lives: Helping Workless Families - Supporting Methodology Paper

	(Model 1)	(Model 2)	(Model 3)
<i>Table 11 continued...</i>			
Equivalised Disposable Income Quintiles 2 (base = bottom quintile)			0.08 (0.07)
3			0.17** (0.07)
4			0.23*** (0.07)
Top quintile			0.20** (0.08)
Constant	-29.20** (13.11)	-18.86 (15.19)	-17.59 (15.14)
<i>Observations</i>	10,186	7,644	7,644
<i>R-squared</i>	0.05	0.07	0.07

Data source: Millennium Cohort Study Robust standard errors, clustered by survey design cluster, in parentheses. All models also adjust for ethnicity of main carer, survey stratum, month-year of child's birth, quadratic in child's age in months at survey interview, country of birth, and are weighted using the wave 5 survey weights. *** p<0.01, ** p<0.05, * p<0.1

Table 12: Association of persistence of household work status with child pro-social behaviours at age 11 years, wave 5 unless otherwise noted. (Millennium Cohort Study)

VARIABLES	(Model 1)	(Model 2)	(Model 3)
Episodic worklessness (base=never workless)	-0.15*** (0.04)	-0.04 (0.05)	0.00 (0.05)
Persistent worklessness (base=never workless)	-0.37*** (0.06)	-0.31*** (0.10)	-0.24** (0.10)
Sex of child (male = 0)	0.37*** (0.02)	0.36*** (0.03)	0.36*** (0.03)
Low qualification household		-0.12 (0.10)	-0.10 (0.09)
English second language		-0.12 (0.08)	-0.13 (0.08)
Teen mother		0.15 (0.10)	0.18* (0.10)
Number of children at birth		0.26** (0.12)	0.27** (0.12)
Number of children at birth, squared		-0.04 (0.03)	-0.04 (0.03)
IMD		0.01 (0.01)	0.00 (0.01)
Low birth weight (< 2500 g)		0.01 (0.07)	0.02 (0.07)
Premature birth		-0.04 (0.03)	-0.04 (0.03)
Unplanned pregnancy		0.02 (0.03)	0.02 (0.03)
Depressed parent at 9 months		-0.03 (0.08)	-0.03 (0.08)
Lone parent household at wave 1		-0.02 (0.07)	0.01 (0.07)
Grandfather of child not in work when parent aged 14		0.05 (0.07)	0.06 (0.07)
At least one carer has severe mental distress (wave 4)		-0.14 (0.10)	-0.14 (0.10)
At least one carer has limiting health condition (wave 4)		0.03 (0.04)	0.03 (0.04)
At least one carer has low life satisfaction (wave 4)		-0.13*** (0.03)	-0.12*** (0.03)
Always lone parent (base=always couple)		-0.16 (0.13)	-0.17 (0.13)
At least one transition (base=always couple)		-0.05 (0.04)	-0.04 (0.04)

Continued on the next page...

Improving Lives: Helping Workless Families - Supporting Methodology Paper

	(Model 1)	(Model 2)	(Model 3)
<i>Table 11 continued...</i>			
Equivalised Disposable Income Quintiles 2 (base = bottom quintile)			0.06 (0.07)
3			0.14* (0.07)
4			0.20*** (0.08)
Top quintile			0.17** (0.08)
Constant	-28.61** (13.10)	-19.08 (15.23)	-17.83 (15.17)
<i>Observations</i>	10,186	7,644	7,644
<i>R-squared</i>	0.05	0.07	0.07

Data source: Millennium Cohort Study Robust standard errors, clustered by survey design cluster, in parentheses. All models also adjust for ethnicity of main carer, survey stratum, month-year of child's birth, quadratic in child's age in months at survey interview, country of birth, and are weighted using the wave 5 survey weights. *** p<0.01, ** p<0.05, * p<0.1

2) Indicator two: parental conflict

Overview of section

This section of the Analysis and Research Pack presents the new parental conflict measures: the proportion of children living in couple-parent families reporting relationship distress, and the proportion of children in separated families who see their non-resident parent regularly. These indicators are supplemented by our latest contextual analysis. In particular, we conducted regression analyses to explore the degree of association between different forms of disadvantage and the relationship distress indicator and how the latter affects parental separation and child conduct problems.

Data sources

The analysis in this section is based on the Understanding Society survey. A full description of this data source is available later in this document.

Underlying sample population

The underlying sample for this section was children who were present in all waves of the survey. Children could only join or leave the sample in subsequent waves if they were born into the sample, or they stopped being a dependent child. For basic descriptive statistics, all children were included in comparisons. For transitions statistics, children who were present in the study for 2 or more consecutive waves were included (unless otherwise specified).

Some of the analysis in the parental conflict section incorporates the use of multiple waves of data to produce charts by age of child, namely:

- Proportion of children in couple-parent families experiencing relationship distress by age of child;
- Proportion of dependent children by family type and age of child;
- Proportion of children experiencing parental separation, by age of the child; and
- Proportion of children living with both of their birth parents, by age of child.

In these charts, the proportions have been calculated by averaging the proportions calculated in 'snapshots' of each wave. Since each snapshot has a similar sample size, this methodology produces very similar results to pooling the data across the waves and calculating proportions.

Weights used and attrition

This was the same as used in Part one: children in workless families. See page 3.

Regression analysis of relationship distress, parental separation and child outcomes

Methodology

The purpose of this analysis was to enable a better understanding of the degree of association between different forms of disadvantage covered in the *Helping Workless Families* paper and relationship distress, and how the latter affects separation and child conduct problems. It confirmed that the elements of relationship distress captured by the new relationship distress measure were indeed associated with separation and child conduct problems even after we controlled for other important forms of disadvantage, family characteristics and demographics which existing research suggests have a role in explaining these outcomes. The various forms of disadvantage accounted for in this analysis are interrelated with relationship distress, making the impact of the latter on separation and child outcomes difficult to disentangle.

Three logistic regression models were considered:-

1. The first model explored the likelihood of children in couple-parent families experiencing relationship distress, pooling data from waves 1, 3 and 5 and controlling for a range of parental characteristics in the survey in addition to a range of parental disadvantages.
2. The second model explored the likelihood of a child experiencing parental separation by pooling the transitions from waves 1 to 2, 3 to 4 and 5 to 6 and running a logistic regression in the same manner, including relationship distress as an independent variable.
3. The third model considered children aged 10-15 who appear in the Youth Questionnaire in waves 1, 3 and 5, modelling the likelihood of a child having conduct problems, defined as having a high Strengths and Difficulties Questionnaire (SDQ) sub-score on 'conduct problems'. We focussed on conduct problems because research suggests that this is one of the main areas of child development affected by exposure to parental conflict.

In all three models, if either parent had unknown information recorded against these disadvantages, then the overall status of the family/child was unknown, and they were dropped from the regression. This is because, as opposed to our regressions on worklessness, this analysis focusses on a lot of information that had high levels of non-response. Also, because this analysis focuses on couples, the sample cannot be biased towards lone parents.

Results from regression analysis

Table 13: Association between parental characteristics and relationship distress (Model 1, Understanding Society survey)

VARIABLES	Coefficients	Odds ratio	Additional likelihood
	RELATIONSHIP DISTRESS		
At least one parent with poor mental health	0.7052*** (0.069)	2.024	102%
At least one parent has a longstanding limiting illness and/or disability	0.3514*** (0.079)	1.421	42%
Household reports signs of problem debt	0.3505*** (0.087)	1.42	42%
Parent has management or professional occupation (SEC1), (base= workless)	-0.7136*** (0.128)	0.49	-51%
Parent has intermediate occupation (SEC2), (base= workless)	-0.4881*** (0.131)	0.639	-36%
Parent has routine occupation (SEC3), (base= workless)	-0.0548 (0.121)	0.947	-5%
Log of equivalised household income	-0.3302*** (0.089)	0.719	-28%
Parents are married	-0.4923*** (0.095)	0.611	-39%
At least one parent is from an ethnic minority group	0.1871* (0.099)	1.206	21%
Difference between parents' qualification levels	0.2492** (0.099)	1.283	28%
Age of youngest parent is under 25 years	-0.138 (0.168)	0.871	-13%
Living with both birth parents	0.2477** (0.122)	1.281	28%
Young child (5 years and under) in family	-0.0114 (0.075)	0.989	-1%
Lives in a large family (3 or more children)	0.1973** (0.079)	1.218	22%
<i>Observations</i>			

Data source: Understanding Society, waves 1, 3 and 5. All errors were clustered on the personal identifier. The model controls for time period. *** p<0.01, ** p<0.05, * p<0.1 Standard Errors are in parentheses. The parental occupational status is determined to be the highest of either parent.

Table 14: Association between parental characteristics and parental separation (Model 2, Understanding Society survey)

VARIABLES	Coefficients	Odds ratio	Additional likelihood
	PARENTAL SEPARATION		
Relationship distress	1.5766*** (0.156)	4.839	384%
At least one parent with poor mental health	0.3458** (0.144)	1.413	41%
At least one parent has a long-standing and limiting illness and/or disability	-0.1329 (0.162)	0.876	-12%
Household reports signs of problem debt	0.627*** (0.157)	1.872	87%
Parent has management or professional occupation (SEC1), (base= workless)	-0.5414** (0.239)	0.582	-42%
Parent has intermediate occupation (SEC2), (base= workless)	-0.0543 (0.255)	0.947	-5%
Parent has routine occupation (SEC3), (base= workless)	-0.411 (0.2332)	0.663	-34%
Log of equivalised household income	0.089 (0.172)	1.093	9%
Parents are married	-0.4042** (0.175)	0.668	-33%
At least one parent is from an ethnic minority group	0.3545* (0.198)	1.425	43%
Difference between parents' qualification levels	-0.3294 (0.202)	0.719	-28%
Age of youngest parent is under 25 years	0.9299*** (0.244)	2.534	153%
Living with both birth parents	-1.7359*** (0.168)	0.176	-82%
Young child (5 years and under) in family	0.2766* (0.152)	1.319	32%
Lives in a large family (3 or more children)	-0.1743 (0.163)	0.84	-16%
<i>Observations</i>	12,963		

Data source: Understanding Society; the model considers transitions from waves 1 to 2, 3 to 4 and 5 to 6. Logistic regressions are implemented. All errors were clustered on the personal identifier. The model control for time period. *** p<0.01, ** p<0.05, * p<0.1. Standard Errors are in parentheses. The parental occupational status of the child is determined to be the highest of either parent.

Table 15: Association between parental characteristics and conduct problems in couple-parent children, (Model 3, Understanding Society survey)

Improving Lives: Helping Workless Families - Supporting Methodology Paper

VARIABLES	Coefficients	Odds ratio	Additional likelihood
	CONDUCT PROBLEMS		
Relationship distress	0.3991*** (0.152)	1.491	49%
At least one parent has poor parental mental health	0.2818*** (0.1)	1.326	33%
At least one parent has a long-standing and limiting illness and/or disability	0.0698 (0.110)	1.072	7%
Household reports signs of problem debt	0.3943*** (0.130)	1.483	48%
Parent has management and/or professional occupation (SEC1), (base= workless)	-0.2033 (0.179)	0.816	-18%
Parent has intermediate occupation (SEC2), (base= workless)	-0.1132 (0.194)	0.875	-13%
Parent has routine occupation (SEC3), (base= workless)	0.133 (0.191)	1.142	14%
Log of equivalised household income	-0.0229 (0.142)	0.977	-2%
Parents are married	0.0965 (0.156)	1.101	10%
At least one parent is from an ethnic minority group	0.1778 (0.147)	1.195	20%
Age of child	-0.0574 (0.028)	0.944	-6%
Sex of child	-0.4854*** -0.100	0.615	-39%
Difference between parents' qualification levels	-0.0105 (0.133)	0.99	-1%
Age of youngest parent is under 25 years	1.6019** (0.761)	4.962	396%
Lives with both birth parents	-0.6836 (0.143)	0.505	-50%
Family has a young child (5 years and under)	0.00955 (0.147)	1.01	1%
Lives in a large family (3 or more children)	0.3339*** (0.112)	1.396	40%
<i>Observations</i>	3,868		

Improving Lives: Helping Workless Families - Supporting Methodology Paper

Data: Understanding Society survey. The model considers children aged 10-15 who appear in the Youth Questionnaire in waves 1, 3 & 5, modelling the likelihood of a child having conduct problems, defined as having a high strengths and Difficulties Questionnaire sub-score on 'conduct problems'. Logistic regressions are implemented. All errors were clustered on the personal identifier. The model controls for time period. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Standard Errors are in parentheses. The parental occupational status is determined to be the highest of either parent.

3) Indicator three: poor parental mental health

Overview of section

This section presents the new poor parental mental health measure and supporting contextual analysis. This is found in the corresponding section of the Analysis and Research Pack (see pages 70 to 78 of the Analysis and Research Pack).

Data sources

The analysis in this section is based on the Understanding Society survey. A full description of this data source is available in 'Description of the main data sources used' section.

Underlying sample population

For the poor parental mental health measure, children were included in our sample if they were present in any of the five most recent waves (2010-2011 to 2014-2015). We have included any child where at least one parent has responded fully to the 12-item General Health Questionnaire on which the measure is based. For over one in five of the children in our sample, information is missing for one of the parents. We have assumed that these individuals are not reporting poor mental health, the result being that the measure may underestimate the overall level slightly, by around three per cent (assuming unknown individuals report symptoms at the same rate as the known population).

There are a variety of possible ways of constructing and weighting this indicator. However, we found that the direction and relative size of year-to-year changes were similar regardless of underlying construction.

- We could either construct this indicator from the perspective of the parents – 'the proportion of parents reporting symptoms of anxiety and/or depression' – or the perspective of the child – 'the proportion of children living with at least one parent reporting symptoms of anxiety and/or depression'. We chose the latter, because we are primarily interested in how many children are affected by poor parental mental health, whereas a parents-based indicator would count every parent, even if the same child was affected twice. Whilst around 1 in 4 children live with at least one parent reporting symptoms of anxiety and/or depression, around 1 in 5 parents report these symptoms. Trends are very similar for both measures.
- As specified above, for over one in five of the children in our sample, information is missing for one of the parents. We could either only include children where both parents provide responses to GHQ-12, or include children where at least one parent responds. Either comes with a bias; excluding any children with at least one unknown parent biases the sample towards lone parents (since it is easier to be included if you only require the response of one parent). Lone parents are more likely to report symptoms of anxiety and/or depression, and so this construction would over estimate overall rates of poor parental mental health. Alternatively, including all children where at least one parent is known is likely to underreport rates, since we are assuming that the 'unknown' parent does not

have poor parental mental health (when we would expect around 1 in 5 to be affected). We have found there to be around four percentage points difference between the two measures, although trends are very similar for both. In conclusion, we have decided to base the measure on children where at least one parent is known. The only children who could be incorrectly classified by this construction are those where one parent answers, but does not report poor mental health, and the other parent has not answered. We would classify these children as *not* living with a parent with poor mental health. If it were the case that the parent who did not answer did have poor mental health, then the family overall would be incorrectly classified. However given that we know the answer for one parent who does not have poor mental health, we know (from analysis) that the likelihood of the other parent reporting poor mental health is reduced. Therefore, the overall bias of this construction is likely to be smaller.

When interpreting any future year-to-year changes in this indicator, different possible constructions of this indicator should be examined to check that trends are robust to any such changes in methodology.

Weights used and attrition

As this analysis was conducted to construct an indicator, and therefore to monitor changes in the prevalence of this issue, we used cross-sectional weights designed to make the results as representative of the UK population as possible.

There are self-completion weights available that adjust for non-response to self-completion questions (the GHQ-12 is part of the self-completion questionnaire in Understanding Society survey). However, since our indicator is constructed from the perspective of the child, we could not use these weights. When we tried constructing the indicator from the perspective of the parent (see above) we did not find a significant difference between results weighted using either the cross-sectional or self-completion weights.

4) Targeted provision to reduce parental conflict

Overview of section

We have used our new Parental Conflict measures (see page 51 of the Analysis and Research Pack) and analysis of workless families (pages 7 to 39), to support development of the targeted provision to reduce parental conflict. This provision aims to help parents across England to reduce the conflict in their relationships with each other, and secure better outcomes for their children. See the Improving Lives: Helping Workless Families document⁴ for further details.

In particular, we have produced a broad estimate of the number of families who are eligible for this provision, across five broad geographic regions in England, referred to as Contract Package Areas. Parents will be eligible for the provision if:-

- their youngest child is 18 or below,
- they are reporting experiencing relationship distress; and
- they are part of a long-term workless family.

This reflects the evidence showing how workless families are disproportionately likely to experience multiple disadvantages, such as relationship distress, that hold them back and impact on child outcomes.

The following table shows our estimates of the numbers of families eligible for this provision.

Table 16: estimates of the number of families eligible for targeted provision

Contract Package Area	Estimates of the numbers of families eligible for targeted provision
	No.
North East	50,000
Central	80,000
South of England	60,000
North West	40,000
London and Home Counties	70,000
England	290,000

Data source: Understanding Society survey (2013-2014), and HMRC child benefit data (2014).

How we derived these estimates

The estimates presented in table 16 are derived from a combination of internal analyses of Understanding Society survey data and published HMRC data on the number of dependent children⁵.

⁴ <https://www.gov.uk/government/publications/improving-lives-helping-workless-families>

⁵ <https://www.gov.uk/government/statistics/personal-tax-credits-children-in-low-income-families-local-measure-2014-snapshot-as-at-31-august-2014-30-september-2016>

Using Understanding Society survey data, we:

- I. Estimated the proportion of children in long-term workless families by Government Office Region, and how this splits between lone-parent and couple-parent families;
- II. Estimated the proportion of children in workless families reporting relationship distress;
- III. Combined II and III to estimate the proportion of children who are in long-term workless families and have reported forms of relationship distress.
- IV. We combined these estimates with HMRC data to estimate the number of dependent children in scope for the contract by Government Office Region. These are combined to arrive at estimates by Contract Package Area.
- V. We divided these estimates by the average number of children per family (according to ONS estimates) to calculate the number of families in scope for the contract by CPA.

Our definitions for this analysis were as follows:

- *A family* is defined as a married, civil partnered or cohabiting couple with or without dependent children, or a lone parent with at least one dependent child.
- *A long-term workless family* is defined as a family where no resident adult was in paid employment for 2 consecutive survey interviews.
- *Dependent children* are defined as children aged between birth to 16 years of age or 17 and/or 18 year olds who are in full-time education and live at home.
- *Relationship distress* was defined as either:
 - (i) couple-parent families where at least one parent gave very negative answers to questions about the quality of the relationship with their partner; or,
 - (ii) separated families where (according to the resident parent) the child does not see the non-resident parent regularly. Existing research and analysis presented in the Analysis and Research Pack show that these families are less likely to have good-quality co-parenting relationships.

See the Parental Conflict section of the Analysis and Research Pack for further information on these new measures. See the opening section (Part one) of the Analysis and Research Pack for our analysis of workless families.

Underlying sample population and weights used

The underlying sample for this section, and weights applied, was the same as for the parental conflict analysis (see section 2 of this document). Estimates were based on 2013-2014 data.

Description of the main data sources used

The Understanding Society survey

Overview of the survey

Understanding Society is a nationwide household survey, following 41,000 households across the UK from 2009-2010 onwards. It captures important information about people’s social and economic circumstances, attitudes, behaviours and their health. The study provides a rich range of information on families and their circumstances over time, enabling a longitudinal picture to be built on disadvantage and worklessness.

Household members aged 16 or older are interviewed and the same individuals are re-interviewed in successive years to see how things have changed. Individuals become eligible for a full interview once they reach the age of 16. A sub-set of 10-15 year old children are also interviewed in the ‘youth survey’. However, questions differ from the main survey.

Full questionnaires of surveys currently being implemented can be found at: <https://www.understandingsociety.ac.uk/documentation/mainstage/questionnaires>

Frequency of data collection

Understanding Society fieldwork is conducted over a two-calendar-year period, with each individual being interviewed on a yearly basis. It’s important to note, however that the periods of waves overlap, and that individual respondents are interviewed around the same time each year on an annual basis. Table 1, below, shows the periods of data included in each wave.

Wave	Year	Calendar Year						
		2009	2010	2011	2012	2013	2014	2015
1	2009-2010	█	█					
2	2010-2011		█	█				
3	2011-2012			█	█			
4	2012-2013				█	█		
5	2013-2014					█	█	
6	2014-2015						█	█

Sample sizes and attrition

As with most longitudinal surveys, attrition reduces the Understanding Society sample size over time. Table 2, below, shows how many cases are available for analysis in each wave.

Original sample sizes available for analysis

Waves	Survey cases
2009-2010	76,185
2010-2011	76,407
2011-2012	60,428
2012-2013	53,486
2013-2014	48,611
2014-2015	42,693

*Not included in our analysis- see below for a full explanation.

Data: Understanding Society survey, 2010-2015

Wave 1 (2009-2010) income information

There are known issues with the income information in the first Understanding Society survey wave covering 2009-2010. See Dr Paul Fisher's paper [Does repeated measurement improve income data quality?](#) (ISER Working Paper Series, 2016-11) for details of why income data on the first wave of Understanding Society are not comparable with subsequent waves and are likely to be of lower quality.

We have therefore excluded the first wave from most analysis presented in this publication – with the exception of parental separations between the first and second wave, as we needed to increase our underlying sample size, and felt that the first wave of data could be trusted in this regard.

National Pupil Database (NPD) linkage

The National Pupil Database (NPD) has recently been joined to Understanding Society data. Linkage was carried out for all young adults (born after 1981) who consented to data linkage at wave 1 and for school-age children (aged 4-15) whose parents consented in Understanding Society at wave 1 and who were living in England.

This information includes information on attainment data for the Early Years Foundation Profile Stage (age 5) and Key Stages 1, 2 and 4 (relating to ages 7, 11 and 16).

As the availability of this information is heavily restricted, we worked with academics in the Understanding Society Policy Unit at the University of Essex, who joined our derived information on disadvantage to the proportion of pupils who had not reached the expected level at the Early Years Foundation Stage (age 5), KS1 (age 7), KS2 (age 11) or KS4 (age 16) - whichever stage was most recent for that child, from 2012/13 or older.

We'd like to thank Dr Nicole Martin and Dr Ricky Kanabar of the Understanding Society Policy Unit at the University of Essex for their help and assistance with this work. In addition to deriving results on educational attainment and joining it to our existing analysis on disadvantage, they offered numerous helpful suggestions on how to conduct our analysis, and how to best use the data available on the Understanding Society survey.

The Millennium Cohort Study (MCS)

The Millennium Cohort Study is a multi-disciplinary research project following the lives of around 19,000 children born in the UK in 2000-01.

The study has been tracking the Millennium children through their early childhood years and plans to follow them into adulthood. It collects information on the children's siblings and parents. MCS's field of enquiry covers such diverse topics as parenting; childcare; school choice; child behaviour and cognitive development; child and parental health; parents' employment and education; income and poverty; housing, neighbourhood and residential mobility; and social capital and ethnicity. We have used data from the first five surveys of MCS cohort members – at age nine months, three, five, seven and eleven years.

Cohort Profile: UK Millennium Cohort Study (MCS) sample

Waves	Age of child	Families	Children
	Years	No.	No.
2010-2011	9 months	18,552	18,818
2011-2012	3	15,590	15,808
2012-2013	5	15,246	15,460
2013-2014	7	13,857	14,043
2014-2015	11	13,287	13,469

Definitions and terminology within the statistics

Understanding Society survey definitions

The following table explains how we have defined parental disadvantages and characteristics as presented in the Analysis and Research Pack.

Variable	Definition
Equivalised household income	<p>The process of equivalisation adjusts annual incomes for the household size and composition—with an adult couple with no children as a reference point. The process assigns a weight of 1 to the first adult in the household, a weight of 0.5 to each additional adult, and a weight of 0.3 to each child (person aged 0-13). Household income decile and quintile is calculated incorporating appropriate survey weights.</p>
Ethnicity	<p>Understanding Society includes the following ethnicity classification which can be combined into other analytical groupings:</p> <ul style="list-style-type: none"> • White: British/English/Scottish/Welsh/Northern Irish • White: Irish • White: Gypsy or Irish Traveller • White: Any other White background • Mixed: White and Black Caribbean • Mixed: White and Black African • Mixed: White and Asian • Mixed: Any other mixed background • Asian/Asian British: Indian • Asian/Asian British: Pakistani • Asian/Asian British: Bangladeshi • Asian/Asian British: Chinese • Asian/Asian British: Any other Asian background • Black/African/Caribbean/Black British: Caribbean • Black/African/Caribbean/Black British: African • Black/African/Caribbean/Black British: Any other Black background • Other Ethnic Group: Arab • Other Ethnic Group: Any other ethnic group <p>We have recoded these into seven ethnic groups which constitute the largest groups in the data, including:</p> <ul style="list-style-type: none"> • White • Mixed • Indian • Pakistani • Bangladeshi • Black African • Black Caribbean • Other <p>We have measured the ethnicities of the parents, because this has more of a direct impact on the employment outcomes of the parents. Where the parents are of different ethnicities they are coded as 'mixed.'</p>

Improving Lives: Helping Workless Families - Supporting Methodology Paper

<p>Long-standing limiting illness or disability</p>	<p>Understanding Society includes the following question to determine whether someone has a longstanding illness or disability:</p> <p><i>“Do you have any long-standing physical or mental impairment, illness or disability? By 'long-standing' I mean anything that has troubled you over a period of at least 12 months or that is likely to trouble you over a period of at least 12 months.”</i></p> <p>This is asked of all individuals over 16, including in proxy interviews (i.e. where someone else is giving information about the respondent). In non-proxy interviews, if a respondent says they have a long-standing illness or disability they are asked:</p> <p><i>“Does this/Do these health problem(s) or disability(ies) mean that you have substantial difficulties with any of these areas of your life? Please read out the numbers from the card next to the ones which apply to you.”</i></p> <p>12 areas are listed: -</p> <p>mobility (moving around at home and walking); lifting, carrying or moving objects; manual dexterity (using your hands to carry out everyday tasks); continence (bladder and bowel control); hearing (apart from using a standard hearing aid); sight (apart from wearing standard glasses); communication or speech problems; memory or ability to concentrate, learn or understand; recognising when you are in physical danger; your physical co-ordination (e.g. balance); difficulties with own personal care (e.g. getting dressed, taking a bath or shower); other health problem or disability.</p> <p>If a respondent answers says they have substantial difficulties with any of these areas, they are said to have a limiting long-standing illness or disability, if they answer none of these, then the long-standing illness or disability is said to be non-limiting.</p>
<p>Low parental qualifications</p>	<p>A child is coded as having parents with low qualifications only if both parents have qualifications below GCSE (A*-C) level or equivalent.</p>
<p>Poor parental mental health</p>	<p>We are measuring the proportion of children living with at least one parent reporting symptoms of anxiety and/or depression. See page 72 of the Analysis and Research Pack for more details.</p>
<p>Problem debt</p>	<p>Defined as being behind with all bills, council tax or behind with paying for housing costs such as rent or mortgage.</p>
<p>Relationship distress</p>	<p>A child is identified as being in relationship distress if either of their parents says that they regret living together all the time or most the time, they consider divorce/separation all the time or most the time, they get on each other's nerves most the time or all the time or they quarrel most the time or all the time.</p>
<p>Social housing</p>	<p>Those renting from a local authority or housing association.</p>
<p>Work class or skill level</p>	<p>Coded such that the highest parent's occupation is taken.</p>

Improving Lives: Helping Workless Families - Supporting Methodology Paper

Entry into worklessness / exit out of worklessness	Entry is calculated only for those who were in a working household. Exit is calculated only for those who were in a workless household. Thus, the total sample for exit (or not) from worklessness is much smaller than entry (or not) into worklessness.
Workless family	Where neither parent in the family was in paid employment
Working- bottom 40 per cent of the income distribution	Where at least one parent is working (defined as above) and child's household income is in the bottom 40 per cent of the income distribution - based on equivalised household income.
Working- top 60 per cent of the income distribution	Where at least one parent is working (defined as above) and child's household income is in the top 60 per cent of the income distribution - based on equivalised household income.
Workless duration (episodic)	Where the household is workless for 1 or 2 years out of 5 years. This is only calculated where we have work data for 5 consecutive waves.
Workless duration (persistent)	Where the household is workless for 3 to 5 years. This is only calculated where we have work data for 5 consecutive waves.

Millennium Cohort Study (MCS) definitions

For definitions of child outcomes, please see page 34 of the Analysis and Research pack.

General limitations of this analysis

Overview

This section outlines some of the limitations to our analysis. These are issues common to most survey-based longitudinal analysis. The main purpose of our longitudinal analysis is to explore and outline differences between various disadvantaged groups, and we are confident that our findings are robust to any of the issues outlined here.

Specific issues

Standard limitations of using survey data

Surveys gather information from a sample rather than from the whole population. The sample is designed carefully to be as representative of the general population as possible, given practical limitations such as time and cost constraints. However, results from sample surveys are always estimates, not precise figures. This means that they are subject to a margin of error (sampling error) which can affect how changes in the numbers should be interpreted, especially in the short-term. Year-on-year movements should be treated with caution. However, the Analysis and Research Pack does not comment on changes unless we are confident they are statistically, and substantively, significant (i.e. they represent a genuine change that is unlikely to be due to sampling error alone).

Surveys are also at risk from a systematic bias due to non-response, when households that had been selected for interview do not respond to the survey. Individuals within households may also be non-responders even if the rest of the household does respond. In an attempt to correct for these biases, the results are generally weighted to adjust for non-response, and we have made clear where we believe results may be biased.

Non-response can also occur where a respondent has given a full interview, but has refused or given a 'don't know' answer to a particular question, which consequently leads to a missing value for that item. In all analysis in the Analysis and Research Pack we exclude results for children where at least one parent gives an unknown value. This may slightly bias results to lone-parent families, as it is easier for those families to give full responses. The only exception is for the Poor Parental Mental Health Indicator, for reasons explained in the corresponding section of this document.

In addition to sampling errors, consideration should also be given to non-sampling errors. Non-sampling errors arise from the introduction of some systematic bias in the sample as compared to the population it is supposed to represent. As well as response bias, such biases include inappropriate definition of the population, misleading questions, data input errors or data handling problems – in fact any factor that might lead to the survey results systematically misrepresenting the population. There is no simple control or measurement for such non-sampling errors, although the risk has been minimised through careful application of the appropriate survey techniques from the questionnaire and sample design stages through to analysis of

results. Unlike other forms of error or non-response outline, it is likely this would be random, and less likely to be related to some underlying characteristics of the individuals interviewed.

Definition of a family in longitudinal analysis

In the Analysis and Research Pack we regularly talk about the proportion of children living in long-term workless families or, for instance, the likelihood that a child lives in a family that became workless. In our longitudinal analysis, we have followed/analysed the changes in children's lives from wave to wave. This because i) we are interested in children's experiences and outcomes and ii) following families would be too complex, because they regularly form or dissolve over time. Therefore, our definition of family is actually based around the adults (parents) that are living with the child in each wave. This means it is technically possible that a child could be living with completely different parents from one wave to the next. So, for instance, a child that lives in a family that became workless could possibly have been living with working parents/guardians, and then moved to live with a different set of workless guardians/parents. Denoting this precisely would be confusing and potentially misleading: (perhaps, for instance "the proportion of children who persistently experience living with workless parents") and, more importantly, will not apply to the overwhelming majority of children in the survey.

Definition of worklessness

Our definition of a workless family is based on whether a child is living in a family where no adult is in paid employment. This is a purely binary indicator, and we have not removed families where, for instance, both parents are retired, or students. Removing these families reduced the proportion of dependent children in workless families by around half a percent and did not affect the nature of our findings.

Transitions between waves

Survey respondents are interviewed annually. Whilst there is some information on changes in characteristics between waves - for instance, employment transitions, and relationship changes - we have chosen not to use this. We have found that, whilst adding further uncertainty/complexity to the analysis (for instance, the number of unknowns), it did not substantively change results or the nature of the findings.

Defining characteristics in longitudinal analysis

We have based the characteristics being analysed on an individual's status in the last wave being considered, using weights from the last wave. This means that their status in any of the previous waves might differ from that of the last, and these changes may influence, for instance, an individual's persistent worklessness status. As an example, consider a child in a workless lone-parent family in the last interview of the period being considered, but who was in a working couple-parent family for the first three interviews. This individual would be classified as in a lone-parent family that had only temporarily experienced worklessness. However, our analysis finds that most families do not experience frequent changes in family structure and work status, so this will not substantively affect results or the nature of our findings.

Analysis of parental separation

In the regression analysis on separation (see parental conflict section), we define a separation to be when the number of adults in the benefit unit (family) changes from

two to one, from one year to the next. This method may miss out a small number of separations, since it is possible a child could experience a parental separation and then a subsequent re-coupling in the period between the waves of the survey, which would not be identified.

Rounding and suppression

Figures are rounded to the nearest percentage point independently and as a result, differences may not sum exactly due to rounding. Any proportions based on a sample population of one hundred or less are suppressed.