

Department for Work and Pensions

Working Paper No 69

Ethnicity and child poverty feasibility study

Lucinda Platt

A study carried out by the Institute for Social and Economic Research 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 577 4

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

Contents

Acknowledgements	v
The Author	vi
Abbreviations.....	vii
Summary	1
1 Measures, sources and issues	3
1.1 Income poverty and composition of poor and not-poor households ...	3
1.2 Number of years of Family Resources Survey to be pooled for poverty analysis	4
1.3 Material deprivation.....	6
1.4 Poverty/employment transitions	7
1.5 Triangulation of results.....	8
1.6 Levels of analysis.....	10
1.7 Generational analysis	10
1.8 Differentiation within the white majority.....	10
1.9 Regional analysis.....	11
1.10 Harmonisation of data, weighting, etc.....	11
Appendix Data sources.....	15
References	33

List of tables

Table A.1	FRS: Households with children and child counts, by ethnic group and study year	17
Table A.2	FRS: Benefit unit counts, by ethnic group and study year	17
Table A.3	FRS: Numbers of lone parent benefit units, by ethnic group and study year	18
Table A.4	FRS: Estimated numbers of children in poverty, by ethnic group and study year	18
Table A.5	FRS 2001/02 to 2003/04, pooled: Distributions of certain key characteristics of children's context, by ethnic group, cell percentages (confidence intervals)	20
Table A.6	FRS 2004/05 and 2005/06, pooled: Deprivation among benefit units with dependent children. Row percentages (confidence intervals)	21
Table A.7	QLFS: All those with dependent children and those unemployed and inactive, numbers by ethnic group	24
Table A.8	HLFS: Households with dependent children and workless households, numbers by ethnic group	24
Table A.9	HLFS 2001-2005, pooled: Non-UK-born among those with dependent children, by ethnic group. Cell percentages (confidence intervals)	25
Table A.10	QLFS 2002-2005, pooled: distribution of unemployment durations for those currently unemployed, those living with dependent children. Row percentages (confidence intervals)	27
Table A.11	MCS: Numbers of main respondents (parent/carer)	29
Table A.12	MCS: Patterns of low income transitions by ethnic group. Row percentages (confidence intervals)	30
Table A.13	MCS: Distribution of deprivation by ethnic group: Row percentages (confidence intervals)	31
References	33

The Author

Lucinda Platt Lucinda is Senior Lecturer in Sociology at the Institute for Social and Economic Research at the University of Essex. She has researched and published widely in the area of ethnic minorities in the UK, as well as on child poverty.

Abbreviations

DWP	Department for Work and Pensions
FACS	Families and Children Study
FRS	Family Resources Survey
HBAI	Households Below Average Income
HLFS	Household Labour Force Survey
HRP	Household Reference Person
LFS	Labour Force Survey
LSYPE	Longitudinal Study of Young People in England
MCS	Millennium Cohort Study
ONS LS	Office for National Statistics Longitudinal Study
QLFS	Quarterly Labour Force Survey
UK	United Kingdom

Summary

This feasibility study was conducted in order to establish the potential and parameters for a research study of child poverty and ethnicity in the United Kingdom. It explored potential data sources that were amenable to analysis by ethnic group and the sorts of questions relating to child poverty and ethnicity that they might fruitfully answer. It was used as the basis on which to agree the coverage of the main stage report on ethnicity and child poverty (Research Report No. 576 due to be published in May 2009). The feasibility study sets out the range of possibilities and options for analysis. The actual analysis carried out from among these possibilities was agreed with Department for Work and Pensions colleagues on the basis of their priorities and interests.

The analysis for this report and the main body of the text were produced in 2007/08. Since that time there have been various changes to data access, new releases of data and development of the priorities of the commissioning team, with the result that not everything presented in this report remained relevant to the subsequent research that it informed. However, because a Research Report for that subsequent work is being published separately (Research Report No. 576 due to be published in May 2009), independently describing which data sources have been used and why, we feel it is appropriate to publish this feasibility study predominantly as it was initially produced, with additions only where the original text may lead to confusion.

1 Measures, sources and issues

1.1 Income poverty and composition of poor and not-poor households

The main source for measuring and evaluating income poverty is the Family Resources Survey (FRS) and its derived Households Below Average Income (HBAI) data set. This can be used to provide three-year rolling averages of poverty rates across ethnic groups, and thus, the basis of any interrogation of poverty differences. Three-year rolling averages are necessary to avoid considering as significant year-on-year fluctuations that derive from sampling variation, given the small samples of the different minority groups in any year. Generally, analysis of ethnic group differences requires the pooling of different years of the survey to increase minority groups sample sizes sufficiently to enable analysis. The FRS also has the potential to illuminate differences in risks of poverty across groups associated with particular characteristics (for example, lone parenthood, large families, workless families) and thus provide initial indications of the extent to which variations in poverty rates are associated with family characteristics (see Platt, 2006).

Table A.5 provides a number of breakdowns of specific characteristics that have been associated with child poverty and their distributions across ethnic groups. Such tabulations allow us some initial 'take' on compositional issues relating to (poor) families with children. However, to explore the compositional issues associated with poor and not-poor households in more detail, and their variation across ethnic groups, requires regression analysis, or exploration of subpopulations (see e.g. Table A.3) – or both. Regression analysis offers the potential to explore the role of between and within-group differences in poverty risks controlling for relevant factors and for the role of such factors in contributing to poverty risks. Similarly, regression analysis can also be used to evaluate the existence of an 'ethnic penalty' in poverty – or unexplained risk over and above that associated with control variables. Propensity score matching could also be a means to evaluate an 'ethnic penalty', but some difficulties with fulfilling the principle of

'common support', that is, the extent to which groups have sufficiently similar sets of characteristics to genuinely match them, have been noted in the past. The FRS data have been structured into various pooled data sets, which take the child as the unit of analysis, and relevant variables for use in multivariate analysis have been constructed from household, benefit unit and child files, covering, for example, household ethnic group, work status within the family, types of benefits received within the benefit unit, number of children, health status of family members, etc. Exploratory regression analyses have been carried out to test for their feasibility, but it has not been within the scope of this feasibility study to evaluate the possibilities for propensity score matching.

As well as direct measures of poverty, the FRS also allows exploration of the role of means-tested benefit receipt in families with children, both as a proxy for poverty and also to investigate the relative shares of income that such benefits make up. A descriptive analysis of the role of means-tested benefits, and Income Support in particular, in terms of their contributions (relative and absolute) to the incomes of poor and not-poor families with children, could be complemented by an investigation of the risks among families from different ethnic groups of being in receipt of such benefits. The level of detail in income variables allow for relatively precise consideration of these issues, which would amplify the existing tabulations on income components published in the FRS reports, and would provide new evidence on the role of such benefits for children living in poor families. Comparisons could also be made with the Labour Force Survey (LFS), which has information on benefits in receipt, to take advantage of the opportunities for triangulation (see below) and the larger sample sizes of the LFS. Indicators of types of benefit receipt divided between disability benefits, Income Support and unemployment benefits have been created for both the individual-level data in the pooled Quarterly Labour Force Survey (QLFS) sample, and for families in the pooled Household Labour Force Survey (HLFS) sample, where they can be analysed at the family level, to correspond to benefit units in the FRS.

1.2 Number of years of Family Resources Survey to be pooled for poverty analysis

Pooling across three years is expected to provide sufficient sample sizes for a range of analyses. (See Appendix, Tables A.1 and A.2.) The advantage of three years' worth of data is its consistency with the three-year averaged poverty rates from HBAI, and substantial numbers from the main minority groups to allow a series of tabulations and regression analyses; and yet sufficiently few years to enable the results to relate to a discrete time period, and thus, potentially, to events or other analyses within that time period, rather than being spread over a long period, which would only enable rather more diffuse connections. An additional advantage of pooling a relatively small number of waves is that the more waves involved, the more year-on-year changes in the data need to be taken account of (for example, changes in scope of the survey, rotation in questions asked, changes in coding of

questions, changes in the question itself or the sequencing of questions, changes in the derivation of derived variables etc.).

Berthoud (1998) pooled just two waves of FRS data to perform his analysis. But he was analysing individuals rather than specifically children (and their families) and the analysis also involved a higher degree of aggregation of ethnic groups than is recommended for this study (for example, combining Pakistanis and Bangladeshis). There were also clear limitations on the detail he could present in terms of income sources or the confidence he could claim for the precision of his estimates. Moreover, he did not employ a distinction between poor and not poor, but instead looked at the average incomes or range of incomes across the ethnic groups. In the proposed study, much of the analysis would be within the population of children defined as poor according to standard low income measures. (See Appendix, Table A.4 for approximate counts.) Thus, three years is considered the minimum for adequate descriptive analysis of child poverty risks, and comparisons between ethnic groups, as well as of comparisons between poor and not poor children within ethnic groups. Three years would also be the most appropriate basis for the majority of regression analyses with the FRS.

Moreover, overlapping sets of three year averages enable some minimal level of discussion of trends.

However, where variables only cover two waves of data, it would clearly be necessary to carry out analysis on just those waves. For example, deprivation measures are currently only available in 2004/05 and 2005/06. As the 2005/06 data were not distributed by the Data Archive till November 2007, it is not expected that the 2006/07 data would be available within the time-frame of the main study. For these analyses, the ethnic groups subject to analysis would need to be restricted to those where sample sizes are sufficiently large (see Table A.6).

The Appendix provides illustrations of various characteristics as they are distributed across children in a pooled sample of three waves (2001/02 to 2003/04). The estimates in Table A.5 and their standard errors are adjusted to take account of response and of clustering in families. The estimates are accompanied by confidence intervals, based on the adjusted standard errors, which allow us to assess the extent to which we can distinguish between groups on some of these key characteristics when we combine three years of data. Characteristics represent just some of those which have been associated with greater poverty risks. For most of the groups, despite some overlapping confidence intervals, the inferences to be drawn and the contrasts or similarities with the white majority are largely clear. The estimates for the children living in Chinese households are subject to wide confidence intervals, though the patterns from the estimates tend to be consistent across measures (and data sources).

In certain cases, for example, in descriptive analysis of components of income in poor and not-poor households and for breakdowns by subpopulations, such as lone parents or those born within or outside the United Kingdom (UK), a larger

number of waves of data in the pooled sample would be preferable to achieve more robust results. (See, for example, Appendix, Table A.3.) In addition, any analysis of Chinese low-income rates would be more robust with the full five years of data. And some analysis of Bangladeshi low income might be difficult with three years' worth of data. The three-year analyses could therefore be supplemented by some analysis carried out with five years' worth of data, if there was particular interest in these instances. In such cases comparison of the larger sample with the results from the three year samples would also provide some indication of the extent of the gains – or the trade-offs from such larger pooled samples.

The pooling of the LFS, both the HLFS and the QLFS, would aim to cover a similar time period as the five-year range of FRS data. That is approximately 2001/02 to 2005/06. There are issues in the changes to the timing of the quarterly extracts in the LFS from the beginning of 2006 which may mean that it is preferable only to pool up to the end of 2005. The implications of the change from seasonal to calendar quarters for pooled analysis is being explored. (See further discussion in the Appendix.) Currently, data sets have been constructed which pool eight biannual HLFS extracts (from the second half of 2001 to the first half of 2005) and 12 QLFS extracts from the second half of 2002 to the middle of 2005, but these could potentially be extended by a couple of quarters to map more closely onto the FRS time period, if the structure of the data allow.

1.3 Material deprivation

There are two sources for analysis of material deprivation: a full suite of questions – including child deprivation questions – in the FRS and a reduced set of questions in wave 2 of the Millenium Cohort Study (MCS). Given that material deprivation questions have only been asked in the FRS since 2004/05 (giving, so far, two years' worth of data), the sample sizes for analysis compare favourably in the MCS for some groups, in particular the Pakistani group (see row two of Table A.11 and rows one and two of Table A.1). On the other hand, design effects – and their reduction of effective sample sizes – will be greater in the MCS, and the range of questions is greater in the FRS. The recommended methodology for constructing deprivation scores could be implemented within analysis of the FRS (Willitts, 2006). Within the FRS, it would be feasible to carry out deprivation analysis on the larger minority groups (Indian, Pakistani, black Caribbean and black African) and limited analysis of Chinese and Bangladeshi, though Table A.6, provides an outline of deprivation across benefit units from all groups for the FRS. Both sources also include some questions on arrears/debt, and the MCS has questions on subjective measures of well-being, which could be compared with the relative deprivation scores. Illustrative analysis of the deprivation measures in the MCS is provided in the Appendix, Table A.13 and in the associated discussion. The mainstage analysis could involve multivariate analysis of deprivation with deprivation score as the dependent variable. Moreover, the distribution of deprivation across poor and non-poor in the FRS (or across different income bands in the MCS), would provide

a means to explore whether the relationship is any more (or less) consistent across any of the ethnic groups (to the extent that the sample sizes allow us to draw firm conclusions).

1.4 Poverty/employment transitions

Potential sources for exploring employment and income durations and transitions for families with children by ethnic group are the LFS and the Office for National Statistics Longitudinal Study (ONS LS) (for unemployment/worklessness) and the MCS (for low income). The Families and Children Study (FACS) is an additional longitudinal source, specifically of families with children, but due to small sample sizes from individual minority groups, it does not offer great potential for the mainstage study.

The income data in the MCS has its limitations – in that overall family income is banded and there is no derived variable summing the data from all the sources about which information is collected.¹ Nevertheless, the bands are relatively fine grained and banded income has shown to stand up reasonably well compared to more precise measures of income (Micklewright and Schnepf, 2007), particularly when it is the dependent variable. Even as the independent variable there are standard ways of treating banded income that appear relatively robust. Importantly, the MCS is one of the few sources that can be informative about low income transitions among families with (young) children and is, therefore, a potentially valuable source. Preliminary analysis has indicated that it has sufficient sample sizes and valid responses to enable analysis of low-income transitions – and that it can provide important new information on this topic which has not up till now been susceptible to analysis by ethnicity, except for aggregated ethnic groups using the FACS, or by using administrative data (see e.g. Platt, 2003). Some preliminary analysis of income transitions in the MCS is shown in the Appendix (see Table A.12 and related discussion). It will be possible to extend this illustrative analysis in a number of ways, still using the banded data, for example, by looking at transitions across more than one band and by using mid-points of the bands as an approximation for family income and employing regression analyses on wave 2 income dependent on wave 1 income. A further extension would be to look at transitions across three waves if the wave 3 data from the MCS were released in time (release expected February 2008).

Employment transitions in families with children can be explored – on very different time-scales – using the ONS LS and the LFS. It would be possible to analyse long-run transitions in household worklessness by ethnic group among those who are young children in 1991 and are still children in 2001, and to compare these with the patterns for 1981-1991. Preliminary sample sizes for the two cohorts are not currently available, but given the overall size of the ONS LS (one per cent of

¹ After this feasibility study was completed, a new version of the MCS was released with a derived income poverty measure enabling direct analysis of poverty within the study.

the population of England and Wales) and its history of use in studies of ethnic minority groups or of differences between ethnic groups (e.g. Platt, 2005; 2007), the sample sizes for this study are predicted to be sufficient for this analysis, particularly for the later cohort and for the larger ethnic groups.

The QLFS has a short-panel element which can potentially be informative about short-run unemployment durations among those living with dependent children. The LFS also has a retrospective question on unemployment duration for those currently unemployed. Both these sources can provide some limited information on differences in unemployment transitions across ethnic groups among those living with dependent children. Table A.10 in the Appendix and the related discussion provide preliminary tabulations of the retrospective information; while the potential for survival analysis using the panel element of the survey and carried out on those observed to enter unemployment within the life of the panel has been explored and is also discussed in the Appendix. The survival analysis could be varied to consider inactivity transitions as a whole; but is probably most meaningful when used to explore individual unemployment dynamics. While it is not the same as exploring poverty dynamics, analysis of unemployment dynamics and the differences between those with and without dependent children as well as between ethnic groups, could provide novel information about the context in which ethnic minority disadvantage and child poverty occurs.

1.5 Triangulation of results

The use of multiple data sources will add to the richness and variety of our knowledge and understanding of child poverty and its experience by ethnic group. This will potentially be a critical contribution of this project – typically analyses of poverty and disadvantage focus on a single data source addressing the questions that can properly be answered with that source (or occasionally matching in additional information to supplement). The use of multiple sources will be highly beneficial in the context where there is no single source which can provide more than a part of the ethnicity and child poverty 'story'. Moreover, where results from analyses are consistent across different sources, this will lead to greater confidence in findings which may, due to small sample sizes or other data problems, be necessarily tentative when taken in isolation. (Conversely, lack of consistency across sources will tend to heighten cautious interpretation of results, or at least require further investigation of inconsistencies.) While, overall, then, the use of multiple sources is recommended – it does come with costs. First, gaining a detailed knowledge of any given data source and its idiosyncrasies, as well as giving a clear account of the source, is a time consuming process, even after the initial familiarisation undertaken during the feasibility study. To the extent that multiple sources are used, this would increase the time spent purely on data considerations in the main stage of the analysis, and result in less time spent specifically on analyses. It would also reduce the likelihood that there would be time for more complex analyses with any given source. Second, it would necessarily add to the complexity of interpreting the findings – there would be likely to be contradictions between

them or different caveats to be stated for different sources. In terms of any statements of findings, the richness of the range of results would be accompanied by a need for multiple clarifications.

On balance, the advantages of triangulation are considered substantially greater than the disadvantages; but there are diminishing returns as the number of studies used for analysis increases beyond a certain point. Of those considered, the following surveys would seem most relevant for the main stage analysis:

- HBAI;
- FRS;
- MCS;
- ONS LS;
- LFS (quarterly, household and semi-panel).

Fewer benefits would be gained from using:

- FACS;
- Longitudinal Study of Young People in England (LSYPE).

In the case of FACS, despite its valuable longitudinal design, the sample sizes from minority ethnic groups are too small to allow analysis without aggregation. For example, in the 2005 wave, there were available for (cross-sectional) analysis 148 mothers from 'black' groups, 232 from 'Asian' groups and 189 from 'other' groups. The authors of the report on the 2005 data note that: *The survey contained only small numbers of mothers from different minority ethnic groups and thus, it is not possible to determine whether there are statistically significant differences between families by ethnicity* (Hoxhallari et al., 2007: Note 6).

In the case of the LSYPE, given that this is a cohort study, it is only informative about those children aged around 14 at wave 1, though there is some information about other household members. It is not informative about children of all ages therefore, and, as the study ages and becomes more valuable for longitudinal analysis, smaller proportions of the sample will be dependent children. In addition, the income variables are limited. As a cohort study, the MCS has been preferred as it captures children in infancy and enables analysis of those either being born into poor families, or where poverty follows shortly the birth of the child at what is regarded as a crucial time in relation to future child outcomes. Despite the limitations of the household income data in the MCS,² the material deprivation and subjective measures in wave 2 are susceptible to cross-sectional analysis, and some analysis of probabilities of transitions into or out of the lower income bands, controlling for household size, could provide information on poverty transitions. (See Appendix, Table A.12.) Moreover, the MCS has the advantage of covering all the countries of the UK compared to the LSYPE which focuses on England only.

² Though see footnote 1.

1.6 Levels of analysis

There are some instances in which using the household, family, benefit unit or the parent, rather than the child, as the unit of analysis might be more appropriate. For example, when looking at income sources within households with children.

1.7 Generational analysis

Distinguishing the experience of minority group children from different generations is clearly of great interest and relevance to this study. It is important to our understanding of ethnic differentials in child poverty to be able to answer such questions as: do rates of child poverty decline with length of time spent in the UK and are they different according to whether the parent(s) are UK-born or not? However, exploring these questions in detail is hampered by lack of country of birth questions in the FRS.

The non-members' file in the ONS LS could be used in order to ascertain whether parents of the LS members are UK-born or not. That information could be used in multivariate analysis and to investigate the extent to which worklessness risks and transitions differ between 'first' and 'second' (or subsequent) generations. Date of arrival in the UK can also be used to refine the generational information. Similarly, country of birth information and date of arrival in UK information in the LFS can be used to disentangle generational differences in that analysis (see Table A.9 and the associated discussion). As the discussion in the Appendix notes, there may be limitations to the analysis of generational differences in some cases where the large majority of the parents are not UK-born. Thus detailed generational analysis, even within the ONS LS and the LFS might have to be limited to Indians, Pakistanis and black Caribbeans, among the minority groups. The MCS introduced country of birth information for respondents in England at Wave 2 and this could be used in regression analysis to determine if it impacts on family poverty risks and whether there are interactions between being non-UK-born and any of the specific ethnic groups.

1.8 Differentiation within the white majority

Lack of country of birth information inhibits differentiating within the white majority in the FRS, since it would most effectively be achieved through investigating grouped countries of birth. Some analysis of the 'white other' category would still be possible. Differentiation within the white population would be most fruitful within the LFS and the ONS LS, where there are large numbers of those who allocate themselves to one of the white groups, but are not UK-born, (for example, in addition to the two per cent of white British with dependent children who are born outside the UK, around two-thirds of the 13,000 'other white' adults with children in the pooled LFS are also born outside the UK). If it were of interest, then there would be scope for differentiating not only between non-UK and UK-born within the overall white majority but also for exploring the role of duration of residence in the UK and for some distinction between countries or regions of birth.

1.9 Regional analysis

Around nine to ten per cent of FRS households are in London and many minority ethnic groups are concentrated in London. Distinguishing London from other regions could, therefore, be of interest and relevant to the analysis. Analysis by other regions – or combinations of regions – is theoretically possible, but is unlikely to be very meaningful given the diversity within regions, as has been shown with previous analysis employing the LFS to explore ethnic employment differences.

Northern Ireland was only included in the FRS in 2002/03. Therefore, any analysis which includes five years of data cannot include Northern Ireland. Moreover, the detailed ethnicity coding excludes Northern Ireland – given the difference in the ethnic group question there. In the aggregate ethnic group classifications, the vast majority of households are attributed to the aggregate ‘white’ group. For example, over three years’ of pooled data from 2002/03 to 2004/05 there were a total of 37 households in Northern Ireland from non-white minorities, and a total of 18 which contained dependent children. Though it is recognised that child poverty rates overall are an issue in Northern Ireland, and though, other things being equal, it would be preferable to undertake overall UK analysis, this is a study of minority ethnicity and child poverty and the focus should be on communicating the results relating to ethnic differences most effectively. It might, therefore, be considered preferable for the FRS analysis to be for Great Britain rather than the whole UK. Similarly, the LFS analysis might benefit from excluding Northern Ireland. In the MCS, the weights are designed either to apply to individual country analysis or to UK analysis. Therefore, UK-level analysis would be most appropriate for this source. The ONS LS only covers England and Wales so would necessarily be limited to coverage of those countries only.

1.10 Harmonisation of data, weighting, etc

There is typically some variation in the coding of variables between waves/years of data, reflecting changes in policy (e.g. relating to benefit eligibility) or to decisions being taken about ways of collecting the data, or in the ways data are named and organised. In some cases this can be adjusted by simply taking account of changes in naming conventions between years; in other cases, variables can be adjusted to be consistent (an example is the coding of qualifications in the LFS). In other cases, variables in different years have different levels of detail and pooling requires a ‘lowest common denominator’ approach where all versions of the variable are made consistent with the least detailed version. An example is the qualifications data across decades in the ONS LS. In further cases, harmonisation is simply not possible due to changes in the meanings or the categories of a variable between waves/years. In such cases, the analysis cannot be performed including such a variable, or the data have to be restricted to a time preceding or post-dating such a change.

A critical example of a variable where harmonisation is not considered feasible is the change to the measurement of ethnic group from the 1991 Census classification to the 2001 Census classification. This change took effect across data sources in 2001. There is evidence that the change in the form of the question prompted people to answer in different ways and though it is possible to make some adjustments to harmonise between 1991 and 2001 categories, this does not come without some loss of information. Given the focus on ethnic group, it would make most sense in the majority of analyses to employ only those data collected subsequent to the introduction of the 2001 categories.

In the ONS LS, however, in order to facilitate comparison between two cohorts, a slightly different approach would need to be used, discussed further below.

There are further issues to be resolved about how the child's – or the child's family's – ethnicity is coded. In terms of the child's own future outcomes, their ethnic group is, of course, potentially highly relevant, but when we are considering whether they are living in poverty, the assumption is that ethnic group differences are associated in some way with family or family context – whether through different types of family composition which vary with the ethnicity of the parent(s), whether through, e.g. labour market discrimination which restricts the opportunities of the parents, or whether through different distributions of skills across ethnic groups which affect parents' opportunities and so on. In this case it becomes, anyway, more relevant to consider the ethnicity of the parents/household. Given, however, that the ethnic group of different adults in the household will not necessarily be the same, there is a variety of ways of attributing ethnic group to the family/household. The standard approach in HBAI is to attribute ethnic group to the household on the basis of the Household Reference Person (HRP). This has much to recommend it in that it provides a simple solution and effectively allocates the 'weight' of ethnicity with the HRP. It is not, however, the only solution. An alternative is to more precisely describe the ethnic composition of the household allowing for households containing more than one ethnic group. However, this inflates the number of ethnic group categories – and given the relatively small numbers of households which contain people from different ethnic groups, it effectively simply loses cases for analysis. A third option is to attribute ethnicity to the child (via household or benefit unit) by prioritising the ethnicity of one member of the household in households where adults have different ethnic groups. This could work effectively for the FRS and the LFS and is the approach used in the illustrations in the Appendix, where the following hierarchy (which could be adapted) is used: mixed dominates the 'other' ethnic groups; Indian dominates mixed; Pakistani dominates Indian; Bangladeshi dominates Pakistani; black Caribbean dominates Bangladeshi (very rare combination); black African dominates black Caribbean; Chinese dominates black African (very rare combination); and households are only allocated to the white group if all adults are white. The rationale is largely practical: and the approach serves to maximise sample sizes of the minority ethnic groups considered, and tends to favour the smaller over the larger minority groups. In addition, it is not clear that it is substantially more arbitrary in terms

of conceptualising the impact of ethnicity than using the HRP. It is also open to adjustment, to test the empirical impact of the given configuration. A further option would be to allocate ethnicity on the basis of a particular relationship to the child – e.g. using the mother's or the father's ethnicity. This could be the preferred solution for the MCS, where there is, by definition, an interview with the main carer, in most cases the mother, and so the main respondent's ethnicity could be adopted as the ethnicity of the family.

In the ONS LS, information on the ethnicity of the relevant LS member is collected in 1991 and 2001. For those who are children in 1981, the information from 1991 or 2001 can be linked back to them at that stage. For those whose persistence (or not) in a workless household is measured at 1991 and 2001, either classification could be used as, by definition, the respondents have to be observed at both time points. For the comparative cohort, whose presence in a workless household is measured at 1981 and 1991, either 1991 ethnic group could be used or information from 2001 could be allocated to them, even though this is not one of the observation points. The disadvantage of the latter approach is that there are likely to be some observed in 1991 who are lost to follow up by 2001. Where 2001 ethnic group information is missing (imputation flags are available in the ONS LS), or the individual has been lost to follow up, the information could be 'topped up' with that from 1991, where there are valid responses and where the categories are reasonably consistent across the two time points (Platt *et al.*, 2005).

It would also be possible to use the reported ethnicity of the non-sample members – the respondents' parents or responsible adults, to allocate ethnicity to the family of the respondent, either at the first observation point or the second observation point, rather than depending on the ethnicity of the child respondent themselves (which, in any case, will have been attributed to them by the adult completing the census form and is not direct self-report). This would lead to greater consistency with the approach proposed for the LFS and the FRS, though, given the differences between the sources (including, at the most basic level, in geography) and in the analysis, it is not clear that this will be especially important. Moreover, it would introduce further complications in that while a single ethnicity will be attributed to the child, according to the approach proposed, the composition of the household (or parents) may change between 1991 and 2001 (or 1981 and 1991) raising different possibilities for allocation of ethnicity. (A similar issue arises in relation to the panel element of the LFS, but given the short duration, it is much less of an issue and ethnic group could be allocated simply on the composition of the family at the first time point.)

Apart from the ONS LS, the data sets also include weights to adjust for non-response and design bias and, in some cases, to provide grossing up to population numbers. Weights will be used in descriptive analyses and in modelling to take account of design effects, oversampling and so on, and to provide standard errors which take account of effective sample sizes, and so provide appropriate confidence intervals for estimates.

Appendix

Data sources

A.1 Family Resources Survey

In the household and in the benefit unit file for 2005/06, (harmonised) ethnic group breakdown is (1) white, (2) mixed, (3) Indian, (4) Pakistani or Bangladeshi, (5) black or black British and (6) other. A more detailed ethnic group breakdown is provided in the adult file for all years – though here there is the issue that the codes do not cover the whole of the United Kingdom (UK). There is no ethnic group information in the child file.

As a result, an ethnic group variable has been created which describes the household (or benefit unit) on the basis of the adults living there. For these purposes, households with members from more than one ethnic group are accorded ethnic group in a hierarchical fashion: mixed dominates white, Indian dominates mixed, Pakistani dominates Indian, Bangladeshi dominates Pakistani, Caribbean dominates Pakistani, black African dominates Caribbean and Chinese dominates black African. Only those which are not accorded one of these minority groups are allocated to the 'other' category. Only those households which are all white (either white British, other white or a mixture of white British and other white) are allocated as white.

The Family Resources Survey (FRS) potentially allows analysis of the following poverty and deprivation measures across ethnic groups: household income from all sources (all years); low income measures from Households Below Average Incomes (HBAI) (all years); worklessness in household or benefit unit (all years); material deprivation (affordability) questions (2004/05 and 2005/06); arrears questions (2004/05 and 2005/06); receipt of means-tested benefits/receipt of Income Support (all years).

The following tables illustrate the sample sizes for the five years on which proposed analysis will be based. Table A.1 provides the numbers of households with dependent children and the number of children by ethnic group; while Table A.2 provides the numbers of benefit units, which may form the unit of analysis for investigation of benefit receipt and material deprivation, though these could also be analysed at the level of the child.

Table A.3 provides the counts for each of the years and by ethnic group of one of the subpopulations which may be of particular interest, lone parents.

Table A.4 provides a breakdown of the samples according to poverty status, as the primary focus of the study will be on those deemed to be in poverty and the characteristics of them and their families.

Table A.1 FRS: Households with children and child counts, by ethnic group and study year

Study number	Year	Number of households	Number of dependent children	Number of households with dependent children	Number of Indian households with dependent children		Number of Pakistani households with dependent children		Number of Bangladeshi households with dependent children		Number of black Caribbean households with dependent children		Number of black African households with dependent children		Number of Chinese households with dependent children	
					Number of dependent children	Number of dependent children	Number of dependent children	Number of dependent children	Number of dependent children	Number of dependent children	Number of dependent children	Number of dependent children	Number of dependent children	Number of dependent children	Number of dependent children	Number of dependent children
5742	2005/06	28,029	8,562	15,413	183 (330)	161 (391)	46 (123)	136 (226)	148 (301)	28 (45)						
5291	2004/05	28,041	8,697	16,012	177 (301)	118 (287)	62 (160)	127 (225)	117 (242)	26 (44)						
5139	2003/04	28,860	8,954	16,396	155 (299)	116 (297)	59 (152)	136 (223)	124 (283)	35 (62)						
4803	2002/03	28,710	9,166	16,866	179 (338)	142 (361)	68 (209)	134 (257)	92 (192)	28 (46)						
4633	2001/02	25,320	8,035	14,751	164 (318)	120 (305)	67 (173)	163 (295)	106 (232)	31 (52)						

Table A.2 FRS: Benefit unit counts, by ethnic group and study year

Study number	Year	Number of benefit units	Number of Indian benefit units with dependent children		Number of Pakistani benefit units with dependent children		Number of Bangladeshi benefit units with dependent children		Number of black Caribbean benefit units with dependent children		Number of black African benefit units with dependent children		Number of Chinese benefit units with dependent children	
			Number of benefit units with dependent children	Number of benefit units with dependent children	Number of benefit units with dependent children	Number of benefit units with dependent children	Number of benefit units with dependent children	Number of benefit units with dependent children	Number of benefit units with dependent children	Number of benefit units with dependent children	Number of benefit units with dependent children	Number of benefit units with dependent children		
5742	2005/06	33,317	8,596	184	165	49	137	148	28					
5291	2004/05	33,202	8,746	180	125	66	128	118	26					
5139	2003/04	34,243	8,996	155	121	61	138	126	35					
4803	2002/03	33,996	9,200	180	146	72	135	92	28					
4633	2001/02	30,037	8,058	165	122	69	164	107	32					

Table A.3 FRS: Numbers of lone parent benefit units, by ethnic group and study year

Study number	Year	Number of lone parent benefit units					Number of black Caribbean lone parent benefit units			Number of black African lone parent benefit units		Number of Chinese lone parent benefit units
		Number of lone parent benefit units	Number of Indian lone parent benefit units	Number of Pakistani lone parent benefit units	Number of Bangladeshi lone parent benefit units	Number of black Caribbean lone parent benefit units	Number of black African lone parent benefit units	Number of black African lone parent benefit units	Number of Chinese lone parent benefit units			
5742	2005/06	2,422	15	33	6	81	70	8				
5291	2004/05	2,448	18	19	10	53	60	2				
5139	2003/04	2,509	23	21	10	77	57	4				
4803	2002/03	2,535	19	23	6	69	46	7				
4633	2001/02	2,186	15	24	16	73	50	4				

Note: Count varies slightly depending on which version of 'family' type used.

Table A.4 FRS: Estimated numbers of children in poverty, by ethnic group and study year

Study number	Year	Number of Indian children in poverty (estimated rate = 32%)		Number of Pakistani children in poverty (estimated rate = 60%)		Number of Bangladeshi children in poverty (estimated rate = 72%)		Number of black Caribbean children in poverty (estimated rate = 37%)		Number of African children in poverty (estimated rate = 56%)		Number of Chinese children in poverty (estimated rate = 25%)†
		Number of Indian children in poverty (estimated rate = 32%)	Number of Pakistani children in poverty (estimated rate = 60%)	Number of Bangladeshi children in poverty (estimated rate = 72%)	Number of black Caribbean children in poverty (estimated rate = 37%)	Number of African children in poverty (estimated rate = 56%)	Number of Chinese children in poverty (estimated rate = 25%)†					
5742	2005/06	106	234	89	84	169	11					
5291	2004/05	96	172	115	83	136	11					
5139	2003/04	96	178	109	83	158	16					
4803	2002/03	108	216	150	95	108	12					
4633	2001/02	102	183	124	109	130	13					

Note: These figures are not derived directly from HBAI; they are simply an estimate based on approximating from known proportions in poverty by group. (Rates for estimates indicated in brackets, and taken from Platt (2006); numbers to which rates are applied are those in Table A.1.)

† Estimates of poverty rates for Chinese children specifically were not given in the source for the other rates due to doubts about their reliability given small cell sizes. A rate of 25 per cent, which is that for the population as a whole has, therefore, been used here.

To illustrate the robustness of results that can be obtained from pooling years, Table A.5 gives some results from a pooled data set of 2001/02 to 2003/04 samples. The results (and the confidence intervals) take account of weighting and of repeat observations of children within households. Distributions across a few key variables are shown both for the households the children live in and the benefit units. The unit of observation in each case is the child. The unweighted counts for each ethnic group on which the distributions are based are also provided. These vary slightly between household-level variables and benefit unit level variables as ethnicity is defined separately for household and the benefit unit. These distributions are for all children. The proposed analysis would compare across groups for poor children and within groups for poor and not poor children. The analysis excludes Northern Ireland cases. Note the large confidence intervals, even with three pooled years of data, for the Chinese children. The table clearly illustrates the high unemployment rates in Pakistani, Bangladeshi, black Caribbean and black African households and benefit units; as well as the large number of children living in single adult black Caribbean and black African households – the picture is enhanced when we focus on benefit units (families) rather than households. We can see that black Caribbean, Chinese and white children tend to live with fewer other children, while Pakistani and Bangladeshi children tend to live in households with more children, with black African children falling somewhere in between. We also can see the relatively high proportions of Bangladeshi and Pakistani children living in a household containing an adult of working age who is long-term sick or disabled.

Table A.6 illustrates material deprivation measures pooling the two years in which they have so far been carried, 2004/05 and 2005/06. The table distinguishes between the adult deprivation measures and those applying specifically to children. A simple count of the number of measures on which the benefit unit is deprived has been created for each benefit unit; and then this has been aggregated to illustrate the patterns across the ethnic groups. Despite large confidence intervals in some instances, it illustrates the differences between levels of deprivation among the white, Indian and Chinese groups on the one hand and the other groups on the other hand.

Table A.5 FRS 2001/02 to 2003/04, pooled: Distributions of certain key characteristics of children's context, by ethnic group, cell percentages (confidence intervals)

Ethnic group	Children in household				Children in household				Children in household				Children in benefit unit with one adult (CI)††		Unweighted number of households (of benefit unit)
	with an unemployed adult†	with sick/disabled working-age adult	with one adult	with two adults	with one adult	with two adults	with three+ adults	with one child	with two children	with three children	with four+ children	in benefit unit with unemployed adult	in benefit unit with one adult	(CI)††	
White	5.4 (5.1-5.8)	22 (22-23)	21 (20-21)	70 (69-71)	9 (9-10)	24 (23-24)	46 (45-47)	21 (20-22)	9 (8-10)	4.2 (3.9-4.5)	24 (24-25)	39,697 (39,751)			
Indian	4.8 (3.0-7.4)	22 (18-26)	8 (6-11)	65 (60-7)	26 (22-31)	21 (18-24)	44 (39-49)	26 (22-31)	9 (6-14)	3.1 (1.8-5.4)	11 (9-15)	955 (944)			
Pakistani	14.0 (10.4-18.6)	25 (30-41)	11 (8-15)	60 (54-65)	29 (24-35)	10 (8-13)	25 (21-29)	31 (26-36)	35 (29-41)	10.5 (7.4-14.7)	16 (13-21)	963 (958)			
Bangladeshi	18.2 (12.9-24.9)	45 (37-53)	9 (6-14)	58 (50-65)	33 (26-41)	10 (7-13)	22 (17-29)	30 (23-38)	38 (30-46)	9.6 (6.1-15)	15 (10.21)	526 (526)			
Black	9.8 (7.1-13.2)	20 (16-24)	41 (36-46)	49 (44-54)	10 (7-13)	29 (26-34)	40 (35-45)	20 (16-25)	11 (7-16)	7.4 (5.1-10.4)	51 (46-56)	775 (766)			
Black African	8.0 (5.4-11.7)	21 (17-27)	37 (32-43)	53 (47-59)	10 (6-15)	16 (13-20)	36 (31-42)	28 (23-34)	19 (14-25)	7.2 (4.7-10.9)	47 (41-53)	707 (701)			
Chinese	2.6 (1.0-7.1)	15 (9-24)	11 (6-20)	76 (64-84)	14 (7-25)	27 (19-36)	51 (40-62)	17 (9-29)	5 (1-20)	1.1 (0.0-4.2)	15 (9-25)	160 (158)			

Note: Weights to correct for non-response applied.

† The unemployed adult need not necessarily be a parent (though it will be the responsible adult where there is only one adult in the household). This measure is distinct from any economically inactive adults within the household.

†† Benefit units have either one or two adults so it is only necessary to specify the percentage with one adult to show the full distribution.

**Table A.6 FRS 2004/05 and 2005/06, pooled: Deprivation among benefit units with dependent children.
Row percentages (confidence intervals)**

	Adult deprivation			Child deprivation		
	Deprived on no counts	Deprived on one or two counts	Deprived on three or more counts	Deprived on no counts	Deprived on one count	Deprived on two or more counts
White British	43 (42-44)	21 (20-22)	36 (35-36)	62 (61-63)	21 (20-22)	17 (16-18)
Indian	48 (43-54)	22 (18-27)	30 (25-35)	55 (49-60)	25 (21-30)	20 (16-25)
Pakistani	21 (16-26)	22 (17-27)	58 (51-64)	28 (23-34)	29 (23-34)	43 (27-50)
Bangladeshi	21 (14-31)	16 (9-26)	63 (52-73)	24 (16-34)	26 (18-36)	51 (40-61)
Black Caribbean	26 (21-32)	21 (17-27)	52 (46-59)	37 (31-44)	31 (25-37)	32 (27-38)
Black African	21 (16-27)	15 (11-20)	64 (57-69)	33 (27-39)	23 (18-28)	44 (38-51)
Chinese	51 (37-65)	26 (16-41)	23 (13-36)	67 (52-79)	18 (9-33)	15 (8-26)

Note: Weights to compensate for non-response have been applied.

A.2 Labour Force Survey

The Labour Force Survey (LFS) does not have information on household income but it does have earnings information for those in employment (or for the two-thirds of those in employment who respond to this question); and it has employment information. In the household version of the LFS this includes information jointly on both responding adults in the household so it is possible to analyse workless families with children and combinations of employment, unemployment and inactivity. Household Labour Force Survey (HLFS) data sets are produced biannually, unlike the quarterly individual-level LFS data sets.

A.2.1 LFS cross-sectional

The LFS has been extensively used to explore ethnic (and migrant) differences in employment outcomes, including earnings, (and related issues, such as job search), by pooling quarters of the data set to increase sample sizes from any given ethnic group. This analysis has tended to focus on individual outcomes (though with controls for children – or sometimes using children to identify selection models), and has tended to use the individual-level Quarterly Labour Force Survey (QLFS). Pooling quarters has become slightly more of a complicated question since 2006, when the timing of the quarters changed from being ‘seasons’ (March-May, June-August, September-November, December-February), to quarters starting in January and ending in December. One option is only to use quarters up to the final ‘seasonal’ one (December 2005-February 2006), or to include a shift by ‘missing out’ December 2005 and moving from the September-November 2005 quarter to the January-March 2006 quarter in the sets of four quarters. It is not anticipated that this would introduce substantial bias in the pooled results.

It is possible, rather than concentrating on individuals, to focus simply on households with children (particularly by making use of the household data set) and compare the probabilities of such households being, e.g. workless, according to the ethnic group of the parent(s). This can range from simple distributions (c.f. comparable analysis of households with long-term sick adults), to models which control for additional household characteristics in exploring the probabilities of being workless. It is also possible to take the children themselves as the units of analysis (though appropriate weighting factors become an issue here, since there are household weights and individual (adult) weights but not weights for the children), and explore their probability of living with an unemployed parent according to ethnic group.

Using the individual-level data set, the analysis could focus on those individuals with dependent children, or dependent children below a particular age, and compare their characteristics with those without children by ethnic group.

Worklessness is clearly not the same as poverty, but LFS-based worklessness analysis could complement the FRS results in relation to workless households and provide an additional source of information. Given the extracts and the large sample sizes

at each quarter, the LFS has the benefit of swiftly producing relatively large sample sizes for minority groups.

Tables A.7 and A.8 show the numbers of individuals with children and the number of households with children in pooled sets of QLFS and HLFS data, respectively. For the QLFS it also shows the numbers unemployed and inactive, by ethnic group, and for the HLFS it shows the numbers of workless households (where all adults are either unemployed or inactive).

Generational issues have also been raised for the analysis. One way of exploring these is to distinguish whether the parents of the children are themselves 'first generation' or 'second or subsequent generation' using country of birth information. Country of birth and date of arrival in the UK for those not born in the UK are both collected in the LFS. Table A.9 illustrates the proportions not UK-born among those with children, by ethnic group. There are distinctive differences in the proportions who are not UK-born, which will affect the options for generational analysis across groups. For example, among Bangladeshis and black Africans, it will be much harder to distinguish generational differences given the high proportions who are not UK-born. In multivariate analysis, it will however be possible to include a UK-born variable. In fact, a variable which also incorporates some information on length of time in the UK is also possible, distinguishing between UK-born, non-UK-born who are long-term residents in the UK, and more recent arrivals. Such a variable has been employed as a control in the longitudinal analysis summarised in Section A.2.2.

Birth outside the UK, can also be used to provide some differentiation within the white groups. Thus as well as the two per cent of white British, the 66 per cent of white other who are not UK-born (also given in Table A.9) can be analysed according to their duration in the UK and their country/region of birth, to examine whether migration appears to have implications for worklessness among those with children.

Table A.7 QLFS: All those with dependent children and those unemployed and inactive, numbers by ethnic group

Period/pooling	Individuals living with dependent children: all/unemployed/inactive					
	Indian	Pakistani	Bangladeshi	Black Caribbean	Black African	Chinese
12 quarters pooled from March-May 2002 to December 2004-February 2005	7,958/ 330/ 1,740	6,803/ 403/ 3,335	2,732/ 219/ 1,520	3,899/ 344/ 892	4,144/ 260/ 1,563	1,244/ 41/ 332
As above but first wave for each respondent only (earnings only collected in waves 1 and 5)	1,412/ 61/ 314	1,212/ 71/ 612	456/ 42/ 248	661/ 57/ 147	707/ 48/ 276	230/ 6/ 70

Table A.8 HILFS: Households with dependent children and workless households, numbers by ethnic group

Period/pooling	All groups: all/workless					
	Indian	Pakistani	Bangladeshi	Black Caribbean	Black African	Chinese
Eight quarters pooled from September-November 2001 to March-May 2005	1,012/ 93	909/ 227	361/ 121	712/ 204	714/ 265	151/ 26

**Table A.9 HILFS 2001-2005, pooled: Non-UK-born among those with dependent children, by ethnic group.
Cell percentages (confidence intervals)**

	White British	White other	Indian	Pakistani	Bangladeshi	Black			Chinese
						Caribbean	Black African	Chinese	
All	2 (2-3)	66 (64-67)	69 (67-71)	66 (64-68)	85 (82-87)	33 (30-36)	85 (83-87)	77 (72-81)	
Men	2 (2-3)	63 (61-65)	72 (69-75)	69 (66-72)	87 (82-90)	37 (32-42)	86 (82-89)	74 (66-80)	
Women	2 (2-3)	68 (66-70)	66 (64-69)	63 (60-66)	83 (79-87)	31 (27-34)	85 (82-88)	80 (73-85)	

Note: Estimates and confidence intervals have been adjusted for non-response and repeat observations on individuals.

A.2.2 LFS longitudinal

The LFS is collected as a short panel, so that respondents remain in the survey for five waves – or quarters. This makes it possible to trace their movements into and out of employment (for example) across the period of 15 months. A longitudinal version of the data set exists linking respondents across the waves. However, it is also possible to make the link within the pooled data, to increase numbers from minority ethnic groups – as with the cross sectional analysis. Such analysis would focus only on those respondents with children – to explore the dynamics of employment for parents, and thus the implications for children. The short run of the panel limits the survival analysis which could be conducted, particularly if only entrants to unemployment (or inactivity) are considered in order to avoid the problem of ‘left censoring’. (In the pooled QLFS data set there were 83 Indian, 138 Pakistani, 87 Bangladeshi, 78 black Caribbean, 63 black African and 16 Chinese men living with children who were observed to enter unemployment during the life of the short panel.) A data set was constructed on the basis of male entrants to unemployment; and their ‘survival’ in unemployment was examined using a discrete-time competing risks model – where an exit could either be to employment or to inactivity. The model was successfully executed (not illustrated here), though it produced no statistically significant distinctions between those from different ethnic groups. This is probably partly due to the small numbers from different ethnic groups, but may also illustrate a genuine lack of differences between ethnic groups once relevant controls have been introduced and over the short period considered. The approach, however, would appear to be a viable one, though amplifying the sample sizes with further data extracts may help to ascertain if the results are convincing.

In addition, there is a variable within the LFS which asks about the duration of unemployment for those who are currently unemployed. This is banded but provides some information about the length of unemployment for the observed unemployed. Table A.10 gives the (unweighted) counts of those with dependent children to whom this variable applies (those observed to be unemployed), and also the proportions within a reduced range of bands, given that long periods were only experienced by a small minority. As well as being informative in its own right, this variable could also be used to correct for the left censoring problem in the unemployment durations analysis.

Table A.10 QLFS 2002-2005, pooled: distribution of unemployment durations for those currently unemployed, those living with dependent children. Row percentages (confidence intervals)

Ethnic group	N (unweighted)	<3 months	3-6 months	6-12 months	>1 year
White British	9,856	48 (47-49)	19 (18-20)	16 (15-17)	17 (16-18)
Indian	330	45 (39-52)	19 (16-24)	16 (11-22)	20 (14-27)
Pakistani	400	35 (29-41)	21 (17-25)	19 (15-24)	25 (20-32)
Bangladeshi	219	33 (26-41)	25 (19-32)	17 (12-23)	26 (18-36)
Black					
Caribbean	344	37 (31-45)	18 (14-24)	18 (13-23)	26 (20-35)
Black African	259	43 (36-50)	15 (11-21)	21 (15-27)	21 (15-29)
Chinese	41	46 (34-63)	29 (17-45)	9 (3-22)	13 (5-28)

Note: Estimates and confidence intervals have been adjusted for non-response and repeat observations on individuals.

A.3 Millennium Cohort Study

The Millennium Cohort Study (MCS) is a study of a sample of children born in 2000/01, who are followed over time as they grow up. It follows in the cohort tradition as exemplified, for example, in the National Child Development Study and the 1970 British Cohort Study (BCS70). It has the potential to be extremely informative about the lifecourses of children from a very early age, through their formative years and into adulthood (assuming it continues that long). There are currently two waves of the MCS survey available: one collected when the cohort members were approximately nine months old and one collected when they were approximately three years old. Main respondents are predominantly mothers, but partners (predominantly fathers) are also interviewed (though a substantial number of those eligible for interview were not contacted at either wave). It is thus informative about children at a particular age and about parents who are in the specific circumstances of caring for young children (though the cohort baby may not be their first – or indeed their last – child). Thus the information is limited to a particular section of children at risk of poverty; but this is potentially a peculiarly interesting group. That is, children in the early stages of life, and where analysis is not complicated by differences in the distribution of children across age ranges, and where parents tend also to congregate in a limited age range.

Though the MCS collects detailed information on earnings (including from second jobs), benefits in receipt and other sources of income, the only household income variables are based on a question that uses income bands. The banded income question covers joint total incomes for couples and own total income for lone parents. It is provided as annual amounts. In MCS wave 1, the two questions were combined into one variable involving just a small number of bands that covers both lone and couple parent households. A comparable variable has been

constructed for wave 2. There are also a couple of questions in MCS wave 1 on ability to save and financial pressure (subjective measures), and some questions on housing conditions that could be used as deprivation measures. At wave 2 there are also: subjective measures of financial welfare; debt questions (whether behind on various sorts of bills); saving questions (whether and what for); questions about whether respondent is managing to pay rent; and material deprivation questions (for both child and self).

The banded income variable has 19 bands at wave 1 and 18 bands at wave 2, giving some level of definition to analysis of moves between different income levels. Different bands are however used for respondents in couples and lone parent respondents. Making the bands consistent across respondents involves a reduction to six bands. (An alternative which can be pursued in the main study is to take the midpoint of the band for each respondent as their actual individual/household income and create an equivalised cut-off on the basis of that – using the FRS to calculate the ‘midpoint’ for the top band.) A simple analysis would be to consider a move from one of the bottom two (aggregate) bands into a higher band (and vice versa) as a low income transition – and this provides some matter for the relatively short period considered here. It is also possible to look at risks of remaining below a particular level of income (controlling for household size). However the variable does suffer from a substantial level of missing data. In MCS wave 2, around 15 per cent of cases both for those in couples and lone parents are coded as ‘not applicable’. This would appear to reflect non-response that has been coded as not applicable, as there is no clear reason why the information should be not applicable in each of the cases. In wave 1 of the MCS ‘not applicable’, ‘refused’ and ‘don’t know’ together amounted to nine per cent of lone parent responses and eight per cent of couple parent responses. The ‘non-response’ reduces the sample sizes available for analysis of this variable at wave 2, though still leaves a substantial number of cases to work with.

Table A.11 MCS: Numbers of main respondents (parent/carer)

Study number	Wave	Total number of families (children)	Number of Indian main respondents	Number of Pakistani main respondents	Number of Bangladeshi main respondents	Number of Black Caribbean main respondents	Number of Black African main respondents	Number of Chinese main respondents
4683	MCS1 (2001)	18,552 (18,819)	480	893	371	264	379	48
5350	MCS2	15,590 (15,808)	379	683	260	181	252	34

Note: The numbers vary between waves 1 and 2, on the one hand because of attrition between the two waves and, on the other, because of addition in wave 2 of children eligible for inclusion in the study (at wave 1) but missed at wave 1 (these amounted to 692 of the 15,590 wave 2 families). For wave 2 analysis of material deprivation, all the 15,590 families can be used (dependent on valid data being available for the relevant questions). For analysis of transitions, only those who are present in both wave 1 and wave 2, 14,898 families, are available for analysis – and again only to the extent that there are valid values on the variables used in analysis. The effective sample sizes represented by the numbers for minority groups are actually substantially smaller, given the sampling design. The application of weights (as in the illustrations below) takes account of these design effects.

As an illustration, Table A.12 shows the proportions experiencing different patterns of transitions between low (defined as roughly in the bottom quarter of the distribution of annual incomes) and higher incomes, by selected ethnic groups. These proportions (and their confidence intervals) have been calculated employing weights to account for the design of the MCS. The table illustrates the potential (and some of the limitations) of the MCS for exploring poverty transitions for those families with young children.

Table A.12 MCS: Patterns of low income transitions by ethnic group. Row percentages (confidence intervals)

Ethnic group	Low income in both waves	Moves up from low to high	Moves down from high to low	Higher income in both waves
White	10 (9-11)	6 (5-7)	7 (6-7)	77 (75-78)
Indian	9 (5-15)	7 (3-13)	9 (5-15)	75 (67-82)
Pakistani	23 (18-28)	18 (13-25)	17 (13-24)	41 (34-49)
Bangladeshi	26 (18-36)	17 (11-26)	20 (13-31)	37 (25-51)
Black Caribbean	26 (17-37)	9 (5-15)	16 (10-24)	49 (37-61)
Black African	19 (12-28)	12 (7-19)	9 (5-17)	60 (45-73)
Chinese	5 (1-25)	–	6 (1-25)	89 (70-97)

Notes: Weights to account for sample design applied.

We can see the relatively high rates of poverty persistence among the Pakistani, Bangladeshi and black Caribbean families (and to a lesser extent the black African families). Regression analysis was subsequently carried out (not illustrated) which looked at the probability of being in the low income bands at wave 2 for those in the low income bands at wave 1 and those not. It indicated an increased probability of low income at wave 2 among Pakistanis and Bangladeshis, both among those on a higher income and those on a lower income at wave 1 – though the risk relative to white counterparts was particularly strong among the group that started on a higher income. This was after various controls for family size, change in family size, couple and lone parent status at both waves and age of respondent had been introduced. Among those who started on a higher income, black Caribbeans also faced an increased risk of a fall into low income, despite the controls for family structure. But poverty persistence was not found to be statistically significantly different from that for white respondents with the same family structure and composition. While the results in terms of poverty persistence for the Bangladeshis and Pakistanis might not be surprising given their known high child poverty rates, the increased risks of falling into poverty from a higher income are interesting as they indicate that there is mobility within the group in terms of income. Sample sizes were sufficient to enable such regression analysis across the main ethnic groups in the sample and to give reasonable confidence in the results, with the exception of the small Chinese group. Future work would aim to refine how the income variables are treated and the modelling of transitions.

In relation to deprivation measures, wave 2 of the MCS holds a suite of nine measures – three relating directly to the child cohort member and the other six relating to the respondent parent. The measures cover whether an item is lacked and if so whether that is due to affordability or not wanting it. The typical cut off for indicating deprivation is at two or more lacks that cannot be afforded – though it is also possible to consider the whole distribution across the measures. Table A.13 indicates those with no or one lack and those ‘deprived’ if the conventional measure of two or more lacks is used. To provide a slightly different distribution, those facing no unaffordable lacks, those facing either one or two and those facing three or more are summarised in the right hand part of the table. The first part of the table shows that all the minority groups, with the exception of Indian and Chinese, are significantly more likely to be deprived than the white majority, with the black Africans facing particularly high risks of deprivation, notably higher than any of the other groups. The second part of the table shows a similar pattern, though the rates facing no deprivation are very similar for black Caribbean, black African and Bangladeshi families. It is among those facing three or more sources of deprivation that the black groups, and particularly the black Africans, show a more extreme situation than the other groups. The small numbers of Chinese mean that attempts to analyse their experience are limited. But the findings are consistent with the income information, which tends to suggest a situation for these families that is more advantaged than any of the other families (including white ones).

**Table A.13 MCS: Distribution of deprivation by ethnic group:
Row percentages (confidence intervals)**

Ethnic group	Deprived on no or 1 measures	Deprived on 2+ measures	Deprived on no measures	Deprived on 1 or 2 measures	Deprived on 3+ measures
White	78 (77-79)	22 (21-23)	65 (63-66)	21 (20-22)	14 (13-15)
Indian	78 (73-82)	22 (18-27)	60 (53-67)	29 (24-35)	11 (8-15)
Pakistani	66 (61-71)	34 (29-39)	44 (37-52)	35 (29-42)	21 (17-25)
Bangladeshi	59 (51-68)	41 (32-49)	37 (30-44)	41 (35-46)	23 (17-29)
Black Caribbean	55 (44-66)	45 (34-56)	38 (30-46)	31 (23-40)	31 (21-42)
Black African	49 (41-58)	51 (42-59)	39 (30-49)	22 (16-30)	38 (30-48)
Chinese	84 (68-93)	16 (7-32)	77 (62-89)	9 (3-24)	13 (5-27)

Note: Weights to correct for design effects applied.

A.4 Office for National Statistics Longitudinal Study

Analysis using the Office for National Statistics Longitudinal Survey (ONS LS) has to go through the process of an application for a specific project or set of analyses. It is not, therefore, possible to give indicative analyses in advance of such an application being approved and work with the data commencing. The ONS LS benefits from large sample sizes and the possibility of looking at long-run changes as study members have their circumstances measured at each census point from 1971-2001. It would, therefore, have substantial potential for investigation of children's transitions between living in a workless household and not living in a workless household (and vice versa) at a ten-year interval.

References

Berthoud, R. (1998) *The Incomes of Ethnic Minorities*. ISER Report 98-1. Colchester, University of Essex, Institute for Social and Economic Research.

Hoxhallari, L. et al. (2007) *Families with children in Britain: findings from the 2005 Families and Children Study (FACS)*. Department for Work and Pensions Research Report No. 424. Leeds, Corporate Document Services.

Micklewright, J. and Schnepf, S. V. (2007) *How Reliable are Income Data Collected with a Single Question?* S3RI Applications & Policy Working Papers, A07/08. Southampton, UK, Southampton Statistical Sciences Research Institute.

Platt, L. (2003) Ethnicity and inequality: British children's experience of means-tested benefits. *Journal of Comparative Family Studies* 34 (3): 357-377.

Platt, L. (2005a) The intergenerational social mobility of minority ethnic groups. *Sociology*, 39(3): 445-461.

Platt, L. (2005b) *Migration and social mobility: The life chances of Britain's minority ethnic communities*. Bristol, The Policy Press.

Platt, L. et al. (2005) Stability and change in ethnic group in England and Wales. *Population Trends* 121: 35-46.

Platt, L. (2006) *Ethnicity and child poverty*. Paper prepared for the Ethnic Minority Employment Task Force, autumn 2006. See <http://www.emetaskforce.gov.uk/>

Willitts, M. (2006) *Measuring child poverty using material deprivation: possible approaches*. Department for Work and Pensions Working Paper No. 28. Leeds, Corporate Document Services.

