Labour Force Survey User Guide
Volume 8: Household and family data
2008

Contents

Section 1: Introduction 2
Section 2: Concepts and definitions 4
Section 3: Differences between the household and person-level datasets 7
Section 4: Weighting the LFS sample of households 8
Section 5: Adjustments for anomalies in family and household data (1992-1995) 12
Section 6: Variables specific to the household datasets 13
Section 7: When to use the household datasets 15
Section 8: General guidance on producing family and household analyses 16
Section 9: Producing family and household analyses: worked examples in SPSS 28
Section 10: Creating family and household-level variables in SuperCROSS 34
Section 11: Publication of family and household statistics 37
Section 12: Improving family and household statistics 39
Annex A: Adjusting estimates of workless households by type and size of household 39
Annex B: References 41

Section 1: Introduction

The Labour Force Survey (LFS) is a survey of households living at private addresses in the UK. The first round of interviews are face-to-face with an exception north of the Caledonian Canal in Scotland where some households are interviewed by telephone. After that each interview takes place by telephone at quarterly intervals. The household address is the main sampling unit, rather than the people who live at the address. Thus, if the occupants change, the address would remain in the sample, and any new occupant(s) interviewed.

It provides information about the UK labour market to evaluate labour market policies. The primary use of the LFS is producing person-level statistics (such as employment, unemployment and economic inactivity levels and rates) broken down by personal characteristics (such as age, sex and ethnicity). As the survey collects information about each member of participating households, there is the alternative to produce person-level statistics broken down by the characteristics of the families and households in which people live. This analysis is important, because a person's family and household circumstances may influence their participation in the labour market. For example, having children or how old their children are, may influence their decision to work, or look for work. In a married or cohabiting couple, one partner in employment may influence the other partner's decision to work or look for work.

By collecting information about each member of participating households, the LFS also provides family and household-level statistics that describe the combined economic activity status of family and household members. For example, the LFS is the principal source of statistics on couples where both partners are working, one partner is working, or neither partner is working. It is also the main source of statistics on 'working' households (where all the adults are working); 'mixed' households (containing both working and non-working adults); and 'workless' households (where none of the adults is working). This analysis is important because of a tendency for households to become polarised into 'working' and 'workless' which affects the distribution of household income and wealth. People living in workless households are more likely to experience poverty and children growing up in workless households are less likely to have favourable outcomes[1].

The LFS is a rich source of information about the labour market status of people, families and households. It has a sample size of around 53,000 households each quarter and collects a wide range of information. Originally it was a survey that focused on people and in the early stages there was little attention given to the information it provides on families and households. The survey design, data collection and processing procedures were chosen to produce the best possible statistics at the person level but, in some respects, that did not coincide with producing the best possible statistics on families and households. ONS has therefore produced separate LFS datasets for family and household-level analyses. This has involved adjusting for some inconsistencies in the data for earlier periods which affected the family and household structure; adding new derived variables for use in family and household-level analyses; and deriving household-level weighting factors.

In autumn 2008, ONS launched new annual local area datasets called the Annual Population Survey (APS) household datasets. They allow production of family and household labour market statistics at local areas and for small sub-groups of the population across the UK.

Derivation of the APS household dataset comes from the Labour Force Survey (LFS) and the APS (person) dataset. The APS dataset is created by combining individuals in waves 1 and 5 from four consecutive LFS quarters to the English, Welsh and Scottish Local Labour Force Surveys (LLFS). The APS household datasets therefore contain results from four different sources. The APS household datasets available for the calendar period of January to December for individuals years from 2004 and the LFS household datasets are available for the period April to June from 1992.

The APS household sample is three times the size of the LFS sample. It contains information collected from a sample of around 160,000 households (300,000 people aged 16 or over), whereas the LFS datasets only has a sample size of around 53,000 households (100,000 people aged 16 or over). The APS household dataset aims to get information on a minimum of 510 economically active persons for each Local Area Authority and Unitary Authority. It is the recommended source for local area data because of the greater sample size. Table 1 compares the sample sizes of the APS (household) and LFS (household) datasets.

Table 1

Year	APS			LFS*		
	Sample Size	Number of households	People aged 16 and over	Sample Size	Number of households	People aged 16 and over
2004	387,026	163,803	306,982	128,524	54,037	101,654
2005	383,771	162,771	305,393	126,587	53,275	100,389
2006	362,901	154,506	289,019	124,108	52,398	98,858
2007	365,926	156,142	292,997	123,715	52,202	98,327

* Using April to June quarter datasets

Volume 1 of the LFS User Guide[2] gives general information about the background and methodology of the LFS with Volume 6 giving information about the APS. This volume gives information that is specific to the LFS and APS household datasets, including information on:

- concepts and definitions used in family and household analyses (see Section 2);
- differences between the household and person-level datasets (Section 3);
- household weighting factors (Section 4);
- adjustments for inconsistencies in the family and household data for 1992 to 1995 (Section 5);
- the types of analyses that the household datasets should be used for (Section 6);
- variables specific to the household datasets (Section 7);
- producing family and household analyses, including worked examples (Sections 8 and 9);
- creating user-defined family and household-level variables (Section 10);
- publication of family and household statistics (Section 11);
- recent and planned improvements to family and household statistics (Section 12).

Section 2: Concepts and definitions

Households and family units
A **household** comprises of a single person, or a group of people living at the same address who have the address as their only or main home. They also share one main meal a day or share the living accommodation (or both). Most households contain one family unit but it is not uncommon for a household to contain two or more family units.

A **family unit** can comprise either a single person, or a married/cohabiting couple, or a married/cohabiting couple and their never-married children who have no children of their own living with them, or a lone parent with such children. Note the LFS concept of a 'family unit' is broader than the concept of a 'family' as defined in other sources. For example, 'one-person family units' as defined in the LFS may be 'persons not in families' in other surveys. Also, LFS family units can include non-dependent 'children' (who may be adults by their age) provided they have never-married and have no children of their own living with them. This may be inconsistent with other sources.

People who live with their parents and are married or have children of their own living with them are treated (with their spouse/children) as being a separate family unit from their parents. Other examples of households that comprise two or more family units include:

- two or more friends living together;
- two or more siblings living together (if neither of their parents is living in the household);
- foster parents and their foster children;
- same-sex cohabiting partners (treated as being in separate family units).

An **extended family unit** includes all related people within a household: partners; parents; children; grandparents; grandchildren; brothers and sisters; relatives by marriage; and other relatives. Foster parents and foster children are not included in the same extended family and are assigned to separate family units within a household. Introduction of the extended family unit into the LFS first happened in 1996, when the survey adopted a 'household matrix' approach.

Household matrix
This design of the matrix method is with the aim of identifying all family units within households; providing more details on the composition of families and households; and providing greater flexibility in the analysis of family and household relationships. Information is collected on the relationship between each pair of household members, and a matrix of relationships is produced from which the number and composition of families is calculated. Originally recording of the relationship of to the head of household took place and interviewers divided individuals to family units according to their own judgement.

Person, family and household-level analyses
Person level analyses count and describe people, for example the number of mothers in employment.

Family level analyses count and describe family units, for example the number of couples where both partners are in employment.

Household level analyses count and describe households, for example the number of households where all the adults are in employment.

Household reference person
The household reference person is the householder, which is the household member who owns the accommodation; or is legally responsible for the rent; or occupies the accommodation as reward of their employment, or through some relationship to its owner who is not a member of the household. If there are joint householders, the one with the highest income is the household reference person. If their income is the same, then the eldest one is the household reference person.

The household reference person (HRP) was introduced into the LFS in 2001, in line with other ONS household surveys, to replace the head of household (HoH).

Head of household
In a household where there is one adult only, that adult is the head of their household (HoH). If there are two adults of the opposite sex living together as a married or cohabiting couple, the husband/male partner is the HoH. Otherwise, the oldest male householder, or the husband/male partner of the oldest female householder, is the HoH. Otherwise, the oldest female householder is the HoH.

Although information about relationship to HoH (RELH06) is still collected in the LFS, information about relationship to HRP (RELHRP6) is generally used instead. Most HoHs are male, because in households comprising either a mixed-sex couple or joint householders of the opposite sex, the male partner/householder is classified as the HoH, regardless of their income and age. To identify the HRP, householder status, income and age are considered. In about 90 per cent of households the HoH and HRP are the same person. In over 90 per cent of cases where they are not the same person, the HRP is the spouse or partner of the HoH. There are more female HRPs than female HoHs and HRPs tend to be slightly younger.

Head of family unit
In a one-person family unit, that person is the head of their family unit. In a mixed-sex couple, the male partner is the head of the family unit. In a lone parent family, the lone parent is the head of the family unit. In a civil partnership, if one of the civil partners is the household reference person (HRP) that partner is the head of the family unit. If neither civil partner is the HRP (because someone else in the household is the HRP), the older partner is the head of the family unit. In same-sex cohabiting couples, the two partners are treated as separate family units and each partner is the head of their family unit.

In a household comprising two or more family units, a person who is the head of their family unit will not necessarily be the head of their household, or the household reference person.

Adults, children and dependent children
Sixteen is the minimum age at which most young people are legally allowed to leave full-time education. It is also the lower age threshold used for the LFS questions to classify respondents' economic activity status and the youngest age at which interviews take place in the survey (other members of the household provide information about people under 16). Therefore, in the LFS, the term 'adult' means 'people aged 16 or over'. However, in the variable CAIND (child/adult indicator) it means 'people who are not dependent children'.

Those never-married, have no children of their own living with them and live with their parents are treated as being part of their parents' family unit. They are coded as being a 'child' of the head of their family unit (RELHFU = 3).

Therefore, in the variables RELHFU, RELHRP6 and RELH06, 'children' includes people who may be adults by their age, as well as people aged under 16. Some of these will be classified as 'dependent children' and others will be 'non-dependent children'. Dependent children are those under 16 years old and those never-married aged 16 to 18 in full-time education. Non-dependent children are those aged 19 or over and those aged 16 to 18 who are married and/or not in full-time education.

Defining dependency is in the context of a family unit or household. This means, for example, that a never-married 16 to 18 year-old in full-time education but 'flown the nest' is not classified as a dependent child. Students living in halls of residence whose parents live in a UK household are an exception because, in the LFS, parents provide information about them as if they were part of their parents' household.

Note the LFS definition of 'dependent child' differs with the definition used in some other sources. For example, the Family Resources Survey defines dependent children as 'all those aged under 16 or an unmarried 16 to 18 year-old in full-time **non-advanced** education'.

Working-age people
The working-age population is men aged 16 to 64 and women aged 16 to 59. This is because 65 and 60 are the minimum ages at which most men and women in the UK can claim a State Retirement Pension.

Pension-age people
The pension-age population is men aged 65 or over and women aged 60 or over. This is because 65 and 60 are the minimum ages at which most men and women in the UK can claim a State Retirement Pension.

Couple
Two adults, living in the same household as each other, who are married or civil partners (spouses), or who live together as such (cohabiting partners).

Working-age couple
A couple where both partners are of working-age.

Working-age household
A household that includes at least one person of working age.

Working household
A working-age household where all the adults (people aged 16 or over) are in employment.

Mixed household
A working-age household containing both working and non-working adults. Includes households with at least one adult in employment and at least one adult who is either unemployed or economically inactive.

Workless household
A working-age household where none of the adults is in employment. Includes households where all adults are unemployed, those where all adults are economically inactive, and those containing both unemployed and economically inactive adults.

Section 3: Differences between the household and person-level datasets

The LFS and APS household datasets differ from the main (person-level) LFS and APS datasets in the following ways:

Weighting

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

The person-level datasets give zero weights to household members with unknown economic activity status. In the household datasets these individuals are weighted equally with the other household members (see Section 4).

Adjustments for anomalies in family and household data

The household datasets for 1992 to 1995 incorporate adjustments for inconsistencies that were found in the family and household data for that period. These adjustments are not included on the equivalent person-level datasets (see Section 5).

Variables

The household datasets include all the variables found on the person-level datasets, except for the earnings variables. They also include several derived variables at the family and household levels, to simplify analysis of the combined economic activity status of family and household members. These variables are not on the person-level datasets (see Section 6).

The household datasets for 1992 to 1995 include variables for family and household type that have been adjusted to correct for the effects of 'missing household members'. They also correct for a problem with the coding of marital status (see Section 5 and Section 6).

Uses

The design of the household datasets is so they give the best estimates at the family and household levels, whereas the design of the person-level datasets is so they give the best estimates at the person level. The weighting procedure inevitably causes differences between the person-level estimates produced from the household datasets and equivalent results produced from the main quarterly datasets.

The main quarterly LFS datasets should be used for most types of analysis at the person level. The household datasets should be used for family and household-level analyses and for person-level analyses involving the characteristics of the families or households in which people live (see Section 7).

The household datasets cannot be used for analyses of earnings, because they do not include an earnings weight (see Section 8).

Availability

Availability of the LFS person-level datasets is for each calendar quarter, whereas availability of the LFS household datasets are for Q2 (April to June) and Q4 (October to December) only (see Section 8). Availability of the APS person-level datasets are on a rolling calendar year basis (January to December, April to March, July to June, October to September. The APS household dataset is only available for the calendar year January to December.

Section 4: Weighting the (i) LFS and (ii) APS sample of households

Introduction

The LFS collects information from a sample of the population. Weighting of the sample ensures the estimates are representative of the whole population and compensates for differential non-response among different sub-groups. Each person in the achieved sample has a weight, which is the number of people that person represents. Before production of the household datasets, the weighting for all LFS results was on a person-level basis. Volume 1 of the LFS User Guide[2] describes this procedure. In summary, the population is split into sub-groups, with population estimates providing the number of people in each sub-group. The weights are calculated by dividing the population in each subgroup by the number of cases in that subgroup. A multi-stage procedure is used: first weighting to population estimates for Local Authority Districts; then to national population estimates by sex and age group; then by sex, region and age group. There are many iterations of the procedure until the results are stable.

Under this procedure age, sex and area of residence affect a person's weighting factor, but not the type of household in which he or she lives, nor the characteristics of the other people in their household. Using person-level weighting factors to weight household-level data may produce a biased distribution of household types, and a biased distribution for sub-groups of people by the type of household in which they live. First, we know from a study which examined the characteristics of LFS non-respondents by using information from the 2001 Census[2], there is under representation of young adults in one-person households in the LFS (being especially difficult to find at home). To compensate for this, those young adults who are in the achieved sample will have a high weighting factor. However, since they are the ones who **are** included in the survey, they will consist disproportionately of young people living in larger households, rather than in one-person households. This results in under representation of one-person households in the weighted household results. Also, although the weighted total numbers of young adults will be correct, there will be bias in their distribution by type of household. Also, since people in the same household will have different weights from each other, the weighting system for the person-level datasets gives estimates of men and women in mixed-sex couples that are inconsistent with each other. Therefore, the design of the household datasets is so the weighting factors are the same for all members of a household.

Calculation of household weighting factors

Various national statistical institutions use the method for building the household weights, referred to as 'integrated weighting' (see Lemaitre and Dufour[3]). The following specifications are defined for the weighted 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 Government Office Region to be consistent with population estimates;

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

Application of these specifications takes place at the same time by using the method known as calibration, or generalised raking. A logit method is used, with the allowable range of weighting factors set to minimise the range, subject to avoiding significant 'bunching' of cases at the boundary.

Note that although the weighting method forces agreement with population totals, it does not fix the weighted numbers of households (nor families), so these are survey estimates. A different sample would produce a different estimate of total households and total families.

The weighting method accounts for disproportionate response by age group and sex and separately by region, but there is no direct action taken for non-response at the household level. So if there is under representation of households consisting of only young men in the responding sample, and if their characteristics differ from young men in other types of household, then weighting will not correct for this bias. Without this type of bias, or with a consistent bias over time, it is likely that period-on-period changes in the weighted household estimates will reflect real changes because the LFS has such a

large sample. However, if response rates fall (or rise) substantially, there is an increased risk that part of any noted change is false.

People with unknown economic activity status (LFS only)
For some households, one or more household members have an unknown economic activity status. This is because they are present and refuse to take part in the survey or because they are absent and other household members feel unable to give a proxy response. The household matrix includes these 'non-respondents' (IOUTCOME = 3) and they are used to derive family and household-type variables. In the person-level datasets, people with unknown economic activity status are 'zero-weighted' (they are ignored in weighting up to population level estimates). However, for family and household-level analyses, they need to be retained in their correct place within the household structure. They are therefore included in the weighting procedure for the household datasets and are given the same weight as the other members of their household. This means there will be some 'not known' cases in any weighted analysis of economic activity status and other variables in the household datasets, either at the person, family or household level.

Imputation for people with unknown economic activity status (APS only)
The APS household datasets aim to provide more robust household and family labour market statistics. To ensure the economic status of all individuals within a household is known a method of 'donor' imputation takes place for those with a 'missing' economic activity status, allowing for analysis of the combined economic status of households.

Donor imputation on the APS household datasets uses the CANadian Edit and Imputation System (CANCEIS) software, developed by Statistics Canada. Its uses include imputing many variables into the 2001 Canadian Census and the imputation process of several ONS surveys. The procedure for imputing the APS household dataset is based on the initial imputation of the variable for economic activity which has the four categories described earlier. Non-responders under the age of 16 years are coded to child and the remaining imputed using the donor imputation method. Donors for cases with no economic activity status are identified using several matching variables. A multinomial logistic regression model identifies a set of matching variables which are age, sex, relationship to household reference person, household type and the economic status from the previous interview. Imputation takes place for non-responders using individuals with similar matching variables, likened to a 'nearest neighbour' form of imputation. Checking of the suitability of the imputation process takes place by comparing the distributions for key variables before and after imputation.

As there are no individuals on the APS household datasets with unknown economic status this source is preferred for producing family and household level analysis.

Comparison of estimates from the household and person-level datasets
Table 4.1 compares estimates for men and women in mixed-sex couples, produced from the LFS household and person-level datasets for April-June 2007. The person-level dataset produces figures for men and women in mixed-sex couples that are inconsistent with each other. According to the person-level dataset, there were 12.335 million husbands but only 11.970 million wives, and there were 2.443 million men but only 2.341 million women cohabiting with a partner of the opposite sex. The equivalent household dataset produces figures for men and women in mixed sex couples that are consistent with each other, because members of the same household have the same weight as each other.

Table 4.1. Men and women in mixed-sex couples: comparison between the household and person-level datasets
United Kingdom, April-June 2008, not seasonally adjusted

Thousands

	Married men	Married women	Cohabiting men	Cohabiting women
Household dataset	12,147	12,147	2,628	2,628
Person-level dataset	12,400	12,088	2,686	2,522

Source: Labour Force Survey

Table 4.2 compares estimates for family units produced from the household and person-level datasets for April-June 2007. The household dataset produces a lower figure for the total number of family units than the person-level dataset. The distributions by type of family unit are similar but one-person family units and married couple families with non-dependent children account for a slightly higher proportion of the total in the household dataset than in the person-level dataset. Also married couple with dependent children and lone mother families with dependent children account for a slightly lower proportion of the total.

Table 4.2. Families by type: comparison between the household and person-level datasets
United Kingdom, April-June 2008, not seasonally adjusted

Thousands and per cent

	Thousands			Per cent		
Type of family unit	Household dataset	Person-level dataset	Diff: hhold minus person	House hold dataset	Person-level dataset	Diff: hhold minus perso n
One-person (male)	5,078	4,792	286	18	17	1
One-person (female)	5,445	5,317	128	19	19	1
Married couple, no children	6,014	6,129	-116	21	22	0
Married couple, non-dep children only	1,556	1,354	202	6	5	1
Married couple, dependent children	4,577	4,917	-339	16	17	-1
Cohab couple, no children	1,506	1,497	9	5	5	0
Cohab couple, non-dep children only	96	76	20	0	0	0
Cohab couple, dependent children	1,027	1,114	-87	4	4	0
Male lone parent, dependent children	167	172	-5	1	1	0
Male lone parent, non-dep children only	176	154	22	1	1	0
Female lone parent, dependent children	1,708	2,014	-306	6	7	-1
Female lone parent, non-dep children only	681	625	56	2	2	0
Same-sex couple, with/without dep children	120	115	5	0	0	0
Civil partners with/without dep children	40	40	-1	0	0	0
Total	**28,190**	**28,316**	**-126**	**100**	**100**	**0**

Source: Labour Force Survey

Table 4.3 compares estimates for households produced from the household and person-level datasets for April-June 2007. The household dataset produces a lower figure for the total number of households than the person-level dataset. The distributions by type of household are similar but married couple households with dependent children and lone parent households with dependent children account for slightly lower proportions of the total in the household dataset than in the person-level dataset. Also married couples with all non-dependent children account for a slightly higher proportion in the household dataset than the person-level dataset.

Table 4.3. Households by type: comparison between the household and person-level datasets
United Kingdom, April-June 2008, not seasonally adjusted

Thousands and per cent

	Thousands			Per cent		
Type of household	Hhold dataset	Person-level dataset	Diff: hhold minus person	Hhold dataset	Person-level dataset	Diff: hhold minus person
One person	7,474	7,648	-174	29	29	0
Two or more persons, all diff. families	859	784	75	3	3	0
Married couple, no children, no others	5,694	5,810	-116	22	22	0
Cohab couple, no children, no others	1,340	1,340	0	5	5	0
Couple, no children, others	278	256	22	1	1	0
Married couple, all dep chld, no others	3,729	4,069	-339	15	15	-1
Cohab. couple, all dep chld, no others	911	988	-77	4	4	0
Marr. couple, dep & non-dep chld, no oth.	638	585	53	2	2	0
Cohab couple, dep & non-dep chld, no oth.	78	75	3	0	0	0
Marr. couple, all non-dep chld, no oth.	1,467	1,281	186	6	5	1
Cohab. couple, all non-dep chld, no oth.	91	75	16	0	0	0
Couple, all dep chld, others	122	131	-9	0	0	0
Couple, dep & non-dep chld, others	22	19	3	0	0	0
Couple, all non-dep chld , others	48	41	7	0	0	0
Lone parent, all dep chld, no others	1,459	1,754	-294	6	7	-1
Lone parent. dep & non-dep chld, no oth.	225	226	-1	1	1	0
Lone parent all non-dep chld, no others	794	735	59	3	3	0
Lone parent, all dep chld, others	77	84	-7	0	0	0
Lone parent, dep & non-dep chld, others	12	11	1	0	0	0
Lone parent, all non-dep chld, others	34	30	4	0	0	0
2 or more fam units, all dep chld.	104	108	-4	0	0	0
2 or more fam units, dep & non-dep chld.	61	52	9	0	0	0
2 or more fam units, all non-dep chld.	29	21	8	0	0	0
2 or more fam units, no children	44	34	10	0	0	0
Same sex couple with/without others	60	57	3	0	0	0
Civil partners, with/without others	39	41	-1	0	0	0
Total	**25,691**	**26,254**	**-563**	**100**	**100**	**0**

Source: Labour Force Survey

Re-weighting

The decennial Census of population is the base for the population figures used in the LFS and APS weighting procedure. Information from the NHS Central Register (for internal migration), the International Passenger Survey (for international migration), and administrative data (for births and deaths) provide data for population estimates between Censuses.

When there are revisions to population estimates, reweighting of the LFS and APS datasets takes place. Since November 2002, ONS has incorporated the latest population figures into the monthly Labour Market Statistics First Release[4], soon after the release of population figures, by using an interim-adjustment reweighting method[5]. Reweighting of the microdata has been less frequent, resulting in estimates from the microdata and those in the Labour Market Statistics First Release being inconsistent. The household datasets for each year back to 1997 have recently been reweighted to incorporate population figures published in August and September 2007. An assessment of the impact of the reweighting exercise, with revisions to family and household series, is in Section 12.

Section 5: Adjustments for anomalies in LFS family and household data (1992-1995)

The original design of the LFS focused on individuals and in the early years of the survey there was little attention given to the information provided on families and households. When users began to draw on this information, they identified several inconsistencies and discontinuities in the family and household data. The cause was changes in the definitions for variables underlying family and household type, and changes and anomalies in applying those definitions. Some of those issues were addressed in the person-level datasets. Production of the household datasets addressed the following issues:

Missing household members

Problems arose in the family and household data when the LFS moved from an annual to a quarterly survey in 1992 and, at the same time, switched from a paper questionnaire to a computerised system for recording the information collected. A few households contained individuals for whom no information could be obtained, either because they were present and refused to take part in the survey or because they were absent and other household members felt unable to answer by proxy. When the LFS moved from an annual to a quarterly survey, a decision was taken that if collection of information for key questions was not possible, to treat the interview for that person as an outright refusal. The algorithms for deriving the family and household type variables, used under the computerised system, ignored any person for whom no information (other than their existence) was gathered and this led to some false results. If the ignored person was married or cohabiting, the other partner was placed as a one-person family unit or as a lone parent, depending on whether they had children. Adopting the 'household matrix approach' in 1996 solved this problem (see Section 8).

Errors in marital status coding

A further problem affected family type coding in the 1995 datasets. Occasionally there was incorrect coding of marital status, resulting in classification of both the man and the woman in a married or cohabiting couple as the head of a male/female lone parent family unit.

Adjustments for missing household members

To compensate for the effects of the problems there are adjustments in the household datasets for 1992 to 1995. The existence of 'missing members' of a household was inferred from the available information on the characteristics of other household members. For some missing people, for example, one of a group of unrelated flat-sharers, this was too difficult. For most cases they could be identified by one of the following criteria: there was no recorded head of household; or the family unit was of the one-person or lone parent type but the head of the family unit was married or cohabiting. In the latter case, the family unit adjusted to a family type code of 'married or cohabiting couple' and the household adjusted to a household type code which was the 'couple household' equivalent of the unadjusted household type. For example, a household with an unadjusted category of 'lone parent with dependent and non-dependent children' would be given an adjusted category of 'married/cohabiting couple with dependent and non-dependent children'. Where the head of the household was missing, adjustments were made according to the unadjusted household type category and the relationship and marital status of the recorded household members. This procedure produced a few cases where, because of an unusual combination of missing household members, the adjusted family type and household type variables were inconsistent. These cases were dealt with by setting either the adjusted family type or household type variable to 'not known'. Occasionally (typically about 100 per dataset), the family unit had no head, or two or more heads, or two or more wives of head. The apparent causes were very varied, and usually it was not possible to identify the most probable real situation, so those cases were dropped.

Section 6: Variables specific to the LFS and APS household datasets

The household datasets include all the variables found on the main (person-level) LFS datasets, except those about earnings. They also include extra variables that are not on the person-level datasets. These fall into five groups:

1. Variables giving the household weight and family and household identifiers:

PHHWT07	Household weight (LFS)
PHHWT07a	Household weight (APS)
FUSERIAL	Number that uniquely identifies a family unit
HSERIAL	Number that uniquely identifies a household

2. For LFS 1992-1995, variables giving adjusted family and household type categories and variables indicating whether family and household type have been adjusted:

TFUADJ	Family type after adjustments
THHADJ	Household type after adjustments
FUCHANGE	Whether adjusted family type differs from unadjusted family type
HHCHANGE	Whether adjusted household type differs from unadjusted household type

3. A pair of family level variables that identify the economic activity status of the head of family unit and the economic activity status of the wife/partner of the head of family unit. The design such to allow analyses of the joint economic activity status of married and cohabiting couples:

HEAHEAD	Economic activity of head of family unit
HEAWIFE	Economic activity of the wife/partner of head of family unit

4. Variables giving the numbers of people in the household for various age groups that are of interest when analysing the economic status of households:

HNWKAGE	Number of working-age people in household
HNPEN	Number of people in household who are of pensionable age
HDPCH4	Number of children in household aged 4 years or less
HDC515	Number of children in household aged between 5 and 15 years
HDPCH18	Number of children in household aged 16 years and over

5. Variables giving the numbers of people in the household in various economic activity status categories, and a variable describing the combined economic activity status of the household:

HNUNEMP	Number of people in household who are unemployed
HNINACT	Number of people in household who are economically inactive
HNDK	Number of people in household whose economic activity status is unknown
HEACOMB	Combined economic activity status of household
HNFTSTUD	Number of people in household who are full-time students
HNOTSTUD	Number of people in household who are NOT full-time students
HNFTIME	Number of people in household who are working full-time
HNPTIME	Number of people in household who are working part-time
HNIWSTU	Number of people in household who are economically inactive, would like to work and are currently students
HNIWSKD	Number of people in household who are economically inactive, would like to work and are currently sick, injured or disabled
HNIWDSC	Number of people in household who are economically inactive, would like to work but are discouraged from seeking work
HNIWFAM	Number of people in household who are economically inactive, would like to work and are looking after family or home
HNIWOTH	Number of people in household who are economically inactive for other reasons and would like to work
HNINOWK	Number of people in household who are economically inactive and do not want to work

For details of these and other LFS variables, see Volume 3 of the LFS User Guide[6].

Section 7: When to use the household datasets

The design of the weighting procedure used in the household datasets is to give the best estimates at the family and household levels. This inevitably results in some differences between the person-level estimates produced from the household datasets and equivalent estimates produced from the main (person-level) datasets. This section explains which dataset should be used in which instance.

Background

The LFS and APS household datasets are more suitable than the person-level datasets for family and household-level analyses because:

- In the household datasets, the weight assigned is the same for all members of a household and individuals with unknown economic activity status do not have zero weights;

- Using person-level weights can produce biased distributions of household types, or biased distributions of particular sub-groups of people by the type of household in which they live. Also, it produces different numbers for the two partners in married and cohabiting couples.

- The household datasets include derived family and household-level variables, to simplify analyses of the combined economic activity status of family and household members. The person-level datasets do not contain these variables;

- The household datasets for 1992 to 1995 incorporate adjustments for anomalies that existed in the family and household data for that period. The person-level datasets do not contain the adjusted family and household type variables (TFUADJ and THHADJ).

General rule

For most person-level analyses use the main (person-level) LFS or APS datasets due to their design in giving the best estimates at the person level. Use the household datasets for family and household-level analyses (analyses that count and describe family units or households). Also use the household datasets for person-level analyses involving the type of family or household in which people live. For example, the household datasets should be used for analyses of:

- lone parents;
- couples (with or without dependent children) by the joint characteristics of the partners;
- people by the type of family or household in which they live;
- children by the characteristics of their parents;
- men and women by married/cohabiting status;
- households by the combined economic activity status of the household;
- working-age people or children by the combined economic activity status of the household in which they live.

For analyses by marital status, the person-level datasets produce different numbers of men and women in mixed-sex couples (there are typically more husbands than wives). Therefore, although analyses by marital status are not strictly analyses by family or household type, the household datasets should be used to ensure consistency of results.

Exceptions

In cases where a set of analyses is to be produced, some of which are at the person level and some are at the family or household level, and where consistency between the analyses is required, the household datasets should be used for all of the analyses rather than mixing results from the household and person-level datasets.

Section 8: General guidance on producing family and household analyses

This section gives general information and guidance that users need to be aware of when producing and interpreting family and household analyses. Section 9 provides guidance based on worked examples. Unless otherwise stated, the variables mentioned below (in capital letters) are for the April-June 2008 household dataset. The equivalent variables for previous years may have different names and categories. For details of these and other LFS variables, see Volume 3 of the LFS User Guide[6].

Periods covered by the household datasets

Before 2006, the household datasets covered seasonal quarters and were available for spring (March-May) quarters back to 1992 and autumn (September-November) quarters back to 1995. In 2006, the LFS moved to report on a calendar quarters, in line with EU requirements. The spring 2006 dataset was the last covering seasonal quarters. The household datasets are now available for Q2 (April-June) and Q4 (October-December) of each year with the Q2 datasets released in August and the Q4 datasets released in February.

ONS has produced a back-series of household datasets on a calendar quarter basis. The calendar-quarter series currently covers:

- Q2 of 1997 to 2008
- Q4 of 2006 to 2008.

ONS hopes to complete the back-series of calendar-quarter household datasets (back to 1992) by the end of 2008. Meanwhile, users should avoid making comparisons between seasonal quarter results for 1992 to 1996 and calendar quarter results for 1997 onwards. This is because:

- the seasonal quarter datasets are weighted to population figures published by ONS in February and March 2003, whereas the calendar quarter datasets are weighted to more recent population figures (published in August and September 2007); and

- seasonal and calendar quarter figures may not be comparable, due to seasonal effects.

Weighting

The weighting variable on the LFS household datasets (PHHWT07) differs from the weighting variable on the LFS person-level datasets (PWT07) in that PHHWT07 is the same for each member of a household and people with unknown economic activity status do not have zero weights. Weighting by PHHWT07 will produce a count of people unless using certain filters to produce a count of families or households. For the APS household datasets the weighting variable is PHHWT07a.

On the SuperCROSS version of the datasets there are two extra weighting variables that provide an alternative method for producing a count of families (FHHWT07) and households (HHWT07).

<u>Family-level analyses</u>

To produce a count of family units, weight by PHHWT07 and filter on RELHFU = 1 (head of family unit). The filter RELHFU = 1 produces a count of family units because each family unit only has one head. SuperCROSS users can use this method, or weight by FHHWT07 (without the filter RELHFU = 1).

<u>Household-level analyses</u>

To produce a count of households, weight by PHHWT07 and filter on RELHRP6 = 0 (household reference person). The filter RELHRP6 = 0 produces a count of households because each household only has one household reference person. SuperCROSS users can use this method, or weight by HHWT07 (without the filter RELHRP6 = 0).

The ‘household reference person’ (HRP) was introduced into the LFS in 2001. To produce household-level analyses for the periods 1996-2000, weight by PHHWT07 and select the head of household. For the period 1992 to 1995, some households with ‘missing household members’ have no recorded head of household (see Section 5). An indicator variable (HHIND) is available, which takes the value 1 for the first person in each household and 0 otherwise. To produce a count of households use the filter HHIND = 1. Table 8.1 summarises which filters to use for household-level analyses for each period back to 1992.

Table 8.1: Filters for producing household-level analyses

Q1 1992 – Q2 1995	HHIND = 1
Q1 1996	RELHOH = 1
Q2 1996 – Q2 2000	RELH96 = 0
Q1 2001 – Q2 2005	RELHRP = 0
Q1 2006 onwards	RELHRP6 = 0

LFS/ APS sample coverage and residency rules

The LFS sample covers people living in private households and NHS accommodation. It also covers students in halls of residence whose parents live in a UK household. People with more than one address are generally counted as being resident at the sampled address if they regard it as their main home in the UK. The following are also included:

- People who normally live there but are temporarily away (on holiday, on business or in hospital, for example), unless they have been living away from the address for six months or more.

- Children aged under 16, even if they are away at boarding or other schools.

Students aged 16 or over are generally treated as being resident at their normal term-time address. However, students in halls of residence are treated as being part of their parents' household. Collection of information about them takes place at their parents' address (usually by proxy, unless the student happens to be at there during the LFS interview).

If a person who usually lives at the sampled address has been away from the address for six months or more, they are not counted as being part of the household. This means that if one of the partners in a married or cohabiting couple has been away from the sampled address for six months or more (working abroad, in hospital or in prison, for example) the remaining partner will be coded either as a one-person family unit or as a lone parent, depending on whether they have children living with them. It also means there may be people whose family type is coded as 'one-person' or 'lone parent' but whose marital status is coded as 'married, living with husband/wife' (MARSTA = 2). The LFS question MARCHK (whether spouse is a household member) was introduced in 1996 to check whether people coded as MARSTA = 2 are legally married and are not legally separated but do not currently live with their spouse. This extra information is used to classify their family unit type. It is not used to classify their marital status, which is based solely on responses to the marital status question.

Before January 2008, LFS interviewers included people who had been living at the address for six months or more, even if they considered their main residence to be elsewhere, and to include people whose address in the UK was temporary while they searched for permanent accommodation, if they were here for more than six months. This six-month residency rule implied that people living in the UK for less than six months should not be included in the survey. The rule was removed from the LFS and other household surveys in January 2008 to improve coverage of short-term migrants. There is no longer a need for a potential respondent to have been living continuously at the address for six months or more. However, interviewers are instructed not to include people living at a temporary address in the UK who are here only for purposes of recreation, holiday, visits to friends and relatives, business, medical treatment or religious pilgrimage and who usually live abroad. The change in the rule has had no significant impact on the estimates and no discontinuity has been established.

Household matrix

Before 1996, interviewers assigned individuals to family units according to their own judgement and only recorded relationships to the head of household. Since 1996, the LFS has used a 'household matrix' approach to collect information about family and household structure. This method is designed to correctly identify all families within households; to provide more details on the composition of families and households; and to provide greater flexibility in the analysis of family and household relationships. It is now 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. Information on the relationship between each pair of household members allows creation a matrix of relationships which is used to derive the number and composition of families and households.

Since the household matrix approach requires recording of certain basic information for all members of the household, and people with no economic activity data are included in the household matrix, it has the additional advantage of eliminating, from 1996 onwards, the problem of 'missing household members' (described in Section 5). The few cases remaining with an apparent inconsistency between family type and marital/cohabitating status are confirmed by a check question (MARCHK) to be genuine cases where the partner has been living away from the residence for six months or more.

The household matrix approach collects more information than before. The matrix records marital status, cohabitation (including same-sex cohabitation); and whether children are dependent or non-dependent, stepchildren, foster children or not. As a result, some new variables were introduced from 1996 and some were revised. For example, the relationship to head of household variable was extended from 10 to 21 categories. The more detailed categories for household and family type distinguish between married and cohabiting couples and between families with no children, with non-dependent children only or with dependent children (with or without non-dependent children). The household matrix data for Q2 1996 may not be fully complete as there are some households (about 3 per cent) where data was imputed (carried forward from the previous quarter) because of circumstantial refusal or non-contact. These households will not have any data in the form of the household matrix.

Residual anomalies in the LFS 1992-1995 household datasets

The household datasets for 1992 to 1995 include adjusted family and household type variables (TFUADJ and (THHADJ) that incorporate corrections for inconsistencies found in the family and household data for that period (see Section 5). For family units and households where adjusted family and household type codes were assigned, it was not possible (because of lack of information) to make changes to the recorded data on relationship to head of family unit or head of household. Therefore, there are some individuals in the 1992 to 1995 datasets whose relationships to other family and household members are inconsistent with their family or household type.

After the adjustments described in Section 5, there remain a few potential sources of minor residual anomalies in the household datasets for 1992 to 1995 that were impossible to resolve. The key potential impact of these residual anomalies is on the distribution of the household combined economic activity status variable (HEACOMB). The most critical indicator derived from this variable is the percentage of workless households. The estimated effect of the residual anomalies is an underestimation of the percentage of workless households of up to 0.07 percentage points. A difference of this size could occasionally cause a slight change in the rounded published figures, which are usually rounded to one decimal point, but is unlikely to affect any inferences.

There are some couple families in the 1992 to 1995 datasets where one of the partners is coded in the marital status variable (MARCON) as 'married' while the other partner is coded as 'cohabiting'. These may be cohabiting couples where one of the partners is married to someone else but it impossible to confirm whether that is the case. The family unit type variable for 1992 to 1995 (TFUADJ) combines married and cohabiting couples into a single category (married/cohabiting). Since there are some couples whose married/cohabiting status is ambiguous, it is not possible to analyse married and cohabiting couples separately for the period 1992 to 1995.

Marital status

The marital status variable (MARSTA) records legal marital status. As mentioned under 'residency rules' above, some people who classify themselves as 'married, living with husband/wife' (MARSTA = 2) are coded in the family type variable (FUTYPE6) as living in either a one-person or lone parent family unit (depending on whether they have children living with them). This is because their spouse has been absent from the address for six months or more and is therefore not counted as a member of the household (MARCHK = 2).

People coded in MARSTA as single (never-married), separated, divorced, widowed or a former civil partner may be part of a cohabiting couple. The variable LIVWTH (whether living together as a couple) identifies whether that is the case.

On the LFS the variable MARDY6 (whether married/cohabiting) is derived from MARSTA, MARCHK and LIVWTH but it should not be used to select people living/not living with a spouse or partner. This is because, in the household datasets, non-respondents (IOUTCOME = 3) are coded as 'unknown' (-9) in MARDY6.

To produce a count of people living with a spouse or partner, weight by PHHWT07(a) and filter on: ((MARSTA = 2, 6 and MARCHK = 1) or (LIVWTH = 1, 3)).

To produce a count of people not living with a spouse or partner, weight by PHHWT07(a) and filter on: ((MARSTA = 1, 3-5, 7-9 and LIVWTH = 2) or (LIVWTH = -9)).

Working-age people

The variable WRKAGE (working age) should not be used to select working-age people. This is because, in the household datasets, non-respondents (IOUTCOME = 3) are coded as 'unknown' (-9) in WRKAGE.

To select working-age people, weight by PHHWT07(a) and filter on: ((SEX = 1 and AGE = 16-64) or (SEX = 2 and AGE = 16-59)).

Children and dependent children

People who live with their parents, are never-married and have no children of their own living with them are treated as being part of their parents' family unit and are coded as a 'child' of the head of the family unit (RELHFU = 3). Therefore, 'children' as defined in the variables RELHFU, RELHRP6 and RELH06 includes people aged 16 or over, as well as children under 16. Some of them will be dependent children and others will be non-dependent children.

Most children live in either a couple or lone parent family unit and are coded as a 'child' of the head of their family unit (RELHFU = 3). However, foster children and other children who live with someone other than their parents (with an older sibling, for example) are treated as being a separate (one-person) family unit and are coded as the head of their family unit (FUTYPE6 = 1,2 and RELHFU = 1).

To produce a count of children under 16 who live with one or both of their parents, weight by PHHWT07(a) and filter on AGE < 16 and RELHFU = 3. To include all children under 16 living in private households, weight by PHHWT07(a) and filter on AGE < 16.

To produce a count of dependent children who live with one or both of their parents, weight by PHHWT07(a) and filter on RELHFU = 3 and CAIND > 1. To include all dependent children living in private households, weight by PHHWT07(a) and filter on CAIND > 1 (to exclude foster children, add the filter RELHRP6 not equal to 5).

Adopted children
Adopted children are treated the same as other children in variables that record relationships to the head of family unit, household reference person or head of household. They are part of their adoptive parents' family unit (RELHFU = 3) and are counted in FDPCH19, AYFL19, HDPCH19 and other variables describing the number and ages of dependent children in their adoptive parents' family and household.

Stepchildren
Before 1996, stepchildren were not distinguished from other children present in the household. From 1996 onwards, stepchild and stepparent are recorded as relationships to household reference person and head of household but stepchild is not recorded as a relationship to head of family unit. Stepchildren are coded as RELHFU = 3 and are counted in FDPCH19, AYFL19, HDPCH19 and other variables describing the number and ages of dependent children in their stepparents' family and household.

Foster children
Foster children are treated as being a separate family unit from their foster parents. The foster child is coded as the head of a one-person family unit (FUTYPE6 = 1,2 and RELHFU = 1). In FDPCH19 and AYFL19 (and other variables describing the number and ages of dependent children in the family unit), the foster child is counted as a dependent child within its own family unit and is not counted as a dependent child within its foster parents' family unit. However, foster children are counted in HDPCH19 and other variables describing the number and ages of dependent children within their foster parents' household. Foster child and foster parent are recorded as relationships to household reference person and head of household.

Other children who do not live with their parents
If a child lives with its grandparents, and neither of its parents lives in the household, the child is generally treated as being part of its grandparents' family unit and is coded as RELHFU = 3. However, if a child lives with another relative, or with a non-relative, it is generally treated as being a separate family unit (RELHFU = 1 and FUTYPE6 = 1,2). In FDPCH19 and AYFL19 (and other variables that describe the number and ages of dependent children in the family unit) the child is counted as a dependent child within its own family unit and is not counted as a dependent child within its carers' family unit. However, such children are counted in HDPCH19 and other variables describing the number and ages of dependent children in their carers' household.

Children in same-sex cohabiting couple and civil partnership families
Before 2006, the family and household type variables did not specify whether same-sex couples had children or not. If any children were present they were not counted in FDPCH19, AYFL19, HDPCH19 or in other variables describing the numbers and ages of dependent children in the family unit or household. Same-sex couple families and households were coded as 'does not apply' (DNA) in such variables. From 2006, children of same-sex cohabiting couples and civil partnerships are counted in FDPCH19, AYFL19, HDPCH19 and other such variables. The variable FUTYPE6 now includes categories that distinguish between same-sex cohabiting couple and civil partnership families with and without dependent children.

The LFS may underestimate numbers of children in same-sex couple and civil partnership families and households because same-sex couples and civil partnerships are recorded only if the respondent mentions, unprompted, that they are part of a same-sex couple when interviewed. The number of same-sex couples and civil partnerships included in the sample may be too small to provide reliable estimates for children in such families and households. Users should be cautious when interpreting results for children in same-sex couple and civil partnership families and households. When producing analyses of children by family or household type, consider grouping same-sex cohabiting couples and civil partnerships together with mixed-sex couples, rather than showing them as a separate category.

Children by the characteristics of their parents
The household datasets are often used for analyses of children by their parents' characteristics. For example, they are the principal source for statistics on children with two parents in employment, one parent in employment, and no parent in employment. When relating the characteristics of children and their parents, family level variables are used in preference to household-level variables. This is because, in a household comprising two or more family units, the children may not belong to the household reference person or head of household.

As mentioned previously, most children live in either a couple or lone parent family and will be coded as a 'child' of the head of their family unit (RELHFU = 3) but foster children (and others who do not live with their parents) are treated as being in a separate family unit (RELHFU = 1 and FUTYPE6 = 1,2). Analyses of children by the characteristics of their parents focus on children who live with one or both of their parents. Foster children (and others who do not live with their parents) are excluded from the analysis by filtering on RELHFU = 3. It is important to bear in mind that the 'children' within a family unit (RELHFU = 3) may be aged 16 or over and some of them will be dependent children while others will be non-dependent children.

For analyses of children under 16 by their parents' characteristics, weight by PHHWT07(a) and filter on RELHFU = 3 and AGE < 16.

For analyses of dependent children by their parents' characteristics, weight by PHHWT07(a) and filter on RELHFU = 3 and CAIND > 1.

Children by the characteristics of the household in which they live

The household datasets are often used for analyses of children by the characteristics of the households in which they live. For example, they are the principal source for statistics on children under 16 who live in a workless household. This type of analysis looks at children by the combined characteristics of the members of their household and does not focus on children who live with their parents (although it is possible to do so). Children under 16 are selected by filtering on AGE < 16. Dependent children are selected by filtering on CAIND > 1. The results will include foster children (and others who do not live with their parents) as well as children who live with natural, adoptive and stepparents. To exclude foster children from the analysis, add the filter RELHRP6 not equal to 5. To exclude all children who live with someone other than their parents, add the filter RELHFU=3.

People with and without dependent children

Analyses of people with and without dependent children are person-level analyses but they involve using variables that describe the families in which people live. Family level variables are used in preference to household-level variables. This is because, in a household comprising two or more family units, the children may not belong to the household reference person or head of household.

To produce a count of people **with** dependent children, weight by PHHWT07(a) and filter on RELHFU = 1,2 and FUTYPE6 = 6,9,10,12,16,19. The results will include adoptive parents and stepparents, as well as natural parents, but foster parents will not be included. This is because adopted and stepchildren are treated as being in the same family unit as their adoptive/stepparents but foster children and foster parents are treated as being separate family units.

People with dependent children should not be selected by filtering on RELHFU = 1,2 and FDPCH19 > 0 (nor by filtering on RELHFU = 1,2 and AYFL19 = 0-18). Otherwise, the results will include foster children (and others who do not live with their parents. Such children are coded as the head of a one-person family unit (RELHFU = 1 and FUTYPE6 = 1,2) and are counted in FDPCH19 and AYFL19 (for their own family unit).

People **without** dependent children are those who do not meet the criteria defined above for people with dependent children. The easiest way to select people without dependent children is to create a user-defined 'parental status' variable in which people with dependent children are defined as RELHFU = 1,2 and FUTYPE6 = 6,9,10,12,16,19, and people without dependent children are defined using an ELSE statement (SPSS) or using the 'drop-through' function (SuperCROSS).

People without dependent children should not be selected by filtering on FUTYPE6 not equal to 6,9,10,12,16,19 (nor by filtering on FDPCH19 = 0, nor by filtering on AYFL19 =19). Otherwise, the results will exclude people who are dependent children themselves. They will also exclude non-dependent children with siblings who are dependent children.

The LFS/ APS results for people without dependent children may include parents whose children live in a separate household from them (separated or divorced parents, for example). It is not possible to identify whether that is the case, because the LFS does not ask people whether they have any children who live in a separate household.

Families with and without dependent children

The household datasets contain several variables that describe the number and ages of the dependent children in the family unit, such as FDPCH19 and AYFL19. Foster children (and others who do not live with their parents) are counted in these variables for their own (one-person) family unit. Therefore, if the filter FDPCH19 > 0 (or AYFL19 = 0-18) is used, without also filtering on family type, the results for 'families with dependent children' will include one-person family units headed by a child. They may also include a few young cohabiting couples without children, who live with the parents of one of the partners, if that partner is coded as a dependent child in CAIND. The best way to select families with or without dependent children is to filter on the categories of FUTYPE6 that are labelled as having dependent children. Note that families in FUTYPE6 categories with and without dependent children may contain non-dependent children.

To produce a count of families **with** dependent children (excluding one-person family units headed by a child) weight by PHHWT07* and filter on RELHFU = 1 and FUTYPE6 = 6,9,10,12,16,19.

To produce a count of families **without** dependent children, either weight by PHHWT07* and filter on RELHFU = 1 and FDPCH19 = 0, or weight by PHHWT07* and filter on RELHFU = 1 and AYFL19 = 19.

* SuperCROSS users can use an alternative method: weight by FHHWT07, without filtering on RELHFU = 1.

Households with and without dependent children

The household datasets contain several variables that describe the number and ages of the dependent children in the household, such as HDPCH19 and AYHL19. Although foster children (and others who do not live with their parents) are not counted in FDPCH19 or other variables describing the number and ages of the dependent children in their foster parents' or carers' family unit, they are counted in HDPCH19 and other variables describing the number and ages of the dependent children within their foster parents' or carers' household.

To produce a count of households **without** dependent children, weight by PHHWT07(a)* and filter on RELHRP6 = 0 and HDPCH19 = 0.

To produce a count of households **with** dependent children, weight by PHHWT07(a)* and filter on RELHRP6 = 0 and HDPCH19 > 0.

* SuperCROSS users can use an alternative method: weight by HHWT07, without filtering on RELHRP6 = 0.

Households with dependent children should not be selected by filtering on the categories of HHTYPE6 that are labelled as having dependent children. This is because there may be dependent children present in HHTYPE6 categories that are not labelled as having dependent children. This is illustrated in **Table 8.2**, which shows a breakdown of households with dependent (HDPCH19 > 0) by type of household.

In the April to June 2008 LFS, there were an estimated 7.465 million households with dependent children, of which:

- 11,000 were classified as 'married or cohabiting couple with no children, with other family units'
- 6,000 were classified as 'same-sex couple or Civil partnership with or without other family units'
- 9,000 were classified as 'two or more people, all different family units'.

Therefore, if households with dependent children are selected by filtering on HHTYPE6 categories that are labelled as having dependent children, the results will undercount the number of households with dependent children by 26,000 (0.4 per cent).

Table 8.2: Households with dependent children by type of household[1]
United Kingdom, April-June 2008, not seasonally adjusted

Thousands and per cent

Household type	HHTYPE6 codes	HDPCH19 > 0 '000s	HDPCH19 > 0 %
Married or cohabiting couple household with dependent children	6-9 ,12, 13	5,501	73.7
Lone parent household with dependent children	15, 16 ,18 ,19	1,773	23.8
Two or more family units with dependent children	21, 22	165	2.2
Couple, no children, other family units	5	11	0.1
Same-sex couple or Civil partnership household with or without others	25, 26	6	0.1
Two or more people, all different families	2	9	0.1
Total	**2, 5-9, 12, 13, 15, 16, 18, 19, 21, 22, 25, 26**	**7,465**	**100.0**

Source: Labour Force Survey household dataset

[1] Children under 16 and those aged 16-18 who are never-married and in full-time education.
[2] The variable HDPCH19 counts the number of dependent children in the household.

Couples and lone parents
Analyses of couples and lone parents are carried out using family-level variables rather than household-level variables. This is because not all couples live in 'couple' households and not all lone parents live in 'lone parent' households. This is illustrated in **Table 8.3**, which shows a breakdown of couple and lone parent families with dependent children by the type of household in which they live.

In April-June 2007 there were an estimated 5.484 million couple families with dependent children, of which 5.402 million lived in a 'couple' household and 82,000 lived a 'household with two or more family units'. There were an estimated 1.86 million lone parent families with dependent children, of which 1.772 million lived in a 'lone parent' household and 89,000 lived in a 'household with two or more family units'.

Therefore, if couples and lone parents with dependent children are selected using HHTYPE6 (rather than FUTYPE6), the results will undercount the numbers of couples and lone parents with dependent children by 82,000 (1 per cent) and 89,000 (5 per cent) respectively.

Table 8.3: Couple and lone parent families with dependent children[1] by the type of household in which they live
United Kingdom, April-June 2007, not seasonally adjusted

Thousands and per cent

	Family unit type (FUTYPE6)		
Household type (HHTYPE6)	Couple[2] with dependent children	Lone parent with dependent children	All families with dependent children
Thousands			
Couple[2] household	5,504	0	5,402
Lone parent household	0	1,771	1,771
Household with two or more family units	106	104	209
Total	**5,610**	**1,875**	**7,486**
Per cent			
Couple[2] household	98	0	74
Lone parent household	0	94	24
Household with two or more family units	2	6	3
Total	**100**	**100**	**100**

Source: LFS household dataset

[1] Children under 16 and those aged 16-18 who are never-married and in full-time education.
[2] Includes married couples, mixed-sex cohabiting couples, same-sex cohabiting couples and Civil partnerships.

Working-age couples
When analysing the joint economic activity status of couples, and when making comparisons between couples with and without dependent children, users may wish to restrict the analysis to couples where one (or both) of the partners is of working age. The best way to do that is to create a pair of user-defined family-level variables: one that identifies whether the head of the family unit is of working age (HEADAGE), and another that identifies whether the wife/partner of the head of family unit is of working age (WIFEAGE). These variables can then be used to select couples where one or both partners are of working age. Guidance on how to create these and other user-defined family-level variables is given in Section 10a (for SuperCROSS users) and Section 10b (for SPSS users).

Note that if the filter HNWKAGE > 0 is used, the results will include couples where one of the partners is of working age and the other is not. They will also include couples where both partners are of pension age, living in a household with someone of working age (with their adult offspring, for example).

If the filter HNPEN = 0 is used, the results will exclude couples where one of the partners is of pension age. They will also exclude couples where both partners are of working-age, living in a household with someone of pension age (with an elderly relative, for example).

Same-sex cohabiting couples and civil partnerships
Categories identifying same-sex cohabiting couples have been included in the family and household type variables since 1996. However, respondents are recorded as being a partner in a same-sex cohabiting couple only if they volunteer this information, unprompted, in answer to the question LIVWTH (whether living together as a couple). The LFS may therefore underestimate the numbers of same-sex cohabiting couple families and households.

In the variable FAMUNIT (family unit number), which is used to derive FUTYPE6, same-sex cohabiting partners are coded into separate family units from each other. Therefore, family-level analyses of same-sex cohabiting couples will produce figures for partners in same-sex couples, rather than figures for same-sex couples.

Additional categories were added to several LFS variables in 2006, following the introduction of civil partnerships. For example, the variable MARSTA (marital status) now includes categories that record whether the respondent is (or has been) a partner in a legally recognised civil partnership. However, respondents are recorded as being (or having been) in a civil partnership only if they volunteer this information, unprompted, in answer to the marital status question. The LFS may therefore underestimate the numbers of civil partnership families and households.

As explained above, it is possible that same-sex cohabiting couples and civil partnerships are under-reported in the LFS. It is also possible that any apparent changes in the numbers of same-sex couples and civil partnerships over time may be due to changes in peoples' propensity to report that they are a part of a same-sex couple or civil partnership when interviewed. The number of same-sex couples and civil partnerships included in the sample may be too small to provide reliable estimates for these and other relatively small subgroups of the population. Users should be cautious when interpreting results for same-sex cohabiting couples and civil partnerships. When producing analyses by family or household type, consider grouping same-sex cohabiting couples and civil partnerships together with mixed-sex married/cohabiting couples, rather than showing them as a separate category.

Combined economic activity status of household
The household datasets are the principal source of statistics on 'working', 'mixed' and 'workless' households. These are produced using the variable HEACOMB (combined economic activity status of household). The filter HNWKAGE > 0 is used to select households with at least one person of working age. The results include households where there are people of state pension age as well as people of working age, but pensioner-only households are excluded.

When interpreting results produced from HEACOMB, it is important to bear in mind that the variable covers the economic activity status of every member of the household aged 16 or over. This includes any dependent children aged 16 or over, any non-dependent children, and any people of pension age present in the household. Therefore, if a household comprises a lone parent or a couple with children aged 16 or over, the combined economic activity status of the household will be affected by whether those children are employed, unemployed or economically inactive, as well as by the economic activity status of their parent(s). If the lone parent or couple is employed but their children aged 16 or over are not employed, the household will be classified as a 'mixed' household (containing both working and non-working adults). Similarly, if an employed working-age person or couple lives with a pensioner who is not employed, the household will be classified as a mixed household.

People, families and households with unknown economic activity status
In some households that participate in the LFS, one or more household members have unknown economic activity status, either because they refuse to participate in the survey, or because they are absent and other household members feel unable to answer by proxy. These non-respondents (IOUTCOME = 3) are included in the household matrix and in the derivation of family and household type variables. In the person-level datasets, people with unknown economic activity status are zero-weighted. In the household datasets, they are given the same weight as the other members of their household (see Section 4).

Where the economic activity status of one or more household members is unknown, the combined economic activity status of the family and household will also be unknown. If no adjustments are made for this, the numbers of people, families and households in the various economic activity status categories may be underestimated. An investigation has been made into the effect that the treatment of households with unknown economic activity status has, particularly on estimates of workless households, and to assess different methods of adjustment. This showed that the characteristics of households with unknown economic activity status are not similar to those of households with known economic activity status overall, but they are similar within each household type category.

ONS has developed a method for adjusting estimates of workless households to compensate for households with unknown economic activity status. The method is demonstrated in Section 9 (see Example 10). It involves dividing all households according to household type (combining together some small, similar categories) and, within each household type category, allocating the 'unknown' households as workless or not in the same proportions as the households with known economic activity status. This method is used to adjust the UK figures for working, mixed and workless households (and working-age people and children in such households) that are published in the 'Work and worklessness among households' First Release[7]. The adjustment method operates at the aggregate level, and is used only for adjusting overall levels of working, mixed and workless households (and working-age people and children in such households). It is not suitable for producing adjusted estimates for subgroups (such as workless households by region, or people in workless households by ethnic group). This is because (a) there are a great many possible subgroups which could be defined, and applying this adjustment methodology would produce adjusted estimates which may not be consistent over different levels and hierarchies of subgroups; and (b) for smaller and/or more specialised subgroups, the numbers of sampled households in the smaller household type categories eventually become small enough to cause volatility in the resulting estimates.

There is no recommended procedure for adjusting figures for people or families to compensate for cases with unknown economic activity status. Figures should be presented unadjusted and percentages should be based on denominators that exclude cases with unknown economic activity status.

Earnings analyses

The household datasets cannot be used for analyses of earnings, either at the person, family or household levels. It would be difficult or impossible (and probably not worthwhile) to attempt to analyse earnings using the household datasets because:

(i) People who take part in the LFS are asked questions about their earnings in the first and last of their five quarterly interviews only. The earnings questions apply to employees only (not to self-employed people); the questions are explicitly voluntary and there is a relatively high level of non-response compared with other questions. This means that each quarterly LFS dataset includes earnings data from less than two-fifths of those respondents who are employees. There has to be an additional stage of weighting and imputation for the LFS earnings variables, to compensate for this.

(ii) Both weighting and imputation need to be done on the basis of person-level variables, including occupation, industry, and full-time/part time status. Either (a) the weighting factors would be different for people in the same household (which is not acceptable for the household datasets), or (b) 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 occupation, industry and full-/part-time categories. Achieving convergence to a solution without negative weighting factors is likely to be either difficult and costly, or impossible.

(iii) Even if a suitable weighting procedure could be developed, information about total family or household earnings would be incomplete for those families and households containing self-employed people, since the LFS earnings questions cover employees only.

Seasonality
Peoples' participation in the labour market is influenced to some extent by events and factors that occur at about the same time every year. For example, employment levels tend to be higher in the summer than in the winter due to seasonal work being available in agriculture, tourism and other industries. Also, there tends to be a large supply of student labour available in the summer to meet the increased demand for temporary workers.

The aggregate person-level LFS statistics in the monthly Labour Market Statistics First Release[4] are seasonally adjusted to remove the effects of seasonal factors. This makes it easier to identify trends and to make comparisons between different periods. The seasonal adjustment process involves estimating and removing seasonal effects. It also adjusts for any 'calendar effects' such as Easter bank holiday dates (which vary from one year to the next), and any differences in the number of working days in each three-month period that would otherwise obscure underlying trends. In order to be able to seasonally-adjust a series, it is necessary to have a long run of data on a consistent basis, covering each quarter of the year.

It would be difficult or impossible to produce seasonally adjusted LFS family and household statistics because the household datasets are produced for Q2 and Q4 only and there are discontinuities in some of the family and household variables. Therefore, when carrying out time-series analyses, or when calculating year-on-year changes, we recommend that comparisons should be made either between Q2 figures only, or between Q4 figures only.

Sampling and non-sampling errors
As with any sample survey, estimates produced from the LFS and APS are subject to a margin or uncertainty. For general information about sampling and non-sampling errors in the LFS, see Volume 1 of the LFS User Guide[2]. Information about sampling variability for the statistics published in the 'Work and worklessness among households' First Release is available on the National Statistics website[8].

Publication thresholds
In general, the smaller the subgroup whose size is being estimated, or from which an estimate is being derived, the less precise that estimate is likely to be. Before 2005, ONS did not publish LFS estimates of less than 10,000, because they are based on small samples and are considered to be unreliable. The policy on data suppression changed in 2005 (as a result of the Freedom of Information Act) and ONS no longer suppresses survey estimates on the grounds of poor reliability. However, the unreliability of estimates based on small samples has not changed. ONS have developed a methodology for producing sampling errors for the APS household dataset and estimates from this dataset are accompanied by their coefficient of variation.

In line with the National Statistics Code of Practice[9], estimates that are based on an un-weighted sample of less than three respondents must not be published[10]. This precaution is designed to protect the confidentiality of survey respondents, since figures based on very small samples are potentially 'disclosive' (it may be possible to identify an individual respondent). For guidance on disclosure control methods to ensure confidentiality of survey respondents in tabular outputs, contact the ONS Statistical Disclosure Control Centre: jane.longhurst@ons.gsi.gov.uk

Section 9: Producing family and household analyses: worked examples

This section demonstrates how to produce family and household analyses. The variables mentioned in the following examples are for the January-December 2007 APS household dataset and April-June 2008 LFS household dataset. Equivalent variables for previous LFS years may have different names and categories. For a description of these and other variables, see Volume 3 of the LFS User Guide[6].

For each of the worked examples, there is a table published in Microsoft Excel format on the National Statistics website[11].

SPSS users

The examples require user-defined family and household-level variables to be created from person-level variables. This section will give instructions for SPSS users (*show in italics*).

SuperCROSS users

Equivalent instructions for producing family and household analysis in SuperCROSS are presented in Section 10. In SuperCROSS, family-level variables have to be converted into person-level variables before they can be used to create person-level variables.

List of worked examples

1. Labour market summary for working-age people by parental status by sex
2. Labour market summary for working-age women by age of youngest dependent child
3. Labour market summary working-age couples with and without dependent children
4. Working-age households by combined economic status of household

Example 1: Labour market summary for working-age people by parental status by sex

This example demonstrates how to produce employment, unemployment and economic inactivity levels and rates for working-age people with and without dependent children. People with dependent children are grouped into: couple mothers, couple fathers, lone mothers and lone fathers. It is important to bear in mind that the sample size for lone fathers is relatively small and results for this subgroup will therefore have relatively wide margins of uncertainty, particularly if broken down by other variables. Users should therefore consider combining lone mothers and lone fathers into a single category (lone parents) instead.

Create a user-defined person-level variable (PARSTAT1) with the following categories:

- People without dependent children
- Married/Cohabiting parents with dependent children
- Lone parents with dependent children

People without dependent children are those who do not meet the criteria defined above for couple parents or lone parents. They should not be defined by filtering on FUTYPE6 not equal to 6,9,10,12,16,19 (nor by filtering on AYFL19 = 19, nor FDPCH19 = 0). Otherwise, the results will exclude people who are dependent children themselves. They will also exclude non-dependent children with siblings who are dependent children.

```
GET
  FILE='Z \AJ08spHC.sav'.

*PARENT RECODE.
COMPUTE PARSTAT1=0.
IF (any(futype6,6,9,16,19) and any(relhfu,1,2)) PARSTAT1=1.
IF (any(futype6,10,12) and relhfu=1) PARSTAT1=2.

VARIABLE LABELS PARSTAT1 'parental status'.
VARIABLE LEVEL PARSTAT1 (NOMINAL).

VALUE LABELS PARSTAT1
0 'People without dependent children'
1 'Married/Cohabiting parent with dependent children'
2 'Lone parent with dependent children'.
```

Filter on people of working-age:

```
*WORKAGE RECODE.
IF (sex=1  & range(age,16,64)) workage =1.
IF (sex=2  & range(age,16,59)) workage =1.

Select if workage=1.

Weight by PHHWT07.
*or Weight by PHHWT07a for APS household.
```

Tabulating, results for LFS April-June 2008 & APS January-December 2007 should match Example 1

```
* General Tables.
TABLES
  /FORMAT BLANK MISSING('.')
  /GBASE=CASES
  /FTOTAL= $t000002 "Total" $t000001 "Total"
  /TABLE=ilodefr + $t000002  BY PARSTAT1 > (STATISTICS) > ( sex > (STATISTICS) + $t000001  )
/TITLE  'Labour market summary for working-age people with and without dependent children'.
```

Example 2: Labour market summary for working-age women by age of youngest dependent child

Create a user-defined variable (MOTHER1) with the following categories:

- Mother: youngest dependent child under 5
- Mother: youngest dependent child 5-10
- Mother: youngest dependent child 11-18

```
GET
  FILE='Z:\AJ08spHC.sav'.

*MOTHERS  RECODE.
IF sex=2 and any(relhfu,1,2) and any(futype6,6,9,12,16,19) and range(ayfl19,0,4) MOTHER1=1.
IF sex=2 and any(relhfu,1,2) and any(futype6,6,9,12,16,19) and range(ayfl19,5,10) MOTHER1=2.
IF sex=2 and any(relhfu,1,2) and any(futype6,6,9,12,16,19) and range(ayfl19,11,18) MOTHER1=3.

VARIABLE LABELS MOTHER1 'mothers'.
VARIABLE LEVEL MOTHER1 (NOMINAL).
VALUE LABELS MOTHER1
        1 'Mother with youngest dependent child under 5'
        2 'Mother with youngest dependent child 5-10'
        3 'Mother with youngest dependent child 11-18'.

*WORKAGE WOMEN RECODE.
IF (sex=2  & range(age,16,59)) workagew =1.

Select if workagew=1.

Weight by PHHWT07.
*or Weight by PHHWT07a for APS household.

* General Tables.
TABLES
  /FORMAT BLANK MISSING('.')
  /GBASE=CASES
  /FTOTAL= $t000002 "Total" $t000001 "Total"
  /TABLE=ilodefr + $t000002  BY MOTHER1 > (STATISTICS) + $t000001
  /TITLE  'Labour market summary for working-age women by age of youngest dependent child'.
```

Results for LFS April-June 2008 & APS January-December 2007 should match Example 2

Example 3: Labour market summary for working-age couples with and without dependent children

This example demonstrates how to analyse the joint economic activity status of married/cohabiting couples with and without dependent children. It produces figures for couples in which both partners are employed, one partner is employed, and neither partner is employed. The method involves running a pair of family-level analyses in which HEAHEAD and HEAWIFE are cross-tabulated. A pair of user-defined family-level variables (HEADAGE and WIFEAGE) is used to select couples where both partners are of working age. This is done by adding the value of a variable for the partner to the head of household so they can be filtered together.

```
GET
  FILE='Z:\AJ08spHC.sav'.

*WORKAGE RECODE.
IF (sex=1  & range(age,16,64)) workage =1.
IF (sex=2  & range(age,16,59)) workage =1.

***Compute working age indicator for the head of the household

Compute HEADAGE=-9.
IF(WORKAGE =1 AND relhfu=1) HEADAGE=1.

VARIABLE LABEL HEADAGE 'Head working age indicator'.
VALUE LABELS HEADAGE 1 'Head of Working-Age' .
VARIABLE LEVEL HEADAGE(NOMINAL).
EXECUTE.
```

Derive a new marker WIFEAGE for head of working-age

```
SORT CASES BY
  fuserial (A) relhfu (D) .

DO IF relhfu=lag(relhfu)-1.
COMPUTE test=1.
END IF.
EXECUTE .

DO IF relhfu=1.
DO IF relhfu=lag(relhfu)-1.
DO IF (lag(fuserial) = fuserial).
COMPUTE wifeage=lag(workage).
END IF.
END IF.
END IF.
EXECUTE .
```

Filter where both the head and wife is of working age

```
Select if Headage=1 & Wifeage=1.

Weight by PHHWT07.
*or Weight by PHHWT07a for APS household.

execute.

* General Tables.
TABLES
  /FORMAT BLANK MISSING('.')
  /GBASE=CASES
  /FTOTAL= $t000001 "Total" $t000002 "Total"
```

```
/TABLE=heahead + $t000001  BY heawife > (STATISTICS) + $t000002  .
```

Results for LFS April-June 2008 & APS January-December 2007 should match Example 3

Similar examples using:
FTPTHEAD and FTPTWIFE identify whether the head and wife/partner of head of family unit are in full-time employment, part-time employment, not in employment or have unknown economic activity status. These two variables can then be used to select couples where both partners work full-time, one works full-time and the other part-time, both work part-time, and so on. The first step is to create a user-defined person-level variable (FTPTSTAT) that identifies whether people are in full-time employment, part-time employment, not in employment or have unknown economic activity status.

DISHEAD and DISWIFE can be created to analyse the joint disability status of husbands and wives or cohabiting partners. Both variables are based on, and have the same categories as, the variable DISCURR (disability status). In the LFS, working-age people are asked the health questions in each of their five quarterly interviews but people of pension age are asked the health questions in their first interview only. This means that, in any given quarter, DISCURR covers all working-age respondents but only a fifth of respondents of pension age. The variables DISHEAD and DISWIFE are therefore designed to cover working-age people only.

ETHHEAD and ETHWIFE can be created to analyse the joint ethnicity of husbands and wives or cohabiting partners. Both variables are based on, and have the same categories as, the variable ETHCEN6 (ethnic group). Sample sizes for couples with certain combinations of ethnicity will be very small. Users should therefore consider combining some of the smaller categories together when using these variables.

Example 4: Working-age households by combined economic status of household
This example demonstrates how to produce figures for ‘working’, ‘mixed’ and ‘workless’ households. When using the LFS household they can be adjusted to compensate for households with unknown economic activity status. The adjustment method involves dividing all households according to household type (combining together some small, similar categories) and, within each household type category, allocating the 'unknown' households as workless or not in the same proportions as households with known economic activity status. This method is used to produce the statistics on working, mixed and workless households that are published in the ‘Work and worklessness among households’ First Release[7].

The adjustment method described below operates at the aggregate level and should be used only to produce adjusted figures for overall levels of working, mixed and workless households. It is not suitable for producing adjusted figures for subgroups (such as workless households by region). This is because the method may produce figures for different categories that are inconsistent when aggregated and because, for smaller and/or more specialised subgroups, the numbers of sampled households in the smaller household type categories eventually become small enough to cause volatility in the resulting estimates. Estimates of workless households by type or size of household can be adjusted, using the method described in **Annex A**.

Estimates produced using the APS household data do not need to be adjusted.

Producing unadjusted estimates

```
GET
  FILE='Z:\AJ08spHC.sav'.

*HHOLD TYPE RECODE.
recode HHTYPE6
(1=1) (2=2) (3,4=3) (5=4) (6,7=5) (8,9=6) (10,11=7) (12 thru 14=8) (15=9) (16=10) (17=11)
(18 thru 20=12) (21 thru 24=13) (25=14) (26=15)
into HHTYPE15.

Variable label HHTYPE15 'Household types condensed '.
Variable level HHTYPE15(nominal).

Value labels HHTYPE15
```

```
1 'One person'
2 'Two or more persons, all different family units'
3 'Couple, no children, no other family units'
4 'Couple, no children, other family units'
5 'Couple, all dependent children, no others'
6 'Couple, dependent  & non-dependent children, no others'
7 'Couple, all non-dependent children, no others'
8 'Couple, dependent & non-dependent children, others'
9 'Lone parent, all dependent children, no others'
10 'Lone parent, dependent & non-dependent children, no others'
11 'Lone parent, all non-dependent children, no others'
12 'Lone parent, dependent or non-dependent children, others'
13 'Two or more family units, depdnt/non-depdnt/no children'
14 'Same sex couple with/without others'
15 'Civil partners, with/without others'.


*COMBINED ECONOMIC RECODE.
recode HEACOMB
(1=1)
(2 thru 4=2)
(5 thru 7=3)
(else=4)
into HEACOMBR.

Variable label HEACOMBR 'working, mixed and workless households'.
Variable level HEACOMBR(nominal).

Value labels HEACOMBR
1 'Working households'
2 'Mixed households'
3 'Workless households'
4 'Household with unknown economic activity status'.

*Restrict to households with at least one person of working age.
*To count number of household filter on household reference person.

Select if HNWKAGE>0 and relhrp6=0.

Weight by PHHWT07.
*or Weight by PHHWT07a for APS household.

* General Tables.
TABLES
  /FORMAT BLANK MISSING('.')
  /GBASE=CASES
  /FTOTAL= $t000001 "Total" $t000002 "Total"
  /TABLE=HHTYPE15 + $t000001  BY HEACOMBR > (STATISTICS) + $t000002
  /TITLE  'Working-age households by combined economic activity status of household'.
```

Adjusting for households with unknown economic activity status: for LFS household only

For each household type category, adjust the figures for working households as follows:

$$\frac{\text{unadjusted working households multiplied by total households}}{\text{(total households minus households with unknown economic status).}}$$

For each household type category, adjust the figures for mixed households as follows:

$$\frac{\text{unadjusted mixed households multiplied by total households}}{\text{(total households minus households with unknown economic status).}}$$

For each household type category, adjust the figures for workless households as follows:

unadjusted workless households multiplied by total households
(total households minus households with unknown economic status).

Results for LFS April-June 2008 & APS January-December 2007 should match Example 42

Variations on example 4: People in workless households, children in workless households

People in working-age households by economic status of the household.
as Example 4, but remove filter by household reference person.
Select of hnwkage>0.

Children in working-age households by economic status of the household.
as Example 4 but add a filter for children aged 0 to 15
Select if range(age,0,15) & hnwkage>0.

Section 10: Creating family and household-level variables in SuperCROSS

This section gives instructions for SuperCROSS users on how to:

- convert family-level variables into person-level variables;
- create user-defined family and household-level variables.

It is important to bear in mind that the household datasets include individuals for whom no information about the characteristic in question is held. It is therefore advisable to include the category 'unknown' when creating user-defined variables and to take the level of 'unknowns' into account when interpreting the results.

1. Converting family-level variables into person-level variables

FUTYPE6p

Open the 'Define Fields' box
Select 'Multiple'
At 'Field Name:' type 'FUTYPE6p'
At 'Create Field For:' select 'Person'
At 'From Associated:' select 'Family'
From the three radio buttons, select 'Copy'
At 'Field to Copy:' select 'FUTYPE6'
Click 'OK'
Recode FUTYPE6p, ignoring the categories 'No matches' and 'Multiple matches'.

AYFL19p

Open the 'Define Fields' box
Select 'Multiple'
At 'Field Name:' type 'AYFL19p'
At 'Create Field For:' select 'Person'
At 'From Associated:' select 'Family'
From the three radio buttons, select 'Copy'
At 'Field to Copy:' select 'AYFL19'
Click 'OK'
Recode AYFL19p, ignoring the categories 'No matches' and 'Multiple matches'.

2. Creating family-level variables for analysing the joint characteristics of couples
To analyse the joint characteristics of husbands and wives or cohabiting partners, create a pair of user-defined variables *at the family unit level*, one holding the characteristic of the head of family unit and the other holding the characteristic of the wife/partner of the head of family unit. These user-defined variables can then be cross-tabulated to give the joint distribution of husbands and wives or cohabiting partners. They can also be used as filter variables to select couples with particular joint characteristics.

HEADAGE and WIFEAGE
This example demonstrates how to create a pair of user-defined variables (HEADAGE and WIFEAGE) that identify whether the head of family unit and the wife/partner of the head of family unit are of working age. These variables can then be used to select couples where one partner is of working age, both partners are of working age, or neither partner is of working age. The first step is to create a user-defined person-level variable (WORKAGE) that identifies whether people are of working age or not. The variable WRKAGE should not be used for this purpose because, in the household datasets, 'non-respondents' (IOUTCOME = 3) are coded as 'unknown' in WRKAGE. There is no need to include the category 'unknown' when creating WORKAGE, because the datasets do not include any people with unknown age or sex.

Create a user-defined variable (WORKAGE) with the following categories:

1. Working-age: ((SEX = 1 and AGE = 16-64) or (SEX = 2 and AGE = 16-59))
2. Pension age: ((SEX = 1 and AGE = 65+) or (SEX = 2 and AGE = 60+))
3. Under 16: (AGE <16)

To create HEADAGE (whether head of family unit is of working age):

Open the ‘Define Fields’ box
Select ‘Multiple’
At ‘Field Name:’ type ‘HEADAGE’
At ‘Create Field For:’ select ‘Family’
At ‘From Associated:’ select ‘Person’
From the three radio buttons, select ‘Copy’
At ‘Field to Copy:’ select ‘WORKAGE’
At ‘Based on Field:’ select RELHFU = 1
Click ‘OK’
Recode HEADAGE, ignoring the categories ‘No matches’ and ‘Multiple matches’.

To create WIFEAGE (whether wife/partner of head of family unit is of working age), follow the procedure for HEADAGE but replace RELHFU = 1 with RELHFU = 2.

Combined economic status by country of birth of household members

HNUKB (number of UK-born adults in household)

Open the ‘Define Fields’ box
Select ‘Multiple’
At ‘Create field for’ select ‘Household level’
At ‘From associated’ select ‘Person level’
From the three radio buttons on the left, select ‘Count’
Set ‘Maximum’ to 20
At ‘Based on criteria’ select AGE >=16 and CRYOX7 = 926 (UK/GB).

HNOSB (number of overseas-born adults in household)

Open the ‘Define Fields’ box
Select ‘Multiple’
At ‘Create field for’ select ‘Household level’
At ‘From associated’ select ‘Person level’
From the three radio buttons on the left, select ‘Count’
Set ‘Maximum’ to 20
At ‘Based on criteria’ select AGE >=16 and CRYOX7 > 0 and CRYOX7 not equal to 926 (UK/GB).

HNCOBDK (number of adults in household with unknown country of birth)

Open the ‘Define Fields’ box
Select ‘Multiple’
At ‘Create field for’ select ‘Household level’
At ‘From associated’ select ‘Person level’
From the three radio buttons on the left, select ‘Count’
Set ‘Maximum’ to 20
At ‘Based on criteria’ select AGE >=16 and CRYOX7 = -8, -9.

To create **HCOB** (country of birth of adult household members):

Open the ‘Define Fields’ box
Select ‘Single’
At ‘Create field for’ select ‘Household level’
Create the following categories:

1. Households where all of the adults are UK-born

HNOSB = 0 and HNCOBDK = 0

2. Households containing both UK-born and overseas-born adults

HNUKB > 0 and HNOSB >0 and HNCOBDK = 0

3. Households where all of the adults are overseas-born

HNUKB = 0 and HNOSB > 0 and HNCOBDK = 0

4. Households where the country of birth of one or more adults is unknown

HNCOBDK > 0.

Section 11: Publication of family and household statistics

The LFS collects a wide range of information about people, families and households. ONS publishes statistical tables showing key family and household series regularly and occasional journal articles presenting more in-depth analyses on specific topics. This section gives a list of publications that include LFS family and household statistics, and explains how to obtain LFS results that are not published regularly.

Work and worklessness among households First Release

A statistical First Release showing the latest figures on working-age households by household economic activity status. The First Release includes statistics on working, mixed and workless households (and working-age people and children living in such households), by household type, region, and ethnicity. The Release also includes employment rates for working-age people by parental status (couple mothers, couple fathers, lone parents and people without dependent children). Publication takes place in August of each year.
http://www.statistics.gov.uk/statbase/Product.asp?vlnk=8552

Work and worklessness among households: time-series tables

A set of tables (in Microsoft Excel format) showing a time-series of the statistics in the 'Work and worklessness among households' First Release. The tables include statistics on working, mixed and workless households (and working-age people and children living in such households), by household type, region and ethnicity. They also include employment rates for working-age people by parental status; and information about sampling variability for key series. An update of the tables takes place each August to coincide with the release.
http://www.statistics.gov.uk/statbase/Product.asp?vlnk=14977

Economic and Labour Market Review

A monthly journal that replaced Labour Market Trends and Economic Trends, in January 2007. Table 6.01 of the online edition gives statistics on working and workless households, workless lone parent households, and working age people and children in workless households. The table is published in Microsoft Excel format and is updated each August:
http://www.statistics.gov.uk/elmr/

Social Trends

An annual publication that draws together social and economic data from a wide range of government departments and other organisations. The chapters on households and families, and on the labour market, include figures produced from the LFS household datasets. Published in hard copy and online (in PDF format with links to tables in Microsoft Excel format):
http://www.statistics.gov.uk/StatBase/Product.asp?vlnk=5748&Pos=&ColRank=1&Rank=422

Journal articles

The journal articles listed below present analyses of LFS family and household data and/or give technical information about the development of the household datasets. Labour Market Trends articles for the period July 2001 to December 2006 are available online in PDF format:
http://www.statistics.gov.uk/statbase/Product.asp?vlnk=550

Earlier editions of Labour Market Trends are available in hard copy only. Please contact:
ons@palgrave.com

- 'Families and work', Labour Market Trends (July 2005): http://www.statistics.gov.uk/cci/article.asp?ID=1192&Pos=9&ColRank=1&Rank=1

- 'Workless households: results from the spring 2004 LFS', Labour Market Trends (November 2004): http://www.statistics.gov.uk/cci/article.asp?ID=1001&Pos=10&ColRank=1&Rank=1

- 'LFS household data: spring 2000 analyses', Labour Market Trends (January 2001)

- 'LFS household data: spring 1999 analyses', Labour Market Trends (January 2000)

- 'Update on Labour Force Survey household datasets', Labour Market Trends (May 1999)

- 'Additions to Labour Force Survey household tables', Labour Market Trends (September 1998)

- 'Production of household datasets from the Labour Force Survey', Labour Market Trends (August 1998)

- 'Analysis of household data from the Labour Force Survey', Labour Market Trends (August 1998)

- 'Workless households, unemployment and economic inactivity', Labour Market Trends (September 1997)

- 'Economic activity of working-age households', Labour Market Trends (September 1997)

- 'Household and family data from the Labour Force Survey: recent improvements in approach', Labour Market Trends (June 1997)

Future articles
Labour Market Trends was replaced by Economic and Labour Market Review (ELMR) in January 2007. Future articles about LFS family and household statistics will be published in ELMR, which is available in hard copy and online:
http://www.statistics.gov.uk/StatBase/Product.asp?vlnk=14692&Pos=&ColRank=1&Rank=144

How to obtain unpublished LFS data
Customers wanting access to LFS data that are not published can either:

- commission a made-to-order tabulation of LFS data (supplied in Excel format);

- receive copies of the LFS datasets on a regular or one off basis, for use on a computer within their own organisation. The datasets are available in SPSS, SAS and SuperCROSS formats;

- access the LFS datasets via the SuperCROSS service. SuperCROSS is a tabulation software package used by ONS. Training is provided by ONS on how to use the system, which is accessed online by external customers.

For information about the above options, please contact the LFS Dataservice:
LFS.Dataservice@ons.gsi.gov.uk
Tel: 01633 655732.

The Data Archive, based at the University of Essex, holds copies of all the LFS datasets. Customers from academic institutions can access the datasets at specially agreed rates. For details, please contact the Data Archive:
Archive-Userservices@essex.ac.uk
www.data-archive.ac.uk
Tel. 01206 872017 / 873574

Section 12: Improving family and household statistics

This section describes recent and planned dvelopments aimed at improving the quality of family and household statistics in terms of their relevance, accuracy, comparability, accessibility, clarity and timeliness.

Re-weighting the household datasets

In recent years, figures produced from the LFS household datasets have been inconsistent with the aggregate (person-level) LFS series published in the monthly Labour Market Statistics First Release[4] in that the household datasets were weighted to population figures published in February and March 2003, while the aggregate LFS series incorporated more recent population figures using an interim adjustment re-weighting methodology[5]. The Q2 (April to June) 2008 LFS household dataset is weighted to the latest population figures (published in August and September 2007) and the datasets for Q2 of 1997 to 2007 have recently been re-weighted to incorporate revised population figures for that period (also published in August and September 2007). In that respect, the household datasets for Q2 1997 to 2008 are now consistent with the latest population figures and with the aggregate LFS series.

The household datasets for 1992 to 1996 are currently still weighted to population figures published by ONS in February and March 2003. This is because re-weighting the full back-series of micro-datasets is a lengthy process. The datasets are therefore re-weighted in order of priority. In line with customers' requirements, the person-level datasets have been re-weighted before the household datasets, and the household datasets for 1997 to 2007 have been re-weighted before the 1992 to 1996 datasets. ONS hopes to release re-weighted household datasets for 1992 to 1996 by the end of 2008.

Annex A: Adjusting estimates of workless households by type and size of household

As explained in Section 8, the LFS household datasets include people, families and households with unknown economic activity status and if no adjustments are made for this, the numbers of working, mixed and workless households (and of people living in such households) may be underestimated. The general adjustment method (demonstrated in Section 9, Example 10) should be used only to adjust figures for the UK as a whole. It is not suitable for adjusting figures for subgroups, such as workless households (or people in workless households) by region or by ethnicity. This is because the general adjustment method may produce figures for different subgroups that are inconsistent with adjusted UK totals when aggregated and because, for smaller subgroups, the numbers of sampled households eventually becomes small enough to cause volatility in the adjusted results. However, figures for workless households (and people in workless households) broken down by type or size of household can be adjusted, using the method described below.

Adjusting figures for household type categories

The recommended method for adjusting estimates of working, mixed and workless households for specific types of household (for the individual categories of HHTYPE6) is as follows.

1. Produce unadjusted and adjusted estimates of working, mixed and workless households by type of household using the method described Section 9 (see Example 4).

2. For each category of the user-defined variable HHTYPE15, calculate a 'workless household adjustment factor' by dividing the adjusted number of workless households by the unadjusted number of workless households. This is demonstrated in Example 4[11]

3. If an adjusted estimate is required for a specific household type that is grouped together with other household types in the user-defined variable HHTYPE15, apply the adjustment factor for the group in which that type of household falls. For example, to adjust figures for 'cohabiting couple households with dependent children only and no other family units' (HHTYPE6 code 7) use the adjustment factor for 'couple households with dependent children only, no other family units' (HHTYPE15 code 5).

Adjusting figures by size of household

A different method is required for adjusting subgroups defined in terms of the number of adults (people aged 16 or over) in the household. The incidence of households with unknown economic status varies considerably between different types of household and (closely related to this) between households with different numbers of adults. This is to be expected, since households are assigned to the 'unknown' economic activity status category if someone in the household has unknown economic activity status either because they are present and refuse to participate in the survey or because they are absent and other household members feel unable to give a proxy response. It is not possible for a household containing only one adult to have unknown economic activity status, because if that person was absent or refused to participate in the survey, then the whole household would be non-responding and would not be included in the survey.

The recommended method is to take an adjustment factor of exactly one for households with one adult and to calculate an adjustment factor for households with more than one adult in such a way that it produces the correct adjusted total number of workless households. This is done by subtracting from the adjusted figure for all workless households the adjusted figure for workless one-adult households, then dividing the result by: unadjusted all workless households minus unadjusted workless one-adult households. Note that 'one-person household' is the only household type category that invariably contains only one person aged 16 or over. Lone parent households may contain dependent and/or non-dependent children aged 16 or over. Therefore, 'one-person household' is the only category for which it is impossible to find households with unknown economic activity status.

If a subgroup is defined in terms of the number of *working-age* adults in the household, the situation is more complicated, because there are some working-age households with only one working-age person that contain one or more people of state pension age. It is possible for such households to have unknown economic activity status, so it is not appropriate to assign to all workless households with one working-age person an adjustment factor of one. Households with one person of working age need to be divided into (a) those with no-one of pension age, which are given an adjustment factor of one, and (b) those with one or more people of pension age, which are given the adjustment factor for households with more than one adult, as calculated above.

NOTE: Adjustments are not required for the APS household datasets and these are recommended for household level analysis over the LFS as they have a bigger sample size and imputation takes place for individuals with unknown economic status.

Annex B: References

1. 'Helping people from workless households into work', National Audit Office (July 2007): http://www.nao.org.uk/publications/nao_reports/06-07/0607609.pdf

2. LFS User Guide Volume 1 (Background and methodology): http://www.statistics.gov.uk/statbase/Product.asp?vlnk=1537

3. 'An integrated method for weighting persons and families', Lemaitre, G. & Dufour, J. (1987), Survey Methodology, Vol.13, pp199-207.

4. Labour Market Statistics First Release: http://www.statistics.gov.uk/statbase/Product.asp?vlnk=1944

5. 'Interim LFS estimates consistent with the 2001 Census', Labour Market Trends (December 2002): http://www.statistics.gov.uk/cci/article.asp?ID=360&Pos=2&ColRank=1&Rank=1

6. LFS User Guide Volume 3 (Details of LFS variables): http://www.statistics.gov.uk/statbase/Product.asp?vlnk=1537

7. Work and worklessness among households First Release: http://www.statistics.gov.uk/statbase/Product.asp?vlnk=8552

8. Sampling variabilities for key family and household series: http://www.statistics.gov.uk/statbase/Product.asp?vlnk=14977.

9. National Statistics Code of Practice: http://www.statistics.gov.uk/about/national_statistics/cop/default.asp

10. Data Access and Confidentiality Protocol: http://www.statistics.gov.uk/about/national_statistics/cop/protocols_published.asp

11. Tables for the worked examples described in Section 9 and Annex 1 of this user guide: http://www.statistics.gov.uk/downloads/theme_labour/All-examples-LFS-APS.xls