

Information paper

Quality and Methodology Information

General details

Title of output:	Families and households
Designation:	National Statistics
Geographic coverage:	UK
Date of last SQR or QMI*:	November 2015
Contact details:	Demographic Analysis Unit at families@ons.gsi.gov.uk or +44 (0)1329 447890.

Executive summary

This overview pulls together important qualitative information and summarises the methods used to compile the output.

This report relates to estimates of [families and households](#) in the UK. These estimates were first published in 2011, and are now published annually (based on 3 months of data in each year). Estimates currently published show figures for 1996 to 2015.

The Families and Households release presents time series of estimates of the number of families by type, people in families by family type and children in families by type. Types of family include married couple families, cohabiting couple families and lone parent families. Tables on household size, household types and people in different household types are also provided. These include estimates of the numbers of people living alone and multi-family households.

The Labour Force Survey (LFS), a social survey of the private resident population in the UK, is used to provide estimates about the UK's families and households. The estimates are published on our website around 4 months after the end of the April to June reference period. Further information about the [questionnaire, methodology and topics covered in the LFS](#) are available on our website.

Quality measures such as; confidence intervals, coefficient of variation and an indication of whether a change is statistically significant or not, is provided alongside the estimates and within the statistical bulletin. These quality measures give users an indication of the robustness of the estimates and an indication of what the estimates should be used for.

Family and household estimates are used to analyse families by family type and to understand changing household structures and sizes. More information on the uses and users of the estimates can be found in the information note [Families and households statistics: User experience](#).

* Quality and Methodology Information (QMI) reports replaced Summary Quality Reports (SQR) from April 2011.

This document contains the following sections:

- Output quality
- About the output
- How the output is created
- Validation and quality assurance
- Concepts and definitions
- Other information, relating to quality trade-offs and user needs
- Sources for further information or advice

Output quality

This document provides a range of information that describes the quality of the output and details any points that should be noted when using the output.

We have developed [guidelines for measuring statistical quality](#) based on the 5 European Statistical System (ESS) quality dimensions. This document addresses these quality dimensions and other important quality characteristics, including:

- relevance
- timeliness and punctuality
- coherence and comparability
- accuracy
- output quality trade-offs
- assessment of user needs and perceptions
- accessibility and clarity

About the output

Relevance

(The degree to which the statistical outputs meet users' needs.)

The Labour Force Survey (LFS) provides a time series of data on families and households back to 1996. It covers people in private households, NHS accommodation and students in halls of residence whose parents live in the UK. Such students are included through proxy interviews with their parents. People who live in other communal establishments, such as prisons and nursing homes, are not covered. LFS interviewers attempt to interview all adults resident at a selected address. There is further information about the LFS in the "Validation and quality assurance" section.

The families and households tables published are:

- Table 1: Families by family type and presence of children
- Table 2: People in families by family type and presence of children
- Table 3: Families with dependent children by family type and number of dependent children
- Table 4: Dependent children in families by family type
- Table 5: Households by size
- Table 6: People living alone by age and sex
- Table 7: Households by type of household and family
- Table 8: People in households by type of household and family

These tables are updated annually. The choice to publish these particular tables is based on analysis of ad hoc requests by users of families and households data during 2010. It is not possible to provide tables to meet all user needs, but the tables published aim to answer the most fundamental and common questions about the number of families in the UK by family type and the number of people in such families, including children. The tables also answer basic questions about the number of households in the UK, what types of family are living in these households and the number of people living in these households. Further information about the strengths and limitations of other sources of data that were considered when developing this output are in the "Comparability and coherence" section. Information on user needs can be found in the "Assessment of user needs and perceptions" section.

Timeliness and punctuality

(Timeliness refers to the lapse of time between publication and the period to which the data refer. Punctuality refers to the gap between planned and actual publication dates.)

Before Families and Households was first published, in April 2011, estimates about the UK's families and households were published in an ad hoc manner, in publications such as [Focus on Families](#) and [Social Trends](#), which are no longer published. The Families and Households release provides important estimates of families and people in families, as well as households and people in households. They are updated annually, around 4 months after the end of the April to June reference period.

The Labour Force Survey (LFS) datasets are reweighted in most years to take account of the latest available population estimates and projections. This means that there are small revisions to previously published estimates. We publish a table to show the [impact of reweighting](#) of the LFS datasets on the families and households estimates. Results from the 2011 Census have now been incorporated into weighting for this publication. These were first published on [28 January 2015](#) and the estimates were revised back to 2002.

If further changes to the estimates are proposed, to methods or classifications for example, we will consult users via our website. We will give users a time-frame in which you can share your views about the proposed changes with us.

There are more details on related releases on the [GOV.UK release calendar](#), which provides 12 months advance notice of release dates. We will also notify you of any changes to the pre-announced release schedule and provide the reasons for the change will be explained, as set out in the [Code of Practice for Official Statistics](#).

How the output is created

The original objective for the output is to provide summary statistics covering the most basic and commonly asked questions about families and households. This includes the number of families, cohabiting couples and families with children in the UK.

The data are produced using the Labour Force Survey (LFS). The LFS is a social survey of the private resident population in the UK. Its purpose is to provide information on the UK labour market, but it includes estimates on a variety of other variables, including family types and household types.

Full details of the background and methodology used to produce the LFS can be found in [Volume 1 of the LFS user guide](#). [Quality and Methodology Information for the LFS](#) is also available. For information on response rates and other quality-related issues for the LFS, including breakdowns of response region and by question, can be found in the [Labour Force Survey quality measures](#).

The LFS household dataset is used for the Families and Households publication. The household dataset differs from the person-level dataset in the way that the weighting is applied to the sample. The household dataset makes sure that the weight given to each member of the same household is the same. The design of the household dataset is such that it gives the best estimate of families and households, as well as people in families and households. The person-level dataset, which is more commonly used for information on the labour market, includes only individual weights, which usually differ between members of the same household.

Estimates based on sample sizes of less than 3 are suppressed in published tables, but this has no impact on the estimates for most years.

Validation and quality assurance

Accuracy

(The degree of closeness between an estimate and the true value.)

The Labour Force Survey (LFS) is a social survey of the private resident population in the UK. Each quarter's (3 month) LFS sample of 41,000 households is made up from 5 "waves"; each wave consists of approximately 8,000 households. Each wave is interviewed in 5 successive quarters, so that in any 1 quarter, 1 wave will be receiving their first interview, 1 wave their second and so on, and 1 wave receiving their fifth and final interview. There is an 80% overlap in the samples for each successive quarter. The samples for the sixth quarter and the first quarter have no common elements. The LFS household datasets are produced twice a year (for the April to June quarter, and the October to December quarter). [Families and households](#) uses the April to June quarter each year.

As estimates of the UK's families and households are based on the LFS, all estimates produced are subject to sampling variability. This is because the sample selected is only 1 of a large number of possible population samples that could have been drawn from the population. Estimates produced from a sample survey vary according to the specific characteristics of the respondents that have been sampled. Some inferences can be made as to the characteristics of non-respondents, and whether non-response has an impact on the quality of the survey results.

Once LFS household datasets are received, quality checks are made on the most commonly used family and household type variables by tabulating 1 variable against another. The checks help to verify that:

- lone parents do not have a partner in the household
- no one aged under 16 are married
- children are not heads of families
- there are no large year-on-year changes

If errors are found, these are referred back to our Social Survey Division for investigation and correction before estimates about families and households are published.

Confidence intervals and coefficient of variation

Confidence intervals are commonly used as indicators of the extent to which the estimate, based on a sample, may differ from the true population value; the larger the confidence interval, the less precise the estimate is. Confidence intervals are shown alongside the estimates in the reference tables. The estimates themselves have been colour coded using the coefficient of variation to indicate statistical robustness as follows:

Coefficient of variation	Description	Format
Coefficient of variation $\leq 5\%$	Estimates are considered precise	None
Coefficient of variation $>5\%$ and $\leq 10\%$	Estimates are considered reasonably precise	Light blue
Coefficient of variation >10 and $\leq 20\%$	Estimates are considered acceptable	Blue
Coefficient of variation $> 20\%$	Estimates are considered unreliable for practical purposes	Dark blue

The coefficient of variation is the ratio of the standard error of the estimate to the estimated value itself; the smaller this value, the more precise the estimate.

Estimates with a coefficient of variation $> 20\%$ are considered unreliable for practical purposes. This means that they should not be used in isolation or used for making policy. They have been included in the tables for completeness (where removing the estimate would mean that the sum of the estimates in each category would not equal the total) and in some cases, because users are specifically interested in particular estimates (such as same-sex couples with dependent children).

Statistical significance

Changes in the estimates from the LFS from one period to the next may occur by chance. In other words, the change may be due to which individuals were selected to answer the survey, and may not represent any real changes in family and household structures.

Statistical tests can be used to determine whether any increases or decreases that we see in the estimates from the LFS are due to chance, or whether they are likely to represent real change. When comparing 2 estimates, a t test is performed which results in the calculation of a 95% confidence interval for the difference between these estimates. If this interval excludes the value 0, then we can conclude that the change is statistically significant and that the difference is very likely to be a real change in family or household structures and not as a result of sampling variation.

The t test divides the difference of the estimates by the square root of the sum of the squared standard errors. The resulting t value needs to be greater than 1.96 to be 95% certain that the estimates are different. The usual standard is to carry out these tests at the 5% level of statistical significance. This means that we would expect only 1 out of 20 statistically significant differences to have occurred purely by chance.

An indication of whether the change in an estimate over time is statistically significant is presented in the [Families and Households statistical bulletin](#). From the year 2015 onwards a t test has been used. Previously the method used checked whether the confidence intervals around the estimates overlap or not, in order to decide whether the change was statistically significant.

The main threats to the accuracy of the data are sampling error and non-sampling error, where non-sampling error includes coverage error, non-response error, measurement error and processing error. Many of the sources of non-sampling error are difficult to measure. Detailed response rates for all waves of the LFS survey are published in the [Performance and Quality Monitoring reports](#).

Coherence and Comparability

(Coherence is the degree to which data that are derived from different sources or methods, but refer to the same topic, are similar. Comparability is the degree to which data can be compared over time and domain, for example, geographic level.)

The data are produced using the Labour Force Survey (LFS), with a published comparable time series going back to 1996.

In 2012, we published [Comparing data sources on families and households](#), which analysed the differences between various administrative, survey and modelled sources of families and household data and the LFS, to explain the strengths and limitations of all sources. The document also provides guidance on the best source of data for different purposes.

Some common themes emerged from the analysis, such as lower estimates of 1 person households in the LFS compared with other non-survey sources, due to non-contact with 1 person households in surveys. Surveys also generally provided higher estimates of civil partners than registration data, for a variety of possible reasons, including people who have formed a civil partnership outside the UK, either before or after civil partnerships became legal in the UK in 2005. However, in general it was reassuring that the estimates are so similar to the other sources, given the different collection methods.

In 2013 we undertook some initial comparisons with the results of the 2011 Census for families and households. The main findings from this comparison are in the [Families and Households 2013 statistical bulletin](#).

Concepts and definitions

(Concepts and definitions describe the legislation governing the output, and a description of the classifications used in the output.)

The definitions used in the output are nationally agreed definitions and consistent with other sources of data (such as those used for the census). No classifications are used in the output. Definitions include:

- a family is a married, civil partnered or cohabiting couple with or without children, or a lone parent with at least 1 child; children may be dependent or non-dependent.
- dependent children are those living with their parent(s) and either
 - (a) aged under 16
 - (b) aged 16 to 18 in full-time education, excluding children aged 16 to 18 who have a spouse, partner or child living in the household
- non-dependent children are sometimes called adult children; they are those living with their parent(s), and either
 - (a) aged 19 or over

(b) aged 16 to 18 who are not in full-time education and have no spouse, partner or child living in the household

- for 1996 to 2010, a household is defined as a person living alone, or a group of people living at the same address who have the address as their only or main residence and either share 1 main meal a day or share living accommodation (or both)
- for 2011 onwards, a household is defined as 1 person living alone, or a group of people (not necessarily related) living at the same address who share cooking facilities and share a living room or sitting room or dining area; to measure the impact of the introduction of the new household definition we developed a set of questions to be asked by interviewers for each address they visited, overall only 2 addresses out of 2,188 productive cases (less than 0.1% of interviewed households) would change classification under the new household definition

Other information

Output quality trade-offs

(Trade-offs are the extent to which different dimensions of quality are balanced against each other.)

The data are produced using the Labour Force Survey (LFS). When deciding the best source of data to use for this publication, we considered using the LFS, the [Annual Population Survey \(APS\)](#), the [General Lifestyle Survey \(GLF\)](#) and the [Integrated Household Survey \(IHS\)](#).

The [APS](#) combines results from the LFS and the English, Welsh and Scottish Labour Force Survey boosts. As such, the APS provides similar information to the LFS and has a larger sample size, so is more commonly used for estimates for smaller geographical areas, such as local authorities, than the LFS. The APS is available annually from 2004 onwards.

The [GLF](#) is a multi-purpose continuous survey that we carry out, it collects information from people living in Great Britain on a range of topics including households, families, marriage and cohabitation. The survey ran continuously between 1971 and 2011, when it finished. It has a smaller sample size than the LFS.

The [IHS](#) is a composite survey combining questions asked on a number of our social surveys to produce a dataset of “core” variables, including those on families and households. In 2009 to 2010 the IHS was created from 6 of our surveys, including the APS. After April 2011, the number of surveys contributing to the IHS dataset fell, as 2 surveys could not continue asking the core questions. The GLF finished at the end of December 2011, and then the IHS comprised only of the Living Cost and Food Survey and APS. In February 2014, we announced the result of a consultation which stopped the core questions being asked in the Living Costs and Food Survey. This effectively ended the IHS as it only consisted of the APS. The IHS had a larger sample size than the LFS, but a shorter time series because data are only available from April 2009 onwards.

We looked at 4 main factors when deciding the best source of data:

Timeliness:

- the LFS is more timely than the APS, IHS and GLF, with the quarterly dataset being available 2 to 3 months after the April to June reference date.
- the IHS dataset is available around 7 months after the January to December reference date
- the equivalent APS dataset is available around 9 months after the January to December reference date
- the equivalent GLF dataset is available 2 years after the January to December reference date

Sample size:

- LFS has a sample size of 43,000 households
- APS 143,000 households
- IHS 153,000 households
- GLF around 8,000 households
- IHS and APS have a larger sample size than the LFS, however the estimates from these 3 sources are still very robust at the UK level

Time series:

- LFS has a longer time series than the APS and IHS, with comparable time series going back to 1996
- APS data are only available from 2004
- IHS data are only available from April 2009 onwards
- GLF data are available from 1971, but stop in 2011

Consistency:

- Data are available from the LFS, APS and IHS on a consistent basis across the UK
- the GLF does not cover Northern Ireland

These factors were evaluated, and on balance it was decided that the LFS was the best source because users were keen to have a long time series and the quality of estimates is robust at UK level. Further, estimates of families and households produced from interviews in the April to June quarter are thought to be representative of the rest of the year.

Assessment of user needs and perceptions

(The processes for finding out about uses and users, and their views on the statistical products.)

An understanding of families and households is crucial for those involved in planning and decision making. In addition, family and social relationships are one of the most important factors contributing to well-being.

The statistics are used by those who want to improve their understanding of the UK's families and households, including:

- policy makers
- MPs
- journalists
- charities
- the private sector
- students
- researchers and academics
- members of the general public

Our team, who produce these statistics, have received feedback from:

- a small selection of people who are known to be interested in families and households were consulted during the initial development of these statistics
- a wide range of users were contacted on the day of publication and asked for feedback
- the UK Statistics Authority, who asked people who use these statistics for their views during the assessment of the statistics
- presentations at conferences, such as the British Society for Population Studies Conference
- people who have contacted us directly

A summary of feedback from the people who use our statistics and our response to specific points is available in a [user experience document](#).

Sources for further information or advice

Accessibility and clarity

(Accessibility is the ease with which users are able to access the data, also reflecting the format that the data are available and the availability of supporting information. Clarity refers to the quality and sufficiency of the release details, illustrations and accompanying advice.)

Our recommended format for accessible content is a combination of HTML web pages for narrative, charts and graphs, with data being provided in usable formats such as CSV and Excel. Our website also offers users the option to download the narrative in PDF format. In some instances other software may be used, or may be available on request. Available formats for content published on our website, but not produced by the ONS, or referenced on our website but stored elsewhere, may vary. For further information please refer to the contact families@ons.gsi.gov.uk or +44 (0)1329 447890.

There is more information about access to the data in:

- [terms and conditions \(for data on the website\)](#)
- [copyright and reuse of published data](#)
- [pre-release access \(including conditions of access\)](#)
- [accessibility](#)

In addition to this Quality and Methodology Information, Basic Quality Information relevant to each release is available in the background notes of the [Families and Households statistical bulletin](#).