

# Household Finances Survey QMI

Quality and Methodology Information (QMI) report for the Household Finances Survey (HFS), detailing the strengths and limitations of the data, methods used, and data uses and users.

Contact:  
Vicky Parker and Georgina  
Thompson  
vicky.parker@ons.gov.uk  
+44 1633 455404

Release date:  
21 July 2023

Next release:  
To be announced

## Table of contents

1. [Output information](#)
2. [About this Quality and Methodology Information \(QMI\) report](#)
3. [Important points](#)
4. [Quality summary](#)
5. [Quality characteristics of the Household Finances Survey data](#)
6. [Methods used to produce the Household Finances Survey \(HFS\)](#)
7. [Other information](#)
8. [Cite this methodology](#)

# 1 . Output information

- Survey name: Household Finances Survey (HFS).
- Frequency: annual.
- How compiled: cross sectional and longitudinal sample survey.
- Geographic coverage: UK.
- Last revised: July 2023.

## 2 . About this Quality and Methodology Information (QMI) report

This QMI report contains information on the quality characteristics of the data (including the European Statistical System's five dimensions of quality) as well as the methods used to create it. More information can be found in our [Quality in official statistics methodology](#).

The information in this report will help you to:

- understand the strengths and limitations of the data
- learn about existing uses and users of the data
- understand the methods used to create the data
- decide suitable uses for the data
- reduce the risk of misusing data

## 3 . Important points

All reasonable attempts have been made to ensure that the data we collect are as accurate as possible. However, there are two potential sources of error that may affect the reliability of estimates and for which no adequate adjustments can be made, known as sampling and non-sampling errors. More information about these can be found in our [Uncertainty and how we measure it for our surveys methodology](#).

The Household Finances Survey (HFS) is a combination of the longitudinal Survey of Living Conditions (SLC) and the cross-sectional Living Costs and Food (LCF) survey.

HFS anonymised datasets will be made available through the UK Data Service, starting with the Financial Year Ending (FYE) 2021 annual dataset, which will be published in July 2023.

## 4 . Quality summary

## Overview

Over the past few years, the Office for National Statistics (ONS) has undertaken a programme of work to transform the production of its household financial statistics.

The Household Finances Survey (HFS) data source was launched in the financial year ending (FYE) 2020 to improve the quality of the ONS's income data. The HFS is a combination of the longitudinal Survey of Living Conditions (SLC) and the cross-sectional Living Costs and Food (LCF) survey. Both the SLC and LCF are conducted throughout the year and across the whole of the UK.

The ONS is responsible for conducting both the SLC and LCF survey in Great Britain. In Northern Ireland, the companion surveys to the Great Britain SLC and LCF are conducted by the Central Survey Unit Northern Ireland Statistics and Research agency (NISRA).

The SLC is an annual survey of approximately 12,000 UK households which provides both longitudinal and cross-sectional data on household income, housing, labour, education, and pensions. Respondents selected for the SLC remain in the survey for six years (or waves), with new wave one addresses introduced each year to maintain the size of the sample.

The LCF is the primary data source for household expenditure statistics. It is an annual survey of approximately 5,000 UK households.

The SLC and the LCF have harmonised questions and together they provide common variables on income, employment, benefits, and pensions. These harmonised variables create the "core" of the HFS output. LCF includes some additional expenditure questions, while SLC contains a policy-relevant module which is regularly reviewed.

The HFS follows a two-stage stratified cluster sampling design for selecting the new SLC panel and the annual LCF sample. In Great Britain, both the SLC and the LCF use the Postcode Address File (PAF) as a sampling frame. In Northern Ireland, a systematic random sample of private addresses is drawn from the Land and Property Services Agency's database. The total achieved sample size is around 17,000 households a year, including:

- April 2021 to March 2022: 17,796 households
- April 2020 to March 2021: 18,182 households

HFS data have been used since 2019 to obtain essential information on household income, its distribution across the population and how it is changing over time (see our [Average household income, UK: financial year ending 2022 datasets](#)). HFS data are used to produce our [Household income inequality bulletin](#) and [Effects of taxes and benefits on household income dataset](#).

On 23 March 2020, the coronavirus (COVID-19) outbreak started officially, and the UK was put into lockdown in an unprecedented attempt to reduce the spread of infection. A full description of the impact of the coronavirus pandemic on HFS data collection can be found in our [Impact of COVID-19 on ONS social survey data methodology](#). Before the pandemic, the mode of collection for both the SLC and the LCF was face-to-face interviewing.

As a result, all interviewing was paused in March 2020 and was restarted in April 2020 as telephone interviewing only. The length of the HFS questionnaire was reduced to optimise for telephone interviewing and lessen burden on respondents. Some questions were removed from the household section, such as material deprivation, expenditure, and health questions. The questions which contribute to important statistical publications produced by the Household Income and Expenditure (HIE) branch at the ONS were retained on both surveys. The main income section provides information on personal and household income from different sources (for example, employment, benefits, self-employment).

From July 2020, during the period of telephone interviewing, a knock-to-nudge exercise was undertaken. This involved interviewers visiting addresses to obtain phone numbers and arrange a telephone appointment with respondents on the doorstep. The interview itself would then take place on telephone.

In September 2022, the SLC and the LCF returned to face-to-face interviewing as the default collection mode and the knock-to-nudge exercise was halted.

The HFS is HIE's principal data source for the period covering FYE 2020 onwards, as well as revised timeseries data covering FYE 2017, FYE 2018 and FYE 2019.

## Uses and users

The HFS data source is used in the production of nation statistics on household disposable income and inequality (HDII), and the effects of taxes and benefits (ETB), and associated statistical bulletins produced by the Household Income and Expenditure Analysis (HIE) branch at the ONS. These statistics provide a greater understanding of how income is distributed among UK households (broken down by the main demographics) and resulting income inequality, as well as how these measures change over time.

Users of the microdata include Think Tanks, Academics, and other Government departments, who conduct income, taxes, and benefits research, for example using micro-simulation models to predict policy changes.

## Strengths and limitations

These income statistics were estimated from the LCF alone until FYE 2019, with a much smaller achieved sample of approximately 5,000 households. The larger HFS sample results in less volatility across the measured income distribution and enables greater precision in headline estimates, increased stability over time, and more granular statistics at regional levels and for different population subgroups.

Overall, the larger sample provided by the HFS leads to greater precision when analysing household income by UK regions. On the other hand, the HFS delivers little demonstrable improvement in precision for estimates of Scotland and Northern Ireland.

# 5 . Quality characteristics of the Household Finances Survey data

## Relevance

Prior to the launch of the HFS data source, extensive research was conducted to ensure there was a relevant and significant improvement to the measurement of household income and inequality statistics.

The HFS is an important source of income-based poverty statistics alongside other sources. A Government Statistical Service (GSS) Income and Earnings Coherence Steering Group was established in 2020, aimed at addressing the coherence and accessibility of income and earnings statistics.

ONS income statistics have been improved by moving from the LCF to the HFS as the primary data source. More detailed information on the combined HFS data source can be found in our [Improving the measurement of household income methodology](#).

See also, a [Review of Income-based poverty statistics](#), produced by the Office for Statistics Regulation.

## Accuracy and reliability

Multiple quality assurance methods ensure that the HFS data are as reliable as possible. These methods are applied during the interview and after collection through outlier detection and comparisons of the data between waves and rounds. All data that are identified as possible errors are investigated and, where appropriate, adjusted.

The longitudinal aspect of the SLC survey means that respondents are revisited in subsequent annual waves, which provides the opportunity to confirm the current round's data against that which has been collected previously.

Surveys, such as the SLC, provide estimates of population characteristics rather than exact measures. In principle, many random samples could be drawn and each would give different results, because of the fact that each sample would be made up of different people, who would give different answers to the questions asked. The spread of these results is the sampling variability, which generally reduces with increasing sample size.

## Coherence and comparability

Major government surveys now use harmonised questions on important topics to ensure comparability of results. Where appropriate, HFS questions are harmonised with other government surveys. A list of harmonised questions is available in the [Harmonisation Programme](#). Further information on the [Government Statistical Service \(GSS\) Harmonisation Strategy](#) can be found on the GSS website.

Up until 2020, when the United Kingdom officially left the European Union, questions were developed to meet Eurostat requirements. Therefore, results from the HFS were internationally comparable. The core questions on the SLC and LCF that make up the HFS are fully harmonised.

## Accessibility and clarity

Our recommended format for accessible content is a combination of HTML webpages 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 us, or referenced on our website but stored elsewhere, may vary.

HFS anonymised datasets will be made available through the UK Data Service, starting with the FYE 2021 annual dataset, which will be published in July 2023.

For further information please email us at [SLCResearch@ONS.gov.uk](mailto:SLCResearch@ONS.gov.uk).

## Timeliness and punctuality

The HFS has been in existence since 2020 and has an annual financial year cycle.

Processed data are supplied to Household Income and Expenditure (HIE) Branch of the ONS annually. Annual files are timetabled to be delivered approximately six months after the end of the reporting period.

The data are used and available in the following releases from our HIE branch:

- [Household income inequality UK: FYE 2022](#), published January 2023
- [All data related to Household income inequality UK: FYE 2022](#), published January 2023
- [Household income inequality UK: FYE 2021](#), published March 2022
- [All data related to Household income inequality UK: FYE 2021](#), published March 2022
- [Average household income UK: FYE 2022](#), published January 2023
- [All data related to Average household income UK: FYE 2022](#), published January 2023
- [Average household income UK: FYE 2021](#), published March 2022
- [All data related to Average household income UK: FYE 2021](#), published March 2022

For more details on related releases, the [Release calendar](#) provides 12 months advance notice of release dates. If there are any changes to the pre-announced release schedule, public attention will be drawn to the change and the reasons for the change will be explained fully at the same time, as set out in the [Code of Practice for Official Statistics](#).

## Concepts and definitions

The HFS utilises a suite of standard concepts and definitions. Some of the important terms are defined in this section.

A household is defined as one 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.

The household reference person (HRP) is defined as the person who:

- owns the household accommodation
- is legally responsible for the rent of the accommodation
- has the household accommodation as an emolument or perquisite
- has the household accommodation by virtue of some relationship to the owner who is not a member of the household

If there are joint householders, the HRP is the one with the highest income; if the income is the same, then the eldest householder is taken. The concept of the HRP replaced the previous concept of the head of household in 2001 to 2002.

In 2001 to 2002, the National Statistics Socio-Economic classification (NS-SEC) was adopted for all official surveys, replacing the social class based on occupation and socio-economic group. The long-term unemployed, which fall into a separate category, are defined as those unemployed and seeking work for 12 months or more.

## Geography

The fieldwork is conducted by ONS in Great Britain and by the Northern Ireland Statistics and Research Agency (NISRA) for the Department of Finance and Personnel in Northern Ireland using largely identical questionnaires. Differences between the two questionnaires reflect the country-specific harmonised standards for ethnicity, nationality and national identity, and the different systems of local taxation used in Great Britain and Northern Ireland.

## Output quality

Prior to publication, HFS data are subjected to a rigorous process of quality assurance. An initial series of automatic checks are applied to raw household and income data as they are collected from respondents and entered onto the questionnaire. These data are further checked by a team of editors within the ONS and missing data are imputed where appropriate.

HIE carry out detailed checks on the HFS data. These checks include an automated process where variables are compared with annual variable guides. This ensures the correct types, naming conventions and acceptable values are provided; checking for missingness is also carried out. Suspicious data are further investigated by the HFS team and may be corrected if found to be in error. Automated checking is carried out at each stage of the production process and outputs are further quality assured using various techniques, for example, year on year analysis to check for dubious variations. Income estimates are compared with similar outputs and any movements in benefits data are assessed against known changes to the benefits system.

## Sampling and non-sampling error

All reasonable attempts have been made to ensure that the data are as accurate as possible. However, there are two potential sources of error that may affect the accuracy of estimates and for which no adequate adjustments can be made: sampling and non-sampling errors.

Sampling error refers to the difference between the results obtained from the sample and the results that would be obtained if the entire population was fully enumerated. The survey estimates are therefore likely to differ from the figures that would have been produced if information had been collected for all households or individuals in the UK. The extent to which survey estimates vary from their population values can be estimated, to a given level of confidence, through the calculation of [confidence intervals](#) via the [standard error](#) of the estimate.

The standard error is a measure of sampling variability, which shows the extent to which the estimates are expected to vary over repeated random sampling. To estimate standard errors correctly, the complexity of the survey design needs to be accounted for.

Additional inaccuracies, which are not related to sampling variability, may occur for reasons such as errors in response and reporting. Inaccuracies of this kind are collectively referred to as non-sampling errors and may occur in a sample survey or a census. The main sources of non-sampling error are:

- response errors such as misleading questions, interviewer bias or respondent misreporting
- bias resulting from non-response, as the characteristics of non-responding persons may differ from responding persons
- data input errors or systematic mistakes in processing the data

Non-sampling errors are difficult to quantify in any collection. However, every effort was made to minimise their effect through careful design and testing of the questionnaire, training of interviewers, and extensive editing and quality control procedures at all stages of data processing. Imputation is another method used to improve accuracy resulting from missing observations in the dataset.

For more information on these terms, see our [Uncertainty and how we measure it for our surveys methodology](#).

## Response

The HFS is conducted with people who volunteer their time to answer questions. The voluntary nature of the survey means that people who do not wish to take part in the survey can refuse to do so. Where possible, we collect the reasons given for not participating in the survey. The sample is designed to ensure that the results of the survey represent the population of the UK. The risk of the survey not being representative may increase with every refusal or non-contact with a sampled household (survey non-response).

In 2020 to 2021, the HFS response rate was 42.5% in Great Britain. The response rate for the SLC was 44.2% for 2020 to 2021, while the LCF was 30.9%. Non-response weighting is applied to help mitigate non-response bias.

# 6 . Methods used to produce the Household Finances Survey (HFS)

## How the output is created

The HFS data source is a combination of data from the six waves of the longitudinal Survey of Living Conditions (SLC) data and the cross-sectional Living Costs and Food (LCF) survey data. Both the SLC and LCF are conducted throughout the year and across the whole of the UK. Both the SLC and LCF surveys are conducted by interviewers.

Together, the SLC and the LCF provide common variables on income, employment, benefits, and pensions. These harmonised variables create the “core” of the HFS output. LCF includes some additional expenditure questions, while SLC contains a policy-relevant module which is regularly reviewed.

HFS data provides essential information on household income, its distribution across the population and how it is changing over time, contributing to one of the most widely accessed bulletins produced by the ONS.

The LCF and SLC are voluntary sample surveys of private households. The basic unit of the survey is the household. The HFS (in line with other government social surveys) uses the harmonised definition of a household: 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.

The household section should be completed for each household and must be answered by the household reference person (HRP) or their spouse or partner. Changes to the household composition between waves are also recorded. This allows new members of the household to be added and ineligible respondents who have left the household to be removed.

Detailed questions are asked about the different sources of income of each adult member of the household. In addition, personal information such as age, sex, marital status, and general health is recorded for each household member.

## Editing

The fieldwork, and the editing and coding of data are carried out on a monthly basis. Interim monthly datasets are compiled and quality assured. Additional quality assurance checks are carried out at the completion of each quarter and quarterly datasets are produced. Once the field work for the calendar year has been completed, further quality assurance is carried out before producing the annual dataset for delivery to the Household Income and Expenditure (HIE) branch.

## Imputation

Imputation is an adjustment process that is used to determine and assign replacement values to resolve problems of missing, invalid, or inconsistent data.

The problem of missing data in the HFS is approached via imputation. Statistical imputation is carried out using a nearest-neighbour imputation method where information from a donor record that has no errors or missing values is used to replace the missing values for a recipient record. In this approach, a donor is selected from a pool of potential donors with similar characteristics based on conditional probabilities.

## Outliers

As part of the data cleaning process during the editing phase, outliers are identified. Outliers are checked through examination of other variables to find evidence to support or inform an edit to the outlier. Amendments are only made to data where sufficient evidence to support an amendment exists.

Additionally imputed data are quality assured to ensure no outliers are used as donors.



## Weighting

Weighting for the annual LCF sample and the new SLC panel follows a procedure commonly used for cross-sectional surveys. All drawn households are first assigned their design (sampling) weight which equals the reciprocal of their selection probability. As not every drawn household will participate in the survey, a nonresponse adjustment factor is then multiplied onto the design weight. This factor is equal to the reciprocal of the household's response probability as obtained from a logistic regression model with survey participation as outcome and output area classification (OAC) and NUTS2 region as predictors. The adjustment therefore corrects for differential nonresponse rates across these two characteristics.

Finally, a calibration procedure is applied which ensures that weights sum up to known totals for specified population subgroups. Mathematically, this is done such that the input weights (nonresponse-adjusted design weights) are modified as little as possible. For HFS calibration, age and sex groups, regional totals at personal and household level, employment and tenure are used as control variables.

For the returning SLC panels, the weighting procedure starts from a "base weight" produced while weighting the sample of the previous year (for the former new panel, the base weight is the final calibrated weight while for the longer-running panels the base weight essentially equals the input to the final calibration stage).

To account for attrition between the previous and the current round, an attrition factor is applied which again is derived from a logistic regression model. In contrast to the model for the new panel, the attrition models contain a much wider range of predictors as information collected in previous rounds can be included. Currently, predictors in the attrition models include, for example, household composition (number of adults and children) as well as age, ethnicity, and health status of the HRP.

After incorporating new births and recent joiners to households, weights are calibrated using the same control groups as for the new panel. Finally, the weights for the new and the returning panels are combined in various ways to produce different sets of outputs, for example weights for the complete HFS sample or for the SLC part only.

It should be noted that a different scheme is applied when weighting LCF as an individual survey rather than as a part of HFS (see our [Living Costs and Food Survey Technical Report \(PDF, 688KB\)](#) for details).

## 7 . Other information

Statistical disclosure control methodology is applied to Household Finances Survey (HFS) data. This ensures that information attributable to an individual or individual organisation is not identifiable in any published outputs. The [Code of Practice for Official Statistics](#) and specifically the Principle on Confidentiality set out practices for how we protect data from being disclosed. The principle includes the statement that the Office for National Statistics (ONS) outputs should "ensure that official statistics do not reveal the identity of an individual or organisation, or any private information relating to them, taking into account other relevant sources of information".

More information can be found in the [National Statistician's guidance on the confidentiality of official statistics](#) and on our [Policy for social survey microdata methodology](#).

## 8 . Cite this methodology

Office for National Statistics (ONS), released 21 July 2023, ONS website, methodology, [Household Finances Survey QMI](#)