HOUSEHOLD AND FAMILY DATA
2001

CONTENTS

SECTION 1 - INTRODUCTION

The Labour Force Survey (LFS) has traditionally been used to provide information on individuals in the labour market. But since the survey gathers information about complete households, data are available for households and families as well as for individuals. This is important information, because the ways in which people behave in relation to the labour market are influenced by their household and family circumstances, and by the characteristics and behaviour of the other members of their family and household. For example, in a married or cohabiting couple, one partner's decision to work or seek work may be influenced by whether the other partner is employed, unemployed or not in the labour force.

Equally important is the effect on the household's economic circumstances and well-being of the combination of economic activities experienced by all the adult members of the household. For example, the increasing tendency for some households to consist either of members who are all working, or of members who are all non-working, in comparison with the pattern more commonly seen in the past of households including someone who is working and someone who is not, has important effects on the distribution of household income and wealth.

The LFS is a unique source of detailed information about the ways that households and families behave in relation to the labour market. Its data on the characteristics of households and families can also be used for more general demographic purposes. The LFS has both high frequency (quarterly) and a relatively large sample size, which means it is a particularly good (and sometimes the only practical) source of household and family data on smaller population sub-groups, for example ethnic minorities.

However, because the LFS was designed and developed as a survey focussing on individuals, in the early stages little attention was given to the information on household and family. The survey design, data collection and processing procedures were chosen to produce the best possible data on individuals, but in some respects this did not coincide with producing the best possible data on households and families. ONS has therefore decided to produce separate LFS datasets especially designed to be suitable for analyses at household and family level.

This has involved making adjustments for some inconsistencies and discontinuities in the data which have in the past affected the recording of household and family structure; adding new derived variables for use in analysing economic activity at household and family level; and deriving household level weighting factors. A series of past datasets was released in 1998, including spring 1990, 1992 1993, 1994; spring and autumn 1995, 1996, 1997; and spring 1998. From the autumn 1998 dataset onwards, the production of household datasets has been incorporated within the regular production programme of LFS data, producing one from each autumn and spring quarter's data.

While this process of adjustment and household level weighting produces datasets which we consider will give the best estimates at household and family level, it inevitably results in some differences from the regular LFS datasets which are designed to give the best estimates at the individual level. **It is therefore recommended that the regular quarterly person level datasets be used for any analysis at individual level, and that the household datasets be used for any analysis involving household or family level data.** For analysis of individual members of households according to household level characteristics, the situation is more complex, and further guidance is given in Section 7 of this volume.

This volume of the User Guide gives information which is specific to the LFS household datasets. It:

- gives definitions of the key household and family concepts;

- describes the inconsistencies and discontinuities which were found in household and family data and the adjustments which have been made to correct or compensate for them in the household datasets;

- describes the improvements to household and family data which were brought about by the introduction of the household matrix approach to gathering data on the household composition;

- describes the process used for deriving household level weighting factors and assesses the effect of this;

- describes the household datasets and the new household and family level variables which have been added;

- explains and demonstrates various aspects of analysing household and family level data.

SECTION 2 - DEFINITIONS

HOUSEHOLD

A household is defined as a single person, or a group of people living at the same address who have the address as their only or main residence and either share one main meal a day or share the living accommodation (or both).

Working-age household

A working-age household is a household which includes at least one person of working age, i.e. a woman aged between 16 and 59 or a man aged between 16 and 64.

LFS FAMILY UNIT

A LFS family unit comprises either a single person, or a married or cohabiting couple on their own, or with their never-married children who have no children of their own, or lone parents with such children.

It should be noted that the term 'family' is often used in the literature in a more restrictive sense than the LFS definition of a family unit. Commonly, single person LFS family units are not described as families but as 'persons not in families'. Also, LFS family units include non-dependent 'children' (who can in fact be adult) provided they are never-married and have no children of their own in the household. This may be inconsistent with some other usage.

DEPENDENT/NON-DEPENDENT CHILDREN

Dependent children are defined as children under 16 years of age, or those aged 16 to 18 who are never-married and in full-time education. Dependency of a child is defined in the context of a family.

ADOPTED AND STEP-CHILDREN

Adopted children are not separately distinguished, and are treated in the same way as other children. Up to 1995, step-children were not separately distinguished. From 1996 onwards, step-child and step-parent are recorded separately as relationships to the head of household and within the household matrix, but step-children are included together with other children in variables describing the number of children in the family unit or household.

FOSTER CHILDREN

Foster children are treated as separate family units to their foster parents. The foster child is considered to be a child of the head of the household (HoH) and included in the number of dependent children in the household but not in the number of dependent children in the HoH's family unit. Also, the foster child will not be recorded in variables identifying the number of dependent children in his or her own family unit.

EXTENDED FAMILY

An extended family includes all people within a household who are related in some way: partners; parents; children; grandparents; grandchildren; brothers and sisters; relatives by marriage; guardians; and other relatives.

SAME SEX COUPLES

In spring 1996, categories identifying same sex cohabiting couples were included in the new household type, household composition and family type variables, though only recorded as such if the information was volunteered. However, the family type variable identifies same sex cohabitees as separate family units. The family type variable does not identify whether children are present in same sex families and thus if any are present they will not be counted in the variables identifying numbers of children in families.

SECTION 3 - ADJUSTMENTS FOR INCONSISTENCIES AND DISCONTINUITIES

Because the LFS was designed and developed as a survey focussing on individuals, in the early stages little attention was given to the information on households and families which was included in the data by virtue of the household-based sample design. As over time users began to draw on this information, a number of inconsistencies and discontinuities in the household and family data were identified, caused by changes over time in the definitions of the variables underlying household and family type, and changes or anomalies in applying these definitions. Some of these have already been addressed in the production of the individual level datasets; some have only recently been addressed, in the production of the household datasets. For completeness, all are described below.

SPRING 1990 HOUSEHOLD DATASET

The dataset for spring 1990 is the only one of the household datasets dating from before the introduction of the quarterly LFS. (See Section 6 below for details of how the household datasets were selected.) In the course of producing this dataset, it was found that the currently available individual level dataset for the year did not include a full household type variable (TYPEHH). Since this was considered a key variable for household datasets, it was re-derived, but the new variable proved to be inconsistent with the pre-existing household composition variable (AGCOCD). Since the original specifications for these variables, and the software to run them, were no longer available, the source of the inconsistency could not be investigated. It was decided that, since the re-derived TYPEHH was consistent with later years (the 1992 to 1994 specifications having been used), and since the AGCOCD variable was very little used compared with TYPEHH, the household dataset should be produced including TYPEHH but not AGCOCD. Note also that there are various differences in how other variables are defined between 1990 and the later datasets – see Volumes 3 and 7 of this User Guide for more details.

FAMILY UNIT ALLOCATION

Up to and including 1995, the coding of family unit was done by LFS interviewers, who allocated all the individuals in a household to one or more family units on the basis of the responses to the questions on relationship to head of household and marital status. In most cases the household and the family unit are the same, but it is not uncommon for a household to contain more than one family unit.

From spring 1992 to spring 1993 there were found to be problems with the allocation of individuals to family units, which arose as follows. From 1984 to 1991 the annual LFS dataset was made up of the continuous LFS (a 15,000 household rotating quarterly panel) and the spring boost of 45,000 households. Both of these systems were based on paper questionnaires for most of that time, and included incomplete interviews with some respondents in otherwise fully responding households. Having incomplete data in this way caused problems for anyone using the grossed up results, so when it was planned (in 1989) to adopt CAI for the continuous panel element, it was agreed to drop from the records all potential respondents for whom no information (other than their existence) could be gathered. When the quarterly LFS was introduced, it followed the design of the continuous panel element (which had been written in BLAISE). The coding check on the assignment of household members to family units - which had been part of the quarterly supplement when it was conducted using paper questionnaires - was discontinued. This was because previous experience had shown that it was very difficult for coders to make amendments with the information available from the interview and only a small number of the discrepancies could be corrected.

One of the chief effects of dropping these checks was that unrelated respondents in a household could be erroneously coded as members of the same family even though they were not coded as married or living together as a couple - two people in the same household, each of whose marital status was 'divorced', could be coded as being in the same family unit, for example. To add to the problem, the derived variable relationship to head of family unit (RELHFU) would have identified both as the head of the family unit.

The situation was more complicated where children were involved. In the example above, if one or both of the divorced people living in the same household had children, then not only would the whole household have been classed as a single family unit, but the woman would have been coded as heading a male lone parent family. When it later emerged that secondary analysts wished to use the quarterly LFS for demographic analysis, the checking of household member assignment to families was reinstated for the summer 1993 survey, and beyond. For earlier quarters it was necessary to re-examine the data and impute values for 'implausible' households; spring 1992 and spring 1993 have been made consistent and family unit information from summer 1993 onwards is reliable. (It should be noted that as the data for spring 1992 and spring 1993 were imputed retrospectively when some additional constraints applied, they will not follow quite the same model as is currently applied). This leaves only summer 1992, autumn 1992, and winter 1992/3 as problematic; there are no plans to correct these, as the 'correction' process is extremely resource-intensive.

ERRORS IN DERIVATION OF HOUSEHOLD AND FAMILY VARIABLES

In June 1995, errors were discovered in the programs producing some of household and family derived variables from spring 1992 to winter 1994/5. A decision was made not to correct this as the effects were thought to be minor and household and family data were little used at that time. A reference to this issue was meant to be included in the User Guide issued in September 1995 but unfortunately this did not happen. The errors were caused by programs accessing the records of previous household members who had left and their records (including age) had been set to zero. The variables affected were: household composition (AGCOCD); household type (TYPEHH); age of oldest dependent child aged under 16 in family (AOFL16); age of oldest dependent child aged under 19 in family (AOFL19); and age of youngest child aged under 19 in family (AYFL19).

In 1997 it became apparent that the problem could have significant effects for some kinds of analysis, and the affected variables were re-derived for the spring quarters of 1992, 1993 and 1994. Since the re-derivation process was resource-intensive, and some non-spring quarters were already known to have other problems affecting household and family data, it was decided not to apply the re-derivation to the non-spring quarters, and to recommend that only the corrected spring quarters from 1992 to 1994 should be used for analyses using the affected household and family unit variables.

Table 3.1 below shows the differences in the household composition variable AGCOCD between the original and revised datasets. The main changes were substantial numbers moving from 2 adults under pension age with no children (category 11) to 2 (married or not married) adults one over pensionable age with no children (categories 7 and 12) and from 3 or more adults with no children (category 17) to 3 or more adults with no children including at least one married or cohabiting couple (category 14).

Table 3.1: Differences between original and revised unweighted datasets for household composition variable

	Category	Difference (+/-)		
		Spring 1992	Spring 1993	Spring 1994
1	1 male over pensionable age with no children	+3	+5	-124
2	1 female over pensionable age with no children	+98	+173	-293
3	1 adult with no children	-71	-178	+417
4	1 adult with one child	0	0	0
5	1 adult with two or more children	+1	0	0
6	2 married adults both under pensionable age with no children	+18	0	0
7	2 married adults one over pensionable age with no children	+1,808	+1,860	+1,768
8	2 married adults with one child	+30	+29	+26
9	2 married adults with two children	+1	+2	+1
10	2 married adults with three or more children	+4	+1	-1
11	2 adults under pensionable age with no children	-2,595	-2,531	-2,410
12	2 adults one under pensionable age with no children	+769	+671	+642
13	2 adults with one or more children	-31	-32	-26
14	3 or more adults with no children including at least one married or cohabiting	+952	+964	+903
15	3 or more adults with one or two children including at least one married or cohabiting couple	+15	+21	+22
16	3 or more adults with three or more children including at least one married or cohabiting couple	-1	0	-1
17	3 or more adults with no children	-955	-964	-903
18	3 or more adults with one or more children	-16	-21	-21
	Total	**3,699**	**3,726**	**3,779**

For the household type variable TYPEHH, 845 cases were affected in spring 1992, 406 in spring 1993 and 348 in spring 1994. For the three age of child variables affected, a number of cases had an incorrect value of zero. When corrected, these cases were allocated across all ages of children. Table 3.2 below shows the number of cases affected.

Table 3.2: Number of cases affected in each quarter

	Spring 1992	Spring 1993	Spring 1994
AOFL16	327	286	240
AOFL19	326	286	233
AYFL19	321	285	245

MISSING HOUSEHOLD MEMBERS

Even after the amendments described above to correct for family unit allocation, the quarterly LFS family unit data were not directly comparable with the corresponding annual LFS data. This is chiefly because, as described above, in the annual survey records were maintained of partial/incomplete interviews, whereas in the quarterly survey it was decided that if no information was given to key questions, then the interview for that person should be treated as an outright refusal. However the algorithms for deriving family and household type which continued to be used in these new circumstances now ignored any person for whom there were no data. If the ignored person were married or cohabiting, the other partner would be classified either as a single person family unit or as a lone parent depending on whether there were children, and the estimates of such family types were therefore inflated. This problem was eventually solved by the introduction of the household matrix approach to gathering household and family

data in spring 1996, which is described in Section 4 below. Adjustments to compensate for its effects have been made to the household datasets for the period 1992 to 1995, as described below.

ERRORS IN MARITAL STATUS CODING

Before describing the adjustments for the 'missing members' problem, one further data problem is described, which these adjustments also deal with. This problem affected family type coding in spring and summer 1995. In some cases, (121 in spring and 137 in summer), marital status was incorrectly coded and this resulted in both the man and the woman of a married or cohabiting couple being classified as the head of a male lone parent family unit.

ADJUSTMENTS FOR MISSING HOUSEHOLD MEMBERS

The existence of 'missing members' of a household had to be inferred from internal information on the characteristics of the recorded household members. For some kinds of possibly missing people, for example one of a group of unrelated flat-sharers, there was no way of doing this, but it was thought that most cases where the missing person had led to distortion of the recorded household or family structure could be identified by one of the following criteria: there was no recorded head of household; or the family unit was of the single person or lone parent type but the head of the family unit was married or cohabiting.

The adjustment procedure was therefore as follows:

If type of family unit (TYPEFU) is one person, male lone head of family or female lone head of family, and the head of family unit is married or cohabiting, then the adjusted type of family unit variable (TFUADJ) is set to male head of married/cohabiting family.
If type of household (TYPEHH) is one person and that person is married or cohabiting and is either the head of household or the wife or cohabitee of the head, then the adjusted type of household variable (THHADJ) is set to a couple household with no children and no other family units.

If type of household (TYPEHH) is one person and that person is a child of the head of household, then the adjusted type of household variable (THHADJ) is set to a lone parent household with all non-dependent children and no other family units.

If type of household (TYPEHH) is one person and that person is anything other than the head of household or the wife, cohabitee or child of the head, then the adjusted type of household variable (THHADJ) is set to two or more persons, all different family units.

If type of household (TYPEHH) is one of the lone parent categories and the head of family unit is married or cohabiting and is either the head of household or the wife or cohabitee of the head, then the adjusted type of household variable (THHADJ) is set to a couple household of the corresponding category as regards type of children and whether or not there are other family units.

This procedure however produced a small number of cases where, because of an unusual combination of missing members, after adjustment the household type and family type variables were inconsistent. These were dealt with by setting either the adjusted household type variable or the adjusted family type variable to 'not known', according to the following procedure:

If after adjustment household type is lone parent, all non-dependent children, no other family units, and family type is one person, then set adjusted family type to not known.

If after adjustment household type is two or more persons, all different family units, and family type is one person, and marital status is married or cohabiting, then set adjusted family type to not known.

If after adjustment household type is lone parent with either non-dependent children or other family units, with or without dependent children, and family type is married/cohabiting couple, then set adjusted household type to not known.

If after adjustment household type is two or more persons, all different family units, and family type is one person or lone parent, and marital status is married or cohabiting, then set adjusted household type and adjusted family type to not known.

The numbers of instances of adjusted household and family type variables are given in table 3.3 below:

Table 3.3: Number of instances of adjusted household and family type variables

Changed from	Changed to	Spring 1992	Spring 1993	Spring 1994	Spring 1995	Autumn 1995
HOUSEHOLDS						
Lone parent	Couple	541	661	556	490	605
One person	Couple	782	776	844	520	673
One person	Lone parent	56	58	52	72	87
One person	2+ persons	79	109	122	146	166
All	DK	36	45	28	27	27
FAMILY UNITS						
One person	Couple	915	902	944	669	822
Lone parent	Couple	566	717	595	532	628
All	DK	278	251	179	247	256

There were also found to be a small number of cases (typically about 100 per dataset) where a family unit had no head, or two or more heads, or two or more wives of head. The apparent causes were very varied, and in most cases it was not possible to identify the most probable real situation, so these cases were dropped. This problem continued to occur after 1995.

It should be noted that the data during the period under consideration do not enable us to distinguish between cases where a spouse is missing from the data, and cases where a spouse is genuinely absent from the household for more than six months, for example working abroad or in prison. The adjustment procedure has therefore been applied to both situations, though it is likely that the second is much rarer.

It should also be noted that, for households and family units where adjusted household and family type codes have been assigned, it has not been possible (owing to lack of information) to make changes to the recorded data on relationship to head of household or head of family unit. In some cases, therefore, these will not be consistent with the adjusted household and family type codes.

After the adjustments described above have been applied, there remain a few minor potential sources of residual problems in the datasets up to 1995 which it has not been possible to address. An assessment of the maximum possible effect of these is given at Annex C.

SECTION 4 - HOUSEHOLD MATRIX IMPROVEMENTS

HOUSEHOLD MATRIX APPROACH

In recent years, there has been increasing interest in household and family data from the LFS, particularly in economic activity data for households and families, and also for general demographic purposes. In spring 1996 a new approach, the 'household matrix', was introduced for the collection of LFS data on household and family structure. This method was designed to correctly identify all families within households; to provide more detailed information on the composition of families and households; and to provide greater flexibility in the analysis of family relationships. It has been adopted as the harmonised approach for government surveys as part of an initiative to standardise questions and introduce common classifications, definitions and standards for government social surveys[1].

In the household matrix approach, information is obtained on the relationship between each pair of household members, and a matrix of relationships is produced from which the number and composition of families and households can be derived automatically. (Previously only relationships to the head of household were recorded, and interviewers allocated individuals to family units according to their own judgement.)

Since this approach requires that certain basic data are recorded for all members of the household, and people with no economic activity data are still included in the household matrix and hence in the derivation of the family and household types, it has the additional advantage of eliminating, from spring 1996 onwards, the problem of 'missing members' described in the previous section. The few cases remaining with an apparent discrepancy between family type and marital/cohabitational status are confirmed by a check question to be genuine cases where the partner has been living away from the residence for more than six months.

The household matrix approach collects more information than before. The matrix records marital status, cohabitation (including same sex cohabitees), whether children are dependent or non-dependent, step, foster or natural. As a result, some new variables were introduced from spring 1996 and some revised. For example, the relationship to head of household variable has been extended from 10 to 21 categories, and the more detailed categories for household and family type distinguish between married and cohabiting couples, and between families with no children, non-dependent children only or dependent children (with or without non-dependent children). More details of the old and new variables are given in Volume 3 of this User Guide.

HOUSEHOLD REFERENCE PERSON

From spring 2001 the LFS is switching from using Head of Household (HOH) to Household Reference Person (HRP) in line with all other ONS household surveys. HOH and HRP are derived in different ways. At present, the majority of heads of a household are male. This is because the HOH is defined to be the male in a couple or where more than one person has equal claim to be HOH and they are of opposite sex. The HRP will take into account accommodation ownership / rental, income and age in order to identify the HRP (see Volume 3 of this User Guide for more details).

Data on HRP have been collected since spring 2000, enabling comparisons to be made between the two definitions. Overall, in around 10% of households the HOH and HRP are different people. In over 90% of such cases the HRP is the spouse or cohabitee of the HOH and therefore the vast majority of changes are from male to female. Also HRPs tend to be slightly younger than HOHs.

Information on HOH will continue to be collected. However, HRP will be used instead of HOH in creating a number of derived variables, such as RELH96 (relationship to head of household) and CAIND (child/adult indicator). Variables derived from RELH96 and CAIND will also be affected. Analyses which have

---

[1] GSS Harmonised Concepts and Questions for Government Social Surveys, ONS, 1996.

previously been based on these variables may change slightly due to the new definition. Also, if analyses which had previously been done by HOH switch to using HRP, results may differ. For example, table 4.1 below shows percentage changes in the numbers of workless and work-rich households by ethnicity of the HOH compared to HRP for spring 2000.

Table 4.1: Percentage changes in the numbers of workless and work-rich households
by ethnicity of the HOH compared to HRP; United Kingdom; spring 2000

Percentage change HRP over HOH	Workless households	Work-rich households	All households
White	0.12%	0.08%	0.09%
Black Caribbean	-0.06%	-6.61%	-4.24%
Black African	0.00%	-8.91%	-3.57%
Other Black	-6.12%	-2.96%	-6.92%
Indian	-1.27%	0.92%	0.17%
Pakistani/Bangladeshi	-0.92%	-3.31%	-1.33%
Chinese	0.00%	3.71%	5.35%
Other	-1.78%	3.10%	-0.23%

Users are advised to use the relationship to HRP variable (RELHRP). However, RELH96 will still be available when it is important to produce a consistent time series of analyses of the HOH. Household level analyses previously produced in SPSS by selecting RELH96 = 0 (head of household) will not be affected by selecting the HRP.

EXTENDED FAMILIES

An additional concept which was introduced into LFS data at the same time as the household matrix is that of the extended family. An extended family includes all people within a household who are related in some way: partners; parents; children; grandparents; grandchildren; siblings; relatives by marriage; guardians; and other relations. (Foster parents and children are not included in the same extended family, and foster children are assigned to separate family units within a household.)

NB It should be noted that the household matrix data for spring 1996 may not be fully complete as there are some households (about 3%) where data has to be imputed (carried forward from the previous quarter) due to circumstantial refusal or non-contact and these households will not have data in the form of the household matrix. Also, a small number of people (about 100) in spring 1996 were found to have no data on age or sex - these individuals were dropped. For cases where age and sex were recorded but region was missing, the region code of the other household members was assigned.

SECTION 5 - WEIGHTING THE LFS SAMPLE OF HOUSEHOLDS

PROBLEMS OF USING INDIVIDUAL LEVEL WEIGHTING FACTORS FOR HOUSEHOLD LEVEL DATA

Until the development of the special-purpose LFS household datasets, the grossing for all LFS data, i.e. the application of weighting factors in order to produce estimates at the level of the whole population, was done entirely on the basis of the individual person. The process, which continues to be used for the individual level LFS datasets, is fully described in Volume 1 of the User Guide, but in summary the population is split into subgroups (or cells) where the number of people in each subgroup is known (based on population estimates). The weights are calculated by dividing the population in that subgroup by the number of cases in the sample. A multi-stage procedure is used: first grossing to population estimates for individual Local Authority Districts; then to national population estimates by sex and age group; then grossing by sex, region and age group. This procedure is iterated until the results are stable.

Under this procedure sex, age and area of residence will affect a person's grossing factor, but not the type of household in which he or she lives, and not the characteristics of the other people in the household. This implies that using the individual grossing factors to gross up household level data may produce a biased distribution of household types, and a biased distribution of particular subgroups of people by the type of household to which they belong.

There are some particularly striking examples of likely biases. First, we know from a study described in Volume 1 of the User Guide which investigated the characteristics of LFS non-respondents using 1991 Census data, that young adults in one person households are particularly under-represented in the LFS, (being especially difficult to find at home). To compensate for the under-representation of young adults, those who are in the sample will have a high weighting factor - but since they are the ones who have **not** dropped out, they will consist disproportionately of young people in larger households, rather than in one person households, with the result that one person households will still be under-represented in the grossed up household data. Equally, although the grossed up total numbers of young adults will be correct, their distribution by the type of household they live in will be biased.

Another aspect of the problem is that, since individuals in the same household are likely to have different weights, the grossing system will produce different grossed up numbers for the two partners in a married or cohabiting couple. Similarly, a joint distribution of, for example, heads of household and their spouses, can produce quite different numbers depending on whether the basis for grossing up is the head or the spouse.

It was therefore decided that the LFS datasets produced specifically for analysis at the household and family level should have weighting factors which were the same for all members of a household.

CALCULATION OF COMMON HOUSEHOLD WEIGHTING FACTORS

The following specifications were defined for the grossed estimates:

(i)   Distribution by sex and age to be consistent with the latest available population estimates, using the following age categories:

0-4, 5-9, 10-15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25-29, 30-34, 35-39, 40-44, 45-49, 50-54, 55-59, 60-64, 65-69, 70-74, 75-79, 80 and over.

(ii)   Distribution by region to be consistent with the latest available population estimates, using the following region categories:

From spring 1990 to spring 1996 (based on Standard Statistical Regions)

Tyne and Wear; Northern - other; Yorkshire and Humberside - metropolitan; Yorkshire and Humberside - other; East Midlands; East Anglia; Inner London; Outer London; South East - other;

South West; West Midlands - metropolitan; West Midlands - other; North West - metropolitan; North West - other; Wales; Strathclyde; Scotland - other; Northern Ireland.

From autumn 1996 onwards the regions used in grossing are based on Government Office Regions (see 'Regrossing the household datasets' section below)

Tyne and Wear; Rest of North East; Greater Manchester; Merseyside; Rest of North West; South Yorkshire; West Yorkshire; Rest of Yorkshire and the Humber; East Midlands; West Midlands; Metropolitan County; Rest of West Midlands; East of England; Inner London; Outer London; South East; South West; Wales; Strathclyde; Rest of Scotland; Northern Ireland.

(iii)   Weights for persons in the same household to be equal.

It was possible to apply these specifications at the same time by using the recently developed methodology known as calibration, or generalised raking. Following Elliot (1997)[2], the software used was CALMAR, a SAS macro produced by INSEE. The logit method was used, with the allowable range of weighting factors set so as to minimise the range, subject to avoiding significant 'bunching' of cases at the boundary.

The possibility of using prior weights to compensate for non-response in characteristics other than those included in the grossing procedure was considered, but it was found that the gain to the sample from movers-in satisfactorily offset the loss through non-response.

Population estimates

The population estimates used to provide the control totals by age, sex and region were the latest available when the household datasets were first produced in 1998. Figures on databases up to spring 1996 were controlled to population data based on ONS's mid-year estimates (MYEs) and autumn 1996 and beyond used control totals based on 1992 population projections. However, in 2000, ONS undertook a regrossing exercise to release revised LFS estimates using more up-to-date population estimates than the existing ones. The effects of this exercise on the household datasets are described below in the section 'Regrossing the household datasets'.

RE-WEIGHTING THE HOUSEHOLD DATASETS

In early 1999, some technical revisions were undertaken to the weighting procedure used to create the household datasets. A problem had been found which caused the weighted distribution of households by type to be distorted for Northern Ireland, with too many one person households, too few households consisting of couples with children, and too high a total number of households. The source of the problem was identified, and to solve it the weighting factors needed to be recalculated using a set of starting weights in which Northern Ireland has a different value from other regions.

The full series of household datasets up to spring 1998 was re-weighted. However, the re-weighting had minimal effects upon the analyses extracted from these datasets, the only exception to this being analyses involving household type carried out on the Northern Ireland region.

REGROSSING THE HOUSEHOLD DATASETS

In April 2000, ONS released regrossed person level LFS data back to autumn 1993 using the most up-to-date population control totals available. The household datasets were also regrossed using the same population figures as were used for the main databases and were released in August 2000. It was only necessary to regross the household datasets back to autumn 1996 as datasets up to spring 1996 were already based on ONS's MYEs. As a result of the regrossing, datasets from autumn 1996 to spring 1998 are based on MYEs; from autumn 1998 to spring 1999, MYEs and 1996-based population control totals; and from autumn 1999, 1996-based population control totals are used. For more information on the regrossing see *Labour Market Trends*, February 2000.

---

[2] Elliot, D (1997) Software to weight and gross survey data. *GSS Methodology Series No 1*. ONS.

The following table 5.1 presents the effect of regrossing the household datasets on spring 1997 to spring 1999 analyses for households by combined economic activity of households. While the numbers of households increased due to regrossing, the rise in numbers of working-age households was greater. The proportions of work-rich households increased by an average of 0.3 per cent across the three spring quarters. The numbers and proportions of workless households, and of working-age people and children in workless households decreased, although the effect was slightly greater for children in workless households. Between spring 1997 and 1999 the average fall in the proportion of children in workless households was 0.3 percentage points.

Table 5.1: Comparison of regrossed and pre-regrossed analyses of households by combined economic Household; United Kingdom; spring 1997-2000

	All households	All working-age households	Work-rich households[a]	Workless households[a,b]	Working age in workless households[a,b,c]	Children in workless households[a,b,]
						Thousands
Spring 1997	24,100	18,309	9,986	3,271	4,719	2,163
Old Spring 1997	24,046	18,229	9,897	3,281	4,732	2,215
Difference	*54*	*80*	*88*	*-10*	*-12*	*-52*
Spring 1998	24,285	18,480	10,227	3,237	4,634	2,156
Old Spring 1998	24,209	18,380	10,109	3,253	4,651	2,226
Difference	*76*	*100*	*117*	*-15*	*-17*	*-69*
Spring 1999	24,383	18,530	10,376	3,156	4,488	2,087
Old Spring 1999	24,309	18,425	10,253	3,175	4,511	2,170
Difference	*74*	*104*	*122*	*-20*	*-23*	*-83*
Spring 2000	24,554	18,663	10,693	3,066	4,318	1,907
						Per cent
Spring 1997	100	100	54.5	17.9	13.2	17.9
Old Spring 1997	100	100	54.3	18.0	13.3	18.2
Difference	*0*	*0*	*0.2*	*-0.1*	*-0.1*	*-0.2*
Spring 1998	100	100	55.3	17.5	12.9	17.9
Old Spring 1998	100	100	55.0	17.7	13.0	18.2
Difference	*0*	*0*	*0.3*	*-0.2*	*-0.1*	*-0.3*
Spring 1999	100	100	56.0	17.0	12.4	17.3
Old Spring 1999	100	100	55.6	17.2	12.6	17.6
Difference	*0*	*0*	*0.3*	*-0.2*	*-0.1*	*-0.3*
Spring 2000	100	100	57.3	16.4	11.9	15.8

a   Adjusted to include estimates for households with unknown economic activity. See Section 8 for more details.
b   A workless household is a household with at least one person of working-age where no-one is in employment. For more definitions see Section 2.
c   Working age refers to men between the ages of 16 and 64 and women between 16 and 59.
d   Children refers to all children under 16.

TREATMENT OF 'MISSING MEMBERS'

As noted in Section 3 above, during the period 1992 to 1995 individuals for whom no data were gathered in otherwise responding households were excluded from the datasets. The adjustment procedure described in Section 3 has, where possible, changed the household and family type variables for households affected by this problem, but it was not possible to make any adjustment for the missing members in the grossing procedure. The weighting factors for these households may therefore be slightly different from what they would have been if data for the missing member had been available.

As noted in Section 4 above, from 1996 onwards individuals for whom no data on economic activity were gathered were still included in the household matrix, and used in the derivation of household and family type variables. In the individual level LFS datasets these individuals are zero-weighted, i.e. they are ignored in grossing up to population level estimates. However, for analyses at household and family level, these individuals need to be retained in their correct place within the household structure. They have therefore been included in their households for the grossing procedure for the household datasets, and have the same weight as the other household members. This means that there will be some 'not known' cases in any grossed analysis of economic activity variables using the household datasets, at either the individual, family or household level.

This reinforces the recommendation that **the regular quarterly person level datasets be used for any analysis at individual level, and that the household datasets be used for any analysis involving household or family level data.**

COMPARISONS BETWEEN HOUSEHOLD ESTIMATES GROSSED AT HOUSEHOLD AND INDIVIDUAL LEVEL

Table 5.2 below show two examples of the grossed distributions by household type when the grossing is done on a household basis as described above, compared with the results when the grossing is done on an individual basis, as described in Volume 1 of the User Guide. Since both the procedures for adjusting and weighting the household datasets, and the categorisations of the household and family type variables, are different for the period from 1996 onwards and the period up to 1995, one example is given from each period, for spring 2000 and spring 1993 respectively.

For 2000, when the only difference is in the weighting factors, the estimated total number of households is a little smaller, and there are some differences in distribution. In particular, the proportions of couples with all dependent children and lone parents with all dependent children are smaller and that of couples with all non-dependent children larger. For 1993, which is also affected by adjusted household type categories, there are some larger (though not very much larger) differences for categories particularly affected by the adjustments. In particular there is a smaller proportion of one person households as well as of couples and lone parents with all dependent children, and a larger proportion of couples with no children as well as of couples with all non-dependent children.

Table 5.2: Comparison of the distribution of household types produced from using the old individual level and new household level weights; United Kingdom; spring 1993 and 2000

Type of household	Old individual level weight[a]		New household level weight		Diff.
	Thousands	Per cent	Thousands	Per Cent	% pts.
Spring 2000					
One person	7,101	28.2	6,993	28.5	0.3
Two or more people, all different family units	645	2.6	749	3.1	0.5
Couple, no children, no other family units	6,930	27.5	6,792	27.7	0.2
Couple, no children, other family units	193	0.8	214	0.9	0.1
Couple, all dependent children, no other family units	5,370	21.3	4,891	19.9	-1.4
Couple, dependent and non-dependent children, no other family	659	2.6	684	2.8	0.2
Couple, all non-dependent children, no other family units	1,408	5.6	1,559	6.3	0.8
Couple, children, other family units	186	0.7	190	0.8	0.0
Lone parent, all dependent children, no other family units	1,576	6.3	1,304	5.3	-0.9
Lone parent, dependent and non-dependent children, no other family units	182	0.7	172	0.7	0.0
Lone parent, all non-dependent children, no other family units	646	2.6	694	2.8	0.3
Lone parent, other family units	109	0.4	107	0.4	0.0
Other	197	0.8	205	0.8	0.1
All household types	**25,202**	**100**	**24,554**	**100**	**..**
Spring 1993					
One person	6,518	27.4	6,258	26.5	-1.0

Table 5.2: Comparison of the distribution of household types produced from using the old individual level and new household level weights; United Kingdom; spring 1993 and 2000

Type of household	Old individual level weight[a]		New household level weight		Diff.
	Thousands	Per cent	Thousands	Per Cent	% pts.
Two or more people, all different family units	530	2.2	617	2.6	0.4
Couple, no children, no other family units	6,268	26.4	6,497	27.5	1.1
Couple, no children, other family units	195	0.8	205	0.9	0.0
Couple, all dependent children, no other family units	5,184	21.8	4,959	21.0	-0.9
Couple, dependent and non-dependent children, no other family	690	2.9	758	3.2	0.3
Couple, all non-dependent children, no other family units	1,603	6.7	1,804	7.6	0.9
Couple, children, other family units	210	0.9	240	1.0	0.1
Lone parent, all dependent children, no other family units	1,205	5.1	1,001	4.2	-0.8
Lone parent, dependent and non-dependent children, no other family units	166	0.7	136	0.6	-0.1
Lone parent, all non-dependent children, no other family units	867	3.6	878	3.7	0.1
Lone parent, other family units	152	0.6	126	0.5	-0.1
Other	158	0.7	159	0.7	0.0
All household types	**23,746**	**100**	**23,656**[b]	**100**	**..**

a   Individual weight of head of household.
b   The total number of households for spring 1993, produced using the new household level weight, includes a very small proportion of cases where it is not possible to assign an adjusted household type.

SECTION 6 - HOUSEHOLD DATASETS

SELECTION OF HOUSEHOLD DATASETS

As noted in Section 3 above, some problems with household and family data during the period 1992 to 1994 have previously been identified and corrections or adjustments made, but because of the extremely resource-intensive nature of the work involved in the correction and adjustment processes, this work was applied only to the spring quarters. It was decided therefore that, for this period, household datasets would be produced for spring quarters only.

For the period before 1992, as one looks further back in time, it becomes increasingly difficult to assess the quality of the household and family data, but there are known to be some sources of discontinuity, and it is a period for which there has been little use made of the household and family data. The available LFS datasets for the period do not now include a detailed household type variable, which would be a key element of most household level analyses. Consultation with major customers indicated that for LFS household datasets there was most interest in the period from 1992 onwards, but that it was regarded as important to have a dataset for 1990, in view of its position in the business cycle.

From 1995 onwards, the consensus of major customers was that a greater frequency than annual was desirable, but that two quarters per year, spring and autumn, would be satisfactory. The following is a summary of the series of available household datasets:

Spring 1990
Spring 1992
Spring 1993
Spring 1994
Spring and autumn quarters from 1995 onwards.

VARIABLES IN THE HOUSEHOLD DATASETS

The household datasets include all the usual variables found in the individual level LFS datasets with the exception of those relating to income, which are not asked in all quarters and have a separate, individual level weighting procedure, and so are only suitable for analysis at individual level.

There are also substantial additions. As well as the family and household variables which have been produced in the past, and which will continue to be included in the household datasets in future, there are several groups of new variables which are now being produced to assist users in producing household and family level analyses. These fall into five groups:

(i)   variables giving the household weight and household and family unit identifiers;

(ii)   variables giving the adjusted household and family type categories, as described in Section 3;

(iii)   variables giving the numbers of people in the household for various age groups which are of importance in analysis of the labour market characteristics of the household:

   •   number of people of working age (i.e. 16 to 59 for women, 16 to 64 for men) in the household;

   •   number of people over working age in the household;

   •   number of children aged 0 to 4 in the household;

   •   number of children aged 5 to 15 in the household;

   •   number of dependent children (i.e. those in full-time education) aged 16 to 18 in the household;

(iv)   variables giving the numbers of people in the household in various economic activity categories of particular interest for analyses of the whole household, and describing the combined economic activity of the household members:

- number of people in employment in the household;

- number of people in full-time employment in the household;

- number of people in part-time employment in the household;

- number of people ILO unemployed in the household;

- number of people economically inactive (neither employed nor ILO unemployed) in the household;

- number of people in the household who are economically inactive but would like to work and are not seeking work because they are students;

- number of people in the household who are economically inactive but would like to work and are not seeking work because they are sick or disabled;

- number of people in the household who are economically inactive but would like to work and are not seeking work because they do not believe there is work available ('discouraged workers');

- number of people in the household who are economically inactive but would like to work and are not seeking work because they are looking after a family;

- number of people in the household who are economically inactive but would like to work and are not seeking work for other reasons;

- number of people in the household who are economically inactive and do not want work;

- number of people in the household whose basic economic activity is not known;

- combined economic activity of the household;

(v)   a pair of variables, defined at the level of the family unit, and giving the economic activity of the head of family unit and wife/partner, to make it easier to analyse the economic activity of married or cohabiting couples:

- economic activity of the head of the household;

- economic activity of the wife of the household.

A detailed description (including their derivation) of these variables is given in Annex A.

DIFFERENCES BETWEEN HOUSEHOLD AND INDIVIDUAL DATASETS

Summarising some of the issues discussed above, the household datasets differ from the regular quarterly individual level datasets in the following ways:

- For 1992 to 1995, the household datasets include adjusted variables for household and family type, which are adjusted for the effects of missing household members during this period, and for the marital status coding problem in 1995.

- For 1996 onwards, the individual level datasets give zero weights to individual household members with no economic activity data, whereas in the household datasets they are weighted equally with the other household members.

- The household datasets include a number of new derived variables at household and family unit level, to facilitate the analysis of the economic activity patterns of whole households.

- The household datasets include common household weighting factors, which are the same for every member of a household, whereas the individual level datasets include only individual weights, which usually differ between members of the same household.

The household datasets are designed to give the best estimates at household and family level, whereas the regular individual level datasets are designed to give the best estimates at the individual level. **It is recommended that the regular quarterly person level datasets be used for any analysis at individual level, and that the household datasets be used for any analysis involving household or family level data**. Guidance on analysis of individuals according to their household and family level characteristics is given in Section 7.

SECTION 7 – WHEN TO USE HOUSEHOLD AND INDIVIDUAL LFS DATASETS

This section provides guidance on which type of LFS dataset, household or individual, is the most appropriate for use in any given situation.

BACKGROUND

The weighting procedure for LFS household datasets differs from that used for the individual datasets in that the weight allocated is the same for all members of a household. Using individual level weights can produce biased distributions of household types, or biased distributions of particular subgroups of people by the type of household to which they belong. Also, it may produce different numbers for the two partners in a couple, or for joint distributions depending on whether the basis for grossing is the head of household or the spouse (see Section 5 for more details). Prior to 1996, the household datasets also contain adjustments for errors and inconsistencies in household type variables (see Section 3 for more details). All this makes the household datasets more suitable for analysis at the household or family level.

As the person level datasets are not suitable for household level analyses, the variables which classify by household or family type, composition or size, (HHTYPE, FUTYPE, HLDCMP, TOTNUM, TOTFU, FMDP, FMNDP, TOTXFU, XFMDC, XFMNDC), are no longer be included in the person level datasets.

LFS household datasets are available for spring and autumn quarters since 1995, and also for spring quarters in 1990 and 1992-4 (see Section 6). They do not contain information on earnings, as no satisfactory basis could be devised for weighting such information at the household level.

GENERAL RULE

In general, the household datasets should be used both for analyses of households and families, and for analyses of individuals where the distribution by household or family type is important. For example analysis of:

- Lone parents;

- Couples by economic activity of both partners;

- People by economic activity by household type;

- Women by whether married/cohabiting and whether dependent children in family.

For analyses by marital status, as mentioned above, the individual datasets give different grossed up numbers in married couples - there are typically more husbands than wives. Therefore, although analysis by marital status is not strictly an analysis by household or family type, again the household datasets should be used in order to ensure consistency of results.

The different weighting means that results will differ from the individual datasets, therefore the household datasets should also be used where consistency is required with other analyses which have been done on the household datasets.

Otherwise the individual datasets should be used. In particular, it is appropriate to do analyses which use the number or ages of dependent or non-dependent children in a household as classificatory variables.

EXCEPTIONS

Some exceptions to these rules follow from what has been said above.

Firstly, the household datasets are not available for all quarters, therefore if figures are required for quarters not covered by the household series the individual datasets will be the only option. If a quarterly time series of figures is required, the individual datasets alone should be used in preference to mixing figures from the two different sources.

Secondly, the household datasets have no income information, therefore for earnings analyses the individual datasets will be the only available source. See below for issues relating to income / earnings and the LFS household datasets.

Thirdly, in cases where a set of analyses are to be produced, some of which are at the individual level and some at the household level, and where consistency between the analyses is required, the household datasets should be used for all the analyses rather than the individual.

INCOME / EARNINGS AND THE LFS HOUSEHOLD DATASETS

The following points illustrate why it is not worthwhile, and probably not feasible, to attempt to include earnings or any other elements of income in the LFS household datasets:

(i)   The income module is only asked in two waves. Also it is explicitly voluntary, so there is additional case non-response. For both these reasons, there has to be an additional stage of weighting for the variables in this module. Also there is substantial item non-response, so imputation is needed.

(ii) For the earnings variables, both weighting and imputation need to be done on the basis of individual level variables, including occupation, industry, and whether working full or part time. This would imply, for the weighting, that either the weighting factors would be different for people in the same household, (which is not acceptable for the household datasets), or there would need to be a very resource-intensive weighting process using pseudo-control totals based on (a large number of) household level variables for the numbers of people in the household in each of the individual occupation, industry and full-/part-time categories. It is likely that this would involve difficulties in achieving convergence to a solution without negative weighting factors.

(iii) As regards benefit income, this is only meaningful at aggregate household level. Also, it is not gathered from people over 70 unless they are in employment, so a high proportion of income from some types of benefit is not recorded. Initial customer consultation on this topic produced no responses, indicating that demand was negligible. Furthermore, income from self-employment is not gathered. Therefore total income is incomplete and would involve weighting simultaneously at household and person level for different components, which is unlikely to be workable.

The above suggest that the only element of income data which is likely to be useful and which has any prospect at all of being feasible at household level is earnings, but it would still be either difficult and costly, or impossible, to produce satisfactory weighting factors.

The potential uses of data on total household or family earnings are in relation to assessing policy needs of, or monitoring effects on, low-income households or families. However, benefit income is significant in the relevant subgroup, so data on earnings alone would not be helpful. The Family Resources Survey (FRS) or the Family Expenditure Survey (FES) would be a more appropriate source. The other potential use would be to analyse simultaneously the earnings of couples, looking jointly at the earnings of husbands and wives. The benefits of this appear to be small relative to the costs.

SECTION 8 - ANALYSES AT HOUSEHOLD LEVEL

ANALYSES AT HOUSEHOLD LEVEL ONLY

Many analyses at household level are simplified by being able to make use of the new household level variables on the household datasets, making production of tables a simple matter of defining frequency distributions or cross-tabulations.

For any analysis at household level only, it is necessary first to select one person from each household (except for Quanvert users[3]). Usually, the simplest way of doing this is to select the head of household[4], i.e. select:
RELTOHOH = 1 (for spring 1990),
RELHOH = 1 (for spring 1996) (For spring 1996, when both RELHOH and RELH96 are available, RELHOH is preferred for this purpose.), or
RELH96 = 0 (from autumn 1996 to autumn 2000).
NB From spring 2001 use the relationship to household reference person variable (RELHRP) (see Section 4 for more details).

However, for the period 1992 to 1995, some of the households with 'missing members' (described in Section 3) have no recorded head of household. For this period an indicator variable HHIND has therefore been produced, which takes the value 1 for the first person in each household, and 0 otherwise. Thus one person per household can be selected using HHIND = 1.

For an analysis at family unit level, the selection of one person from each family unit can be done by selecting the head of family unit, using RELHFU = 1. The very few cases of a missing head of family unit were dropped from the household datasets, as described in Section 3.

Then, if a simple distribution is required, and the characteristic is described by an existing household level variable, defining the appropriate frequency distribution is all that is required. (**But it is necessary to remember to weight by the household weight if this is not a default of the software being used.**) For example, if a distribution by household economic activity is required, this is provided by a frequency distribution using the new variable HEACOMB. Table 8.1 shows a distribution of this variable over a number of years.

Table 8.1: Working-age households by combined economic activity of household; United Kingdom; 1990 and

	All household	All employed	Employed + unemploye unemploye	Employed + inactive	Employed + unemploye + inactive	All unemployed	Unemployed + inactive	All inactive
								Thousands
Spring 1990	17,023	9,059	802	4,499	255	362	419	1,628
Spring 1992	17,601	8,877	977	4,364	340	526	628	1,889
Spring 1993	17,882	9,121	982	4,163	333	607	691	1,985
Spring 1994	18,178	9,441	873	4,174	299	639	643	2,109
Spring 1995	18,407	9,780	780	4,143	258	621	574	2,251
Spring 1996	18,204	9,686	718	4,090	267	578	515	2,351
Spring 1997	18,309	9,986	649	4,173	230	496	442	2,333
Spring 1998	18,480	10,227	583	4,238	194	447	365	2,425
Spring 1999	18,530	10,376	598	4,199	201	430	360	2,366
Spring 2000	18,663	10,693	558	4,148	198	385	316	2,365

Figures have been adjusted to include estimates for households with unknown economic activity. See Section 8 – Treatment of households with unknown economic activity, for details of the adjustment procedure that should be used.

---

[3] The exception to this is for household level analyses produced from the spring 1996 quanvert household dataset. Users should filter on RELHOH = head as the file is incorrectly grossed to everyone and not just the head.
[4] From spring 2001 the LFS is switching from using Head of Household (HOH) to Household Reference Person (HRP) in line with all other ONS household surveys (see Section 4 for more details).

To show this distribution separately for subgroups of households defined by an existing household level variable, a cross-tabulation of HEACOMB and the other variable, for example HHTYPE or HNWKAGE, is all that is needed.

SELECTING PARTICULAR SUBGROUPS OF HOUSEHOLDS

Many analyses of economic activity need to be focussed upon the population of working age. For analysis of complete households, there are a number of possible ways of selecting those of working-age, but for general use the recommendation[5] is to select all households which include at least one person of working age. This can easily be done by selecting households with HNWKAGE > 0.

There is also particular interest in some individual categories or groups of categories of households as described by the household level economic activity. These are listed below and can be identified using the HEACOMB variable as shown:

- households with no-one in employment ('workless households'): HEACOMB = 5, 6 or 7;

- households with all adults in employment ('work-rich households'): HEACOMB = 1;

- households with someone unemployed: HEACOMB = 2, 4, 5 or 6.

The table below shows the distributions of these subcategories by household type.

Table 8.2: Economic activity of working-age households by type of household; United Kingdom; spring 2000

Type of household	Workless households	With all in employment	With someone unemployed
			Thousands
One person	986	2,438	246
Two or more people, all different family units	148	317	77
Couple, no children, no other family units	655	2,918	167
Couple, no children, other family units	28	48	16
Couple, all dependent children, no other family units	283	3,035	327
Couple, dependent and non-dependent children, no other family units	21	262	121
Couple, all non-dependent children, no other family units	81	711	163
Couple, children, other family units	14	30	26
Lone parent, all dependent children, no other family units	641	567	122
Lone parent, dependent and non-dependent children, no other family	26	56	39
Lone parent, all non-dependent children, no other family units	134	223	101
Lone parent, other family units	32	23	20
Other	18	65	31
Total all household types	**3,066**	**10,693**	**1,457**

Figures have been adjusted to include estimates for households with unknown economic activity. See Section 8 – Treatment of households with unknown economic activity, for details of the adjustment procedure that should be used.

Some other examples of subcategories likely to be of interest are:

- households which contain dependent children (HDPCH19 > 0, using an already established variable);

- households containing pre-school children (HDPCH4 > 0);

- households containing someone who is in employment (HNEMP > 0);

- households consisting entirely of students (HNOTSTUD = 0);

- households containing someone who is economically inactive and would like work, but is not seeking work because of a belief that there is none available ('discouraged worker') (HNIWDSC > 0);

---

[5] For more details see the article on Economic Activity of Working-age Households in *Labour Market Trends,* September 1997.

- households containing someone who is working part-time and no-one who is working full-time (HNFTIME = 0 and HNPTIME > 0).

ANALYSIS OF ECONOMIC ACTIVITY OF COUPLES

It is often of interest to analyse jointly the economic activity of partners in a married or cohabiting couple. To facilitate this, variables have been created at family unit level for the economic activity of the head of the family unit (HEAHEAD) and the economic activity of the wife/cohabiting partner of the head (HEAWIFE). To analyse these jointly, it is first necessary to select one record per family unit and to select family units where there is a wife/partner - this can be done simultaneously by selecting RELHFU = 2. Then a simple cross-tabulation of HEAHEAD with HEAWIFE (specifying weighting by the household weight if necessary) produces the required information, such as that shown in table 8.3.

Table 8.3: Economic activity status of working-age couples; United Kingdom; spring 2000

	Female partner			
	Employed	Unemployed	Inactive	**All**
Male partner				**Per cent**
Employed	63.7	1.5	17.0	82.2
Unemployed	1.3	0.3	1.5	3.1
Inactive	4.6	0.2	10.0	14.7
All	69.6	2.0	28.4	100

Percentages exclude couples where either one or both partner's economic activity status is unknown.

WORKING-AGE COUPLES

For some topics, especially those concerned with labour market characteristics, it may be more useful to confine the analysis to couples of working age. (Table 8.3 has been done on this basis.) The simplest way of doing this is to restrict the tabulation to households of working age using HNWKAGE > 0, making use of the existing variable for the number of persons of working age in the household. This will include households where there is someone of pensionable age as well as someone of working age - to restrict the analysis only to people of working age use the filter HNPEN = 0 (i.e. no people of pensionable age in the household). Note that using HNWKAGE > 0 may include some couples both of pensionable age living in households where there is also an adult of working age, though the number of such cases will be small. If you need to analyse exactly those couples where at least one partner is of working age, then it will be necessary to create a pair of family unit level variables, (for guidance see Annex B), which indicate whether the head and wife respectively of the family unit are of working age, and select on both of these variables.

CREATING NEW HOUSEHOLD AND FAMILY LEVEL VARIABLES

Many users need to create variables at household and family unit level on topics of particular interest to them, but which are not of such widespread interest to other LFS users. Annex B gives guidance on how to create some of the commonly needed kinds of variable - separate versions are given for users of SPSS and Quanvert software. The annex covers:

- joint characteristics of married/cohabiting couples;

- joint characteristics of children and parents;

- household level variables counting people with a specified characteristic;

- household level variables describing the combined characteristics of the members.

TIPS ON PRODUCING HOUSEHOLD AND FAMILY TABLES

The following tables are worked examples for producing some household and family analyses. ONS intends to further develop this list of examples.

For further advice please contact the Labour Market Division helpline: 020 7533 6094, or the Data Service: lfs.dataservice@ons.gov.uk.

1.   Lone parents:

- These are produced from the household datasets.
- Select / filter on RELHFU = 1 (head of family unit) and FUTYPE = 10 or 12 (male or female lone parents with dependent children).
- Tables produced from Quanvert need to be run at the person level.

E.g. 1: Lone parents by sex

Autumn(Sep. 2000 to Nov. 2000)
sex : Sex
Level person
Weighted
Filter: RELHFU Head of family unit
and (FUTYPE Male lone parent with dep children
    Or Female lone par with dep children)

	Thousands
Base	1,624
Male	170
Female	1,454
NA	*
DNA	*

*   Sample size too small for reliable estimate.

---

2.   Economic activity status of women/mothers with dependent children:

- These are produced from the person level datasets.
- Select / filter on SEX = female, RELHFU = 1 or 2 (head of family unit or wife/partner of head of family unit) and AYFL19 = 0-18.
  - NB Selecting / filtering on RELHFU = 1 or 2 ensures that only mothers with dependent children are selected for the analyses. Omitting this filter would mean that women of working age who are themselves children of the family unit would also be included.
- Tables produced from Quanvert need to be run at the person level.
- A figure for women with no children can be obtained by subtracting the number of women with children from all women.

E.g. 2: Women/mothers with dependent children by economic activity

Winter(Dec. 2000 to Feb. 2001)

Economic activity (Ilodefa)	Economic activity (Ilodefa)
Level person	Level person
Weight matrix intwt	Weight matrix intwt
Filter: SEX Female	Filter: SEX Female
and AYFL19 0.0 : 18.0	

Economic activity of women with dependent children		Economic activity of all women		Economic activity of women without dependent children by subtraction	
	Thousands		**Thousands**		**Thousands**
All persons aged	8,698	All persons aged 16+	23,892	All persons aged 16+	15,194
Economically active	5,862	Economically active	13,143	Economically active	7,282
In employment	5,553	In employment	12,576	In employment	7,022
ILO unemployed	309	ILO unemployed	568	ILO unemployed	259
Inactiv	2,837	Inactive	10,749	Inactive	7,912

3. Number of children in the family by family type:

- These are produced from the household datasets.
- For Quanvert analyses only: as FDPCH19 is a numeric variable this needs to be re-created as an axis variable with the required categories and labels in order to produce a cross-tabulation.
- Select / filter on AGE < 19 and cross-tabulate FUTYPE by FDPCH19 (axis version for table produced from Quanvert) and run at the person level.
- It is preferable to use FDPCH19 as this is the derived variable, rather than CAIND.

Table 3: Number of children in family by family type

Autumn(Sep. 2000 to Nov. 2000)
futype : Type of family unit
Level person
Weighted
Filter: AGE < 19.0

	All families	No. of dependent children			
		None	1 to 4	5 or more	Other
					Thousand
Base	14,229	457	13,406	365	*
1 person - male[1]	66	43	23	*	*
1 person - female[1]	72	52	20	*	*
Same sex family unit	*	*	*	*	*
Married couple with no children	*	*	*	*	*
Married couple, non-dep children only	226	226	*	*	*
Married couple with dep children	9,598	*	9,329	269	*
Cohab couple with no children	26	26	*	*	*
Cohab couple with non-dep children only	20	20	*	*	*
Cohab couple with dep children	1,350	*	1,321	29	*
Male lone parent with dep children	275	*	273	*	*
Male lone par with non-dep children only	26	26	*	*	*
Female lone par with dep children	2,504	*	2,439	65	*
Female lone par, non-dep chldren only	61	61	*	*	*
NA	*	*	*	*	*
DNA	*	*	*	*	*

[1] Also includes foster children and others who are counted as separate family units.

* Sample size too small for reliable estimate.

4. Number of children in families with no-one in employment:

- This is produced from the household datasets.
- This analysis looks at children in families rather than households.
- The following example gives the economic activity of parents with dependent children.
- Children will be in either couple or lone parent families. As there is no family level variable for economic activity, the figures have to be produced from running couples and lone parents analyses separately.
- For Quanvert analyses only: both the couple and lone parent family analyses involve producing a cross-tabulation using FDPCH19. Therefore Quanvert analyses should begin by re-creating FDPCH19 as an axis variable with the required categories and labels. This is needed for the cross-tabulations that follow which should be run at the person level in Quanvert.
- For children in couple families: select / filter on RELHFU = 1 (head of family unit), FUTYPE = 6 or 9 (married or cohabiting couples with dependent children).
    - cross-tabulate FDPCH19 = 1-10 (axis version for table produced from Quanvert) by HEAHEAD by HEAWIFE (i.e. creating a cross-tabs for each category of FDPCH19). Then sum the categories of the children from each table by multiplying up (i.e. 2 children x 2, 3 children x 3 etc.) only for those families where neither parent is in work (i.e. unemployed and inactive categories).
- For children in lone parent families: select / filter on RELHFU = 1 (head of family unit), FUTYPE = 10 or 12 (male or female lone parents with dependent children).
    - cross-tabulate ILODEFA by FDPCH19 (axis version for table produced from Quanvert). Sum as above where the economic activity of the parent is either unemployed or inactive.
- Sum the number of children from lone parent and couple families where the parents are not in employment.

E.g. 4: Number of children in workless families

4a: Couples
Autumn(Sep. 2000 to Nov. 2000)
Heahead : Economic activity of head of family unit
heawife : Economic activity of wife of family unit
Level person
Weighted
Filter: RELHFU Head of family unit
and (FUTYPE Married couple with dep children or Cohab couple with dep children)

Base	All	In emp- loyment	ILO Unemp- loyed	Inactive	NA	DNA
						Thousands
Base	5,776	4,060	136	1,518	*	62
Employed	5,081	3,773	110	1,156	*	42
Unemployed	190	68	15	105	*	*
Inactive	367	126	*	230	*	*
NA	*	*	*	*	*	*
DNA	138	93	*	27	*	15
One child						
Base	2,235	1,647	50	510	*	28
Employed	1,959	1,521	42	378	*	18
Unemployed	67	27	*	35	*	*
Inactive	149	60	*	85	*	*
NA	*	*	*	*	*	*
DNA	59	38	*	11	*	*
Two children						
Base	2,505	1,825	53	604	*	23
Employed	2,249	1,709	45	479	*	17
Unemployed	71	27	*	39	*	*
Inactive	130	49	*	77	*	*
NA	*	*	*	*	*	*
DNA	55	41	*	*	*	*

E.g. 4: Number of children in workless families

4a: Couples
Autumn(Sep. 2000 to Nov. 2000)
Heahead : Economic activity of head of family unit
heawife : Economic activity of wife of family unit
Level person
Weighted
Filter: RELHFU Head of family unit
and (FUTYPE Married couple with dep children or Cohab couple with dep children)

Base	All	In emp-loyment	ILO Unemp-loyed	Inactive	NA	DNA
						Thousands
Three children						
Base	805	492	24	281	*	*
Employed	696	456	17	216	*	*
Unemployed	35	11	*	19	*	*
Inactive	55	13	*	40	*	*
NA	*	*	*	*	*	*
DNA	19	12	*	*	*	*

Etc. for remaining categories of FDPCH19

* Sample size too small for reliable estimate.

4b: Lone parents

Autumn(Sep. 2000 to Nov. 2000)
Ilodefa: Economic activity
Level person
Weighted
Filter: RELHFU Head of family unit

and (FUTYPE Male lone parent with dep children or Female dep children)

4c:
Calculation of number of children workless families (by

	All persons	Economic-ally active	In employ-ment	ILO unem-ployed	Inactive	No of children in family	Couples	Lone parents
					Thousands			Thousands
Base	1,624	922	806	116	689	Total	758	1,426
One	875	536	471	66	330	1	128	395
Two	525	296	258	38	225	2	248	526
Three	163	72	64	*	90	3	195	298
Four	49	16	13	*	33	4	105	146
Five	*	*	*	*	*	5	50	33
Six	*	*	*	*	*	6	24	19
Seven	*	*	*	*	*	7	*	*
Eight	*	*	*	*	*	8	*	*
Nine	*	*	*	*	*	9	*	*
Ten	*	*	*	*	*			
Other	*	*	*	*	*			

Total no. of children in workless families = 2,183
(Total of couple and lone parents columns in table 4c above)

* Sample size too small for reliable estimate.

5. Number of workless households by duration:

- This is produced from the household datasets.
- The analysis below uses the standard definition of a workless household, that is, households with at least one person of working age with no-one in employment.
- Create a variable at the household level for the number of people in the household who have left their last job in a specified time period. Select / filter on the appropriate categories of WNLEFT (together

with na and dna categories). This will need to be a numeric variable for Quanvert analyses created using the 'inc' command which produces '0s' and '1s'.

- Produce a table by selecting / filtering on the required categories of the numeric variable, HNWKAGE > 0 and HNEMP = 0.

- Optional: this analysis could also be done by excluding those households where all those of working age in the household are full-time students. Select / filter on HNFTSTUD = 0 when running the table.

Table 5: Workless households by duration

Number of working age households that have been workless for 3 years or
Spring (March 1994 to May 1994)
Level hhold
Weighted
Filter: HHDURUN 0.0   numeric variable created using WNLEFT
and HNWKAGE > 0.0
and HNEMP 0.0
and HNFTSTUD 0.0

	Thousands
Base	**1,559**
All employed	*
All employed or unemployed	*
All employed or inactive	*
All employed, unemployed or inactive	*
All unemployed	220
All unemployed or inactive	150
All inactive	1,189
DNA	*
At least one person without an economic activity	*
NA	*

*   Sample size too small for reliable estimate.

SOME POINTS TO BE WARY ABOUT

Treatment of individuals with no economic activity data

As is described more fully in Section 3, from 1992 to 1995 household or family members for whom no data on economic activity were gathered have no person level records. The derivation of household and family unit type did not therefore take into account such 'missing members', but the household datasets include adjusted household and family type variables which, as far as possible, do take account of them. It has not however been possible to make consistent adjustments to the relationship variables for the other, fully recorded members. There are therefore for 1992 to 1995 some households where not every member has an individual record, and some individuals whose reported relationships are inconsistent with their household or family type. For 1996 onwards the household and family type variables take account of such individuals, and they have person level records, but with most variables unstated. The numbers affected form a small proportion of the whole, but they may have relatively large effects on some small subgroups. The appropriate treatment of these cases when creating new household or family unit level variables depends on the nature and purpose of each new variable.

Publication threshold

Any category where the estimated number of households, (or persons or family units, whichever is the level at which the variable is defined), is less than 10,000, is based on too small a sample to be reliable. Care is needed especially when tabulating by several variables together, or when creating new variables from several other variables.

Analysis of parents and children

Analyses of parents and children should be done at family unit level, since children in a household may not be the children of the household head.

'Children' in a family unit may be the adult offspring of the head of the family unit. Selecting non-adults can be done either on the basis of age, or by restricting the analysis to 'dependent children'. Household type categories distinguish between those containing dependent children and those containing only non-dependent children. The series showing children in workless households published in the LFS Quarterly Supplement relates to children under 16, so that there is no overlap with the series showing people of working age. If a similar analysis showing all dependent children is needed, it can be done by selecting individuals with CAIND > 1 instead of AGE < 16, or by summing HDPCH19 over the relevant households. (But it should be noted that approaching half of all dependent children aged 16 to 18 are in employment, so including this group has the effect of decreasing the proportion of children in workless households.)

Treatment of unknown economic activity for household datasets

Analysis of households

For the household datasets to be suitable for all kinds of household and family level analyses, it is important that households with unknown economic activity (described in Section 5) are given weighting factors in the same way as the other households. However, if no account is taken of the unknown cases, then there may be understatement in the estimation of the numbers and proportions of households in particular categories of household economic activity.

An investigation was made into the effect that the treatment of households with unknown economic activity has on the estimates, particularly of workless households, and to assess different methods of adjustment. This showed that the characteristics of 'unknown' households were not similar to those of 'known' households overall, but that this was a reasonable assumption **within each household type category**.

The adjustment methodology which is used for producing the regularly published series on workless households in the *LFS Quarterly Supplement* is therefore to divide all households according to household type, (combining together some small, similar categories), and **within each household type category** to allocate the 'unknown' households, (or adults or children in unknown households, as appropriate), as workless or not in the same proportions as the households (or adults or children in them) with known economic activity. This methodology has been extended to estimate overall levels of all combined household economic activity states (e.g. households with all persons employed, households with all persons inactive). Adjustment factors for these states are available upon request. See Annex E, Table A.E1 as an example.

This adjustment method operates at the aggregate level, and has only been used for estimating **overall levels** of workless households and adults and children in workless households. It is not suitable for producing adjusted estimates for subgroups. This is because there are a great many possible subgroups which could be defined, and applying this adjustment methodology would produce adjusted estimates which might well not be consistent over different levels and hierarchies of subgroups; and also because for smaller and/or more specialised subgroups the numbers of sample households in the smaller household type categories eventually become small enough to cause volatility in the resulting estimates. Due to increasing demands for adjusted estimates for subgroups, a simple, consistent method was developed and is described in Annex E. This annex also describes the on-going methodological work to improve the treatment of households with unknown economic activity.

Analysis of individuals and families

There is no adjustment procedure for the analysis of individuals (e.g. women's economic activity status by number of dependent children) or families (e.g. economic activity status of couples). Figures are presented unadjusted and percentages are based on totals excluding cases with unknown economic activity.

For example, the proportion of couples where both partners were in work in spring 2000 was 63.7% and is calculated as follows:

Data in thousands

No. of couples – both partners in work	/	(Total no. of working-age couples	-	couples with unknown economic activity) x 100
7,550	/	(12,338	-	487) x 100

SECTION 9 – PUBLICATION OF LFS HOUSEHOLD DATA

REGULAR PUBLICATION OF LFS HOUSEHOLD DATA

- LFS Quarterly Supplement – updated biannually with spring or autumn household data
- LFS Historical Supplement – generally updated annually with spring household data
- StatBase – updated annually with spring household data
- Labour Market Trends (pink pages) – updates for spring or autumn household data are published biannually
- Biannual press release presenting the latest figures for key series from the LFS household datasets.
- Analyses from the household datasets have also appeared in the annual publications, Social and Regional Trends and the Social Focus series.

LABOUR MARKET TRENDS ARTICLES RELATING TO HOUSEHOLD AND FAMILY ANALYSES FROM THE LFS

- Data on household and families from the Labour Force Survey	March 1997
- Household and family data from the Labour Force Survey: recent improvements in approach (illustrative data)	June 1997
- Economic activity of working-age households	September 1997
- Workless households, unemployment and economic inactivity	September 1997
- Analysis of household data from the Labour Force Survey	August 1998
- Production of household datasets from the Labour Force Survey	August 1998
- Additions to Labour Force Survey household tables	September 1998
- LFS household data: spring 1998 analyses	January 1999
- Update on Labour Force Survey household datasets	May 1999
- LFS household data: spring 1999 analyses	January 2000
- LFS household data: spring 2000 analyses	January 2001

See Volume 1 (Background and Methodology) of this User Guide for more information on the dissemination and publication of LFS data.

ANNEX A: DETAILS OF NEW HOUSEHOLD AND FAMILY LEVEL VARIABLES

NOTE: These household and family level variables are specific to the household datasets. The frequency for variables is for all household datasets unless otherwise stated.

1. WEIGHT AND IDENTIFIER VARIABLES

HHWT2 – Household weight used in the regrossing exercise

Use this variable to weight to give number of **persons**.

FREQUENCY: This is for the period autumn 1996 onwards.

To give the number of **households**, select head of household using the variable:

RELH96 = 0 from autumn 1996 to autumn 2000.

Note: From spring 2001 use the relationship to household reference person variable (RELHRP) (see Section 4 for more details).

To give the number of **family units**, select head of family unit using the variable:

RELHFU = 1.

NEWHHWT - Household weight used prior to the regrossing exercise in 2000

Use this variable to weight to give number of **persons**.

To give the number of **households**, select head of household using the variable:

RELTOHOH = 1 for spring 1990;

HHIND = 1 for spring 1992 to autumn 1995;

RELHOH = 1 for spring 1996;

To give the number of **family units**, select head of family unit using the variable:

RELHFU = 1.

HHIND - Indicator variable for household level analysis for 1992 to 1995 (not in Quanvert)

Takes the value 1 for the person with the lowest value of RECNO in a household, and 0 otherwise.

FREQUENCY: This is for the period spring 1992 to autumn 1995 only.

HSERIAL - Number that uniquely identifies a household (not in Quanvert)

HSERIAL = SERNO for spring 1990.

HSERIAL = (QUOTA*10000000) + (WEEK*100000) + (THISWV*10000) + (ADD*100) + HHLD from spring 1992 to spring 2000.

HSERIAL = (QUOTA*1000000000) + (WEEK*10000000) + (W1YR*1000000) + (QRTR*100000) + (ADD*1000) + (WAVFND*100) + HHLD from autumn 2000 onwards (see Annex D for more details).

FUSERIAL - Number that uniquely identifies a family unit (not in Quanvert)

FUSERIAL = (SERNO*100) + FAMUNIT for spring 1990.

FUSERIAL = (QUOTA*1000000000) + (WEEK*10000000) + (THISWV*1000000) + (ADD*10000) + (HHLD*100) + FAMUNIT from spring 1992 to spring 2000.

FUSERIAL = (QUOTA*100000000000) + (WEEK*1000000000) + (W1YR*100000000) + (QRTR*10000000) + (ADD*100000) + (WAVFND*10000) + (HHLD*100) + FAMUNIT from autumn 2000 onwards (see Annex D for more details).

2. ADJUSTED HOUSEHOLD AND FAMILY TYPE

THHADJ

Household type after adjustments described in Section 3. This variable has the same categories as TYPEHH and should be used as the household type.

FREQUENCY: This is for the period spring 1992 to autumn 1995 only.

TFUADJ

Family type after adjustments described in Section 3. This variable has the same categories as TYPEFU and should be used as the family type.

FREQUENCY: This is for the period spring 1992 to autumn 1995 only.

HHCHANGE

Indicates whether the adjusted household type THHADJ is different from the unadjusted household type TYPEHH.
1 indicates the household type has been adjusted
0 indicates no adjustment has been made.

FREQUENCY: This is for the period spring 1992 to autumn 1995 only.

FUCHANGE

Indicates whether the adjusted family type TFUADJ is different from the unadjusted family type TYPEFU.
1 indicates the family type has been adjusted
0 indicates no adjustment has been made.

FREQUENCY: This is for the period spring 1992 to autumn 1995 only.

3. HOUSEHOLD COMPOSITION BY AGE

HNWKAGE - Number of people in the household who are of working age

Working age is between 16 and 59 years for females and 16 and 64 years for males.

HNPEN - Number of people in the household who are of pensionable age

Pension age is 60 years and over for females and 65 years and over for males.

HDPCH4 - Number of children in the household aged 4 years or less

Based on a person satisfying the criteria AGE <= 4 and CAIND =
2   Child of head of household and head of family unit
3   Child of other family
4   Child of head of household but not head of family unit

NOTE: CAIND identifies dependent children, i.e. aged < 16 or aged 16 to 18 in full-time education and never married.

HDC515 - Number of children in the household aged between 5 and 15 years

Based on a person satisfying the criteria AGE >= 5 and AGE <= 15 and CAIND =
2   Child of head of household and head of family unit
3   Child of other family
4   Child of head of household but not head of family unit

NOTE: CAIND identifies dependent children, i.e. aged < 16 or aged 16 to 18 in full-time education and never married.

HDPCH18 - Number of children in the household aged 16 years and over

Based on a person satisfying the criteria AGE >= 16 and CAIND =
2   Child of head of household and head of family unit
3   Child of other family
4   Child of head of household but not head of family unit

NOTE: CAIND identifies dependent children, i.e. aged < 16 or aged 16 to 18 in full-time education and never married.

4. HOUSEHOLD COMPOSITION BY ECONOMIC ACTIVITY

HNDK - Number of people in the household whose basic economic activity is not known

Based on a person satisfying the criteria:

IOUTCOME = 3 and FACTORSP = 0 for spring 1990

INECACA = -9 from spring 1996

NOTE: This variable is not available for spring 1992 to autumn 1995.

HNEMP - Number of people in the household who are employed

Based on a person satisfying the criteria ECAR =
1   Employee
2   Self-employed
3   Employment status not stated
4   YTS - employer based
5   YTS - college based
6   ET employer/practical training
7   ET college based
8   Other schemes

FREQUENCY: spring 1990

NOTE: Set to unknown (-9) for HNDK > 0.
Replaces variables NOEMHH and NOSCHH.

OR

Based on a person satisfying the criteria INECACA =
1   Employee
2   Self-employed
3   Government employment and training programs
4   Unpaid family worker

FREQUENCY: From spring 1992 onwards.

NOTE: Set to unknown (-9) if THHADJ or TFUADJ unknown (-9), or HHCHANGE or FUCHANGE = 1, for spring 1992 to autumn 1995; or HNDK > 0 from spring 1996.

HNUNEMP - Number of people in the household who are unemployed

Based on a person satisfying the criteria ECAR =
9   Seeking work
10   Temporarily sick
11   On holiday
12   Waiting to start new job
13   Waiting for job application results

FREQUENCY: spring 1990

NOTE: Set to unknown (-9) for HNDK > 0.
Replaces variable NOUNHH.

OR

Based on a person satisfying the criteria INECACA =
5   ILO unemployed

FREQUENCY: From spring 1992.

NOTE: Set to unknown (-9) if THHADJ or TFUADJ unknown (-9), or HHCHANGE or FUCHANGE = 1, for spring 1992 to autumn 1995; or HNDK > 0 from spring 1996.

HNINACT - Number of people in the household who are inactive

Based on a person satisfying the criteria ECAR =
14   Full time student - not working or seeking work
15   Full time student - seeking but must complete education
16   Other student
17   Man aged 70 or over, woman aged 65 or over
18   Long term sick or disabled
19   Looking after family/home
20   Retired
21   Does not want/need employment
22   Believes no jobs available
23   Not yet starting looking for employment
24   Other reason not looking for employment
25   No reason given not looking for employment

FREQUENCY: spring 1990

NOTE: Set to unknown (-9) for HNDK > 0.
Replaces variable NOEIHH.

OR

Based on a person satisfying the criteria INECACA =
6   Inactive - seeking, unavailable, student
7   Inactive - seeking, unavailable, looking after family, home
8   Inactive - seeking, unavailable, temporarily sick or injured
9   Inactive - seeking, unavailable, long-term sick or disabled
10   Inactive - seeking, unavailable, other reason
11   Inactive - seeking, unavailable, no reason given
12   Inactive - not seeking, would like work, waiting results of job application
13   Inactive - not seeking, would like work, student
14   Inactive - not seeking, would like work, looking after family, home

15 Inactive - not seeking, would like work, temporarily sick or injured
16 Inactive - not seeking, would like work, long-term sick or disabled
17 Inactive - not seeking, would like work, believes no job available
18 Inactive - not seeking, would like work, not started looking
19 Inactive - not seeking, would like work, not looked
20 Inactive - not seeking, would like work, no reason
21 Inactive - not seeking, not like work, waiting results of job application
22 Inactive - not seeking, not like work, student
23 Inactive - not seeking, not like work, looking after family, home
24 Inactive - not seeking, not like work, temporarily sick or injured
25 Inactive - not seeking, not like work, long-term sick or disabled
26 Inactive - not seeking, not like work, not need or want job
27 Inactive - not seeking, not like work, retired
28 Inactive - not seeking, not like work, other reason
29 Inactive - not seeking, not like work, no reason given

FREQUENCY: From spring 1992 onwards.

NOTE: Set to unknown (-9) if THHADJ or TFUADJ unknown (-9), or HHCHANGE or FUCHANGE = 1, for spring 1992 to autumn 1995; or HNDK > 0 from spring 1996.

HEACOMB - Household economic activity variable

1 All persons in the household are employed
2 All persons in the household are either employed or unemployed
3 All persons in the household are either employed or inactive
4 All persons in the household are either employed, unemployed or inactive
5 All persons in the household are unemployed
6 All persons in the household are either unemployed or inactive
7 All persons in the household are inactive
-9 Unknown

NOTE: Based on a combination of the variables HNEMP, HNUNEMP and HNINACT. Set to unknown (-9) if THHADJ or TFUADJ unknown (-9), or HHCHANGE or FUCHANGE = 1, for spring 1992 to autumn 1995; or HNDK > 0 from spring 1996.

HNFTSTUD - Number of people in the household who are full-time students

Based on a person satisfying the criteria CURED =
1 At school
2 Sandwich course
3 Full-time at university or college

NOTE: From spring 1992 to winter 1996 this also included full-time at polytechnic.

HNOTSTUD - Number of people in the household who are not full-time students

All persons in the household who are not classified as a full-time student as defined above. Also includes CURED not stated.

HNFTIME - Number of people in the household who are working full-time

Based on a person satisfying the criteria

FTPTWORK = 1 and ECAR = 1,2,3,4,5,6,7 or 8 for spring 1990;

FTPTWK = 1 and INECACA = 1,2,3 or 4 for spring 1992 to spring 1994;

FTPT = 1 and INECACA = 1,2,3 or 4 from spring 1995.

HNPTIME - Number of people in the household who are working part-time

Based on a person satisfying the criteria

FTPTWORK = 2 and ECAR = 1,2,3,4,5,6,7 or 8 for spring 1990;

FTPTWK = 2 and INECACA = 1,2,3 or 4 for spring 1992 to spring 1994;

FTPT = 2 and INECACA = 1,2,3 or 4 from spring 1995.

Note that the following 6 variables are not available for spring 1990 because information on inactivity was more limited at that time.

HNIWSTU - Number of people in the household who are inactive and would like to work and are currently students

Based on a person satisfying the criteria INECACA =
6   Inactive - seeking, unavailable, student
13   Inactive - not seeking, would like work, student

FREQUENCY: From spring 1992.

NOTE: Set to unknown (-9) if THHADJ or TFUADJ unknown (-9), or HHCHANGE or FUCHANGE = 1, for spring 1992 to autumn 1995; or HNDK > 0 from spring 1996.

HNIWSKD - Number of people in the household who are inactive and would like to work and are currently sick, injured or disabled

Based on a person satisfying the criteria INECACA =
8   Inactive - seeking, unavailable, temporarily sick or injured
9   Inactive - seeking, unavailable, long-term sick or disabled
15   Inactive - not seeking, would like work, temporarily sick or injured
16   Inactive - not seeking, would like work, long-term sick or disabled

FREQUENCY: From spring 1992.

NOTE: Set to unknown (-9) if THHADJ or TFUADJ unknown (-9), or HHCHANGE or FUCHANGE = 1, for spring 1992 to autumn 1995; or HNDK > 0 from spring 1996.

HNIWDSC - Number of people in the household who are inactive and would like to work but discouraged from seeking work

Based on a person satisfying the criteria INECACA =
17   Inactive - not seeking, would like work, believes no job available

FREQUENCY: From spring 1992.

NOTE: Set to unknown (-9) if THHADJ or TFUADJ unknown (-9), or HHCHANGE or FUCHANGE = 1, for spring 1992 to autumn 1995; or HNDK > 0 from spring 1996.

HNIWFAM - Number of people in the household who are inactive and would like to work but looking after family or home

Based on a person satisfying the criteria INECACA =
7   Inactive - seeking, unavailable, looking after family, home
14   Inactive - not seeking, would like work, looking after family, home

FREQUENCY: From spring 1992.

NOTE: Set to unknown (-9) if THHADJ or TFUADJ unknown (-9), or HHCHANGE or FUCHANGE = 1, for spring 1992 to autumn 1995; or HNDK > 0 from spring 1996.

HNIWOTH - Number of people in the household who are inactive for other reasons but would like to work

Based on a person satisfying the criteria INECACA =
10   Inactive - seeking, unavailable, other reason
11   Inactive - seeking, unavailable, no reason given
12   Inactive - not seeking, would like work, waiting results of job application
18   Inactive - not seeking, would like work, not started looking
19   Inactive - not seeking, would like work, not looked
20   Inactive - not seeking, would like work, no reason

FREQUENCY: From spring 1992.

NOTE: Set to unknown (-9) if THHADJ or TFUADJ unknown (-9), or HHCHANGE or FUCHANGE = 1, for spring 1992 to autumn 1995; or HNDK > 0 from spring 1996.

HNINOWK - Number of people in the household who are inactive and do not want to work

Based on a person satisfying the criteria INECACA =
21   Inactive - not seeking, not like work, waiting results of job application
22   Inactive - not seeking, not like work, student
23   Inactive - not seeking, not like work, looking after family, home
24   Inactive - not seeking, not like work, temporarily sick or injured
25   Inactive - not seeking, not like work, long-term sick or disabled
26   Inactive - not seeking, not like work, not need or want job
27   Inactive - not seeking, not like work, retired
28   Inactive - not seeking, not like work, other reason
29   Inactive - not seeking, not like work, no reason given

FREQUENCY: From spring 1992.

NOTE: Set to unknown (-9) if THHADJ or TFUADJ unknown (-9), or HHCHANGE or FUCHANGE = 1, for spring 1992 to autumn 1995; or HNDK > 0 from spring 1996.

5. FAMILY UNIT VARIABLES

HEAHEAD - Economic activity of the head of the family unit

    1   Employed
    2   Unemployed
    3   Inactive

NOTE: Set to unknown (-9) for ECAR = -9 for spring 1990; THHADJ or TFUADJ unknown (-9) or FUCHANGE = 1 for spring 1992 to autumn 1995; or INECACA = -9 from spring 1996.
Can take the value 0 if head of family unit is less than 16 years old, (as can be the case for foster children).

HEAWIFE - Economic activity of the wife of the family unit

    1   Employed
    2   Unemployed
    3   Inactive

NOTE: Set to unknown (-9) for ECAR = -9 for spring 1990; THHADJ or TFUADJ unknown (-9) or FUCHANGE = 1 for spring 1992 to autumn 1995; or INECACA = -9 from spring 1996.

This variable is only coded where wife/partner of head of family unit is present.

ANNEX B(I): GUIDANCE ON CREATING HOUSEHOLD AND FAMILY UNIT LEVEL VARIABLES

SPSS version

Joint characteristics of married/cohabiting couples

For looking at the characteristics of husbands and wives or cohabitees together, the simplest approach is to create two variables *at family unit level*, one holding the characteristic of the husband and the other holding that of the wife, similar to the variables HEAHEAD and HEAWIFE which have already been created for the broad economic activity of husbands and wives. These pairs of variables can then be cross-tabulated at family unit level to give the joint distribution of husbands and wives, or further cross-tabulated and/or selected for more elaborate analyses.

First, to create the two variables for the husband's and wife's characteristics **based on an existing variable**, the procedure is shown below using the existing variable for the ethnic group in broad groupings (ETHNICA). The variables created are ETHHEAD for the ethnic group of the head of family unit, and ETHWIFE for the ethnic group of the wife/cohabitee of the head of family unit - both have the same categories as ETHNICA.

```
GET FILE='hhmm98.sav'
/KEEP RELHFU ETHNICA FUSERIAL.

COMPUTE ETHHEAD=-9.

SELECT IF (RELHFU EQ 1).

IF (ETHNICA EQ 0) ETHHEAD=0.
IF (ETHNICA EQ 1) ETHHEAD=1.
IF (ETHNICA EQ 2) ETHHEAD=2.
IF (ETHNICA EQ 3) ETHHEAD=3.
IF (ETHNICA EQ 4) ETHHEAD=4.

SORT CASES BY FUSERIAL.

SAVE OUTFILE='mm98head.sav'.

GET FILE='hhmm98.sav'
/KEEP RELHFU ETHNICA FUSERIAL.

COMPUTE ETHWIFE=-9.

SELECT IF (RELHFU EQ 2).

IF (ETHNICA EQ 0) ETHWIFE=0.
IF (ETHNICA EQ 1) ETHWIFE=1.
IF (ETHNICA EQ 2) ETHWIFE=2.
IF (ETHNICA EQ 3) ETHWIFE=3.
IF (ETHNICA EQ 4) ETHWIFE=4.

SORT CASES BY FUSERIAL.

SAVE OUTFILE='mm98wife.sav'.

MATCH FILES FILE='mm98head.sav'
/TABLE='mm98wife.sav'
/BY FUSERIAL.

SAVE OUTFILE='mm98headwife.sav'.

MATCH FILES FILE='hhmm98.sav'
/TABLE='mm98headwife.sav'
/BY FUSERIAL.
```

```
VALUE LABELS ETHHEAD ETHWIFE 0 'White'
                             1 'Black'
                             2 'Indian'
                             3 'Pakistani/Bangladeshi'
                              4 'Mixed/Other'.
SAVE OUTFILE='hhmm98add1.sav'.
```

When analysing these variables, one record per family unit must be selected, e.g. by RELHFU = 1, (unless the analysis is to be of all individual family members, in which case all records should be retained - see below for an example of this.) If the analysis is to be restricted to couple family units, (i.e. to exclude family units with a head and no wife/cohabitee), then select by RELHFU = 2, (wife/cohabitee of head of family unit). If this is not done, the heads of family unit who do not have a wife/ cohabitee will appear in the tabulation, though they will have DNA for the wife's category.

Next, to create the two variables for the husband's and wife's characteristics **combining information from two (or more) existing variables**, and perhaps combining some categories, the procedure is shown below, using as an example a pair of new variables derived from ETHNICA and country of birth (CRY), with categories: white born in UK; white not born in UK; not white born in UK; not white not born in UK.

```
GET FILE='hhmm98.sav'
/KEEP RELHFU ETHNICA CRY FUSERIAL.

COMPUTE ETBHEAD=-9.

SELECT IF (RELHFU EQ 1).

IF (ETHNICA EQ 0 AND CRY EQ 1) ETBHEAD=1.
IF (ETHNICA EQ 0 AND CRY GT 1) ETBHEAD=2.
IF (ETHNICA GT 0 AND CRY EQ 1) ETBHEAD=3.
IF (ETHNICA GT 0 AND CRY GT 1) ETBHEAD=4.

SORT CASES BY FUSERIAL.

SAVE OUTFILE='mm98hed.sav'.

GET FILE='hhmm98.sav'
/KEEP RELHFU ETHNICA CRY FUSERIAL.

COMPUTE ETBWIFE=-9.

SELECT IF (RELHFU EQ 2).

IF (ETHNICA EQ 0 AND CRY EQ 1) ETBWIFE=1.
IF (ETHNICA EQ 0 AND CRY GT 1) ETBWIFE=2.
IF (ETHNICA GT 0 AND CRY EQ 1) ETBWIFE=3.
IF (ETHNICA GT 0 AND CRY GT 1) ETBWIFE=4.

SORT CASES BY FUSERIAL.

SAVE OUTFILE='mm98wif.sav'.

MATCH FILES FILE='mm98hed.sav'
/TABLE='mm98wif.sav'
/BY FUSERIAL.

SAVE OUTFILE='mm98hedwif.sav'.

MATCH FILES FILE='hhmm98.sav'
/TABLE='mm98hedwif.sav'
/BY FUSERIAL.

VALUE LABELS ETBHEAD ETBWIFE 1 'White born in UK'
                            2 'White not born in UK'
                            3 'Not white born in UK'
                            4 'Not white not born in UK'.

SAVE OUTFILE='hhmm98add2.sav'.
```

Joint characteristics of children and parents

One major use of the kind of family level variables described above is to analyse children in terms of their own and their parents' characteristics. This can include both simple estimates of the numbers of children in families with particular combinations of parental characteristics, and cross-tabulations of characteristics of both children and parents. (But note that very small sample sizes are quickly reached by putting together multiple categorisations - LFS estimates of less than 10,000 are regarded as based on too small a sample size to be reliable.)

It is important to use **family unit level variables** in relating the characteristics of parents and children. Household level variables are not usually appropriate, since the children in a household may not be children of the head of the household. But there is a complication even in using family unit level data, since the LFS definition of a family unit includes any 'children' of the head of the family unit who are never-married and without children of their own in the household - there is no age restriction. So to analyse children of a family unit who are not adults, it is necessary to filter on the age or some other suitable indicator variable. The two main possibilities are to filter on AGE < 16, thereby restricting the analysis to children aged under 16; or to filter on CAIND > 1, thus selecting the 'child' categories - this will restrict the analysis to 'dependent children', defined in the LFS as children aged under 16, or those aged 16 to 18 who are never-married and in full-time education.

As an example, to estimate the number of children whose parents fall into each category of ETHNICA, tabulate by the variables ETHHEAD and ETHWIFE defined above, select on AGE or CAIND to restrict the analysis to children of the requisite definition, and tabulate all cases, (do not restrict to those with RELHFU = 1).

To cross-tabulate also by a characteristic of the children, for example their ethnic group, tabulate by ETHHEAD, ETHWIFE and ETHNICA, filter on AGE or CAIND to select children, and tabulate all cases.

To cross-tabulate by an additional characteristic of the parents, for example their ages, it will be necessary to create another pair of family unit level variables, one each for the age of the head and of the wife, using the methods described above, and add these variables to those in the tabulation; (**and to check that the sample sizes are still large enough**). Or instead of creating several variables for different characteristics of the parents, it is possible to create a single more elaborate pair of variables which combine categories of several different characteristics - but this approach is less versatile and flexible.

Household level variables counting people with a specified characteristic

Turning to analyses at household level, these often need a variable which counts the number of people in the household with a specific characteristic. For the main labour market characteristics a number of such variables are already provided in the household datasets, for example the number of adults in a household who are employed, or unemployed, or in various types of inactivity, or employed full-time or part-time. For users who need similar variables based on other characteristics, the procedure is shown below, using number of people aged 16 and over in the black ethnic group as an example.

```
GET FILE='hhmm98.sav'
/KEEP HSERIAL ETHNICA AGE.

COMPUTE ETHBL=0.

IF (ETHNICA EQ 1 AND AGE GE 16)  ETHBL=1.

SORT CASES BY HSERIAL.

AGGREGATE OUTFILE=*
/BREAK=HSERIAL
/HNETHBL=SUM(ETHBL).

SAVE OUTFILE='mm98ethb.sav'.

MATCH FILES FILE='hhmm98.sav'
/TABLE='mm98ethb.sav'
/BY HSERIAL.

SAVE OUTFILE='hhmm98add3.sav'.
```

Household level variables describing the combined characteristics of the members

The 'count' variables described above can also be used to construct household level variables which describe the combined characteristics of the members, in a similar way to the combined household economic activity variable HEACOMB which is already available in the household datasets. For example, suppose we have a set of 'count' variables similar to HNETHBL (the number of black adults) used as an example above, but for each category of ETHNICA, so that there are also variables HNETHWH for the number of white adults, HNETHIN for Indians, and so on. Then these can be used to create a variable for the combined ethnic group category of the household, with categories: only white; only black; . . . . only mixed/other; white + black; white + Indian; . . . ; Pakistani/Bangladeshi + mixed; and so on. But it is already clear that even with five categories in the original variable, the combined variable is both unmanageable and likely to produce unacceptably small sample sizes. Three categories in the original variable is about the maximum practicable, unless there is extensive grouping of the combined variable categories. The example below uses three ethnic group categories - white, black and other - in creating a variable HETHCOM for the combined ethnic group composition of the adults in the household. It has seven categories: only white; only black; only other; white + black; white + other; black + other; white + black + other. It is created as follows, using the count variable HNETHBL described above, and similar variables HNETHWH and HNETHOT for the numbers of white and other ethnic group adults respectively. (Note that this can be done at the level of the individual record, as shown below, or at the aggregated household level at the same time as creating the 'count' variables as shown above.)

```
GET FILE='hhmm98.sav'.

COMPUTE HETHCOM=0.

IF (HNETHBL GT 0 AND HNETHWH EQ 0 AND HNETHOT EQ 0) HETHCOM=1.
IF (HNETHBL EQ 0 AND HNETHWH GT 0 AND HNETHOT EQ 0) HETHCOM=2.
IF (HNETHBL EQ 0 AND HNETHWH EQ 0 AND HNETHOT GT 0) HETHCOM=3.
IF (HNETHBL GT 0 AND HNETHWH GT 0 AND HNETHOT EQ 0) HETHCOM=4.
IF (HNETHBL GT 0 AND HNETHWH EQ 0 AND HNETHOT GT 0) HETHCOM=5.
IF (HNETHBL EQ 0 AND HNETHWH GT 0 AND HNETHOT GT 0) HETHCOM=6.
IF (HNETHBL GT 0 AND HNETHWH GT 0 AND HNETHOT GT 0) HETHCOM=7.

VALUE LABELS HETHCOM  1 'Only white'
                      2 'Only black'
                      3 'Only other'
                      4 'White + black'
                    5 'White + other'
                      6 'Black + other'
                      7 'White +black + other'.

SAVE OUTFILE='hhmm98add4.sav'.
```

ANNEX B(II): GUIDANCE ON CREATING HOUSEHOLD AND FAMILY UNIT LEVEL VARIABLES

Quanvert version

Joint characteristics of married/cohabiting couples

For looking at the characteristics of husbands and wives or cohabitees together, the simplest approach is to create two variables *at family unit level*, one holding the characteristic of the husband and the other holding that of the wife, similar to the variables HEAHEAD and HEAWIFE which have already been created for the broad economic activity of husbands and wives. These pairs of variables can then be cross-tabulated at family unit level to give the joint distribution of husbands and wives, or further cross-tabulated and/or selected for more elaborate analyses.

First, to create the two variables for the husband's and wife's characteristics **based on an existing variable**, the procedure is shown below using the existing variable for the ethnic group in broad groupings (ETHNICA). The variables created are ETHHEAD for the ethnic group of the head of family unit, and ETHWIFE for the ethnic group of the wife/cohabitee of the head of family unit - both have the same categories as ETHNICA. The Quanvert prompt is shown in ordinary type, and the user's response is shown in **bold**.

```
Command:  c
New axis name:  ethhead
Enter overall conditions for this axis, if any -
Filter:  relhfu
Which elements of relhfu:  head
            Head of family unit
                                                        Or:  <rtn>
And:  <rtn>
Enter titles for this axis, if any -
Title:  Ethnic group of head of FU
Title:  <rtn>
Enter heading for this axis, if any -
Heading:  ethhead
Enter element definitions according to the prompts -
Element 1 - enter element type:  a
Existing axis name:  ethnica
(B) Base - action:  t
8 elements taken as they are
 ---- End of axis ethnica
Element 9 - enter element type:  <rtn>
Enter column headings, if any -
            Line 1: <rtn>
Level:  fmunit
```

(Quanvert then starts processing the new variable, and eventually produces the following prompt, at which point start specifying the second of the pair of variables, relating to the wife)

```
New axis name:  ethwife
Enter overall conditions for this axis, if any -
Filter:  relhfu
Which elements of relhfu:  wife
            Wife or cohabitee of head of family unit
                                                        Or:  <rtn>
And:  <rtn>
Enter titles for this axis, if any -
Title:  Ethnic group of wife of head of FU
Title:  <rtn>
Enter heading for this axis, if any -
Heading:  ethwife
Enter element definitions according to the prompts -
Element 1 - enter element type:  a
Existing axis name:  ethnica
(B) Base - action:  t
```

8 elements taken as they are
 ---- End of axis ethnica
Element 9 - enter element type: **<rtn>**
Enter column headings, if any -
            Line 1: **<rtn>**
Level: **fmunit**

(Quanvert then starts processing the second variable)

Notes: (i) <rtn> means just press the Return key.
        (ii) It is optional to give a title and heading for the new variable, but it is strongly advised in order to keep track of the new variables by making them easily identifiable.

When analysing these variables, the level must be set to **fmunit**, (unless the analysis is to be of all individual family members, in which case the level should be person - see below for an example of this.) If the analysis is to be restricted to couple family units, (i.e. to exclude family units with a head and no wife/cohabitee), then set a filter for the tabulation of **relhfu** restricted to **wife**. If this is not done, the heads of family unit who do not have a wife/ cohabitee will appear in the tabulation, though they will have DNA for the wife's category.

Next, to create the two variables for the husband's and wife's characteristics **combining information from two (or more) existing variables**, and perhaps combining some categories, there are two possible approaches. One is to create directly a pair of variables at family unit level, one each for the head of family unit and the wife of head, with the categories derived by combining the existing variables as required. The procedure is shown below, using as an example a pair of new variables derived from ETHNICA and country of birth (CRY), with categories: white born in UK; white not born in UK; not white born in UK; not white not born in UK.

Command: **c**
New axis name: **etbhead**
Enter overall conditions for this axis, if any -
Filter: **relhfu**
Which elements of relhfu: **head**
        Head of family unit
                                          Or: **<rtn>**
And: **<rtn>**
Enter titles for this axis, if any -
Title: **Ethnic group and country of birth of head of FU**
Title: **<rtn>**
Enter heading for this axis, if any -
Heading: **etbhead**
Enter element definitions according to the prompts -
Element 1 - enter element type: **b**
Enter conditions for this row, if any -
Filter: **<rtn>**
Side text: **base**
Enter any options: **<rtn>**
Element 1 defined
Element 2 - enter element type: **r**
Enter conditions for this row, if any -
Filter: **ethnica**
Which elements of ethnica: **white**
        White
                                            Or: **<rtn>**
And: **cry**
Which elements of cry: **UK**
        UK, British
                                            Or: **<rtn>**
And: **<rtn>**
Side text: **White UK-born**
Enter any options: **<rtn>**
Element 2 defined
Element 3 - enter element type: **r**
Enter conditions for this row, if any -

Filter: **ethnica**
Which elements of ethnica: **white**
      White
                                                          Or: **<rtn>**
And: **cry**
Which elements of cry: |**UK**
      Not UK, British
                                                          Or: **<rtn>**
And: **<rtn>**
Side text: **White not UK-born**
Enter any options: **<rtn>**
Element 3 defined
Element 4 - enter element type: **r**
Enter conditions for this row, if any -
Filter: **ethnica**
Which elements of ethnica: |**white**
      Not White
                                                          Or: **<rtn>**
And: **cry**
Which elements of cry: **UK**
      UK, British
                                                          Or: **<rtn>**
And: **<rtn>**
Side text: **Not white UK-born**
Enter any options: **<rtn>**
Element 4 defined
Element 5 - enter element type: **r**
Enter conditions for this row, if any -
Filter: **ethnica**
Which elements of ethnica: |**white**
      Not White
                                                          Or: **<rtn>**
And: **cry**
Which elements of cry: |**UK**
      Not UK, British
                                                           Or: **<rtn>**
And: **<rtn>**
Side text: **Not white not UK-born**
Enter any options: **<rtn>**
Element 5 defined
Element 6 - enter element type: **<rtn>**
Enter column headings, if any -
             Line 1: **<rtn>**
Level: **fmunit**

(Quanvert then starts processing the head of FU variable. The variable for wife of head of FU is the same except for the changes to the overall filter and the titles and headings which are shown in the previous example.)

An alternative approach is to first create a new axis variable at person level, with categories derived by combining the existing variables as required, and then using this new variable to create the family unit level variables for the head of family unit and the wife of head. The first step is done as for the above example of ETBHEAD, but with no overall condition at the beginning, and with the level set to **person** instead of famunit at the end. The second step is done as in the first example of ETHHEAD and ETHWIFE, replacing ETHNICA with the newly created variable.

Joint characteristics of children and parents

One major use of the kind of family level variables described above is to analyse children in terms of their own and their parents' characteristics. This can include both simple estimates of the numbers of children in families with particular combinations of parental characteristics, and cross-tabulations of characteristics of both children and parents. (But note that very small sample sizes are quickly reached by putting together multiple categorisations - LFS estimates of less than 10,000 are regarded as based on too small a sample size to be reliable.)

It is important to use **family unit level variables** in relating the characteristics of parents and children. Household level variables are not usually appropriate, since the children in a household may not be children of the head of the household. But there is a complication even in using family unit level data, since the LFS definition of a family unit includes any 'children' of the head of the family unit who are never-married and without children of their own in the household - there is no age restriction. So to analyse children of a family unit who are not adults, it is necessary to filter on the age or some other suitable indicator variable. The two main possibilities are to filter on AGE < 16, thereby restricting the analysis to children aged under 16; or to filter on CAIND, selecting the elements which include 'child' - this will restrict the analysis to 'dependent children', defined in the LFS as children aged under 16, or those aged 16 to 18 who are never-married and in full-time education.

As an example, to estimate the number of children whose parents fall into each category of ETHNICA, tabulate by the variables ETHHEAD and ETHWIFE defined above, filter on AGE or CAIND to restrict the analysis to children of the requisite definition, and set the level to person.

To cross-tabulate also by a characteristic of the children, for example their ethnic group, tabulate by ETHHEAD, ETHWIFE and ETHNICA, filter on AGE or CAIND to select children, and set the level of analysis to person.

To cross-tabulate by an additional characteristic of the parents, for example their ages, it will be necessary to create another pair of family unit level variables, one each for the age of the head and of the wife, using the methods described above, and add these variables to those in the tabulation; (and to **check that the sample sizes are still large enough**). Or instead of creating several variables for different characteristics of the parents, it is possible to create a single more elaborate pair of variables which combine categories of several different characteristics - but this approach is less versatile and flexible.

Household level variables counting people with a specified characteristic

Turning to analyses at household level, these often need a variable which counts the number of people in the household with a specific characteristic. For the main labour market characteristics a number of such variables are already provided in the household datasets, for example the number of adults in a household who are employed, or unemployed, or in various types of inactivity, or employed full-time or part-time. For users who need similar variables based on other characteristics, the procedure in Quanvert is first to create a numeric variable for the number of people in the household, filtering on the characteristic of interest. This numeric variable can be used for subsequent filtering, including the creation of further variables, but if it is desired to tabulate by the numeric variable, an axis variable has to be created based on it. The procedure is shown below, using number of people aged 16 and over in the black ethnic group as an example.

First the creation of the numeric variable:

> Command: **inc**
> New numeric variable name: **hnethbl**
> Enter filters if any at lower level ('fmunit', 'person')
> Filter: **ethnica**
> Which elements of ethnica: **black**
>   - Black
>   Or: **<rtn>**
> And: **age**
> Enter value or range for age: **>15**
>   Or: **<rtn>**
> And: **<rtn>**
> Enter units if any: **<rtn>**
> At which higher level is hnethbl to be created ('fmunit', 'hhold')
> Level: **hhold**
>
> *(Quanvert then starts processing the new variable)*

Then the creation of an axis variable based on the numeric variable:

Command: **c**
New axis name: **hethbl**
Enter overall conditions for this axis, if any -
Filter: <**rtn**>
Enter titles for this axis, if any -
Title: **No. adults of black ethnic group in household**
Title: **<rtn>**
Enter heading for this axis, if any -
Heading: **hethbl**
Enter element definitions according to the prompts -
Element 1 - enter element type: **b**
Enter conditions for this row, if any -
Filter: **<rtn>**
Side text: **base**
Enter any options: **<rtn>**
Element 1 defined
Element 2 - enter element type: **r**
Enter conditions for this row, if any -
Filter: **hnethbl**
Enter value or range for hnethbl: **0**

                                                    Or: **<rtn>**

And: **<rtn>**
Side text: **0**
Enter any options: **<rtn>**
Element 2 defined
Element 3 - enter element type: **r**
Enter conditions for this row, if any -
Filter: **hnethbl**
Enter value or range for hnethbl: **1**

                                                    Or: **<rtn>**

And: **<rtn>**
Side text: **1**
Enter any options: **<rtn>**
Element 3 defined
Element 4 - enter element type: **r**
Enter conditions for this row, if any -
Filter: **hnethbl**
Enter value or range for hnethbl: **2**

                                                    Or: **<rtn>**

And: **<rtn>**
Side text: **2**
Enter any options: **<rtn>**
Element 4 defined
Element 5 - enter element type: **r**
Enter conditions for this row, if any -
Filter: **hnethbl**
Enter value or range for hnethbl: **>2**

                                                    Or: **<rtn>**

And: **<rtn>**
Side text: **3+**
Enter any options: **<rtn>**
Element 5 defined
Element 6 - enter element type: **<rtn>**
Enter column headings, if any -
                 Line 1: **<rtn>**

(Quanvert then starts processing the new variable)

Or it is possible in a similar way to create an axis variable which simply describes whether or not there are any people with the specified characteristic(s) in the household, by letting the numeric variable take the values 0 or > 0.

Household level variables describing the combined characteristics of the members

The 'count' variables described above can also be used to construct household level variables which describe the combined characteristics of the members, in a similar way to the combined household economic activity variable HEACOMB which is already available in the household datasets. For example, suppose we have a set of 'count' variables similar to HNETHBL (the number of black adults) used as an example above, but for each category of ETHNICA, so that there are also variables HNETHWH for the number of white adults, HNETHIN for Indians, and so on. Then these can be used to create a variable for the combined ethnic group category of the household, with categories: only white; only black; . . . . only mixed/other; white + black; white + Indian; . . . ; Pakistani/Bangladeshi + mixed; and so on. But it is already clear that even with five categories in the original variable, the combined variable is both unmanageable and likely to produce unacceptably small sample sizes. Three categories in the original variable is about the maximum practicable, unless there is extensive grouping of the combined variable categories. The example below uses three ethnic group categories - white, black and other - in creating a variable HETHCOM for the combined ethnic group composition of the adults in the household. It has seven categories: only white; only black; only other; white + black; white + other; black + other; white + black + other. It is created as follows, using the numeric count variable HNETHBL described above, and similar variables HNETHWH and HNETHOT for the numbers of white and other ethnic group adults respectively:

```
Command:  c
New axis name:  hethcom
Enter overall conditions for this axis, if any -
Filter:  <rtn>
Enter titles for this axis, if any -
Title:  Combined ethnic group of household
Title:  <rtn>
Enter heading for this axis, if any -
Heading:  hethcom
Enter element definitions according to the prompts -
Element 1 - enter element type:  b
Enter conditions for this row, if any -
Filter:  <rtn>
Side text:  base
Enter any options:  <rtn>
Element 1 defined
Element 2 - enter element type:  r
Enter conditions for this row, if any -
Filter:  hnethwh
Enter value or range for hnethwh:  > 0
                                                              Or:  <rtn>

And:  hnethbl
Enter value or range for hnethbl:  0
                                                              Or:  <rtn>

And:  hnethot
Enter value or range for hnethbl:  0
                                                              Or:  <rtn>

And:  <rtn>
Side text:  only white
Enter any options:  <rtn>
Element 2 defined
Element 3 - enter element type:  r
Enter conditions for this row, if any -
Filter:  hnethwh
Enter value or range for hnethwh:  0
                                                              Or:  <rtn>

And:  hnethbl
Enter value or range for hnethbl:  > 0
                                                              Or:  <rtn>

And:  hnethot
Enter value or range for hnethbl:  0
                                                              Or:  <rtn>

And:  <rtn>
Side text:  only black
Enter any options:  <rtn>
Element 3 defined
```

Element 4 - enter element type: **r**
Enter conditions for this row, if any -
Filter: **hnethwh**
Enter value or range for hnethwh: **0**

Or: **<rtn>**

And: **hnethbl**
Enter value or range for hnethbl: **0**

Or: **<rtn>**

And: **hnethot**
Enter value or range for hnethbl: **> 0**

Or: **<rtn>**

And: **<rtn>**
Side text: **only other**
Enter any options: **<rtn>**
Element 4 defined
Element 5 - enter element type: **r**
Enter conditions for this row, if any -
Filter: **hnethwh**
Enter value or range for hnethwh: **> 0**

Or: **<rtn>**

And: **hnethbl**
Enter value or range for hnethbl: **> 0**

Or: **<rtn>**

And: **hnethot**
Enter value or range for hnethbl: **0**

Or: **<rtn>**

And: **<rtn>**
Side text: **white + black**
Enter any options: **<rtn>**
Element 5 defined
Element 6 - enter element type: **r**
Enter conditions for this row, if any -
Filter: **hnethwh**
Enter value or range for hnethwh: **> 0**

Or: **<rtn>**

And: **hnethbl**
Enter value or range for hnethbl: **0**

Or: **<rtn>**

And: **hnethot**
Enter value or range for hnethbl: **> 0**

Or: **<rtn>**

And: **<rtn>**
Side text: **white + other**
Enter any options: **<rtn>**
Element 6 defined

Element 7 - enter element type: **r**
Enter conditions for this row, if any -
Filter: **hnethwh**
Enter value or range for hnethwh: **0**

Or: **<rtn>**

And: **hnethbl**
Enter value or range for hnethbl: **> 0**

Or: **<rtn>**

And: **hnethot**
Enter value or range for hnethbl: **> 0**

Or: **<rtn>**

And: **<rtn>**
Side text: **black + other**
Enter any options: **<rtn>**
Element 7 defined
Element 8 - enter element type: **r**
Enter conditions for this row, if any -
Filter: **hnethwh**
Enter value or range for hnethwh: **> 0**

Or: **<rtn>**

And: **hnethbl**

Enter value or range for hnethbl: **> 0**

Or: **\<rtn\>**

And: **hnethot**
Enter value or range for hnethbl: **> 0**

Or: **\<rtn\>**

And: **\<rtn\>**
Side text: **white + black + other**
Enter any options: **\<rtn\>**
Element 8 defined
Element 9 - enter element type: **\<rtn\>**
Enter column headings, if any -
                    Line 1: **\<rtn\>**

(Quanvert then starts processing the new variable)

ANNEX C: ASSESSMENT OF EFFECTS OF RESIDUAL PROBLEMS IN DATASETS UP TO 1995

The key potential effect of all the residual problems described below is on the distribution of the household combined economic activity variable. The most critical indicator derived from this variable is the percentage of workless households. The magnitude of the effects is therefore assessed in relation to this indicator.

Missing household members during 1992 to 1995

After adjusted household and family type variables have been created, there are three known possible sources of residual error:

(i) Some of the adjusted households may have had missing members who were genuinely missing, (e.g. working abroad, or in prison). The proportion of households of this type can be estimated from the proportion of households which still have an apparent inconsistency between household type and the characteristics of the household members after the introduction of the household matrix. This is about 0.5% (in 1996), almost all of the 'missing spouse' type, either in a couple household with no children (0.3%) or a couple household with children (0.2%). The correct household economic activity of these households is that of the spouse who was still present. If we take as a starting point the economic activity distributions of complete households of the relevant types, and assume that in two-thirds of the households in the employed plus inactive/unemployed category the remaining spouse is inactive/unemployed, the overall percentage of workless households for those with genuinely missing members is estimated to have been about 27% (in 1997). With an adjusted household type, however, these households would have been given an unstated economic activity category, and would therefore have been assumed to be distributed as for households with known economic activity of the same household type, (for the purpose of estimating the overall percentage of workless households, adjusted for unknown cases). The percentage of workless couple households is about 12%, some 15 percentage points less than the estimated average level for the genuine missing members households, so that the possible effect of this problem would be to decrease the estimated level of workless households by 0.5 x 0.15 percentage points or 0.07 percentage points.

(ii) Some cases of missing household members would not have been identified by the internal inconsistency criteria. Comparing the percentages of households with unknown economic activity for 1995 and 1996 suggests that about 1 percent of households may be unidentified cases of missing members in 1992 to 1995. The most probable types of households to be involved, *viz* missing flat-sharers who are not the head of household, and missing non-dependent children, are ones where the household economic activity without the missing member is likely to be the same as that with all members present, and also where workless households constitute a relatively small proportion. The effect of this problem is therefore assessed as negligible.

(iii) There are about 100 households in each dataset where, because of a variety of errors in recording the family unit and household structure, it did not prove possible to derive consistent adjusted family and household type categories, and the households were therefore dropped from the sample. This constitutes under 0.2% of all households. The individual errors involved are very varied, and there is no reason to suppose that they differ systematically from other missing member cases. On this basis, if they had been included in the dataset, their proportion of workless households would have been estimated as that for all households, and since they were dropped, this was in effect what was done. The effect of this problem is therefore assessed as negligible.

The spring 1990 dataset

The survey design and fieldwork arrangements in 1990 are expected to have resulted in household members with unstated economic activity (as in 1996 onwards), rather than household members being entirely missed from the dataset (as in 1992 to 1995). However, the percentage of households with unknown economic activity for 1990 is similar to those for the 1992 to 1995 period, and smaller than those for 1996 onwards. It is likely that the percentage of unknowns would have been smaller in 1990 than later

because of fewer flat-sharing households, and larger households in general, but the size of the difference suggests that there may also have been some element of missing household members. If so, the most likely people to be missed would be unrelated household members, and the same arguments as in (ii) above apply, and the problem is assessed to have a negligible effect.

Residual data problems

It is possible that there still exist some problems arising from errors in the recording of family and household structure during the period before these data were expected to be used for analysis. However, in view of the extensive analyses of household and family structure which have been undertaken in the course of producing the household datasets, it seems unlikely that any major problems could have escaped detection.

Summary of possible effects

The total estimated effect of possible residual problems is an underestimation of the percentage of workless households of up to 0.07 percentage points during the period 1992 to 1995. A difference of this level could in some cases tip the rounding of published figures which are usually to one decimal point, but would be unlikely to affect any inferences.

ANNEX D: REVISIONS TO THE DERIVATION OF HSERIAL AND FUSERIAL

HSERIAL and FUSERIAL are identifier variables specific to the household datasets. Their purpose is to uniquely identify households and family units respectively. They are used to produce the household datasets and to create household and family level variables. From autumn 2000 the derivation of HSERIAL and FUSERIAL has been revised.

During the regrossing of the household datasets, a problem was identified with HSERIAL and FUSERIAL, which revealed that a small proportion of households and family units were not being uniquely identified. This allowed some households / families to be assigned the same HSERIAL and FUSERIAL numbers.

To correct the problem the combination of system variables that were previously used to create HSERIAL and FUSERIAL has been changed (see Annex A for the derivation of the variables). This prevents the circumstances occurring that causes duplication of HSERIAL.

This problem was corrected on the autumn 1996 to spring 2000 datasets within the scope of the regrossing project. A digit 9 was added to the end of the serial numbers for all members of one of the duplicate households.

ANNEX E: PROCEDURE FOR ADJUSTING FOR HOUSEHOLDS WITH UNKNOWN ECONOMIC ACTIVITY IN ESTIMATES OF WORKLESS HOUSEHOLDS FOR SUBPOPULATIONS

Introduction

As described in the main text, the adjustment methodology used for producing the regularly published series on workless households in the *LFS Quarterly Supplement* is to divide all households according to household type, (combining together some small, similar categories), and **within each household type category** to allocate the 'unknown' households, (or adults or children in unknown households, as appropriate), as workless or not in the same proportions as the households (or adults or children in them) with known economic activity. Table A.E1 below shows the calculations involved, using spring 2000 as an example. (In accordance with previous practice, workless households are defined as working-age households, i.e. households containing at least one person of working age, in which no-one is in employment.)

This adjustment method operates at the aggregate level, and is only used for estimating **overall levels** of workless households and adults and children in workless households. It is not suitable for producing adjusted estimates for subgroups, because it may produce estimates for different categorisations which are inconsistent when aggregated, and because for smaller and/or more specialised subgroups the numbers of sample households in the smaller household type categories eventually become small enough to cause volatility in the resulting estimates. However, there are increasing demands for adjusted estimates for subgroups, and this annex describes a simple, consistent method for producing them.

It is important that the method should have the following characteristics:

- is straightforward and quick to apply;
- produces consistent results, so that adjusted estimates for subgroups add up to the adjusted estimate for their total;
- does not depend on using possibly volatile small cells;
- does not distort the estimates for workless households.

We know from previous investigations that the incidence of households with unknown economic activity varies considerably between different household types, and, (closely related to this), between households with different numbers of adults. This is what we would expect from the circumstances in which a household is assigned to the unknown economic activity category. This happens if there is any person in the household for whom economic activity data cannot be obtained, either because they are absent from the dwelling and the respondent does not feel able to give a proxy response, or they are present and refuse to participate in the survey. Every person aged over 16 is asked for information on their economic activity, and so it is not possible for a household containing only one person over 16, (described in this context as an adult), to have unknown economic activity, because if that one person were absent or refused, then the whole household would be non-responding.

Table 2A
Adjustment of estimates of workless households[1], UK, spring 1990, spring 1992 to spring 2004, autumn 1995 to autumn 2003.

Household type[2]	Codes of grouped household type categories spr 1990, spr 1992 to aut 1995	Codes of grouped household type categories spr 1996 onwards	spr 1990	spr 1992	spr 1993	spr 1994	spr 1995	aut 1995	spr 1996	aut 1996	spr 1997	aut 1997	spr 1998	aut 1998	spr 1999	aut 1999	spr 2000	aut 2000	spr 2001	aut 2001	spr 2002	aut 2002	spr 2003	aut 2003	spr 2004
1 person	1	1									1.0000	1.0000	1.0002	1.0001	1.0004	1.0000	1.0003	1.0002	1.0003	1.0001	1.0001	1.0003	1.0006	1.0010	1.0003
2 or more persons, all different family units	2	2									1.4020	1.4411	1.4680	1.4932	1.5045	1.4923	1.4584	1.6121	1.4952	1.4817	1.4389	1.6142	1.5488	1.6013	1.5513
Couple, no children, no other family units	3	3,4									1.0341	1.0368	1.0344	1.0369	1.0384	1.0388	1.0362	1.0362	1.0347	1.0387	1.0444	1.0404	1.0500	1.0491	1.0462
Couple, no children, other family units	4	5									1.1982	1.1963	1.2587	1.2044	1.2185	1.2408	1.2121	1.2586	1.2043	1.1998	1.2509	1.3021	1.3176	1.2637	1.3021
Couple, all dependent children, no other family units	5	6,7									1.0271	1.0267	1.0266	1.0252	1.0286	1.0265	1.0299	1.0286	1.0306	1.0292	1.0295	1.0321	1.0370	1.0357	1.0342
Couple, dependent & non-dependent children, no other family units	6	8,9									1.0513	1.0607	1.0509	1.0656	1.0664	1.0707	1.0715	1.0712	1.0891	1.0910	1.0806	1.0734	1.1040	1.0936	1.0936
Couple, all non-dependent children, no other family units	7	10,11									1.0952	1.0921	1.0888	1.0728	1.0951	1.0997	1.1024	1.0995	1.1048	1.1003	1.1027	1.1189	1.1294	1.1288	1.1230
Couple, children, other family units	8-10	12-14									1.1885	1.1476	1.1612	1.1468	1.1598	1.1599	1.2029	1.1806	1.1756	1.1836	1.1998	1.2303	1.1602	1.1567	1.2380
Lone parent, all dependent children, no other family units	11	15									1.0023	1.0015	1.0022	1.0003	1.0013	1.0014	1.0017	1.0017	1.0026	1.0016	1.0018	1.0025	1.0017	1.0026	1.0012
Lone parent, dependent & non-dependent children, no other family units	12	16									1.0491	1.0543	1.0607	1.0504	1.0645	1.0778	1.0705	1.0511	1.0619	1.0634	1.0671	1.0661	1.0862	1.0581	1.0526
Lone parent, all non-dependent children, no other family units	13	17									1.0733	1.0632	1.0802	1.0872	1.0855	1.0793	1.0755	1.0814	1.0869	1.0672	1.0773	1.0769	1.1016	1.0964	1.1042
Lone parent, other family units	14-16	18-20									1.1819	1.1440	1.1639	1.1493	1.2099	1.1353	1.2231	1.2440	1.1436	1.1864	1.1827	1.1663	1.1316	1.2044	1.2023
2 or more family units and others	17-20	21-25									1.2422	1.1778	1.1516	1.1989	1.2194	1.2330	1.2197	1.2385	1.2287	1.1812	1.2188	1.2017	1.1908	1.2158	1.2245
Total all household types											**1.0373**	**1.0373**	**1.0391**	**1.0389**	**1.0409**	**1.0368**	**1.0384**	**1.0439**	**1.0412**	**1.0384**	**1.0392**	**1.0423**	**1.0469**	**1.0465**	**1.0464**
Workless households with one adult											1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
Workless households with more than one adult											1.0775	1.0788	1.0830	1.0853	1.0888	1.0817	1.0839	1.0955	1.0892	1.0852	1.0872	1.0966	1.1062	1.1056	1.1025

[1] A workless household is a household with at least one person of working age where no-one is in employment.

[2] The household type variable used for spring 1990 is TYPEHH, THHADJ for spring 1992 to autumn 1995 and HHTYPE for spring 1996 onwards.

Table 2B
Adjustment of estimates of working-age people[1] in workless households[2], UK, spring 1990, spring 1992 to spring 2004, autumn 1995 to autumn 2003.

Household type[3]	Codes of grouped household type categories spr 1990, spr 1992 to aut 1995	Codes of grouped household type categories spr 1996 onwards	Adjustment factor																						
			spr 1990	spr 1992	spr 1993	spr 1994	spr 1995	aut 1995	spr 1996	aut 1996	spr 1997	aut 1997	spr 1998	aut 1998	spr 1999	aut 1999	spr 2000	aut 2000	spr 2001	aut 2001	spr 2002	aut 2002	spr 2003	aut 2003	spr 2004
1 person	1	1									1.0000	1.0000	1.0002	1.0001	1.0004	1.0000	1.0003	1.0002	1.0003	1.0001	1.0001	1.0003	1.0006	1.0010	1.0003
2 or more persons, all different family units	2	2									1.5161	1.6088	1.6128	1.6957	1.7021	1.6733	1.5977	1.8321	1.6410	1.6722	1.5738	1.8607	1.7508	1.8678	1.7445
Couple, no children, no other family units	3	3,4									1.0352	1.0382	1.0354	1.0383	1.0398	1.0401	1.0375	1.0369	1.0366	1.0404	1.0458	1.0415	1.0521	1.0510	1.0479
Couple, no children, other family units	4	5									1.2678	1.2677	1.3835	1.3027	1.3247	1.3425	1.2790	1.4042	1.2806	1.2451	1.3599	1.4233	1.4549	1.3617	1.4092
Couple, all dependent children, no other family units	5	6,7									1.0272	1.0265	1.0266	1.0252	1.0284	1.0264	1.0292	1.0284	1.0299	1.0285	1.0285	1.0316	1.0363	1.0349	1.0336
Couple, dependent & non-dependent children, no other family units	6	8,9									1.0527	1.0609	1.0492	1.0633	1.0669	1.0726	1.0710	1.0724	1.0885	1.0909	1.0802	1.0735	1.1003	1.0933	1.0918
Couple, all non-dependent children, no other family units	7	10,11									1.1035	1.0966	1.0960	1.0808	1.1022	1.1062	1.1131	1.1057	1.1131	1.1083	1.1110	1.1289	1.1405	1.1428	1.1361
Couple, children, other family units	8-10	12-14									1.2130	1.1690	1.1803	1.1645	1.1678	1.1651	1.2265	1.2188	1.1990	1.2281	1.2356	1.2733	1.1713	1.1813	1.2938
Lone parent, all dependent children, no other family units	11	15									1.0033	1.0028	1.0044	1.0005	1.0024	1.0029	1.0025	1.0026	1.0045	1.0024	1.0029	1.0047	1.0034	1.0035	1.0015
Lone parent, dependent & non-dependent children, no other family units	12	16									1.0477	1.0529	1.0634	1.0482	1.0699	1.0833	1.0737	1.0607	1.0639	1.0633	1.0723	1.0704	1.0902	1.0614	1.0581
Lone parent, all non-dependent children, no other family units	13	17									1.0864	1.0734	1.0926	1.0991	1.0921	1.0879	1.0888	1.1008	1.1076	1.0738	1.0909	1.0908	1.1141	1.1079	1.1258
Lone parent, other family units	14-16	18-20									1.2178	1.1955	1.2057	1.1933	1.2530	1.1599	1.2598	1.2926	1.1824	1.2560	1.2156	1.2207	1.1772	1.2528	1.2502
2 or more family units and others	17-20	21-25									1.2810	1.2175	1.1781	1.2241	1.2745	1.3116	1.2709	1.2870	1.2843	1.2252	1.2634	1.2389	1.2216	1.2775	1.2988
Total all household types											**1.0627**	**1.0626**	**1.0681**	**1.0687**	**1.0713**	**1.0637**	**1.0659**	**1.0765**	**1.0705**	**1.0652**	**1.0667**	**1.0756**	**1.0846**	**1.0814**	**1.0807**
Working-age people in workless households with one adult											1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
Working-age people in workless households with more than one adult											1.0996	1.1016	1.1103	1.1149	1.1172	1.1070	1.1093	1.1269	1.1162	1.1098	1.1118	1.1302	1.1438	1.1385	1.1346

[1] Working-age refers to men between the ages of 16 and 64 and women between 16 and 59.

[2] A workless household is a household with at least one person of working age where no-one is in employment.

[3] The household type variable used for spring 1990 is TYPEHH, THHADJ for spring 1992 to autumn 1995 and HHTYPE for spring 1996 onwards.

Table 2C

Adjustment of estimates of children[1] in workless households[2], UK, spring 1990, spring 1992 to spring 2004, autumn 1995 to autumn 2003.

Household type[3]	Codes of grouped household type categories spr 1990, spr 1992 to aut 1995	Codes of grouped household type categories spr 1996 onwards	Adjustment factor																						
			spr 1990	spr 1992	spr 1993	spr 1994	spr 1995	aut 1995	spr 1996	aut 1996	spr 1997	aut 1997	spr 1998	aut 1998	spr 1999	aut 1999	spr 2000	aut 2000	spr 2001	aut 2001	spr 2002	aut 2002	spr 2003	aut 2003	spr 2004
2 or more persons, all different family units	2	2									1.1625	1.1573	1.1821	1.0000	1.0000	1.0000	1.0809	1.0000	1.0000	1.0000	1.1671	1.1300	1.3415	1.0000	1.1171
Couple, no children, other family units	4	5									1.0000	1.2489	1.0000	-	1.0000	1.0000	1.0314	-	1.0000	1.1160	1.2221	1.0809	1.0305	1.0268	1.0000
Couple, all dependent children, no other family units	5	6,7									1.0273	1.0264	1.0259	1.0239	1.0274	1.0252	1.0291	1.0290	1.0310	1.0300	1.0302	1.0315	1.0372	1.0362	1.0345
Couple, dependent & non-dependent children, no other family units	6	8,9									1.0543	1.0680	1.0696	1.0774	1.0792	1.0819	1.0798	1.0752	1.1123	1.0994	1.0939	1.0820	1.1347	1.1080	1.1252
Couple, children, other family units	8-10	12-14									1.1583	1.1278	1.1518	1.1133	1.1410	1.1472	1.2168	1.1768	1.1860	1.1355	1.1807	1.2203	1.1428	1.1542	1.1775
Lone parent, all dependent children, no other family units	11	15									1.0015	1.0002	1.0009	1.0002	1.0008	1.0004	1.0009	1.0002	1.0015	1.0010	1.0013	1.0025	1.0000	1.0022	1.0006
Lone parent, dependent & non-dependent children, no other family units	12	16									1.0549	1.0477	1.0475	1.0449	1.0450	1.0810	1.0766	1.0592	1.0692	1.0814	1.0815	1.0753	1.0602	1.0528	1.0419
Lone parent, other family units	14-16	18-20									1.1279	1.1024	1.1564	1.1090	1.2036	1.1224	1.1780	1.2740	1.1084	1.0864	1.1420	1.1304	1.1170	1.1590	1.1384
2 or more family units and others	17-20	21-25									1.2329	1.1917	1.1269	1.2215	1.2037	1.2197	1.1980	1.1794	1.1969	1.1691	1.1427	1.1620	1.1857	1.2158	1.2397
Total all household types											**1.0197**	**1.0180**	**1.0172**	**1.0142**	**1.0192**	**1.0185**	**1.0202**	**1.0211**	**1.0192**	**1.0174**	**1.0200**	**1.0203**	**1.0191**	**1.0225**	**1.0196**
Children in workless households with one adult											1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
Children in workless households with more than one adult											1.0450	1.0428	1.0418	1.0381	1.0485	1.0477	1.0521	1.0533	1.0498	1.0456	1.0503	1.0517	1.0520	1.0601	1.0505

[1] Children refers to all children under 16.

[2] A workless household is a household with at least one person of working age where no-one is in employment.

[3] The household type variable used for spring 1990 is TYPEHH, THHADJ for spring 1992 to autumn 1995 and HHTYPE for spring 1996 onwards.

General method

The Labour Force Survey User Guide: household and family data, Volume 8 (1999) described a general adjustment method for adjusting subgroups of workless households defined by characteristics which are **not** closely associated with household type or number of adults. It was found that the unknown household economic activity rate varied across regions and therefore the application of the general method would not be suitable for producing regional estimates. As this problem might extend to other subgroups, the general method should not be used to adjust subgroups of workless households. Figures should be presented as unadjusted and percentages based on totals excluding those households with unknown economic activity.

For example, the proportion of workless households in Wales in spring 2000 was 21.9% and is calculated as follows:

Data in thousands

No. of workless hhlds in Wales	/	(Total no. of working-age hhlds in Wales	-	hhlds with unknown economic activity) x 100
191	/	(904	-	30) x 100

Method for household type categories

The method for adjusting subgroups defined in terms of one or more specific household types, is as follows. For each specified household type apply the adjustment factor which has been obtained for that household type for the UK as a whole for the appropriate dataset. For example, if we require the adjusted estimate of workless households which are lone parent households with dependent children only and no other family units, for spring 2000, the adjustment factor is **1.0017** (taking the data from the appropriate row of Table A.E1 for that household type):

Data in thousands
        641 / (97 + 11 + 532) = **1.0017**

If an adjusted estimate is required for a household type which is grouped together with other household types in Table A.E1, then the proposed method is to apply the adjustment factor derived from the appropriate group of household types. For example, if we require the adjusted estimate of workless households which are cohabiting couple households with dependent children only and no other family units, for spring 2000, the adjustment factor is **1.0300** (taking the data from the appropriate row of Table A.E1 for couple households with dependent children only and no other family units, since cohabiting couples form a subgroup of all couples with the same characteristics):

Data in thousands
        283 / (18 + 100 + 157) = **1.0300**

NB Subgroups defined in terms of household type further analysed by another variable, such as region (i.e. variables which would have previously been associated with the general method), should be presented as unadjusted due to the issues raised in the above 'General method' section.

Method for household size categories

A different method is required for adjusting subgroups defined in terms of the number of adults in the household, i.e. the number of people over 16. The proposed method is to take an adjustment factor of exactly one for households with one adult, and to calculate the adjustment factor for households with more than one adult in such a way that it produces the correct adjusted total of workless households. This is done by subtracting from the total of adjusted numbers of workless households the number of one-adult workless households, and dividing that by the unadjusted total number of workless households minus the number of one-adult workless households. For example, for spring 2000, the number of workless households with one adult is 1.589 million. The adjustment factor for workless households with more than one adult is then **1.0855**:

Data in thousands
        (3,066 - 1,589) / (2,950 - 1,589) = **1.0855**

It is important to note that the only household type category which invariably contains only one adult (defined as person over 16) is that of one-person households. Lone person households may contain non-dependent children, who while they are the never-married offspring (i.e. 'children') of the lone parent, are not children in age, since they are aged over 16. Even households containing only dependent children may include some who are over 16, since dependent children include, as well as children under 16, those aged 16 to 18 who are in full-time education. Thus of the categories of households in Table A.E1, only in the first category is it theoretically impossible to find households with unknown economic activity. (And even here, there is a very small number of such technically impossible cases in the table – however, these arise from 2 sample persons out of some 150,000, and make a negligible difference to the estimates, so their effect may reasonably be ignored.)

If a subgroup is defined in terms of the number of adults of working age, (i.e. those aged 16 to 59 for women or 64 for men), the situation is more complicated, because there are some working-age households with only one adult of working age which also include one or more adults of pensionable age. It is possible for such a household to have unknown economic activity, so it is not appropriate to assign to all workless households with one person of working age an adjustment factor of one. The category of households with one person of working age will need to be divided into those with no-one of pensionable age, which will be given an adjustment factor of one, and those with one or more people of pensionable age, which will be given the adjustment factor for households with more than one adult, as described in the previous paragraph.

NB Subgroups defined in terms of household size further analysed by another variable, such as region (i.e. variables which would have previously been associated with the general method), should be presented as unadjusted due to the issues raised in the above 'General method' section.

Adjustment factors

For the convenience of users, the adjustment factors for all the household datasets produced to date, for workless households, and working-age adults and children under 16 in workless households, have been calculated and are given in Tables A.E2.1-2.3 which follow.

Methodological development for the treatment of households with unknown economic activity

The household datasets are increasingly being used for more complex analyses of the combined household economic activity status at both the household and family level. Consequently there is increasing demand for a more robust and 'all-encompassing' method than the adjustment factor approach for compensating for households with unknown economic activity. In consultation with users, a project is being conducted into the feasibility of using imputation to overcome the problems caused by households with unknown economic activity. If a satisfactory basis can be found for imputing the economic activity status of the unknown individuals, this will enable the combined activity of all households to be described, thereby removing the need for the current adjustment procedure. Users will be informed of the outcome of the project.

Table A.E2.1

Adjustment of estimates of workless households[1], UK, spring 1990, spring 1992 to spring 2001, autumn 1995 to autumn 2000.

Household type[2]	Codes of grouped household type categories		Adjustment factors																
	spr 1990, spr 1992 to aut 1995	spr 1996 onwards	spr 1990	spr 1992	spr 1993	spr 1994	spr 1995	aut 1995	spr 1996	aut 1996	spr 1997	aut 1997	spr 1998	aut 1998	spr 1999	aut 1999	spr 2000	aut 2000	spr 2001
1 person	1	1	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0004	1.0005	1.0000	1.0000	1.0002	1.0001	1.0004	1.0000	1.0003	1.0002	1.0003
2 or more persons, all different family units	2	2	1.1861	1.0595	1.0818	1.0885	1.1010	1.1214	1.3802	1.4724	1.4106	1.4534	1.4757	1.5069	1.5087	1.5024	1.4658	1.6160	1.4987
Couple, no children, no other family units	3	3,4	1.0150	1.0460	1.0422	1.0509	1.0363	1.0450	1.0227	1.0300	1.0345	1.0371	1.0347	1.0372	1.0388	1.0393	1.0366	1.0365	1.0351
Couple, no children, other family units	4	5	1.1139	1.0000	1.0000	1.0000	1.0000	1.0000	1.1777	1.1812	1.2043	1.2028	1.2674	1.2121	1.2265	1.2436	1.2153	1.2636	1.2111
Couple, all dependent children, no other family units	5	6,7	1.0099	1.0162	1.0208	1.0224	1.0203	1.0247	1.0181	1.0221	1.0271	1.0267	1.0266	1.0255	1.0287	1.0266	1.0300	1.0287	1.0308
Couple, dependent & non-dependent children, no other family units	6	8,9	1.0330	1.0296	1.0475	1.0287	1.0287	1.0262	1.0357	1.0557	1.0514	1.0609	1.0512	1.0651	1.0659	1.0709	1.0713	1.0712	1.0897
Couple, all non-dependent children, no other family units	7	10,11	1.0442	1.0391	1.0445	1.0333	1.0207	1.0314	1.0627	1.0698	1.0957	1.0924	1.0899	1.0733	1.0958	1.1009	1.1028	1.1007	1.1052
Couple, children, other family units	8-10	12-14	1.0669	1.1132	1.0991	1.0451	1.1211	1.1152	1.1056	1.1707	1.1926	1.1509	1.1647	1.1513	1.1610	1.1590	1.2042	1.1826	1.1774
Lone parent, all dependent children, no other family units	11	15	1.0017	1.0000	1.0000	1.0000	1.0000	1.0000	1.0102	1.0008	1.0023	1.0016	1.0024	1.0003	1.0014	1.0015	1.0017	1.0018	1.0026
Lone parent, dependent & non-dependent children, no other family units	12	16	1.0293	1.0000	1.0000	1.0000	1.0000	1.0000	1.0425	1.0394	1.0490	1.0556	1.0611	1.0509	1.0654	1.0793	1.0697	1.0514	1.0624
Lone parent, all non-dependent children, no other family units	13	17	1.0401	1.0265	1.0284	1.0275	1.0381	1.0484	1.0494	1.0534	1.0738	1.0642	1.0818	1.0884	1.0870	1.0815	1.0757	1.0826	1.0875
Lone parent, other family units	14-16	18-20	1.0691	1.0000	1.0000	1.0000	1.0000	1.0000	1.2209	1.1875	1.1849	1.1484	1.1717	1.1538	1.2158	1.1371	1.2316	1.2579	1.1465
2 or more family units and others	17-20	21-25	1.0582	1.0000	1.0000	1.0000	1.0000	1.0000	1.1659	1.2167	1.2479	1.1831	1.1537	1.2019	1.2246	1.2366	1.2223	1.2395	1.2350
Total all household types			**1.0197**	**1.0185**	**1.0201**	**1.0206**	**1.0180**	**1.0216**	**1.0330**	**1.0356**	**1.0383**	**1.0386**	**1.0404**	**1.0403**	**1.0422**	**1.0380**	**1.0395**	**1.0450**	**1.0421**
Workless households with one adult			1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
Workless households with more than one adult			1.0376	1.0340	1.0373	1.0397	1.0361	1.0439	1.0686	1.0743	1.0793	1.0810	1.0852	1.0878	1.0908	1.0836	1.0855	1.0973	1.0907

[1] A workless household is a household with at least one person of working age where no-one is in employment.

Table A.E2.2

Adjustment of estimates of working-age people[1] in workless households[2], UK, spring 1990, spring 1992 to spring 2001, autumn 1995 to autumn 2000.

Household type[3]	Codes of grouped household type categories		Adjustment factors																
	spr 1990, spr 1992 to aut 1995	spr 1996 onwards	spr 1990	spr 1992	spr 1993	spr 1994	spr 1995	aut 1995	spr 1996	aut 1996	spr 1997	aut 1997	spr 1998	aut 1998	spr 1999	aut 1999	spr 2000	aut 2000	spr 2001
1 person	1	1	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0004	1.0005	1.0000	1.0000	1.0002	1.0001	1.0004	1.0000	1.0003	1.0002	1.0003
2 or more persons, all different family units	2	2	1.2691	1.0284	1.0373	1.0403	1.0459	1.0560	1.5096	1.6616	1.5249	1.6217	1.6220	1.7129	1.7048	1.6876	1.6053	1.8341	1.6457
Couple, no children, no other family units	3	3,4	1.0150	1.0249	1.0229	1.0275	1.0196	1.0243	1.0240	1.0315	1.0355	1.0385	1.0357	1.0386	1.0402	1.0405	1.0379	1.0372	1.0369
Couple, no children, other family units	4	5	1.1649	1.0000	1.0000	1.0000	1.0000	1.0000	1.2484	1.2405	1.2758	1.2751	1.3965	1.3149	1.3376	1.3438	1.2836	1.4085	1.2895
Couple, all dependent children, no other family units	5	6,7	1.0099	1.0089	1.0116	1.0121	1.0112	1.0131	1.0184	1.0221	1.0272	1.0265	1.0266	1.0255	1.0285	1.0265	1.0293	1.0285	1.0300
Couple, dependent & non-dependent children, no other family units	6	8,9	1.0349	1.0195	1.0310	1.0209	1.0189	1.0181	1.0351	1.0569	1.0528	1.0612	1.0495	1.0629	1.0662	1.0727	1.0709	1.0724	1.0891
Couple, all non-dependent children, no other family units	7	10,11	1.0479	1.0276	1.0307	1.0227	1.0150	1.0204	1.0683	1.0746	1.1041	1.0973	1.0972	1.0812	1.1031	1.1075	1.1137	1.1072	1.1137
Couple, children, other family units	8-10	12-14	1.0798	1.0922	1.0817	1.0316	1.0963	1.0847	1.1193	1.1947	1.2173	1.1734	1.1852	1.1704	1.1678	1.1641	1.2280	1.2210	1.2011
Lone parent, all dependent children, no other family units	11	15	1.0028	1.0000	1.0000	1.0000	1.0000	1.0000	1.0165	1.0017	1.0033	1.0031	1.0048	1.0005	1.0025	1.0031	1.0026	1.0028	1.0045
Lone parent, dependent & non-dependent children, no other family units	12	16	1.0312	1.0000	1.0000	1.0000	1.0000	1.0000	1.0435	1.0417	1.0479	1.0540	1.0637	1.0482	1.0700	1.0852	1.0721	1.0613	1.0641
Lone parent, all non-dependent children, no other family units	13	17	1.0511	1.0149	1.0158	1.0156	1.0220	1.0279	1.0584	1.0667	1.0872	1.0745	1.0945	1.1002	1.0934	1.0902	1.0887	1.1024	1.1086
Lone parent, other family units	14-16	18-20	1.0864	1.0000	1.0000	1.0000	1.0000	1.0000	1.2485	1.2410	1.2212	1.2012	1.2163	1.1998	1.2609	1.1615	1.2678	1.3109	1.1835
2 or more family units and others	17-20	21-25	1.0676	1.0000	1.0000	1.0000	1.0000	1.0000	1.1894	1.2622	1.2879	1.2245	1.1817	1.2278	1.2800	1.3163	1.2742	1.2886	1.2917
Total all household types			**1.0342**	**1.0131**	**1.0146**	**1.0141**	**1.0129**	**1.0152**	**1.0570**	**1.0617**	**1.0647**	**1.0649**	**1.0705**	**1.0715**	**1.0733**	**1.0658**	**1.0678**	**1.0786**	**1.0721**
Working-age people in workless households with one adult			1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
Working-age people in workless households with more than one adult			1.0519	1.0191	1.0214	1.0212	1.0197	1.0237	1.0907	1.0987	1.1022	1.1047	1.1134	1.1188	1.1197	1.1098	1.1116	1.1296	1.1183

Table A.E2.3

Adjustment of estimates of children[1] in workless households[2], UK, spring 1990, spring 1992 to spring 2001, autumn 1995 to autumn 2000.

	Codes of grouped household type categories		Adjustment factors																
	spr 1990, spr 1992 to aut 1995	spr 1996 onwards	spr 1990	spr 1992	spr 1993	spr 1994	spr 1995	aut 1995	spr 1996	aut 1996	spr 1997	aut 1997	spr 1998	aut 1998	spr 1999	aut 1999	spr 2000	aut 2000	spr 2001
Household type[3]																			
2 or more persons, all different family units	2	2	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0839	1.0000	1.1754	1.1772	1.1968	1.0000	1.0000	1.0000	1.0927	1.0000	1.0000
Couple, no children, other family units	4	5	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.1310	1.1353	1.0000	1.2391	1.0000	-	1.0000	1.0000	1.0299	-	1.0000
Couple, all dependent children, no other family units	5	6,7	1.0097	1.0151	1.0194	1.0211	1.0190	1.0239	1.0170	1.0203	1.0272	1.0264	1.0259	1.0241	1.0275	1.0252	1.0292	1.0291	1.0311
Couple, dependent & non-dependent children, no other family units	6	8,9	1.0335	1.0499	1.0638	1.0289	1.0352	1.0294	1.0316	1.0594	1.0543	1.0686	1.0704	1.0770	1.0786	1.0826	1.0797	1.0751	1.1131
Couple, children, other family units	8-10	12-14	1.0524	1.1159	1.1046	1.0513	1.1057	1.1247	1.1131	1.1549	1.1617	1.1290	1.1549	1.1150	1.1427	1.1472	1.2195	1.1790	1.1858
Lone parent, all dependent children, no other family units	11	15	1.0010	1.0000	1.0000	1.0000	1.0000	1.0000	1.0020	1.0003	1.0015	1.0002	1.0009	1.0002	1.0008	1.0005	1.0009	1.0002	1.0015
Lone parent, dependent & non-dependent children, no other family units	12	16	1.0326	1.0000	1.0000	1.0000	1.0000	1.0000	1.0436	1.0476	1.0540	1.0490	1.0482	1.0453	1.0455	1.0820	1.0768	1.0589	1.0706
Lone parent, other family units	14-16	18-20	1.0564	1.0000	1.0000	1.0000	1.0000	1.0000	1.1913	1.1245	1.1332	1.1062	1.1639	1.1105	1.2044	1.1240	1.1882	1.2830	1.1110
2 or more family units and others	17-20	21-25	1.0523	1.0000	1.0000	1.0000	1.0000	1.0000	1.1483	1.1702	1.2400	1.1987	1.1271	1.2263	1.2136	1.2278	1.2005	1.1797	1.2048
Total all household types			**1.0076**	**1.0097**	**1.0118**	**1.0109**	**1.0095**	**1.0122**	**1.0153**	**1.0159**	**1.0204**	**1.0187**	**1.0180**	**1.0149**	**1.0199**	**1.0194**	**1.0212**	**1.0219**	**1.0201**
Children in workless households with one adult			1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
Children in workless households with more than one adult			1.0163	1.0175	1.0216	1.0205	1.0196	1.0249	1.0326	1.0346	1.0449	1.0429	1.0418	1.0379	1.0481	1.0476	1.0519	1.0526	1.0497

[1] Children refers to all children under 16.

[2] A workless household is a household with at least one person of working age where no-one is in employment.

[3] The household type variable used for spring 1990 is TYPEHH, THHADJ for spring 1992 to autumn 1995 and HHTYPE for spring 1996 onwards.

INDEX FOR ANNEX A: DETAILS OF NEW HOUSEHOLD AND FAMILY LEVEL VARIABLES