Office for
National Statistics

Living Costs and Food Survey

Volume A:

Introduction

April 2015 to March 2016

Great Britain and Northern Ireland

ONS Social Surveys

Office for National Statistics

March 2017

Copyright and reproduction

CONTENTS

If you have any queries relating to the 2015-16 LCF database please contact the customer response team on 01633 45 5678 or socialsurveys@ons.gov.uk

Introduction

This volume provides a brief overview of the Living Costs and Food Survey User Documentation, together with information on database and survey definitions, as well as background to the 2015-16 survey. The details it contains are about information resulting from the Living Costs and Food Survey (LCF) for the financial year April 2015 to March 2016. The move from calendar to financial year comes following user demand.

Guide to the 2015-16 User Documentation

Structure of the User Documentation

Volume A – *Introduction*

> Introduction
> Guide to the 2015-16 User Documentation
> The structure of the 2015-16 database
> Data flow chart
> Database definitions

Volume B - *Household Questionnaire*

> Explanatory notes
> Household questions

Volume C - *Income Questionnaire*

> Explanatory note
> Income questions

Volume D - *Expenditure Codes*

> Explanatory note
> Part 1 Expenditure codes

Volume E - *The Raw Database*

> Explanatory note
> Contents
> Part 1 Raw table definitions
> Part 2 Raw variables in variable name order
> Part 3 Raw variables coding frame in variable name order
> Appendix A

Volume F - *The Derived Database*

> Explanatory note
> Contents
> Anonymisation
> Part 1 Derived table definitions
> Part 2 Derived variables in variable name order
> Part 3 Specification of product codes
> Part 4 Coding frames for derived codes

Volume G – *Derived Variable Flowcharts*

      Contents
      Glossary
      Flowcharts – Person level
      Flowcharts – Household level
      Flowcharts – Loan transactions
      Flowcharts – Hire purchase transactions
      Product codes – Person level
      Product codes – Household level

Volume H - *Database Changes*

      Contents
      Explanatory note
      Part 1 New raw variables for 2015-16
      Part 2 2014 raw variables that have been deleted for 2015-16
      Part 3 2014 raw variables that have changed for 2015-16
      Part 4 New derived variables for 2015-16
      Part 5 2014 derived variables that have been deleted for 2015-16
      Part 6 2014 derived variables that have changed for 2015-16
      Part 7 Major changes (1991 to 2015-16)

Use of the 2015-16 User Documentation

Explanatory notes on the use of the volumes of the documentation are given in each volume. The following gives a summary only. See 'Definitions' (For more information see the ['Living Costs and Food Survey Technical Report'](#)).

Volume A – *Introduction*

Guide to the 2015-16 User Documentation - This volume begins with a guide to the contents and the use of the User Documentation

The structure of the 2015-16 database – This section of the volume outlines the different parts of the database.

Database definitions - Provides definitions of the terms used within the documentation.

Volume B – *Household Questionnaire*

The first part of the LCF questionnaire collects information about households; that is to say that the majority of the questions are asked at a household-level, with the household reference person typically responding on behalf of the household as a whole. The household questionnaire includes questions on a range of subjects which are of interest to the survey's users, including family relationships, ethnicity, employment details and the ownership of household durables. It is also the source of all expenditure information not recorded in the diary; principally that which concerns regular payments typically made by all households and large, infrequently purchased items such as vehicles, package holidays and home improvements.

Volume C – *Income Questionnaire*

The income questionnaire follows on immediately from the household questionnaire and collects the key person-level variables used on the survey. The principal components of the LCF income questionnaire are the sections covering income from employment, benefits and assets. These together form an overview of the total income received by each household, as well as each household member individually.

Volumes B and C outline the questions exactly as they appear within the computer-assisted personal interviewing (CAPI) program and are then asked in the LCF interview. The Blaise program ensures that the correct questionnaire routing is followed through the interview and applies range and consistency error checks where necessary. The Blaise-code and translated routing for each individual question in the household and income questionnaires are shown in Volume B and C respectively. This can be used to both trace the conditions under which a particular question or set of questions is asked and also to illustrate the overall flow and interdependence of the LCF interview as a whole.

Volume D – *Expenditure Codes*

In the Living Costs and Food Survey, information about expenditure is collected at a detailed level. It is collected through both the diary (which respondents fill in daily for a fortnight) and through the Household Questionnaire.

Part 1 Expenditure Codes gives an indication of the types of items to be found under each expenditure code and provides a look-up table between the EFS codes (e-codes) and the COICOP-plus c-codes.

Volume E – *The Raw Database*

This volume of the User Documentation describes the raw database. The raw database contains data 'as received' as well as derived variables calculated within the questionnaire. The volume is in three parts, as outlined below.

Part 1 – Raw table definitions
The raw variables are grouped into three data files (or tables); this part of the volume defines these raw files and gives information on the three sources of the data (diary, household and income questionnaire).

Part 2 – Raw variable list
This is a variable list, listing all raw variables in alphabetical order, and including information on the variable name, description and table name. This list allows details of a variable (such as the file in which it is found) to be looked up if the variable name is known. A column is also present to identify variables that are present in the UK Data Archive anonymised dataset.

Part 3 – Raw variable coding frame
This lists coding frames for each of the variables in the questionnaires.

Appendix A – Standard period codes.

Volume F *- The Derived Database*

This volume of the User Documentation, in four parts, describes the derived database. A summary of anonymisation, particularly in relation to council tax, is also given following the explanatory notes.

Part 1 – Derived table definitions
The derived variables are also grouped into separate files (or tables) and this part of the volume defines these files.

Part 2 – Derived variable descriptions
All derived variables are listed, including information on the variable name, description, file name (i.e. the table in which the variable is held) and coding frame (where relevant). A column is also present to identify variables that are present in the UK Data Archive anonymised dataset.

Part 3 – P-codes
Product codes (or p-codes) are aggregated codes, some at household level and some at person level. Part 3 gives descriptions of these codes as an aid to understanding their structure.

Part 4 – Coding frames
This part of the document contains a list of the coding frames associated with the derived variables, including an index of those most used.

Volumes G – *Derived Variable Flowcharts*

Volume G of the documentation contains flowcharts and tables that demonstrate how the variables in the database are derived.

Flowcharts for person level, household level, loan transaction and hire purchase transaction variables are included, as well as tables for person level and household level Product Codes. These have been produced on tables rather than flowcharts as they are simply the sum of other variables and do not have any routing to flowchart.

Volumes H *- Database Changes*

Volume H of the documentation describes the changes that have taken place in the 2015-16 database compared with 2014.

Part 1 - contains the new raw variables for 2015-16.

Part 2 - contains 2014 raw variables that have been deleted for 2015-16.

Part 3 - contains 2014 raw variables that have changed for 2015-16.

Part 4 - contains new derived variables for 2015-16.

Part 5 - contains 2014 derived variables that have been deleted for 2015-16.

Part 6 - contains 2014 derived variables that have changed for 2015-16.

Part 7 - Major changes, highlights some of the more important definitional changes for 2015-16.

There are two other items of documentation that can be provided to users (if required):

- The LCF Diary (adult and youth)
- LCF Interview Prompt Cards

The Structure of the 2015-16 database

For definitions of the terms used here, please see Part 3 – Database definitions.

Outline of the structure of the LCF database 2015-16

In the LCF database, variables are grouped to form a number of different SPSS DATA FILES. The raw files contain data as received and the derived files contain processed data. There are 5 anonymised SPSS data files. These are:

1. Rawper
2. Rawhh
3. DVPer
4. DVHH
5. DV_Set89

Definitions of the raw database are to be found in Volume E of the User Documentation.

All derived file names are pre-fixed "DV" and definitions are to be found in Volume F of the User Documentation.

The data flow

Items recorded in the two-week diary are recorded in the raw diary file. At this stage, every separate instance of expenditure by every person is represented by a separate "record" or "row" and is assigned an expenditure code that indicates the type of item purchased and the type of transaction (e.g. cash, credit card, business expense). See Volume D for information on expenditure codes. Data at this level are not provided in the user dataset.

In DV_SET89, some expenditure relating to the expenditure codes but collected through the questionnaire is added to that collected through the diary. In this file, there is only one row per expenditure code. This represents the average weekly expenditure for each person on items in that code. Excluded are entries from the diary that relate to types of purchase covered by individual questions in the questionnaire. Included is pocket money given to a child who has completed a youth diary. This file feeds into various C-CODE files. In these, the expenditure codes become VARIABLES and are pre-fixed with 'C'.

Database definitions

The raw database

The raw database contains data "as received". At this stage, weekly equivalent values have been calculated for questionnaire data (see note on 'weekly equivalent values') and some imputation has been carried out. Care should be taken when using raw data.

The raw database is defined in Volume E of the User Documentation.

The derived database

The derived database consists of variables that have been adjusted in some way in relation to the state in which they were received. This adjustment might involve only a change in name, or it might involve some calculation or aggregation. The variables are grouped in files which have DV in the name. Examples are DVHH or DV_SET89.

The main variable prefixes and suffixes are briefly described below, but more information on the derived database is given in Volume F of the User Documentation.

'A' and 'B' codes

A-codes and B-codes are aggregate codes, some at household level and some at person level. They are all derived codes and so are listed in Volume F (The Derived Database) of the User Documentation. That they are pre-fixed 'a' or 'b' has no significance.

Anonymisation

Variables which could lead to the identification of a household or an individual are not released to users.

An 'anonymised' version of sensitive raw or derived variables is provided in the user dataset. Variables which have been anonymised can be identified because they are suffixed 'p' (those which are not have been detailed in the appropriate User Documentation). Un-anonymised versions are, in the main, suffixed 'u' and are not accessible to external users of the LCF. Raw diary information is not provided to any external users. The principles of anonymisation, particularly in relation to council tax, are explained in Volume F of the User Documentation. All variables that are released as part of the UK Data Archive anonymised dataset are identified alongside the variable lists in Volumes E and F.

Codes suffixed 'c'

These codes represent the expenditure of children, aged between 7 and 15, who completed a diary. The codes are aggregated to household level. Variables suffixed 'c' are only available as expenditure or product codes e.g. p522c.

Codes suffixed 't'

These codes represent total expenditure by children and adults. They are aggregated to household level by adding the original (adult) variable to those suffixed 'c', (i.e. p522 + p522c = p522t). Variables suffixed 't' are only available as expenditure or product codes.

Codes suffixed 'l'

These codes represent expenditure in nominated large supermarkets. Variables suffixed 'l' are only available as expenditure codes for food, alcohol, petrol, some household goods and regular purchases. For expenditure on clothing the suffixes are 'x' for selected clothing chains, 'y' for large supermarkets and 'z' for charity shops.

Codes suffixed 'w'

These codes represent expenditure for goods and services ordered via the internet. Variables suffixed 'w' are only available as expenditure codes for certain goods and services ordered via the internet.

Coding frame

If a variable has a coding frame, it means that it is allowed to take on only certain permitted values (i.e. the variable is not a 'free' field).

The name of a coding frame is usually the same as the name of the variable using it, however, the more commonly used coding frames have their own names. If a variable has a coding frame it is named in the final column of part 2 of Volume F of the User Documentation. The coding frames are listed in alphabetical order in Part 4 (Coding Frames) of the User Documentation.

Diary

The diary is completed by every spender. It is completed daily for 14 days and all expenditure in that period is recorded.

See 'Expenditure' in the [technical report](#) for further information on the collection of data through the diary.

Expenditure codes

Each expenditure amount collected through the diary and some collected through the questionnaire are given a code which represents the type of item that was purchased. The LCF expenditure codes differ from those used for the Family Expenditure Survey (FES), therefore, care needs to be taken when comparing LCF and FES data. The LCF codes are based on the European standard classification, COICOP (Classification of Individual Consumption by Purpose).

Variables representing household aggregates of the expenditure codes are called C-codes (or C-variables): the variable name is the expenditure code pre-fixed with 'C'. A complete listing of the expenditure codes and more details on their use can be found in Volume D (Expenditure Codes) of the User Documentation.

Filename

The filename is the database reference name of the file, referring to individual SPSS datafiles.

In the LCF database, the data are stored in a number of different files. A single file may contain information on a number of variables that are related to one another in some way, such as the file called 'DVLOANS_SET86' which contains information about loans, or variables which are at the same level of detail, such as DVPER which contains the main person level derived variables.

Conceptually, a file can be perceived as consisting of rows and columns: the 'column headings' are the names of the variables contained in that table and each 'row' represents an individual 'record'. A 'row' or 'record' is uniquely labelled with, for example, case (household number), person (person number), or itemnum (item number). If both case and person are used, such as in the file DVPER, then there is a separate row in the table not only for every household, but for every person within that household.

FS Codes

FS codes are used only for the preparation of the publication 'Family Spending'.

Household questionnaire – Volume B

The household questionnaire must be completed for every adult in the household.

The household questionnaire collects information on household characteristics and the characteristics of the individuals living there. Also collected through the household questionnaire is information on mortgages, rents, bills, insurance policies, loans, hire purchase, standing orders, season tickets and other irregular expenditure such as that on cars or home improvements. Information on the irregular purchases is often obtained through retrospective recall (i.e. the question may be 'when did you last purchase a ....?'). Expenditure on retrospective recall is converted to a weekly equivalent value. That is, if the recall period is one year, then the weekly equivalent value is calculated by dividing by 52.

See the household questionnaire for details on questions and recall periods.

Income questionnaire – Volume C

The income questionnaire must be completed for EVERY adult in the household. The income questionnaire collects information on the income of the individuals in the household and the household as a whole. See the technical report for further information on income.

Period codes

Where expenditure amounts are not given on a weekly equivalent basis, a weekly equivalent is calculated by dividing the amount given by the appropriate period. The standard period codes can be found at the end of Part 1 of this volume.

Product codes (p-codes)

Product codes (or p-codes) are aggregate codes, some at household level and some at person level. Part 3 of Volume F of the User Documentation gives descriptions of these codes as an aid to understanding their structure, along with more detailed explanatory notes on their use.

Variable name

The variable name (sometimes just written 'variable' in the documentation) is the database reference name of the variable. Variable descriptions can be found in Part 2 of both Volume E and Volume F of the User Documentation.

Variable description

The variable description is a brief description of the variable. Variable lists ordered by variable description can be found in Part 3 of both Volume E and Volume F of the User Documentation.

Weekly equivalent values (wev) - codes prefixed 'DV'

So that expenditure from different sources can be aggregated, it is necessary for all expenditure amounts to be on an equivalent basis. In order to achieve this, all expenditure amounts collected through the questionnaire are converted to WEEKLY EQUIVALENT VALUES. In the dataset, weekly equivalent values can be identified because they are pre-fixed 'DV'. Information on weekly equivalent values is included as part of the questionnaire. Weekly equivalent values are RAW variables and so are listed in Volume E of the User Documentation.

Further Information

Further information regarding the Living Costs and Food Survey can be found on the Office for National Statistics (ONS) website.

The Living Costs and Food Survey technical report and the Quality and methodology section of Family Spending provide further detailed information regarding the methodology used to produce the estimates presented in Family Spending.