

# Quality assurance assessment of the Living Costs and Food Survey

Quality assurance (QA) assessment of the end-to-end processes in compiling Living Costs and Food (LCF) survey data.

Contact:  
Hilary Mainwaring  
lcf\_enquiries@ons.gov.uk  
+44 1633 455915

Release date:  
21 December 2022

Next release:  
To be announced

## Table of contents

1. [Introduction](#)
2. [Quality assurance of administrative data](#)
3. [Quality assurance assessment for compiling Living Costs and Food Survey data](#)
4. [Factors that can affect data quality and cause bias](#)
5. [Safeguards that minimise risk](#)
6. [Future developments](#)
7. [Cite this methodology](#)

# 1 . Introduction

The Living Costs and Food Survey (LCF) is an annual survey collecting information on expenditure on goods and services from households throughout the UK. The survey also gathers information about the income of household members.

One of the main purposes for carrying out surveys of the expenditure of households is to define the "basket of goods", which provides weighting information for the [Consumer Price Indices \(CPI\) and the Retail Prices Index \(RPI\)](#).

Income and expenditure is used in our analysis of how [taxes and benefits affect household income](#), with spending data being used to compile [national estimates of household final consumption expenditure](#), which feed into the national accounts. Regional LCF data are among the sources of regional estimates of consumer spending and other regional statistics.

As the most significant survey on household spending in the UK, the main results are published by the Office for National Statistics (ONS) in our annual [Family spending in the UK bulletin](#).

Alongside uses within ONS, LCF data are widely used by other organisations, including:

- HM Treasury (HMT) and HM Revenue and Customs (HMRC), which use expenditure and income data to study how taxes and benefits affect household incomes and to analyse the effects of policy in these areas
- the Department for Environment, Food and Rural Affairs (Defra), which sponsors the collection of specialist food data in its [Family Food publication](#) and other reports
- those involved in policy making, for example in the areas of housing and transport
- independent research institutes, academic researchers, and business and market researchers

To assure ourselves and users of the quality of our statistics, this report provides a detailed summary of the quality assurance of the end-to-end processes in compiling LCF data. This includes recent developments and future improvements.

## 2 . Quality assurance of administrative data

The [Quality Assurance for Administrative Data \(QAAD\)](#) was introduced in 2015 by the UK Statistics Authority. It contains guidance for statistics producers to review their quality assurance of administrative data used to produce official statistics against the [Code of Practice](#). While the framework is designed for managing risks arising from administrative datasets, the principles can also be applied to survey data. The [Administrative Data Quality Assurance toolkit \(PDF, 243 KB\)](#) provides guidance for meeting assurance levels, either "basic", "enhanced" or "comprehensive" using:

- Quality Management Actions
- a Quality Assurance Matrix
- a Risk and Profile Matrix

This report aims to apply the relevant principles of the QAAD toolkit to provide a comprehensive level of assurance for the data collected and processed within the Living Costs and Food Survey (LCF). The focus is a comprehensive assurance level, meaning that we require a detailed understanding of the operational context in which data are collected, including sources of bias, error, and mismeasurement. This level provides reassurance and transparency regarding the quality of the data and provides users with a better understanding of the reliability and accuracy of using these data in their own statistics.

## **3 . Quality assurance assessment for compiling Living Costs and Food Survey data**

### **Sampling**

The Living Costs and Food Survey (LCF) is a voluntary sample survey of private UK households, focussing on the collection of household spending and income. The sample is drawn statistically, with details of the size, frame and design described within our [Living Costs and Food Survey Quality and Methodology Information \(QMI\)](#) report.

### **Data collection: electronic questionnaire and paper-based diary**

#### **Electronic questionnaire**

Households at sampled addresses are visited by fieldwork interviewers and asked to take part in a face-to-face questionnaire. Interviews are conducted through computer assisted personal interviewing (CAPI) using laptops and the Blaise software package. For details of changes to data collection during the pandemic, see the [Coronavirus \(COVID-19\) pandemic section](#). Interview questions are grouped in thematic blocks recording household demographics, expenditure, and income.

#### **Quality assurance principles, standards and checks**

Each question response is diligently recorded by interviewers, ensuring any discrepancy or questionable answers are clarified and checked at the input stage. In addition, quality assurance checks are built into the CAPI and Blaise software, meaning suspicious responses are automatically flagged during the interview. This allows interviewers to seek direct, immediate clarification and feedback with the respondent.

#### **Paper-based diary**

Diary data collection takes place after household interviewing has finished and asks each individual aged 16 years and over to record all expenditure in the 14 days that follow the interview. Diary spending typically covers areas such as food and drink, clothing and ad-hoc spending on pastimes and leisure. Children aged between 7 and 15 years are asked to keep a simplified diary. For details of changes to diary collection during the pandemic, see the [Coronavirus \(COVID-19\) pandemic section](#).

#### **Quality assurance principles, standards and checks**

Diaries are given to respondents with full instructions from field interviewers. Interviewers check in with the respondents after a few days to ensure the process of recording daily expenditure is fully understood. On completion, interviewers again check the diary to ensure it is fully completed before transmitting it for the next processing stage.

#### **Coding, editing and imputation**

Coders enter diary details into an electronic version using Blaise software. Editors then carry out a series of quality assurance procedures including imputation.

## Quality assurance principles, standards and checks

Coding from the diary component of the LCF is manual and highly skilled. Coders check on the completeness of data, that is, that each item of expenditure is coded to one of the coding categories ([classification of individual consumption by purpose \(COICOP\)](#)) and that the cost of each item is recorded. At this stage, coders may visit the websites of retailers and manufactures to clarify incomplete diary information. Additional checks on range and consistency are built into the Blaise diary (see our [Living Costs and Food Survey technical report: April 2015 to March 2016 \(PDF, 689 KB\)](#) for more information).

Editors open and review every single household questionnaire by following the same route (through the questionnaire) as the respondent(s). This process ensures that every survey response field is examined for "flags" (entries that are potentially in contravention of a pre-coded "business rule"), which are reviewed and considered for edit. All notes and comments recorded in the questionnaire by interviewers are viewed and resolved with relevant data amendments.

This approach means that no entry or query is unscrutinised. This has a positive effect on the quality assurance process. Decisions on handling anomalous responses are guided by established good practice, as set out in detailed documentation with collaboration between interviewers, coders, editors and when necessary, research teams.

In addition, monthly and quarterly coded and edited LCF data are shared with specialist data users, for example food data specialists from Defra, who contribute to the quality assurance of diary and interviewer data. This means the data are processed with expert insight.

Imputation is used to determine and assign replacement values when respondents are unable to provide a response, improving accuracy and reducing non-response bias resulting from missing data. Replacement data are sourced:

- from external (non-LCF) reference tables published elsewhere (mortgage imputation tables based on interest rates, tables and rules on eligibility to state benefits)
- by referencing tables based on LCF data from previous years showing average amounts according to household income
- using information collected elsewhere in the questionnaire or referring back to interviewers

The LCF also accepts missing diaries if the diary of the main diary keeper is present. Missing diaries are imputed, improving data quality through receiving diary data from a person in another responding household with matching characteristics of age, employment status and relationship to the household reference person. In the financial year ending (FYE) 2021, 43 households had imputed diaries, accounting for 1% of responding households (see Table 6 in our [Living Costs and Food Survey: technical report data tables](#)).

## Data processing and statistical analysis

Following coding, editing and imputation, interview and diary data are merged and aggregated, and derived variables essential for analysis are produced.

## Quality assurance principles, standards and checks

Data validation is performed at this stage to ensure consistency between household, individual and diary level data. Other checks include analysing time series data to identify odd movements, examining outliers, extreme values and other data inconsistencies.

Examples of data checks include unit costs that are too high or low, missing shop codes and expenditure coded to inappropriate shops.

Cycles of data derivation and validation are performed to reflect editing because of quality assurance procedures.

## Weighting

Since the survey year ending March 1999, the survey has been weighted to reduce the effect of non-response bias. A detailed description of the weighting process and its effects can be found in [Living Costs and Food Survey technical report: April 2015 to March 2016 \(PDF, 689 KB\)](#). Details of additional calibration constraints to account for collection mode changes because of the coronavirus pandemic are detailed in our [Living Costs and Food Survey technical report: financial year ending March 2021](#).

## Knowledge sharing and collaborative relationships

Curiosity meetings are held with internal Office for National Statistics (ONS) users to share findings about and interpretations of the weighted LCF data. Meetings include knowledge sharing from experts from each part of the data compilation process. This ensures that any changes and their impact are widely communicated and discussed by statistics producers, adding an additional layer of quality assurance. Micro-level unweighted and weighted data are also shared alongside aggregated data.

To strengthen engagement across the LCF user community, the LCF steering group terms of reference, membership list and purpose were reviewed. This has increased our understanding of the value of statistics created using LCF data and allows a wide range of views and uses to be used to focus any ongoing or future development.

## Quality assurance principles, standards and checks

Wider data sharing provides opportunities for users outside the processing teams to contribute to quality assurance. Collaboration between producers and users encourages discussions on data strengths and limitations and improved understanding of data quality and impact on statistics.

# 4 . Factors that can affect data quality and cause bias

## Coronavirus (COVID-19) pandemic

Survey interviewing was paused from 17 March 2020 until mid-April 2020. A strategy to move data collection from face-to-face to telephone interviewing was therefore rapidly implemented to ensure survey data collection could continue, while protecting respondent and field interviewer safety.

To limit the impact on data quality, we:

- shortened the questionnaire to reduce respondent burden and impact of change on response quality ([see Technical report: financial year ending \(FYE\) 2021 questionnaire changes](#))
- adapted the questions for telephone interviewing; for example, rewording questions using showcards and adding standardised prompts to questions that used to require showcards
- asked respondents during the two-week diary period to provide copies of receipts (electronic or paper), with interviewers recording non-receipt-based expenditure via regular telephone calls
- increased sample sizes
- reviewed unconditional incentive values (see [Cash incentive increase](#))
- changed how we contacted respondents
- added an additional housing tenure calibration control to the derivation of FYE 2021 weights to adjust for oversampling bias related to changes in the demographics of the Living Costs and Food Survey (LCF) respondents caused by switching mode from face-to-face to telephone

See our [Impact of COVID-19 on Office for National Statistics \(ONS\) social survey data collection methodology](#) for more details.

## Declining response

The LCF has experienced a declining response rate, which continues to be an ongoing issue. Current response levels are detailed in our [Living Costs and Food Survey technical report: financial year ending March 2021](#).

Declining response is not unique to the LCF or the UK. International comparisons show similar decline, with a number of reasons linked to this.

Respondent buy in is a challenge because of the extra diary component in the LCF compared with other surveys. There are many opportunities for respondents to decline taking part. Interviewers require a lot of skill to sell the survey, maintain rapport and build motivation to complete the diary.

Time and depth of the survey is also a challenge. Respondent burden for LCF arises from length and complexity of the questionnaire in addition to the demands of completing the diary; looking for bills or receipts causes lapse in concentration.

## Other factors affecting data quality

### Questionnaire change process

The questionnaire change process is critical, to ensure questionnaire developments are efficiently implemented with consideration of harmonisation requirements for all surveys included within the [Household Finances Survey \(HFS\)](#). A robust change process system is required, ensuring changes are raised centrally. This allows stakeholders to assess impact, approve and collaborate on development of questionnaire content and implementation of changes including to software.

### Diary fatigue

Respondents record less expenditure towards the end of week two, indicating an ongoing challenge to collect accurate expenditure for the complete time period.

## Statistical processing modules used up to the financial year ending (FYE) 2021 datasets

Processes were manual and labour intensive, lacking in clarity, repetitive in nature and at high risk of human error. They were written in unsupported software, which involved including inefficient processing code and a lack of transparency.

These processes depended upon various network drives containing multiple folder structures which held many versions of data and information. This increased the risk of error.

Limitations within the statistical processing modules meant identifying, resolving and preventing errors required much greater skill and effort.

## 5 . Safeguards that minimise risk

To mitigate the effects of challenges, a number of initiatives have been implemented, some of which have been adapted as a result of coronavirus (COVID-19) limitations.

Although the declining response rate is recognised as a major challenge within the Living Costs and Food Survey (LCF), there are processes in place to minimise the impact upon data quality.

### Cash incentive increase

Prior to the coronavirus pandemic, a trial was conducted to understand whether increasing the value of the conditional incentive would result in a higher response rate without negatively affecting data quality. The trial concluded that offering a £40 incentive increased response rates by 6 and 4 percentage points, compared with £20 and £30, respectively. In April 2020, a trial was planned to test the impact on response rates of offering £40 versus £50 per full interview and diary completed. Given the unplanned mode change in April 2020, the decision was made to not run this trial but to increase the incentive, with all responding adults receiving a £50 voucher.

## **Introduction of knock-to-nudge**

With the coronavirus pandemic forcing a mode change from face-to-face to telephone interviewing, knock-to-nudge was introduced. This involved interviewers on the doorstep encouraging those invited to participate in a survey to provide their telephone number or arrange an appointment for a telephone interview. Following the implementation of this incentive, an increase in response rates was seen.

Read more about the [Impact of COVID-19 on ONS social survey data collection](#).

## **Improvement in the recruitment process through improvement of interviewer skills**

Many respondents that take part do so as a result of the skill of the interviewer explaining the value of their contribution. Skills and training of interviewers able to communicate the value of their contribution was identified as a major reason respondents agree to take part by interview managers.

## **Optimal calling patterns identified and analysed**

Analysis of optimal calling patterns enables the interviewers to reach potential respondents at times where they are most likely to respond and agree to take part in the survey.

## **Other safeguards to minimise risk**

### **Financial surveys questionnaire co-ordination**

The [Household Finances Survey \(HFS\)](#) teams identified the need for collaboration when considering questionnaire updates as a result of areas of questionnaire harmonisation and crossover between topics for those surveys collecting household finances data. Questionnaire changes are proposed and discussed at the financial surveys questionnaire co-ordination group, a group consisting of representatives from HFS teams. This means that questionnaire change proposals can be managed collaboratively, increasing the transparency, understanding and management of risks related to potential questionnaire changes. Quarterly horizon-scanning sessions are held, allowing HFS teams and ONS statistics producers to discuss policy impacts and initiatives and the potential need for questionnaire changes, including assessing and agreeing implementation options across the Household Finances surveys.

## **Project to improve LCF data processing methods and systems**

### **Improved design and transparency of statistical processing modules**

In 2021, a project was started to produce process flow maps detailing the existing LCF statistical processing journey, essentially identifying "pain points". This transparency, combined with expert methodological knowledge and best practice of statistical processing of household finances data, enabled the development of recommendations for strengthening the LCF statistical processing design and governance. Creation of variable dictionaries was automated, and specifications of the calculation of derived variables was embedded within the code. This allowed for the automation of the production of data user's documentation, increasing the transparency and usability of LCF data.

### **Redesigned efficient modules for data extraction, aggregation and statistical processing**

Using agreed recommendations for LCF statistical processing design and governance, existing modules were redesigned into automated pipelines, using open-source programming languages (R and Python) and [Reproducible Analytical Pipelines \(RAP\) principles](#). Testing criteria was agreed and actioned, including parallel runs of new versus old modules with investigation, resolution and sign off of any inconsistencies. This reduces the risk of error, by minimising the manual steps and automating the quality assurance processes (including enhancing the automated detection of rogue or missing data or miscalculated derived variables). This strengthened the robustness of the code through increased efficiency and transparency.

## Creation of a service handbook

Instructions documenting the redesigned LCF statistical processing system were produced in the form of a service handbook. This provides processing teams with documentation for operating the system, as well as describing previous, current and next stage procedures reducing the risk of error. This is contained in a single location for an easy access.

## 6 . Future developments

The Living Costs and Food Survey is undergoing continuous improvement and is one of three Office for National Statistics (ONS) Household Financial Surveys included within the Household Financial Statistics Transformation (HFST) project. This project has been tasked with creating an ambitious solution by which higher quality, more timely and more granular household financial statistics will be provided. This will ultimately offer an in-depth understanding of income, wealth, spending and financial well-being across the UK population.

The ambition is to consolidate the three surveys, and make greater use of non-survey data, maximising the granularity of the data at the lowest possible cost, while allowing new links to be drawn between different elements of household behaviour. These can then contribute to multiple statistical outputs, minimising duplication and maximising consistency.

One of the initial milestones for this project is an enhancement to the current process for collected detailed expenditure data using the diary element of the survey. The aim of this is to reduce interviewer and coder burden, leading to increased data quality. The earliest implementation of the improvement is April 2023. Alongside this, options for development of a respondent-facing digital collection tool to eventually supplement the current diary approach are being explored.

As part of the longer term aims of the project, we will be using existing research such as the ONS-commissioned external review performed by NatCen to focus on a high-level design for expenditure data collection. We will also be collaborating with other national statistical institutes (NSIs) and keeping up to date with international best practice to accomplish the wider aims of the HFST project. User needs are being collected and considered throughout to inform the project's shape. We recently published our [Transforming the ONS's household financial statistics consultation \(PDF, 763 KB\)](#).

## 7 . Cite this methodology

Office for National Statistics (ONS), released 21 December 2022, ONS website, methodology, [Quality assurance assessment of the Living Costs and Food Survey](#)