Following families over time and across studies: towards a generic definition of family composition for use in longitudinal analysis of children's development

Pierre Walthery and Ian Plewis

February 2015

Acknowledgements

This work was supported by a grant from the Economic and Social Research Council to the quantitative node of the National Centre for E-Social Science (RES-149-25-1084). We would like to thank Nick Buck, Jane Elliott and Rachel Rosenberg for helpful comments on an earlier version of this report. We are, however, solely responsible for any errors in this final version.

Introduction

The e-Stat project (which ran from September 2009 until December 2012) was one of the eight nodes of ESRC's Digital Social Research Programme, whose aim is to use the recent advances in digital technologies to augment the impact of research in the social sciences. The project resulted from collaboration between the universities of Bristol, Stirling, Southampton and Manchester and brought together social scientists, statisticians and programmers. Its goals were to empower users of quantitative social science by developing tools that enable collaborative research, in terms of streamlining access and interoperability between advanced statistical packages, but also by developing tools that improve the documentation and sharing of research processes, results and techniques across the social science community.

This paper presents the outcomes of a feasibility study within the e-STAT project whose goal was to design a transferable concept of *family composition* to be used in research that, for example, focuses on children's development; to explore its potential for use with the main UK longitudinal studies; and to discuss its implementation in two key longitudinal studies (the British Household Panel Survey and the Millennium Cohort Study).

It is organised as follows. In section 1, the background to the project and a portable definition of family are presented. Section 2 reviews the longitudinal studies considered in the project, and briefly discusses the potential for operationalising family composition over time in the main UK longitudinal studies. Sections 3 and 4 present the results of its implementation in the Millennium Cohort Study (MCS) and the British Household Panel Survey (BHPS). Section 5 provides a summary and a limited comparison of the results produced with the two datasets although the project did not have detailed analysis as one of its aims. Finally, a conclusion highlights pathways for possible future work.

1. Background

The UK social science community has a wealth of accessible longitudinal data. They are crucial tools for investigating complex policy-relevant issues known to have a longitudinal dimension such as the relationship between family events and children's psychological development and educational achievements. However, efforts to develop standardised classifications have focused on social class, occupation, education and ethnicity whereas family and household composition, and especially their evolution over time, do not seem to have attracted the same attention. It appears that harmonising the practice of researchers in social science for this topic is considered to be of secondary importance with published work characterised by heterogeneous uses of family definitions. Thus, for example, the

very detailed discussions about harmonisation in Hoffmeyer-Zlotnik and Harkness (2005) have very little to say about family composition, even in a cross-sectional context. As a result, the potential to corroborate and triangulate research findings in a significant number of policy-relevant areas remains untapped.

Against this background, the objective of the feasibility study was to provide an operational concept of `family', consisting of a set of variables that could be derived from most if not all the main UK longitudinal surveys, to implement it in two major longitudinal studies (MCS and BHPS), and to sketch its potential for implementation in other longitudinal studies.

Such a portable concept of family operationalised in `ready to use' variables should provide summary information about family composition at any wave of these studies, as well as between key developmental stages. These variables could, for example, contribute to defining a shared framework for carrying out longitudinal analyses of children's achievement over time or when analysing changes in family income equivalised by family composition. They could also be used as a platform on which more complex indicators of family circumstances could be built, in studies where the raw data is available.

The longitudinal studies reviewed in this project are:

- The 1958 National Child Development Study (NCDS)
- The 1970 British Cohort Study (BCS70)
- The Millennium Cohort Study (MCS)
- The British Household Panel Survey (BHPS)
- The Longitudinal Study of Young People in England (LSYPE)
- The Families and Children Study (FACS)
- The Avon Longitudinal Study of Parents and Children (ALSPAC)

Although not the focus of this project, family composition over time could also be investigated in studies where the microdata might not be directly accessible to researchers such as in the ONS Longitudinal Study (ONS LS) as well as in international studies, such as the European Union Statistics on Income and Living Conditions (EU-SILC).

One of the main challenges of operationalising family composition over time lies in finding an acceptable balance between portability and completeness. In order to maintain comparability between all of the surveys reviewed, their uneven degree of refinement needs to be tackled. This means that the scarcity of information available on family relationships in earlier surveys such as NCDS

and BCS70 -- especially the former -- needs to be taken into account in order to define the lowest common denominator. This will come at the cost of leaving aside some of the additional information present in more recent studies, such as MCS and BHPS. Users can choose to add their own dimensions to the basic definitions.

Besides testing the implementation of this generic concept of family, the specific goals of the case studies are to provide potential users with detailed information on how the family variables were derived, descriptive information about their distributions in the two datasets used, syntax files that researchers could amend to fit their own purposes and finally, datasets containing the derived variables, that could be merged with other datasets from each study. We consider birth cohorts of children and so our analyses are generally based on one focus child in each family.

1.1 Defining families

This sub-section sets out some of the issues relevant to family composition. Given the diversity of twenty first century living arrangements, a concept that is more encompassing than the traditional family is needed. The concept of marital/non-marital community of life or community of living used in the German Mikrozensus (*Lebensgemeinschaft*) is an example of an alternative approach (Lengerer and Klein, 2005). It refers to couples with or without children as well as non-couple based living arrangements, the focus being on people who live together rather than on the nature of the relationship between them. Whether or not the main carers for the child are married is not given particular attention.

We aim to create a bridge between the traditional concept of family (i.e. made up of biological parents and siblings) and the household (all those living under the same roof) by considering four groups, which may be combined in different ways. We start from studies where only families and households that include at least one child at the first wave of data collection are considered, given that this pilot project is targeted at studies where children are followed over time.

Within the household where a child lives, we can consider a basic group made up of two generations of individuals related by birth or adoption: one or two biological (or adoptive) parents living under the same roof with at least one of their children. A household can also include members of a second group: one or several people sharing a direct tie (that is neither birth nor adoption) with a member of the basic group. Typically, these are partners of one of the biological parents but could also be a third generation of individuals related by birth or adoption to members of the first group. The household may also include a subsidiary group made up of individuals who may or may not be related to members of the basic group (such as step-children or other more distant relatives/extended family). Finally,

outside the household there can be former members of the basic group who have moved out, typically estranged partners and one or several of their children.

Within this framework, additional elements are required in order to gain a more complete picture of the relationships between and within these different groups. One needs to know the gender and age of group members, especially the children, so that the order and the gap between siblings can be inferred. Further refinements are also possible that would, for instance, take into account the time spent by a child with a non-cohabiting biological parent or siblings.

This relatively simple analytical framework can be used to describe the household/family nexus at a given point in time in existing longitudinal studies, and also to monitor its evolution over time. It can account for complex combinations. There are cases which fall outside this definition, however, such as when both biological parents are missing from a household or when a child enters foster care or is adopted by a new family.

We can now translate this analytical framework into a set of indicators and define a wish list of variables that could be created.

1.2 Cross-sectional indicators

1. Household and family sizes
    i. Number of biological parents and siblings
    ii. Step-parents and number of their biological siblings, number of half-siblings
    iii. Number of other relatives
    iv. Current household size
2. Age of family members
    i. Age of cohabitating biological parents and siblings
    ii. Age of half/step-siblings
    iii. Age of a step-parent / parent's partner
    iv. Ages of non-cohabitating siblings and biological parents
3. Family type
    i. Both biological parents + children
    ii. One biological parent + children, no partner
    iii. One biological parent + children, with partner
    iv. Lone parent, biological children , partner, partner's children
    v. i to iv + other relatives
    vi. Other

1.3 Longitudinal indicators

Users should be able to measure wave to wave changes in the size and composition of the groups set out in 1.1. However, it is also useful to build longitudinal indicators, summarising change over a period of time or to capture wave to wave differences in family composition. Thus:

1. No changes

2. Someone left

   i. Biological father
   ii. Biological mother
   iii. Biological siblings
   iv. Half/step-siblings
   v. Step-parent
   vi. Other relative

3. Someone entered

   i. Birth of biological sibling
   ii. Birth of half-sibling
   iii. Step-sibling
   iv. Step-parent
   v. Other relative

4. Number of partnership changes

5. Why someone left (death, separation etc.)

6. Age of child when change occurred.

Within the limited framework of this project, we were able to retain only a subset of the indicators listed above, and to focus on two main goals. The first was to identify at each survey wave whether biological parents and siblings live in the same household as the focus child, as well as whether other cohabitees, such as half or step-siblings or other family and relatives (including a parent's partner) are also present. The second was to provide information about their evolution over time, in particular when a biological parent leaves the household, as well as the age of the focus child when such an event occurs.

2. Data availability

In this section, the main characteristics of the datasets used are outlined as well as their potential in terms of the implementation of the cross-sectional and longitudinal indicators presented in the previous section.

2.1 National Child Development Study (NCDS)

The sample for NCDS consists of all children (n = 17634) born in Great Britain in a single week in March 1958[1]. Originally concerned with mortality at or around birth, a broader range of questions was gradually added at subsequent waves in areas such as health, education, social and economic circumstances. Nine subsequent waves of data have been collected so far: at ages 7, 11, 16, 23, 33, 42, 46, 50 and 55.

Probably as a result of the dominant social norms at the time the NCDS was designed, variables relating to family composition in the early waves of NCDS are rather rudimentary. Basic information about family breakup is nevertheless available. There is partnership history data for female cohort members aged 16 and more who were partnered at waves 5 or 6 (Kallis, 2006). Details of the sample composition from birth to age 42 can be found in Plewis et al. (2004).

2.2 British Cohort Study (BCS70)

The sample for BCS70 consists of all children (n = 17287) born in Great Britain in one week in April 1970. Like NCDS, it was initially conceived as a medical study but the number of topics covered expanded with time. Eight subsequent waves of data have been collected: at ages 5, 10, 16, 26, 30, 34, 38 and 42. Although an improvement over NCDS in terms of data on family composition, information provided by BCS70 up to age five is nevertheless limited; for example, it is not known whether the father was present at birth nor between birth and the age of five. In some cases the reasons for his absence (divorce/death) are available. Some of the information not available at the first wave can be derived from retrospective questions, mostly asked at age 16 (when sample loss was severe). Details of the sample composition from birth to age 30 can be found in Plewis et al. (2004).

2.3 Millennium Cohort Study (MCS)

The MCS is a survey of children born between October 2000 and September 2001 (England and Wales), and November 2000 and January 2002 (Scotland and Northern Ireland). The sample (n = 18552), which was recruited through records of those in receipt of what was a universal benefit paid to parents called

[1] The sample was augmented by immigrants during the school years in order to maximize sample size throughout the whole cohort. This is also true for BCS70.

Child Benefit, is disproportionately stratified and clustered at the level of electoral ward. The stratification was based on UK country and the characteristics of the sampled wards. More details of the MCS sample design can be found in Plewis (2007). The data collection for the study takes place in the home and involves face-to-face interviews with multiple informants in each family. Interviews have been sought with up to two co-resident parents at every wave and, from wave two onwards, the child has also been asked to participate directly though the nature and extent of their participation has changed at each wave as the children get older. By 2012, there were five waves of the study: at 9 months (2001-2), age 3 (2003-4), age 5 (2006), age 7 (2008) and age 11 (2012). A further group of children eligible for inclusion in the wave one sample, but not interviewed then, joined the MCS between the first and second waves, and the information related to them needs to be backward-imputed using household grid information available at wave 2, as well as retrospective information available at waves 3 and 4. Given issues with item non-response that require further investigation, these children were not included in the case study described in the next section. Information about the extent and correlates of attrition up to wave three can be found in Plewis (2007), Plewis et al. (2008) and Ketende (2010).

2.4 British Household Panel Survey (BHPS)

The BHPS is a longitudinal survey of private households with a two stage stratified design set up initially in Great Britain only and extended later on to Northern Ireland. It has been running since 1991 (n = 10264 individuals) and includes 18 yearly waves. At wave 19 (in 2009), the BHPS became part of a new major longitudinal study in the UK: *Understanding Society*. The BHPS interviews are made up of a core set of questions, asked at each wave, and non-core modules on themes such as social class and health care that are rotated. Given its time span spreading over almost 20 years, the BHPS represents a valuable tool for analyses over time through which cohort and time effects can be modelled together. See Lynn (2006) for a quality profile of the first 13 waves of BHPS; Uhrig (2008) for analyses of sample loss up to wave 14; Taylor (2010) for the most recent and comprehensive user manual. The value of BHPS for analyses of family composition is explored in section 4.

2.5 Avon Longitudinal Study of Parents and Children (ALSPAC)

The ALSPAC study is a survey (n = 14541) of all pregnant mothers and their child resident in the Avon district of SW England (excluding the City of Bath) who were due to deliver between April 1 1991 and December 31 1992. Recruitment was voluntary, based on hospital visits. Postal questionnaires were used. The most recent wave of data was collected in 2009 when 12344 (85%) of the original pregnancies remained in the sample. A total of 22 waves of data are available, consisting either of data about the mother or the child itself.

Respondents are the mothers of the cohort child (sometimes their partner, whether biological father of the child or not), and no other household members are interviewed. The household grid allows other children, step/half-siblings and parents to be identified, as well as other relatives in the household. Information about the household grid has been collected regularly, at least at every other wave.

2.6 Families and Children Study (FACS)

The FACS is a panel study of British families with dependent children receiving Child Benefit with a stratified two stage design. It started in 1999 and ten waves of data had been collected by 2008. The initial sample only included lone parents and families on low income, and was subsequently extended to all families with dependent children in 2001. The stated objectives of the survey are to study the effects of work incentive measures, the effects of policy on families' living standards, and changes in family circumstances over time. See Philo et al. (2010) for more details.

At each wave, a family file, which contains all information about the main carer for the dependent children and their partner, and a child file, with information about each dependent child in the family unit, are constructed.

2.7 Longitudinal Study of Young People in England (LSYPE)

The LSYPE (now known as 'Next Steps') is a yearly panel survey of children aged 13 and 14 in 2004 (born between September 1 1989 and August 31 1990) and attending maintained schools, independent schools and pupil referral units (n = 15770). By 2011, seven waves of data had been collected (see Barnes et al., 2011 for more details). LSYPE has a stratified multi-stage design.

There are restrictions on the use of the LSYPE household grid files. Limited retrospective information about family circumstances prior to age 13 is available, including episodes of lone parenthood, changes in partnership status of the biological parent, and age of the child when this occurred. Summarised details about current household and family composition are available at each wave, but in such a way that it would be difficult to trace movements of biological or step/half-siblings outside the household.

2.8 Summary: potential of current longitudinal studies

Table 1 provides an overview of the characteristics that are the most relevant for this project in the seven surveys reviewed. The main message from this table is that all of the studies reviewed have some potential for the implementation of some of the generic family composition variables detailed in section 1. Most studies allow the identification of a minimum set of relatives cohabiting with the

focus child, consisting of biological, half- and step-siblings and parents. The two exceptions are NCDS and BCS70. In the former case, which represents the most rudimentary of all the studies reviewed, only biological parents and siblings are identified, whereas in the latter the presence of step-parents was also recorded[2].

The time frame of some of the studies allows results to be compared for completed cohorts -- children from birth until they reach age 16. This is so for NCDS, BCS70, ALSPAC, and the first cohort of children created from the BHPS (if data from the first wave of *Understanding Society* is added).

In the following two sections, we focus in more depth on the potential offered by MCS and BHPS for this project. The former represents the most recent and elaborate of the British birth cohort surveys, the latter is one of the most widely used longitudinal studies.

[2] In NCDS a household grid was filled in at age 7 (1965) and can be viewed on page 4 of the parental questionnaire (http://www.cls.ioe.ac.uk/shared/get-file.ashx?id=34&itemtype=document). Only some of the resulting data is available in the dataset.

Table 1: Summary characteristics of the reviewed longitudinal studies

	Population	Years	Waves	Household relationships	Frequency	Child outcomes	Comments
NCDS	Children born in GB in 1958	1958, ongoing	10	Biological parents/ siblings	Age 0, 7, 11, 16, 23, 33, 42, 46, 50 and 55 years	Yes	No information about events between waves during childhood.
BCS70	Children born in GB in 1970	1970, ongoing	9	Biological parents/ siblings; age at event	Ages 0, 5, 10, 16, 26, 30, 34, 38, 42 years	Yes	Retrospective data prior to age 5
MCS	Children born in the UK in 2000/2001	2000, ongoing	5	Biological, half/step-siblings \ parents; relatives \ partners; age at event	Ages 9 months, 3, 5, 7, 11 years	Yes	-
LSYPE/ Next Steps	Children in year 9 (aged 13 & 14) in 2004 in England	2004 to 2010	7	Biological, half/step-siblings \ parents; relatives \ partners; age at event	Yearly	Yes	Retrospective data prior to age 13; restricted access to household grid
BHPS	Households in GB[1]	1991/ 2009[2]	18	Biological, half/step-siblings \ parents; relatives \ partners	Yearly	At some waves	Cohorts need to be derived from the data
ALSPAC	Children born in Avon in 1991/1992	31/12/1991-2009	22	Biological, half/step-siblings \ parents; relatives \ partners	Every other wave	Yes	Data not currently available through the UK Data Archive
FACS	Families in GB receiving Child Benefit	1999 to 2008	10	Biological, half/step-siblings \ parents; relatives \ partners	Yearly	No	Waves 1 and 2 restricted to low income families

Notes

[1] Northern Ireland added in 2001.

[2] Incorporated into Understanding Society in 2009.

3. The MCS case study

This section presents the methods used in the implementation of the family composition variables in MCS as well as a few preliminary results. MCS builds upon the expertise gathered in the two earlier cohort surveys, NCDS and BCS70 and, as far as the relationships within a household are concerned, provides more detailed data than the previous two studies. The section is structured as follows. First, we present a summary of the information available to users interested in families and their evolution over time. The next sub-section introduces the generic family composition variables developed in the project and their rationale, while the last presents some descriptive statistics.

3.1 Families in the MCS

Information about any household member in MCS is recorded in a separate dataset (mcsX_household_grid.dta) at each wave (X). Table 2 presents a sample of the variables available at each wave in the household datasets. Household grid data can be merged with other MCS datasets at the family level using the MCSID variable .

Table 2: Main variables in MCS household grid

Name	Label
xHCNUM00	Cohort member number
xHCSEX00	Cohort member sex
xHCDBM00	Cohort member date of birth (month)
xHCDBY00	Cohort member date of birth (year)
xHPNUM00	Person number
xHPSEX00	Person sex
xHPDBM00	Person date of birth (month)
xHPDBY00	Person date of birth (year)
xHPRES00	Present in wave x
xHPAGE00	Person age (years)
xHCREL00	Relationship to cohort member
xHPRELA0...	Relationship to person 1...
xHPRELR0	Relationship to person 18

Note

X: wave number

It is possible to follow individual household members using the cohort ID number (MCSID) and the Person number (xHPNUM00). Multiple births can be identified using xHCNUM. In this study one cohort member per family was considered.

Information included in the household grid allows users to identify a wide variety of relationships. Table 3 shows the relationships between each household member and the cohort child that can be identified using the xHCREL00 variables. In addition, starting from waves 3 and 4, retrospective variables about former household members who left at previous waves are available (Table 4).

Table 3: Main relationships within MCS households

Code	Label
7	Biological parent
8	Adoptive parent
9	Foster parent
10	Step-parent/partner of parent
11	Biological brother/sister
12	Half-brother/sister
13	Step-brother/sister
14	Adopted brother/sister
15	Foster brother/sister
16	Grandchild
17	Grandparent
18	Nanny/au pair
19	Other relative
20	Other non-relative

Note

Categories of the xHCREL00 variable in MCS (Waves 1-4). Categories 1-6 are not relevant to cohort members who are children and are omitted; codes for missing values are also omitted;

Table 4: Retrospective variables in MCS at waves 3 and 4

Name	**Label**
xHPSTY00	When did person start living with cohort member (year)
xHPSTM00	When did person start living with cohort member (month)
xHPWHP00	What happened to person
xHPDCY00	When did person die (year)
xHPDCM00	When did person die (month)
xHPSPY00	When did person stop living with cohort member (year)
xHPSPM00	When did person stop living with cohort member (month)

3.2 Derivation strategy

Given that the MCS household grid datasets include observations about each member of the household, we use these variables rather than the existing household-level derived variables also available in the Parent Interview dataset. Although this makes derivation of the family composition variables more complex, it allows potential users to tailor them more easily to their needs. Users interested in following specific (non-cohort child) household members are advised to use the relevant xPNUM variables.

Derivation consisted of constructing binary indicators recording the presence of each relevant household member at a given wave, and merging them into a single family relationship variable. Longitudinal indicators were built by recording wave to wave differences in the cross-sectional variables. The STATA code for the constructed indicators is labelled 'Data 1' along with this report on the CMIST website.

3.3 Descriptive statistics

Here we present some frequency tables for the family composition variables, focusing first on indicators available at each wave, then on some longitudinal summary variables. Table 5 presents estimates of the percentages of each type of family member in the household of the MCS cohort child: biological parent, siblings, half and step-siblings, other relatives. These results show expected trends: the number of siblings, irrespective of their type, increases with time, as does the number of families with just one biological parent whose percentage almost doubled between 9 months and age 7. The results are comparable with those given in Calderwood (2007, 2008, 2010).

Table 5: Family composition variables (%) in MCS, waves 1 to 4

	W1: 9 months		W2: 3 years		W3: 5 years		W4: 7 years	
	Number of biological parents							
None	*	n.a.	0.7	*0.5 - 0.8*	*	*n.a.*	*	*n.a.*
1	14.2	*13.3 - 15.1*	17.9	*16.9 - 18.9*	22.6	*21.4 - 23.8*	25.5	*24.2 - 26.8*
2	85.8	*84.8 - 86.7*	81.4	*80.4 - 82.4*	77.1	*75.9 - 78.3*	74.0	*72.8 - 75.3*
	Number of biological siblings in the household							
None	51.7	*50.9 - 52.6*	33.5	*32.6 - 34.4*	24.8	*24.0 - 25.6*	21.1	*20.3 - 21.9*
1	33.2	*32.4 - 34.1*	46.1	*45.2 – 47.0*	48.6	*47.7 - 49.6*	48.0	*47.0 – 49.0*
2	10.8	*10.3 - 11.4*	14.8	*14. 1- 15.4*	19.3	*18.6 - 20.1*	22.3	*21.4 - 23.1*
3 +	4.2	*3.9 - 4.5*	5.6	*5.2-6.0*	7.3	*6.9 - 7.8*	8.7	*8.1 - 9.2*
	Number of half or step-siblings in the household							
None	87.9	*87.3 - 88.4*	87.1	*86.5 - 87.7*	85.5	*84.8 - 86.1*	83.7	*82.9 - 84.4*
1	7.4	*6.9 - 7.8*	8.2	*7.7 - 8.8*	9.5	*9.0 - 10.1*	10.8	*10.1 - 11.3*
2	3.5	*3.2 - 3.8*	3.4	*3.1 - 3.8*	3.8	*3.4 - 4.1*	4.3	*4.0 - 4.8*
3 +	1.3	*1.1 - 1.5*	1.2	*1.1 - 1.5*	1.2	*1.0 - 1.4*	1.3	*1.1 - 1.6*
	Number of other relatives (inc. partner of a biological parent)							
None	91.9	*91.5 - 92.4*	94.2	*93.8 - 94.6*	93.8	*93.4 - 94.3*	89.4	*88.8 – 90.0*
1	2.8	*2.5 - 3.1*	3	*2.8 - 3.4*	3.8	*3.5 - 4.2*	8.2	*7.7 - 8.7*
2	2.4	*2.1 - 2.6*	1.5	*1.3 - 1.8*	1.5	*1.3 - 1.7*	1.5	*1.3 - 1.8*
3 +	2.9	*2.7 - 3.2*	1.2	*1.0 - 1.4*	0.9	*0.7 -1.0*	0.9	*0.7 – 1.0*
N	18552		15583		15240		13851	

Notes

1. Excludes second and higher multiple births.
2. All % weighted by the sample design weights ('weight2').
3. * - < 0.5%
4. 95% CI in italics, allowing for sample design using *svy* commands in STATA.

Table 6 looks at the same indicators from another angle, and provides the distribution of cohort children by broad family type: two biological parent families, with or without siblings, and lone biological parent families subdivided into additional categories according to whether a partner, step - children or additional members live together in the household with the cohort child.

Table 6: Combined family composition variable (%) in MCS, waves 1 to 4

Type	W1: 9 months		W2: 3 years		W3: 5 years		W4: 7 years	
	Two biological parent families							
Only child	42.1	*41.3 – 43.0*	22.5	*21.7 - 23.3*	13.3	*12.7 – 14.0*	10.1	*9.5 - 10.7*
Siblings	43.7	*42.8 - 44.5*	59.2	*58.3 - 60.1*	63.8	*62.9 - 64.7*	64.8	*63.9 - 65.8*
	Lone parent families							
Single	13.9	*13.4 - 14.5*	15.6	*15.0 - 16.2*	18.2	*17.5 – 19.0*	18.9	*18.2 - 19.7*
+ partner	*	*n.a.*	1.2	*1.0 - 1.4*	1.2	*1.0 - 1.4*	2.3	*2.0 - 2.6*
Half/stepchildren	*	*n.a.*	2.4	*2.1 – 2.7*	3.5	*3.1 – 3.8*	3.4	*3.0 - 3.8*
Other	*	*n.a.*	0.6	*0.5 - 0.8*	*	*n.a.*	*	*n.a.*
N	18552		15583		15240		13851	

See notes to Table 5.

As already mentioned, longitudinal indicators are of two types: those computed from the cross-sectional variables, and those resulting from retrospective questions. Table 7 presents estimates of percentages and confidence intervals for the former type. It is important to note that these are based on children present at all four waves.

Table 7: Family changes in MCS by wave 4

Type	**%**	**95% CI**
Whether a biological parent absent at least at one wave		
Father	25.0	*24.1 - 25.9*
Mother	1.2	*1.0 - 1.5*
Number of changes in family type		
None	81.9	*81.0 - 82.7*
One	15.8	*15.0 - 16.5*
Two or more	2.3	*2.0 - 2.8*
Whether change at any wave in the number of :		
Biological siblings	43.9	*42.9 - 45.0*
Other siblings	8.2	*7.7 - 8.8*
Other relatives	11.7	*11.1 - 12.4*
N	11716	

Table 8 is derived from the data used to construct Table 6 and gives the wave by wave estimated transition probabilities of a change in the number of biological parents.

Table 8: Between wave transitions in MCS (%): number of biological parents in the household

	W1 – W2	W2 – W3	W3 – W4	W1 – W3	W2 – W4	W1 – W4
1 – 1	9.9 *[9.4 - 10.4]*	15.4 *[14.7 - 16.1]*	19.6 *[18.8 - 20.4]*	10.1 *[9.6 - 10.7]*	14 *[13.4 - 14.7]*	9.3 *[8.8 - 9.9]*
1 – 2	2.3 *[2.1 - 2.6]*	1.3 *[1.1 - 1.5]*	1.4 *[1.2 - 1.6]*	2.3 *[2.1 - 2.6]*	1.8 *[1.6 - 2.1]*	2.4 *[2.2 - 2.7]*
2 - 1	6.9 *[6.4 - 7.4]*	5.5 *[5.0 - 5.9]*	4.4 *[4.0 - 4.8]*	11.8 *[11.2 - 12.4]*	9.1 *[8.6 - 9.7]*	14.7 *[14.0 - 15.4]*
2 - 2	80.5 *[79.8 - 81.2]*	77.3 *[76.4 - 78.0]*	74.2 *[73.3 - 75.1]*	75.6 *[74.8 - 76.4]*	74.4 *[73.5 - 75.3]*	73.2 *[72.4 - 74.1]*
N	14891	13796	13213	14672	12653	13357

Note

Based on number of children in each pair of waves, ignoring cases where there were no parents at either wave.

3.4 Additional information available in the MCS

As can be seen in Table 2, the MCS offers the opportunity to add indicators which identify more precisely the relationship of household members to the cohort child, such as grandparents, step, foster or half-siblings. This is also true of the age of siblings, which enables users to compute birth orders and age differences between children. Interested users should, however, be aware of the small numbers of observations sometimes involved.

Retrospective variables at waves 3 and 4 allow users to identify the year and the main reason why a particular household member left (Including death). See Table 9 for one use of these data. However, the potential of this variable is limited by item non-response, and the fact that it has not been made consistent with other variables in the household grid so far. Information about the presence of one or two carers and alternative typologies of families is also available in the parent interview.

Table 9: Age of cohort child (%) when father left the household, MCS

	MCS	
	Age 7	
Did not leave	76.4	*75.5 - 77.3*
Absent at birth – up to under one year	7.7	*7.2 - 8.3*
1-3 years	6.2	*5.7 - 6.7*
4-6 years	6.3	*5.8 - 6.8*
7+ years	3.3	*3.0 - 3.7*
N	11280	

4. The BHPS case study

This section describes the implementation of the family composition variables using BHPS data. Besides creating family indicators similar to those developed for MCS, using the BHPS in the framework of the feasibility study required pooling several waves together in order to create cohorts of newborn children that could subsequently be followed over time. Although this comes at the cost of cohort size, greatly reduced by contrast with the birth cohort studies, potential users of the variables will gain access to a wealth of variables available in BHPS. Depending on the sample size needed, users can also opt to create larger cohorts spanning longer periods of time.

This section is structured in the same fashion as the previous one: first, we present a brief introduction to the information available in the BHPS about families and their evolution over time. We then discuss derivation issues before presenting frequency tables of the family composition variables.

4.1 Families and children in BHPS

Generating data for families in BHPS is a more complex task than it is in the UK birth cohort studies, since it is households and their members that are the focus of the study, rather than children and families. As a result, cohorts consisting of children of the same age from several consecutive waves need to be pooled together in order to enable comparisons with the cohort studies.

Although BHPS is a survey of households, there are three types of individual members in BHPS, all of whom can be identified across waves using their unique person number (PNO):

- Original sample members (OSM): anyone who was in the household at wave 1 (in 1991), as well as biological child(ren) born to them at subsequent waves, who become OSM in their own right.
- Permanent sample members (PSM): usually the biological parent of an OSM (i.e. the partner of an OSM who moved into her/his household after wave one).
- Temporary sample members (TSM) who share a household with an OSM.

OSM and PSM are followed over time, in the latter case since their initial inclusion in the sample. However, TSM and their children (typically the partner of an OSM and their child not born to an OSM) are not followed, and therefore step-siblings of children who are OSM cannot be followed once they have left the household. Finally, a cross-wave identifier of the mother of a child OSM is available from Wave D onwards.

In each of the first ten waves of BHPS, between 118 and 186 children aged under one year are present. Two cohorts were created by combining five waves together. The first cohort was made up of the first five waves (waves 1-5, 1991-1995) which were subsequently followed for a further 13 waves. The second cohort consisted of the next five waves (waves 6-10, 1996-2000), which were followed for a further eight waves. Table 10 provides a detailed overview of the composition of these two cohorts.

Table 10: Sample size by wave, cohorts 1 and 2 in BHPS

Cohort	Wave	Age		% missing
		Under 1 year	14/9 years	
1: waves 1 to 5	A-1991	186	119	36
	B-1992	118	73	38
	C-1993	145	93	36
	D-1994	150	96	36
	E-1995	138	92	33
	Total	737	493	36
2: waves 6 to 10	F-1996	118	91	23
	G-1997	131	101	23
	H-1998	125	94	25
	I-1999	132	102	23
	J-2000	126	99	21
	Total	632	487	23

Notes

1. Number of infants (waves 1991-2000, ignoring booster samples) and number with data from all waves remaining at ages 14 (1st cohort) and 9 (2nd cohort).

4.2 Relationships within the household in BHPS

Table 11 shows the relationships to a cohort child that can be identified within households in the BHPS, using xREL (x is wave identifier). The categories of relationships are almost as detailed as in MCS, with the exception that BHPS does not permit the identification of adoptive parents or siblings in the household grid who are therefore treated as biological[3]. Partners of a biological parent are automatically coded as step-parents.

Table 11: Relationships within BHPS households

Code	Label
10	Biological brother/sister
11	Other brother/sister
13	Biological parent
14	Other parent
16	Any grandparent
18	Any cousin
19	Any aunt/uncle
20	Any nephew/niece
21	Any other relative
24	Unrelated sharer
25	Step-parent
28	Half-sibling
30	Other

Note

BHPS, household grid in the xREL variables (xEGOALT records). Categories not relevant to children cohort members and codes for missing values are omitted.

4.3 Operationalisation and data

The BHPS is split into several datasets (also called *records*) at each wave. The ones directly relevant to this paper are xHHRESP which contains information about the household grid, xINDALL where basic information about all members of the household is gathered, including those not qualifying for interview -- which is the case for children under 16 - and xEGOALT, which derives information about

[3] Adoptive parents in the BHPS are included with 'other parents' and can be identified by a code for adoptive child. This was not used given the limited scope of this study.

each pair of relationships within households. The main individual respondent records, xINDRESP are not of direct interest here since only adults aged 16 and above were interviewed, unless users want to incorporate additional information about relatives of the cohort children. Finally, xHHSAMP is also used in order to add information about primary sampling units and strata, necessary in order to compute accurate confidence intervals, and about the weights (although these were not used in the following tables).

The strategy for building the family variables was as follows. At each `initial' wave -- in which children aged under one in December of the year of interview are selected -- the household datasets are merged with the respective xINDALL files, and only children under one are retained. These are then merged to the xEGOALT record, where household-level variables were constructed which identify every relationship between the cohort child and any other household member. The wave 1/age 0 dataset of the cohort is then merged to the subsequent 13 or 8 waves of data, following the same logic. Each time a new wave of data is merged, the household relationship indicators are built. Finally, the cohorts are assembled by pooling five waves together (waves A-E and F-J in Table 10). At the last stage, the longitudinal indicators are computed. It should be noted that, as biological parents may leave the household where a cohort child was born, households do not necessarily remain the same over time. As in MCS, only children selected at wave 1 are included in the dataset on all occasions at which information about them was collected. The STATA code for the constructed indicators is labelled 'Data 2' along with this report on the CMIST website.

4.4 Descriptive statistics

Table 12 presents the distribution of the family composition variables derived from the two cohorts.

Table 12: Family composition variables (%) in BHPS, cohorts 1 and 2

	Cohort 1 (1991-1995)						Cohort 2 (1996-2000)					
	Age 0		Age 7		Age 14		Age 0		Age 7		Age 9	
Number of biological parents												
None	*	n.a.	*	n.a.	*	n.a.	*	n.a.	*	n.a.	*	n.a.
1	12.3	*9.3 - 15.4*	22.8	*19.2 - 26.4*	36.4	*31.7 - 41.0*	9.7	*7.2 - 12.1*	20.4	*16.8 - 24.1*	23.4	*9.5 - 27.4*
2	87.5	*84.5 - 90.5*	76.7	*73.1 - 80.3*	63.4	*58.8 - 68.1*	90.3	*87.9 - 92.8*	79.6	*75.9 - 83.2*	76.6	*72.6 - 80.5*
Number of biological siblings												
None	47.1	*43.5 - 50.7*	18.9	*15.5 - 22.2*	18.2	*14.7 - 21.7*	53.0	*49.4 - 56.6*	15.2	*12.1 - 18.4*	14.0	*11.0 - 17.0*
1	35.1	*32.1 - 38.2*	51.6	*47.1 - 56.2*	51.3	*46.6 - 57.1*	33.2	*28.7 - 36.8*	55.0	*50.1 - 60.0*	54.3	*49.4 - 59.2*
2	12.1	*9.9 - 14.3*	21.8	*18.0 - 25.6*	23.5	*19.4 - 27.6*	10.0	*7.8 - 12.1*	23.4	*19.1 - 27.7*	23.6	*19.4 - 27.8*
3 +	5.7	*3.7 - 7.7*	7.7	*4.9 - 10.5*	7.6	*4.4 - 9.6*	3.8	*2.0 - 5.6*	6.8	*3.9 - 8.7*	8.1	*5.1 - 11.0*
Number of other siblings												
None	92.0	*89.8 - 94.2*	90.7	*88.1 - 93.2*	91.5	*88.9 - 94.1*	89.2	*86.4 - 92.1*	90.5	*87.8 - 93.3*	91.2	*88.3 - 94.0*
1	4.5	*3.0 - 6.0*	6.6	*4.4 - 8.7*	6.1	*3.9 - 8.2*	6.6	*4.4 - 8.9*	6.5	*4.1 - 8.9*	6.0	*3.6 - 8.3*
2	2.7	*1.3 - 4.2*	2.6	*1.2 - 4.1*	1.7	*0.4 - 3.0*	2.1	*0.8 - 3.3*	2.6	*1.1 - 4.1*	2.7	*1.1 - 4.2*
3	0.8	*0.4 - 1.7*	*	*n.a.*	0.8	*0.0 - 1.7*	2.1	*0.8 - 3.3*	*	*n.a.*	*	*n.a.*
Number of other relatives												
None	92.9	*90.9 - 95.0*	89.2	*86.4 - 91.9*	83.5	*80.0 - 87.2*	93.2	*91.1 - 95.3*	90.7	*88.1 - 93.4*	87.7	*84.6 - 90.8*
1	2.7	*1.6 - 3.9*	9.5	*6.9 - 12.1*	15.9	*12.3 - 19.5*	2.7	*1.2 - 4.1*	8.0	*5.6 - 10.4*	10.7	*7.8 - 13.7*
2	1.6	*0. -2.5*	1.0	*0.2 - 1.8*	*	*n.a.*	1.6	*0.6 - 2.5*	0.6	*0.0 - 1.2*	1.2	*0.2 - 2.1*
3 +	2.7	*1.4 - 4.1*	*	*n.a.*	*	*n.a.*	2.5	*1.2 - 3.9*	0.7	*0.0 - 1.6*	*	*n.a.*
N	737		610		528		632		538		521	

Notes

1. * - < 0.5%

2. 95% CI in italics, allowing for clustering but not stratification using *svy* commands in STATA.

The limited sample size in the two cohorts results in wide confidence intervals for most estimates in Table 12. Nevertheless, it is apparent that the estimated percentages remain close to each other in the two cohorts, suggesting that the results are robust. Table 13 also shows reasonable consistency across the two cohorts although the second cohort has fewer lone parent families at ages 0 and 7.

Table 13: Combined family composition variables (%) in BHPS, cohorts 1 and 2

Cohort 1 (1991-1995)							Cohort 2 (1996-2000)					
	Age 0		Age 7		Age 14		Age 0		Age 7		Age 9	
Two biological parent families												
Only	39.9	*36.4 - 43.3*	11.8	*9.0 - 14.6*	8.3	*5.7 - 11.0*	49.1	*45.4 - 52.7*	10.2	*7.6 - 12.8*	8.8	*6.3 - 11.3*
Siblings	47.6	*43.8 - 51.5*	64.9	*60.7 - 69.2*	55.1	*50.3 - 59.9*	41.3	*37.5 - 45.1*	69.3	*65.2 - 73.4*	67.8	*63.5 - 72.0*
Lone parent families												
Single	12.3	*9.3 - 15.4*	16.7	*13.7 - 19.7*	22.2	*18.2 - 26.1*	9.2	*6.8 - 11.6*	14.3	*11.1 - 17.5*	15.4	*11.9 - 18.8*
Partner	*	*n.a.*	6.6	*4.4 - 8.7*	14.4	*10.9 - 17.8*	0.5	*0.0 - 1.2*	6.1	*4.0 - 8.3*	8.1	*5.5 - 10.7*
N	737		610		528		632		538		521	

Table 14 also shows consistency across the two cohorts. For example, about two thirds of the cohort children are living with siblings at age seven and about 15% of them are in single parent families at that age.

Table 14: Family changes (%), BHPS cohorts 1 and 2

	Cohort 1 (1991-1995)				Cohort 2 (1996-2000)			
	Age 7		Age 14		Age 7		Age 9	
	Whether a biological parent was absent at least at one wave							
Father	22.4	*18.2 - 26.6*	36.3	*31.5 - 41.2*	21.3	*17.4 - 25.3*	24.4	*20.4 - 28.5*
Mother	1.5	*0.4 - 2.6*	3.8	*1.9 - 5.7*	1.0	*0.1 - 1.9*	1.2	*0.3 - 2.2*
	Number of changes in family type (Lone parent vs couple)							
None	82.0	*78.1 - 85.4*	66.6	*62.0 - 71.2*	84.2	*80.4 - 87.4*	80.9	*77.2 - 84.6*
One	15.6	*12.5 - 19.4*	28.3	*23.8 - 32.8*	12.1	*9.4 - 15.4*	14.4	*11.2 - 17.5*
Two or more	2.3	*1.3 - 4.2*	5.1	*3.1 - 7.1*	3.7	*2.3 - 6.0*	4.7	*2.7 - 6.8*
	Whether a change occurred at any wave in the number of:							
Biological siblings	45.5	*41.1 - 49.9*	54.5	*50.0 - 59.1*	51.1	*46.9 - 55.4*	53.6	*49.3 - 57.9*
Other siblings	6.8	*4.6 - 9.8*	16.3	*12.8 - 19.8*	8.0	*5.5 - 11.5*	11.1	*7.6 - 14.5*
Other relatives	14.6	*11.3 - 18.6*	29.1	*22.6 - 31.5*	16.2	*12.9 - 20.1*	20.3	*16.3 - 24.4*
N	473				487			

We see in Table 14 that the percentages of mother and father absence at age seven are similar for the two cohorts. Tables 15 and 16 give the transition percentages for the two cohorts with a suggestion that there is more stability in cohort two than in cohort one.

Table 15: Between wave transitions (%) in the number of biological parents in the household, BHPS cohort 1

	W1 – W3	**W3 – W5**	**W5 – W7**	**W1 – W5**	**W3 – W7**	**W1 – W7**	**W1 – W14**
1 – 1	9.1 *(6.6 - 12.4)*	12.9 *(10.0 - 16.6)*	17.0 *(13.7 -20.8)*	8.0 *(5.8 - 11.1)*	12.2 *(9.6 - 15.7)*	7.1 *(5.1 - 9.6)*	6.8 *(4.8 – 9.6)*
1 – 2	2.3 *(1.4 - 3.7)*	1.9 *(1.1 - 3.4)*	1.2 *(0.6 - 2.4)*	2.8 *(1.8 - 4.5)*	2.2 *(1.3 - 3.7)*	3.1 *(2.0 - 4.9)*	2.5 *(1.3 – 4.4)*
2 - 1	6.8 (*5.0 - 9.3)*	7.0 *(5.3 - 9.3)*	5.2 *(3.7 - 7.3)*	11.8 *(9.4 - 14.9)*	10.4 *(8.3 - 13.2)*	15.7 *(12.9-19.1)*	29.6 *(25.4 – 34.1)*
2 - 2	81.7 *(77.6 -85.1)*	77.8 *(73.8 - 81.4)*	76.0 *(72.2 -79.5)*	77.0 *(72.8 -80.7)*	74.6 *(70.6 -78.2)*	73.6 *(69.6 -77.3)*	61.0 *(56.3 – 65.5*)
n	660	627	601	634	602	610	528

Table 16: Between wave transitions (%) in the number of biological parents in the household, BHPS cohort 2

	W1 – W3	W3 – W5	W5 – W7	W1 – W5	W3 – W7	W1 – W7	W1 – W9
1 – 1	7.4 *(5.4 - 10.0)*	10.1 *(7.6 - 13.3)*	14.7 *(11.7 - 18.4)*	7.3 *(5.3 - 10.0)*	9.4 *(7.0 - 12.6)*	7.1 *(5.1 - 9.8)*	7.1 *(5.1 - 9.9)*
1 – 2	1.2 *(0.6 - 2.5)*	0.9 *(0.4 - 2.2)*	**(n.a.)*	0.9 *(0.4 - 2.1)*	0.9 *(0.4 - 2.2)*	0.9 *(0.4 - 2.2)*	1.0 *(0.4 - 2.3)*
2 - 1	4.4 *(3.0 - 6.5)*	5.8 *(4.1 - 8.0)*	5.5 *(3.9 - 7.7)*	8.7 *(6.6 - 11.5)*	11.1 *(8.6 - 14.2)*	13.4 *(10.6 - 16.8)*	16.3 *(13.3 - 19.9)*
2 - 2	86.7 *(83.2 - 89.5)*	83.0 *(79.4 - 86.1)*	79.4 *(75.4 - 83.0)*	83.1 *(79.3 - 86.2)*	78.3 *(74.3 - 81.9)*	78.6 *(74.6 - 82.2)*	75.6 *(71.4 - 80.3)*
n	585	554	530	561	531	538	521

Table 17 shows that the percentage of fathers absent at birth was similar for the two cohorts and suggests that the propensity for fathers to leave the household initially increases after birth but then declines after age seven. A more detailed survival analysis model would, however, be needed to confirm this.

Table 17: Age of child (%) when father left the household, BHPS cohorts 1 and 2

	Cohort 1 (1991-1995)		Cohort 2 (1996-2000)	
	Age 14		Age 9	
Did not leave	63.6	*58.8 - 68.5*	75.6	*71.5 - 79.6*
Absent at birth	8.9	*5.9 - 11.8*	8.0	*5.4 - 10.7*
Under 1 year	2.3	*1.0 - 2.7*	1.2	*0.3 - 2.2*
1-3 years	7.4	*4.8 - 10.0*	7.2	*5.0 - 9.4*
4-6 years	7.4	*5.1 - 9.7*	7.6	*5.3 - 9.9*
7-9 years	4.9	*2.9 - 6.8*	0.4	*0.1 - 1.6*
10-14 years	5.5	*3.4 - 7.6*	-	
N	473		487	

4.5 Additional information available in the BHPS

The BHPS contains a wealth of variables on many different topics. A comprehensive list of the variables related to children and childhood is available on the website of the Institute for Social and Economic Research (ISER). In order to be able to use this additional information, users will need the Person Number (PNO) of a child's biological parent and use it to merge the cohort datasets with either the xCHILD or xINDRESP variables at the relevant waves.

5. Comparing results from MCS and BHPS

This final section summarises the results obtained from MCS (infants and children aged seven) and BHPS (infants and children aged seven in each cohort) followed by some limited comparisons across the two studies, focusing on the core indicators of family type and their change over time.

For MCS, the percentage of co-resident biological parents decreases at each wave; the percentages of biological and other siblings both increase at each wave whereas the percentages of other relatives show little change. By age seven, about a quarter of fathers are absent at least once (including at birth) and 18% of families experience at least one change in family type. The transition probabilities increase with wave for staying in a single parent family but decline for transitions into and out of single parent families. Similar results are obtained for BHPS. However, the percentage of other relatives rises more sharply, especially just one other relative, and transitions out of two parent families tend to increase with age. Further comparisons are brought out in Tables 18 and 19 with BHPS appearing to have fewer lone parents. This could be due to the different research design of the two studies, which, in the case of the MCS, may make it easier to retain lone parents engaged with the study.

It should be borne in mind that some of the differences, both across ages and between studies, are likely to be due to missing data. Both MCS and BHPS suffer from attrition with evidence from MCS suggesting that lone parents were more likely to be lost from the study after wave one (Plewis et al., 2008). It is also important to recognise that members of the second BHPS cohort are survivors (in terms of attrition) for at least five waves longer than members of the first cohort and they might therefore have different socio-economic characteristics. More detailed analyses than is possible here would need to take differential attrition into account.

Table 18: Family composition variables (%) in MCS and BHPS, cohorts 1 and 2

	MCS		BHPS (1991-1995)		BHPS (1996-2000)	
Two biological parent families						
	Under 1	Age 7	Under 1	Age 7	Under 1	Age 7
Only child	40.1	9.9	39.9	11.8	49.1	10.2
Siblings	44.8	64.2	47.6	64.9	41.3	69.3
Lone parent families						
Single	14.0	19.6	12.3	16.7	9.2	14.3
Partner	*	2.5	*	6.6	0.5	6.1
Stepchildren	*	3.4	n.a.	n.a.	n.a.	n.a.
Other	*	*	n.a.	n.a.	n.a.	n.a.
N	18552	13857	737	610	632	538

Comparable results were found with the longitudinal variables shown in Table 16, with between about 27% of children having experienced the absence of their biological father at least at one wave by the age of 7 or 9. The fact that the results for the MCS and the second BHPS cohort are very similar should be read as additional evidence of an attrition issue with the latter cohort -- it could be expected that at the next wave of the MCS, by the time the children have reached the age of nine, these figures will have moved closer to those from the first BHPS cohort, that is with more children with a biological parent absent at least at one wave, and less of them who did not experience any change in their family type.

Table 19: Whether a biological parent was absent for at least at one wave (%), MCS and BHPS

Type	MCS	BHPS (1991-1995)	BHPS (1996-2000)
	Age 7	Age 7	Age 7
Father	25.2	22.4	21.3
Mother	1.3	1.5	1.0
N	11721	473	487

.

Conclusion

This report has highlighted the potential and limits as well as a few results of a harmonisation project that involved cross-sectional and longitudinal indicators of family composition in the main longitudinal studies available in the UK. The main conclusion that can be drawn is that the results are encouraging, both in terms of the actual feasibility of an implementation of common indicators in heterogeneous studies, and in terms of the distribution of observations within the categories of the family composition variables created.

Further work is nevertheless needed. Implementing the family composition variables in additional studies would be necessary in order to confirm the robustness of the results described in this report, especially in older studies with data quality issues such as the NCDS and the BCS70. Additional refinement of the family composition indicators, for instance by following partnership histories of lone parents, could augment the existing set of variables, as well as providing more detail about biological and step-siblings in households.

References

Barnes, M., Green, R. and Ross, A. (2011) *Understanding Vulnerable Young People: Analysis from the Longitudinal Study of Young People in England*. Research Report DFE_RR118. London: Department for Education.

Calderwood, L. (2007) Family Demographics. In Hansen, K. and Joshi, H. (Eds) *MCS, Second Survey: A User's Guide to Initial Findings*. London: Institute of Education.

Calderwood, L. (2008) Family Demographics. In Hansen, K. and Joshi, H. (Eds*) MCS, Third Survey: A User's Guide to Initial Findings*. London: Institute of Education.

Calderwood, L. (2010) Family Demographics. In Hansen, K., Jones, E., Joshi, H. and Budge, D. (Eds) *MCS, Fourth Survey: A User's Guide to Initial Findings*. London: Institute of Education.

Hoffmeyer - Zlotnik, J. H. P. and Harkness, J. A. (Eds) (2005) *Methodological Aspects in Cross-national Research*. ZUMA Spezial Band 11. Mannheim: ZUMA.

Kallis, C. (2006) *Partnership Histories in NCDS 5 and NCDS 6.* CLS Cohort Studies Research Note 5. London: Institute of Education.

Ketende, S. (Ed.) (2010) *The Millennium Cohort Study: Technical Report on Response.* 3rd. Edition. London: Institute of Education.

Lengerer, A. and Klein, T. (2007) Der Langfristige Wandel Partnerschaftlicher Lebensformen im Spiegel des Mikrozensus. *Wirtschaft und Statistik*, 04, 433-447.

Lynn, P. (Ed.) (2006) *Quality Profile: British Household Panel Survey. Version 2.0: waves 1 to 13: 1991 – 2003*. Colchester: University of Essex.

Philo, D., Hoxhallari, L., Conolly, A. and Barnes, M. (2010) *Families and Children Study (FACS) Waves 1 to 10 User Guide (1999 – 2008)*. London: National Centre for Social Research.

Plewis, I., Calderwood, L., Hawkes, D. and Nathan, G. (2004) *Changes in the NCDS and BCS70 Populations and Samples over Time.* CLS Technical Report. London: Institute of Education.

Plewis, I. (Ed.) (2007) *The Millennium Cohort Study: Technical Report on Sampling 4th Edition*. London: Institute of Education.

Plewis, I., Ketende, S., Joshi, H. and Hughes, G. (2008) The Contribution of Household Mobility to Sample Loss in a Birth Cohort Study: Evidence from the First Two Waves of the UK Millennium Cohort Study. *Journal of Official Statistics*, 24, 365-385.

Taylor, M. F. (Ed.) (2010) *British Household Panel Survey User Manual Volume A: Introduction, Technical Report and Appendices.* Colchester: University of Essex.

Uhrig, S. C. N. (2008) *The Nature and Causes of Attrition in the British Household Panel Study*. ISER Working Paper 2008-05. Colchester: University of Essex.