

A guide to using Scottish and UK datasets

A guide to using the Labour Force Survey (LFS) and the Annual Population Survey (APS)

The Labour Force Survey (LFS) began in 1973 as an annual survey, and became the quarterly labour force survey (QLFS) since March 1992. The questionnaire comprises a 'core' of questions which are included in every survey, together with some 'non-core' questions which vary from quarter to quarter. It is the largest household survey in the UK.

A Multi-stage stratified random sample is used. See the document "Labour Force Survey User Guide – [Volume 1: Background and Methodology](#)" for details.

It includes the topics:

- Employment and unemployment
- Economic inactivity
- Occupation
- Training
- Hours of work
- Personal characteristics

The Annual Population Survey (APS) is a combined statistical survey of households in Great Britain which is conducted quarterly by the Office for National Statistics (ONS). It combines individuals in waves one and five from four consecutive LFS quarters. The size of the total APS sample is approximately 500,000 people. The APS contains an enhanced range of variables than the LFS. In particular, more variables are provided on ethnic group, health and gender.

The survey has a panel element, in which the selected households are interviewed annually over four waves, in the same way as for the LFS boosts. They then leave the survey and are replaced by other households.

Multi-stage stratified random sample is used.

Survey years and sample sizes

Labour Force Survey (LFS)

Repeated cross-sectional study:

- Biennially from 1973-1983.
- Between 1984 and 1991 the survey was carried out annually and consisted of a quarterly survey conducted throughout the year and a 'boost' survey in the spring quarter.
- From 1992 quarterly data were made available, with a quarterly sample size approximately equivalent to that of the previous annual data.

- From December 1994, data gathering for Northern Ireland moved to a full quarterly cycle to match the rest of the country, so the QLFS then covered the whole of the UK.
- The LFS uses a rotational sampling design, whereby a household, once initially selected for interview, is retained in the sample for a total of five consecutive quarters. The interviews are scheduled to take place exactly 13 weeks apart, so that the fifth interview takes place one year on from the first. There is therefore an 80% overlap in the samples for each successive survey.

The LFS has around 44,000 responding households in every quarter. The LFS household specific response rates for the whole of the UK are now around 60% (including information for earlier waves rolled forward for refusals) down from around 70% during the start of 2000. In Scotland the response rate was 55.8% in the Q4 2013 LFS. LFS response rates and other quality-related issues are available in the LFS [Performance and Quality Monitoring Reports](#).

Annual Population Survey (APS)

- The first APS data were published for the period January to December 2004.
- Subsequently, APS data are published quarterly with each publication covering a year's data. That is, data for April 2004 to March 2005 was published in September 2005, data for July 2004 to June 2005 in December 2005, etc.
- In 2006, the structure of the Annual Population Survey reverted to the same as the Annual Local Area Labour Force Survey (that is waves 1 and 5 of the quarterly LFS). However, the name 'Annual Population Survey' has been retained and the data continue to be published four times a year (and all questions are now based on the complete database).

Useful websites

Office for National Statistics <http://www.ons.gov.uk/>. There is no separate part of the site that deals with the LFS data but there is a Theme Page called "Labour Market" (which can be found after clicking on the "Browse by Theme" tab on the left hand top corner). This page has a tab called "Data Tables" where labour market statistic tables are listed. Data on Scotland can be identified by the name of the file, for instance: "Regional Labour Market: HI11 - Headline Indicators for Scotland, April 2014 ([Excel sheet 2353Kb](#))".

Under the tab "Guidance and methodology" there is a category called "Specific Methodology". After clicking one gets directed to a page called "Labour Market Methodology" which includes valuable information, including [user guides](#).

The Labour Force Survey Performance and Quality Monitoring Reports (PQM) page contains [a list of all existing PQM reports](#) for each LFS from 2003 onward.

[This page](#) provides an overview of the wide variety of Labour Force Survey (LFS) quality measures which are available

Nomis <https://www.nomisweb.co.uk/>, the official labour market statistics web from ONS. Registration is not required. The site provides both summary statistics and detailed statistics with the possibility to download the data. Clicking on Advanced query in the “detailed statistics” box on the right top hand corner leads to the opening of a window where one can choose from a list of data sources, including the LFS. Clicking on “Annual Population Survey/Labour Force Survey” opens up a series of datasets. FFS Data are not seasonally adjusted and are only available from 1995 to 2005. For more recent data are available from the APS which includes LFS data.

UK Data Service <http://www.ukdataservice.ac.uk/> and type, using quotation marks, in the search bar “Labour Force Survey” or “Annual Population Survey”. The first on the list leads you to a general LFS/APS page, with documentation and in depth information, including the User Guides which have detailed information on the surveys. In each page, under the heading “Data Access”, there is a list of all available datasets. The “+” sign also reveals whether the data can be explored online though NESSTAR (only relevant for the LFS data).

Questionnaire structure

LFS

The questionnaire comprises a “core” of questions which are included in every quarter of the survey, together with “non-core” questions which are not asked every quarter. Some “non-core” questions are only asked in one or two quarters per year, e.g., the majority of the questions on a respondent’s employment pattern are only asked in the spring quarter. Other “non-core” questions do not appear every year, but are included in the survey every 2 or 3 years, such as questions on childcare that are asked every 2 years, and questions on regional mobility every 3 years. Some “core” questions are only asked at the first interview as they relate to characteristics that do not change over time (e.g. sex, ethnic group). Other core questions in the LFS have the information from the previous wave rotated into the next quarter. Other core questions have been asked each quarter without reference to previous answer, such as the question on whether the respondent is on a paid job or does paid or unpaid overtime.

APS

The APS included all the data of the ALALFS, but also included a further sample boost in more urban areas of England – known as the APS(B). This survey interview for the APS(B) is shorter than the main LFS, however it does include many of the core questions on the LFS and its enhancements.

Missing values

For many quarters responses may be rolled forward (for one quarter only) if a respondent is unavailable. This is referred to as 'imputation'. If in the case of non-core questions (which are not asked in every quarter) a respondent is unavailable in the latest quarter, then the variable will be coded as DNA (Does Not Apply – there will be no data to 'roll forwards'). A procedure has

been established to separate these 'nonresponding' DNAs from 'genuine' DNAs. Most questions also allow "Don't know" answers. For more information, see Sections 5 and 12 of LFS [User Guide Volume 1](#).

Survey weights

The weights are based on the age and sex structures of the sample and of the population.

LFS

The person-weight variable on LFS datasets is usually named 'pwtXX', where 'XX' refers to the year in which the population totals were projected (see section 10.4 of [Volume 1](#) of the LFS User Guide). A household weight and an income weight can also be found on some datasets and may be more appropriate for some types of analysis; these are described briefly in sections 10.5 and 10.6 of [Volume 1](#) of the LFS User Guide.

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 (see [Volume 8](#) of the LFS User Guides)

- **Family-level analyses.** 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.
- **Household-level analyses.** 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. 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.

APS

For the Local Area (annual) Databases (LADB) survey weights are used to improve the 'weighted totals' at the local area level. This is done by using mid-year population estimates for local authorities and taking account of local authority and unitary populations as well as the age (by five-year age groups) and sex structures of the sample and population.

Annual Population Survey Reweighting 2011: The APS data back to 2009 are currently being

reweighted to 2011 population figures and are in the process of being redeposited at the Archive. Users should note that datasets prior to 2009 will remain weighted to earlier population figures; please check the individual datasets for details.

End User Licence, Special Licence and Secure Access data

LFS

As explained [here](#), from the January-March 2003 quarter, a Special Licence (SL) version of the QLFS data is also available in addition to the version made available under the standard End User Licence (EUL). The SL version contains extra variables, and therefore is subject to more restrictive access conditions. Prospective users of the SL version will need to complete an extra application form and demonstrate to the data owners exactly why they need access to the extra variables, in order to get permission to use that version.

More comprehensive versions of the QLFS datasets are also available via the UK Data Service Secure Access system. These datasets include further additional, detailed variables not included in either the EUL or SL versions. They are subject to further access restrictions.

APS

There are two versions of each APS dataset. One is available under the standard End User Licence (EUL) agreement, and the other is a Special Licence (SL) version. The SL version contains more detailed variables.

The EUL version contains less detailed variables. For example, the lowest geography is Government Office Region, only banded age is available, only 3-digit SOC is available for main, second and last job, and only industry division for main, second and last job. Users are advised to first obtain the standard EUL version of the data to see if they are sufficient for the research requirements.