
UK Data Service



Teaching Dataset

Living Costs and Food Survey, 2013

User Guide

Author: Margherita Ceraolo/UKDS

Version: 2

Date: 1 July 2019

Contents

About this User Guide	3
Introduction to the Living Costs and Food Survey	3
Main Topics	4
The Datasets	4
The Sample Design and Response Rate	5
Weighting the Dataset	5
Other Useful Sources of Information	5
Introduction to the LCF 2013: Unrestricted Access Teaching Dataset	6
Details of variables	7
Variable names and suffixes:	8
Codes suffixed 'r'	8
Codes suffixed 't'	8
Anonymisation 'p'	9
Downloading the Data	9
Appendix I: Data Dictionary	10
Appendix II: SPSS Syntax for variables derived for the LCF 2013: Unrestricted Access Teaching Dataset	16
Appendix III: P550tp Total Expenditure, by Adults and Children	19

About this User Guide

This is a user guide for the Living Costs and Food Survey (LCF), 2013: Unrestricted Access Teaching Dataset. This teaching dataset has been created for teaching related purposes only such as for teaching basic statistics or for students to write reports for coursework. Those wishing to use the Living Costs and Food Survey for research purposes are advised to use the original [SN 7702 Living Costs and Food Survey, 2013](#) that can be downloaded from the UK Data Service. This LCF 2013 Unrestricted Access Teaching Dataset contains a reduced number of variables from the original LCF 2013. It gives students the opportunity to gain experience of using nationally representative UK data.

Introduction to the Living Costs and Food Survey

The LCF collects information on spending patterns and the cost of living that reflects household budgets across the country. The survey is conducted throughout the year across the whole of the UK and is the most significant consumer survey undertaken in the UK. The LCF is a voluntary sample survey of private households¹. Each individual aged 16 and over in the household visited is asked to keep diary records of daily expenditure for two weeks.

The LCF provides information for the Retail Prices Index, National Accounts estimates of household expenditure, the analysis of the effect of taxes and benefits and trends in nutrition. The results, however, are multi-purpose, providing an invaluable supply of economic and social data. The survey is carried out by the Office for National Statistics (ONS) in Great Britain and by the Northern Ireland Statistics and Research Agency (NISRA) in Northern Ireland.

Data from the LCF survey are anonymised and deposited with the [UK Data Service](#).

¹ A household comprises one person or a group of people who have the accommodation as their only or main residence and (for a group): either share at least one meal a day or share the living accommodation, that is, a living room or sitting room.

Main Topics

Household questionnaire:

The first part of the LCF questionnaire collects information about households; 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 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 LCF 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.

Individual 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.

LCF Diary:

In addition to the two questionnaires, each individual aged 16 years and over in the household is asked to keep diary records of daily expenditure for two weeks.

The Datasets

In the LCF dataset, variables are grouped to form a number of different data files. The raw files contain the data as received. The derived data file (DV) consists of variables that have been adjusted in some way in relation to the state in which they were received. The LCF, 2013: Unrestricted Access Teaching Dataset is based on the derived household file “2013_dvhh_ukanon.sav” and includes only derived variables.

The Sample Design and Response Rate

The LCF sample for Great Britain is a multi-stage stratified random sample with clustering. It is drawn from the Small Users file of the Postcode Address File (PAF) – the Post Office’s list of addresses. All Scottish offshore islands and the Isles of Scilly are excluded from the sample because of excessive interview travel costs.

The Northern Ireland sample is drawn as a random sample of addresses from the Land and Property Services Agency list.

The total sample size for LCF 2013 is 5,144. Of these only 151 households are from Northern Ireland.

Weighting the Dataset

Respondents to the survey did not have an equal chance of selection. There were also differences in response rates between groups. To adjust for the unequal selection probabilities and non-response, users will usually need to use the appropriate weight when analysing the LCF.

This dataset includes two weight variables.

- **weighta** – is the original annual weight used to reduce the effect of non-response bias and produce population totals and means. This weight is designed to gross the frequencies to the population of England and Wales. A ‘grossing weight’ is useful for calculating the prevalence of phenomena at the nation level. However, users should be aware that grossing weights can cause errors when calculating the precision of estimates
- **weightar** – This weight has been calculated for this teaching dataset. It is the original annual weight scaled to have a mean of 1. This weight variable will be particularly useful for new users exploring statistics in statistical packages such as SPSS.

Other Useful Sources of Information

For a brief overview of the 2013 LCF user documentation, together with information on database, survey definitions, weighting, as well as background to the survey please see the [LCF 2013 documentation](#) on the UK Data Service website, which includes the following:

- Volume A - Introduction
- Volume B: The Household Questionnaire User Guide
- Volume F: Derived Variable User Guide

Introduction to the LCF 2013: Unrestricted Access Teaching Dataset

The LCF 2013: Unrestricted Access Teaching Dataset contains 14 variables which include information on total expenditure as well as standard socio-demographic variables. The variables have been chosen to enable new users to explore the data and to use it apply some basic statistical techniques such as linear regression analysis.

Details of the variables are given below including a data dictionary which lists information about variable names, values, labels, missing values and frequency. The name of the variables and their labels remain the same or very close to the original LCF data set. However, for statistical disclosure reasons, some variables have been recoded and their level of detail reduced. Variables that differ from the original LCF have been listed in Table 2. The syntax for how these variables were recoded is in Appendix II. In addition a new ID variable was created in order to avoid linking the Unrestricted Access Teaching Dataset with the original dataset.

The total expenditure variable is a continuous variable and has been included in the Unrestricted Access Teaching Dataset to enable users to carry out linear regression analysis. This variable is collected through both the diary (which respondents fill in daily for a fortnight) and through the Household Questionnaire. The coding structure is based on the European classification, 'The Classification Of Individual Consumption by Purpose' (COICOP) and includes the following categories²:

- 1 Food and Non-Alcoholic Beverages
- 2 Alcoholic Beverages, Tobacco and Narcotics
- 3 Clothing and Footwear
- 4 Housing, Water, Electricity, Gas and Other Fuels
- 5 Furnishings, Household Equipment and Routine Maintenance of the House
- 6 Health
- 7 Transport
- 8 Communication
- 9 Recreation
- 10 (A) Education
- 11 (B) Restaurant and Hotels
- 12 (C) Miscellaneous Goods and Services
- 20 (K) Non-Consumption Expenditure

Details of how the expenditure variable was derived can be found in Appendix III.

² For more details, see the Excel document among the [LCF 2013 documentation](#) on the UK Data Service website: 7702_volume_d_expenditure_codes_2013.xls.

Details of variables

Table 1: Variable list

Observations: 5,144

Variables: 14

Variable name	Variable label	Variable type
casenew	Randomly generated case number	Scalar
weighta	Annual weight	Scalar
P550tpr	Total expenditure, by adults & children (top-coded)	Scalar
P344pr	Gross normal weekly household income (top-coded)	Scalar
P425r	Main source of household income	Nominal
A172	Internet connection in household	Nominal
A093r	Economic position of Household Reference Person	Nominal
A094r	NS-SEC 3 class of Household Reference Person	Nominal
A121r	Tenure type	Nominal
SexHRP	Sex of Household Reference Person	Nominal
A049r	Number of persons in household	Ordinal
G018r	Number of adults in household	Ordinal
G019r	Number of children in household	Ordinal
Gorx	Government Office Region - modified	Nominal
weightar	Weight (rescaled)	Scalar
maininc	Main source of household income (recoded, P425-1)	Nominal
income	Income	Scalar
expenditure	Total expenditure (top coded, formerly P550tpr)	Scalar
hhsiz	Household size, number of people in household (recoded)formerly A049r	Nominal

Variable names and suffixes:

The variable names are written in form of codes and are based on the original LCF 2013 dataset. Below is a brief explanation how to read these codes.

Codes suffixed ‘r’

When creating the unrestricted teaching dataset a few variables have been recoded in order to reduce their level of detail for reasons of confidentiality. These variables are suffixed “r” (to keep the original name) and are listed in table 2. SPSS syntax on how they were recoded is given in Appendix II.

Table 2: List of recoded variables

New Recoded	Variable Label	Original variable
P550tpr	Total expenditure	P550tpr (top-coded at 96th percentile)
P344pr	Household income	P344pr (top-coded at 96th percentile)
maininc	Main source of household income	recoded (P425-1)
expenditure	Total expenditure	P550tpr top coded
hhsz	Household size, number of people in household	recoded, formerly A049r

P425r	(2 cat.)	HH income	P425	(6 cat.)
A093r	(4 cat.)	Economic position of HRP	A093	(7 cat.)
A094r	(5 cat.)	NS-SEC 3 of HRP	A094	(12 cat.)
A121r	(3 cat.)	Tenure type	A121	(8 cat.)
A049r	(5 cat.)	No. of persons in HH	A049	(8 cat.)
G018r	(3 cat.)	No. of adults in HH	G018	(6 cat.)
G019r	(3 cat.)	No. of children in HH	G019	(7 cat.)

Codes suffixed ‘t’

Variables with a “t” as a suffix represent total expenditure by children and adults (P550tp). They are aggregated to household level by adding the original (adult) variable to those suffixed ‘c’, (i.e. $P522 + P522c = P522t$).

Anonymisation ‘p’

Variables which are suffixed ‘p’ have been anonymised. This has been done to avoid identification of a household or an individual. Also, a new random ID variable has been created which replaced the “case” variable in the original dataset. This has been done to avoid linking individuals in the LCF Unrestricted Access Teaching Dataset with the original LCF 2013 data.

Downloading the Data

Unlike other government surveys the LCF 2013: Unrestricted Access Teaching Dataset does not require any registration and can be downloaded via the UK Data Service website:
www.ukdataservice.ac.uk.

The dataset is available in two formats: SPSS and STATA.

Appendix I: Data Dictionary

casenew

N	Valid	5144
	Missing	0
Mean		7487.70
Std. Deviation		4309.792
Range		14998
	50	7509.50
Percentiles	75	11205.50
	90	13429.50

weighta

N	Valid	5144
	Missing	0
Range		24.50

P550tpr

N	Valid	5144
	Missing	0
Mean		479.7584
Std. Deviation		292.36523
Range		1144.48
	50	419.9035
Percentiles	75	645.1281
	90	934.4759

P344pr

N	Valid	5144
	Missing	0
Mean		620.4336
Std. Deviation		359.15568
Range		1184.99
	50	563.1500
Percentiles	75	928.1385
	90	1184.9900

P425r

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1 Earned income	2841	55.2	55.2	55.2
	2 Other income	2303	44.8	44.8	100.0
	Total	5144	100.0	100.0	

A172

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1 Yes	4232	82.3	82.3	82.3
	2 No	912	17.7	17.7	100.0
	Total	5144	100.0	100.0	

A093r

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1 Full-time working	2440	47.4	47.4	47.4
	2 Part-time working	536	10.4	10.4	57.9
	3 Unemployed and work related Government Training Programmes	164	3.2	3.2	61.0
	4 Economically inactive	2004	39.0	39.0	100.0
	Total	5144	100.0	100.0	

A094r

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1 Higher managerial, administrative and professional occupations	1589	30.9	30.9	30.9
	2 Intermediate occupations	691	13.4	13.4	44.3
	3 Routine and manual occupations	1045	20.3	20.3	64.6
	4 Never worked and long term unemployed, students and occupation not stated	200	3.9	3.9	68.5
	5 Not classified for other reasons	1619	31.5	31.5	100.0
	Total	5144	100.0	100.0	

A121r

	Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1 Public rented	880	17.1	17.1
	2 Private rented	798	15.5	32.6
	3 Owned	3466	67.4	100.0
	Total	5144	100.0	

SexHRP

	Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1 Male	3156	61.4	61.4
	2 Female	1988	38.6	100.0
	Total	5144	100.0	

A049r

	Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1 1 person	1434	27.9	27.9
	2 2 person	1926	37.4	65.3
	3 3 person	779	15.1	80.5
	4 4 person	670	13.0	93.5
	5 5 person or more	335	6.5	100.0
	Total	5144	100.0	

G018r

	Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1 1 adult	1729	33.6	33.6
	2 2 adults	2826	54.9	88.5
	3 3 adults	422	8.2	96.8
	4 4 and more adults	167	3.2	100.0
	Total	5144	100.0	

G019r

	Frequency	Percent	Valid Percent	Cumulative Percent
Valid 1 No children	3605	70.1	70.1	70.1
2 One child	654	12.7	12.7	82.8
3 Two or more children	885	17.2	17.2	100.0
Total	5144	100.0	100.0	

Gorx

	Frequency	Percent	Valid Percent	Cumulative Percent
Valid 1 North East	251	4.9	4.9	4.9
2 North West and Merseyside	585	11.4	11.4	16.3
3 Yorkshire and the Humber	462	9.0	9.0	25.2
4 East Midlands	424	8.2	8.2	33.5
5 West Midlands	526	10.2	10.2	43.7
6 Eastern	497	9.7	9.7	53.4
7 London	480	9.3	9.3	62.7
8 South East	681	13.2	13.2	75.9
9 South West	429	8.3	8.3	84.3
10 Wales	246	4.8	4.8	89.1
11 Scotland	412	8.0	8.0	97.1
12 Northern Ireland	151	2.9	2.9	100.0
Total	5144	100.0	100.0	

weightar

N	Valid	5144
	Missing	0
Mean		1.0000
Std. Deviation		.34172
Minimum		.00
Maximum		4.70
Percentiles	50	.9428
	75	1.1437
	90	1.3919

maininc

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	0 earnings	2841	55.2	55.2	55.2
	1 other source	2303	44.8	44.8	100.0
	Total	5144	100.0	100.0	

Income

N	Valid	5144
	Missing	0
Mean		620.4336
Std. Deviation		359.15568
Minimum		.00
Maximum		1184.99
Percentiles	50	563.1500
	75	928.1385
	90	1184.9900

expenditure

N	Valid	5144
	Missing	0
Mean		479.7584
Std. Deviation		292.36523
Minimum		30.52
Maximum		1175.00
Percentiles	50	419.9034
	75	645.1281
	90	934.4759

Hhsize

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1	1434	27.9	27.9	27.9
	2	1926	37.4	37.4	65.3
	3	779	15.1	15.1	80.5
	4	670	13.0	13.0	93.5
	5	335	6.5	6.5	100.0
	Total	5144	100.0	100.0	

Appendix II: SPSS Syntax for variables derived for the LCF 2013: Unrestricted Access Teaching Dataset

Recode “Main source of household income” P425 into P425r

```
RECODE P425 (0=2) (1 thru 2=1) (3 thru 6=2) INTO P425r.  
VARIABLE LABELS P425r 'Main source of household income (recoded)'.  
VALUE LABELS P425r 1 'Earned income' 2 'Other income'.  
EXECUTE.
```

Recode “Economic position of Household Reference Person” A093 into A093r

```
RECODE A093 (3=2) (1 thru 2=1) (4 thru 5=3) (6 thru 7=4) INTO A093r.  
VARIABLE LABELS A093r 'Economic position of Household Reference Person (recoded)'.  
VALUE LABELS A093r 1 'Full-time working ' 2 'Part-time working' 3  
'Unemployed and work related Government Training Programmes'  
4 'Economically inactive'.  
EXECUTE.
```

Recode “NS - SEC 12 Class of Household Reference Person” A094 into A094r

```
RECODE A094 (12=5) (1 thru 3=1) (4 thru 5=2) (6 thru 8=3) (9 thru 11=4)  
INTO A094r.  
VARIABLE LABELS A094r 'NS-SEC 3 Class of Household Reference Person  
(recoded from NS-SEC 12)'.  
VALUE LABELS A094r 1 'Higher managerial, administrative and professional  
occupations' 2 'Intermediate occupations' 3 'Routine and manual  
occupations' 4 'Never worked and long term unemployed, students and  
occupation not stated' 5 'Not classified for other reasons'.  
EXECUTE.
```


Recode “Tenure - type” A121 into A121r

```
RECODE A121 (1 thru 2=1) (3 thru 4=2) (5 thru 7=3) (8=2) INTO A121r.  
VARIABLE LABELS A121r 'Tenure - type (recoded)'.  
VALUE LABELS A121r 1 'Public rented' 2 'Private rented' 3 'Owned'.  
EXECUTE.
```

Recode “Household size; number of persons in HH” A049 into A049r

```
RECODE A049 (1=1) (2=2) (3=3) (4=4) (4 thru 9=5) INTO A049r.  
VARIABLE LABELS A049r ' Household size; number of persons in HH  
(recoded)'.  
VALUE LABELS A049r 1 '1 person' 2 '2 person' 3 '3 person' 4 '4 person' 5  
'5 person and more'.  
EXECUTE.
```

Recode “Number of adults” G018 into G018r

```
RECODE G018 (1=1) (2=2) (3=3) (4 thru Highest=4) INTO G018r.  
VARIABLE LABELS G018r 'Number of adults (recoded)'.  
VALUE LABELS G018r 1 '1 adult' 2 '2 adults' 3 '3  
adults' 4 '4 and more adults'.  
EXECUTE.
```

Compute new rescaled weight with mean 1

```
COMPUTE weightar=weighta/5.2181.  
EXECUTE.  
VARIABLE LABELS weightar 'Weight (rescaled)'.
```

Recode "'Number of children" G019 into G019r

```
RECODE G019 (0=1) (1=2) (2 thru Highest=3) INTO G019r.  
VARIABLE LABELS G019r 'Number of children (recoded)'.  
VALUE LABELS G019r 1 'No children' 2 'One child' 3 'Two or more  
children'.  
EXECUTE.
```

**Top-code the expenditure and gross income variables at the
96th percentile.**

```
RECODE P550tp (1175 thru Highest=1175) (ELSE=Copy) INTO P550tpr.  
VARIABLE LABELS P550tpr ' Total expenditure (top coded)'.  
EXECUTE.  
  
RECODE p344pr (1184.99 thru Highest=1184.99) (ELSE=Copy) INTO P344pr.  
VARIABLE LABELS P344pr 'Gross normal weekly household income - top-  
coded'.  
EXECUTE.
```

**Compute a new ID variable with random numbers (from 0-15000) and then
sort dataset by the new ID variable descending.**

```
COMPUTE casenew=TRUNC (UNIFORM(15000)) +1.  
VARIABLE LABELS casenew 'New random ID number'.  
EXECUTE.  
SORT CASES BY casenew (D).  
EXECUTE.
```

Appendix III: P550tp Total Expenditure, by Adults and Children

Below is a brief summary of how the variable **P550tp Total Expenditure, by Adults and Children** was derived. For more detailed breakdown please see the [Living Costs and Food Survey 2013 Volume F: Derived Variable User guide](#).

P550tp Total expenditure, by adults and children (anonymised data)³:- sum of codes P536tp, P537t, P538t, P539t, P540t, P541t, P542t, P543t, P544t, P545t, P546t, P547t, P548t, P549t (see note 7).

This includes the following items:

P536tp ONS housing (net), expenditure by adults and children (anonymised data):- sum of codes P536p and P536cp (see note 7). (as P515tp)

P537t ONS fuel, light and power, expenditure by adults and children - (negative answers possible):- sum of codes P537 and P537c (as P517t)

P538t ONS food, expenditure by adults and children:- sum of codes B260t, P118t (see note 7). (as P518t)

P539t ONS alcoholic drink, expenditure by adults and children:- sum of codes P539 and P539c (see note 7). (as P519t)

P540t ONS tobacco, expenditure by adults and children:- sum of codes P540 + P540c (see note 7). (as P520t)

P541t ONS clothing and footwear, expenditure by adults and children:- sum of codes P541 + P541c (see note 7). (as P521t)

P542t ONS household goods, expenditure by adults and children:- sum of codes P542 and P542c (see note 7) (as P522t).

P543t ONS household services, expenditure by adults and children:- sum of codes P543 and P543c (see note 7).

P544t ONS personal goods and services, expenditure by adults and children:- sum of codes P544 and P544c (see note 7). (as P524t)

P545t ONS motoring expenditure, by adults and children (negative values possible):- sum of code P545 and P545c. (as P525t)

³ P550p is the anonymised total expenditure variable which is released to other users and has been recalculated to include anonymised council tax, water and sewerage charges.

P546t **ONS fares and other travel costs, total expenditure by adults and children:-**
sum of codes B216, B217, B218, B219, B248, P134t (see note 7). (as P526t)

P547t **ONS leisure goods, expenditure by adults and children:-** sum of codes P547 and
P547c (see note 7). (as P527t)

P548t **ONS leisure services, expenditure by adults and children:-** sum of codes P548
and P548c (see note 7). (as P528t)

P549t **Miscellaneous, expenditure by adults and children:-** sum of codes P126t (see
note 7) and B237 + B238.

Note 7) Product codes P515 to P528, P530, P531, P535 to P550 include payments made via
credit, HP, standing order, or deductions from current pay (because they use codes P116 to
P121, P126, P128, P130 to P136 and P153).