

Survey sampling for Family Food

Last updated: December 2015

As the Living Costs and Food Survey is a sample of households and not a census of the whole population, the results are liable to differ to some degree from those that would have been obtained if every single household had been covered. The main factors that affect the reliability of the survey results are sampling variability, non-response bias and some incorrect reporting of certain items of expenditure and income. This note explains the sampling frame, the response rate and accuracy, and how non-response bias is corrected for using sample-based weighting and census population-based weighting.

Background

The figures in Family Food are sourced from The Living Costs and Food Survey run by the Office for National Statistics. One element of the survey - The Family Food Module collects detailed quantity and expenditure information on food and drink household purchases and itemised lists of food and drink eating out purchases for use by Defra.

The Office for National Statistics has overall project management and financial responsibility for the survey while Defra sponsors the specialist food data.

Sampling frame

The Living Costs and Food Survey (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 – the Post Office’s list of addresses. The Northern Ireland sample is drawn as a random sample of addresses from the Land and Property Services Agency list.

The survey is a voluntary sample survey of private households run at household level. The survey is continuous, interviews being spread evenly over the year to ensure that seasonal effects are covered. Each household member over the age of seven keeps a diary of all their expenditure over a 2 week period. The diaries record expenditure and quantities of purchases of food and drink rather than consumption of food and drink.

In 2014 the survey collected the diaries of 121250 people within 5144 households across the United Kingdom.

Response rate

Response for this survey for 2014 was 48% for Great Britain and 60% for Northern Ireland. Response rates for each year since the survey began are shown in Table 1 these rates are in line with other major Government surveys.

Table 1: Response rate for the survey

Year	Response rate	
	Great Britain	Northern Ireland
2001-02	62%	51%
2002-03	58%	56%
2003-04	58%	58%
2004-05	57%	52%
2005-06	57%	50%
2006	55%	54%
2007	53%	54%
2008	51%	54%
2009	50%	56%
2010	50%	59%
2011	54%	62%
2012	52%	57%
2013	48%	61%
2014	48%	60%

Under-reporting is a problem with all dietary surveys but is considered to be lower in the Family Food Module. Its focus on all expenditure with most food items collected from till receipts reduces the scope for under-reporting of household purchases.

Survey errors

Great care is taken in collecting information from households and comprehensive checks are applied during processing, so that errors in recording and processing are minimised. The main factors that affect the reliability of the survey results are sampling variability, non-response bias and some incorrect reporting of certain items of expenditure and income. In order to maximise response, interviewers make at least four separate calls, and sometimes many more, at different times of day on households which are difficult to contact.

Further information is available from the ONS in the report [Family Spending](#), Appendix B: Methodology.

Non-response bias

The households which decline to respond to the survey tend to differ in some respects from those which co-operate. It is therefore possible that their patterns of expenditure and income also differ. The survey data has been weighted to reduce the effect of non-response bias and produce population totals and means. The weights are produced in two stages. First, the data are weighted to compensate for non-response (sample-based weighting). Second, the sample distribution is weighted so that it matches the population distribution in terms of region, age group and sex (population-based weighting). Since 2007, the survey has used population projections from the 2001 Census. From 2007 the non-response classes and weights have also been updated using 2001 Census data. This change in the method of weighting causes very few changes in quantities and expenditure greater than +/-0.5% and no changes in nutrient intakes of more than 0.2%.

Sample-based weighting using the Census

Weighting for non-response involves giving each respondent a weight so that they represent the non-respondents that are similar to them in terms of the survey characteristics. The LCF uses results from the Census-linked study of non-respondents to carry out non-response weighting.

The Census-linked study matched Census addresses with the sampled addresses of some of the large continuous surveys, of which this is one. In this way it was possible to match the address details of these respondents as well as the non-respondents with corresponding information gathered from the Census for the same address. The information collected during the 1991 Census matching work was then used to identify the types of households that were being underrepresented in the survey.

A combination of household variables were analysed to identify which characteristics were most significant in distinguishing between responding and non-responding households. These characteristics were used to produce ten weighting classes with different response rates. For example, in 2005–06, households within each of the weighting classes were assigned a non-response weight, based on the inverse of their response rate. A group with a low response rate was therefore given a high initial weight. The ten weighting classes depend on government office region, metropolitan area, household composition, marital status and sex of household reference person.

Population-based weighting

The second stage of the weighting adjusts the non-response weights so that weighted totals match population totals. As the LCF sample is based on private households, the population totals used in the weighting need to relate to people living in private households. The LCF uses population projections from the 2001 Census. These estimates exclude residents of institutions not covered by the LCF, i.e. those living in bed-and-breakfast accommodation, hostels, residential homes and other institutions. The existing non-response weights were calibrated so that weighted totals matched population totals, for males and females in different age groups and for regions. An important feature of the population-based weighting is that it is done by adjusting the factors for households not individuals.

The weighting is carried out separately for each quarter of the survey. The main reason is that sample sizes vary from quarter to quarter more than in the past. This is due to re-issuing addresses where there had been no contact or a refusal to a new interviewer after an interval of a few months, which results in more interviews in the later quarters of the year than in the first quarter. Quarterly weighting therefore counteracts any bias from the uneven spread of interviews through the year. Quarterly weighting also results in small sample numbers in some of the age/sex categories that were used in previous years. The categories have therefore been widened slightly to avoid this.