What is weighting?

UK Data Service

Author: UK Data Service
Updated: March 2014
Version: 2.0

We would like to acknowledge the late Alasdair Crockett, UK Data Archive and Institute for Social and Economic Research, with whom this guide originated.

With thanks to Professor Natalie Shlomo for her valuable input.

Contents

List of Figures

List of Tables

1. What is weighting?

Many social survey datasets available through the UK Data Service include weighting variables, or weights.

Data producers calculate weights to make the data better represent the population it is designed to represent. A weighting variable assigns a value to each case in the dataset to indicate how much they should be represented in the analysis. Weights can be applied in statistical packages such as SPSS and Stata and can be vital for ensuring that results relate to the population of interest.

This guide explains the main reasons for using weights, how weights work and how to use weighting variables in statistical analyses.

2. Why weight data?

Weights are primarily used to compensate for factors that can make the sample data unrepresentative of the population.

The aim of survey research is usually to make generalisations about a wider population such as the population of the UK. However, the design of a survey and survey non-response commonly lead to some groups being over-represented in the sample. If we do not adjust for survey design and non-response, our estimates of population characteristics can be inaccurate, or biased. The potential for estimates to be biased is an important issue in survey research. Most social survey datasets include weighting variables and the problem applies to descriptive statistics such as means and totals, as well as estimates from more complex statistical models.

Additionally, weights can be used to scale the sample to the population that it was designed to represent (these type of weights are discussed further in section 3.1.5 Scale or grossing weights).

2.1. Survey design

The design of surveys, especially large-scale social surveys, often means that not all units (i.e. individuals or households) in the population have an equal chance of being selected to take part.

When selection probabilities are unequal, the individuals or households with higher probabilities of being selected become over-represented within the sample. Unless we adjust the data to account for the unequal selection probabilities, the sample will not reflect the population and estimates of population characteristics will be biased.

2.1.1. Why are the selection probabilities unequal?

Survey designs often include unequal selection probabilities to increase the number of cases representing specific groups in the population. This practice is frequently used in relation to groups comprising a small proportion of the population such as ethnic minorities. The aim is that by increasing the number of cases in the sample we can make estimates relating to the group more precise.

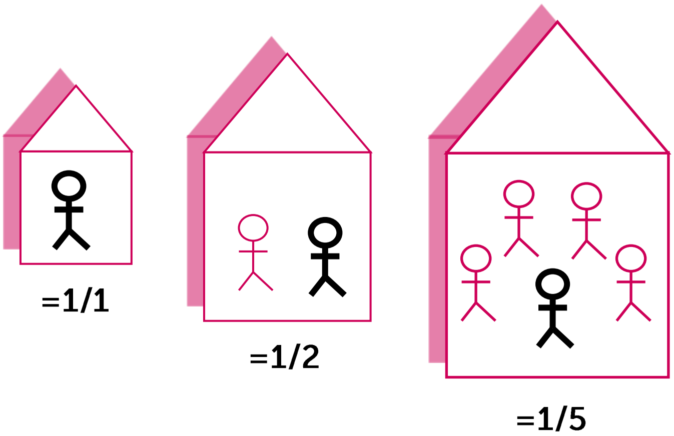
Example: English Housing Survey (EHS), 2010-11
The *English Housing Survey (EHS) 2010-11* oversamples households renting from local authorities or housing associations. Being able to obtain more precise estimates about less common housing types enhances a study of housing. However, over-representing households

in socially rented accommodation will bias estimates of the distribution of different types of tenure in England and related estimates such as the calculation of average rent.

Additionally, selection probabilities can be unequal as a result of the options and resources available to survey designers for sampling a population. For example, survey designers wishing to interview one person per household might only have a list of addresses. When using a list of addresses as a sampling frame, the probability of an individual being selected depends upon how many people live in the household.

Error! Reference source not found. illustrates how selection probabilities vary by household size. In a single person household, the probability of the person living there being selected for interview is 1, whilst individuals in two-person households have a probability of 1/2 and in a five person household 1/5. So, the higher probabilities of being selected mean the sample would over-represent individuals living in smaller households.

Figure 1: Selection probabilities vary by household size



Unequal selection probabilities are a common feature of UK surveys because the most commonly used sampling frame, the Postcode Address File (PAF), is a list of addresses. Since the number of dwellings, households and individuals at each address is not indicated, the PAF cannot be used to draw a simple random sample of households or individuals.

Example: British Social Attitudes Survey, 2010

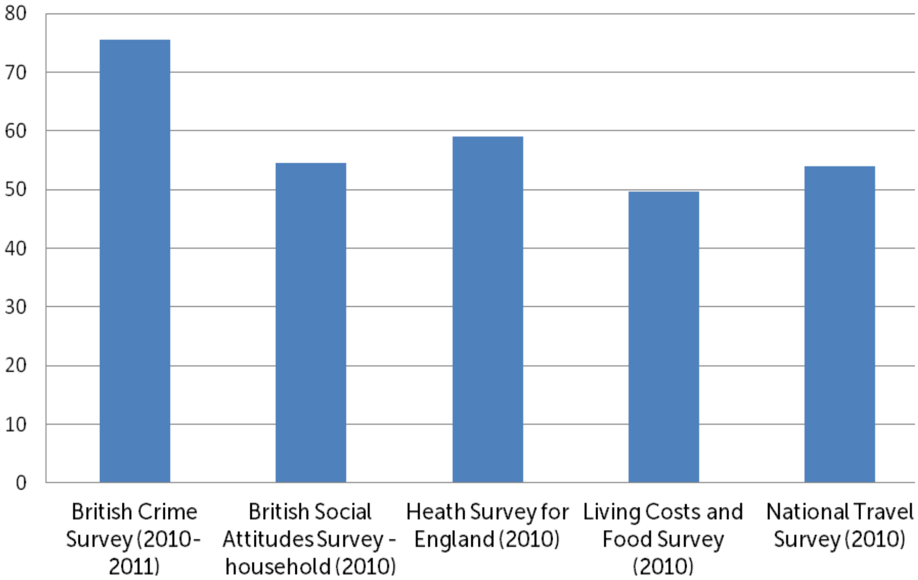
The *British Social Attitudes Survey, 2010* obtained a sample of the adult population in Britain by first selecting a sample of addresses from the PAF. At each selected address, the survey interviewer establishes the number of households, and when there are several, randomly selects one. The interviewer then establishes how many adults live in the household and, when there is more than one, selects an individual to interview at random. The weight included in the dataset adjusts the sample to make those living in households that are larger and where there are multiple households at an address count more in the analysis.

2.2. Non-response

Weights can also be used to limit bias resulting from non-response.

Non-response occurs when selected units do not take part in the survey; typically because individuals refuse or cannot be contacted. Data producers work hard to minimise non-response, for example, by visiting an address more than once and at different times of the day. However, as shown in Figure 2, large-scale social surveys in the UK commonly have response rates between 50 and 70 per cent.

Figure 2: Response rates from a selection of large scale surveys in 2010 (%)



Non-response presents a problem for researchers when the probability of responding relates to the research interests and target variables of the survey. Studies suggest that non-response is not random. Response rates vary by the social, demographic or economic characteristics that interest many social researchers such as age, gender and area deprivation (Groves et al, 2009).

Example: British Crime Survey, 2010-2011
The Technical Report for the *British Crime Survey, 2010-2011* (Fitzpatrick & Grant, 2011) highlights how response to the survey was lower in London and inner city areas, areas classified under the Acorn classification as 'urban prosperity' and where housing was in poor condition. Crucially, since an individual's experience and views of crime are likely to depend upon the area where they live, under-representing those living in certain area types is highly likely to bias estimates relating to crime.

3. How do weights work?

In unweighted analyses each record or 'case' counts as one observation. By contrast, in weighted analyses each case can represent any number of observations. The number of observations you wish the case to count as is given in a weighting variable (e.g. 'weight' shown below).

Here's a very simple example:

Case Number	Age	Weight
1	10	2
2	30	1

The unweighted mean age of the two case data above is

(10 + 30)/2 = 20.

When the weight is applied, case number one counts as two observations, not one, so the mean is now

(2*10) + (1*30)/(2+1) = (10 + 10 + 30) /3 = 16.67.

3.1. Different types of weights

The final survey weight is typically a multiplication of several weighting adjustments, which are compensating for factors such as the survey design, non-response, and aligning population proportions. The different types of weighting adjustment are described below.

3.1.2. Design weights

Weights that adjust for the unequal selection probabilities due to the survey design tend to be called design or probability weights.

The survey design is under the control of the survey designer and therefore selection probabilities are known. These known selection probabilities can be used to weight the data by calculating a weight for each case. The general formula for calculating the weights is 1 divided by the probability of selection due to the survey design.

Table 1 illustrates the basic calculation for the single person, three person and five person households from example from Figure 1.

Table 1: Calculating design weights to adjust for household size

Living in…	Probability of being selected	Inverse of selection probability	Weight
Single person household	1/1	1/(1/1)	1
Three person	1/2	1/(1/2)	2
Five person	1/5	1/(1/5)	5

In a single person household the probability of being selected is 1/1, or 1 and a design weight in this case would be calculated as 1/(1/1), which is 1.

In a three person household the probability of an individual being selected is 1/3, so here a design weight for this case would be 1/(1/3), which is 3.

Similarly, in a five person household the probability of an individual being selected reduces to 1/5 and the design weight would be 1/(1/5), which is 5.

These weights would therefore adjust for the fact that the person in the three person house was three times less likely to be included in the sample than the individual living alone and the person in the five person household was five times less likely to be included in the sample.

To illustrate the adjustment made by weighting, Table 2 shows the unweighted and weighted distributions for two variables from the 2010 British Social Attitudes Survey (BSA). The first variable indicates the number of people living in the household. Reflecting the higher probability of being selected, the unweighted data suggests the proportion of people living in single person households is far greater than the weighted data. Equally, the proportion living in three person households becomes greater when the data has been weighted.

Table 2: Effects of weighting on two questions from the BSA survey

	Unweighted percentage	Weighted percentage
Q46: Number living in household, including respondent		
1	30.2	17.8
2	33.7	34.3
3	15.8	19.7
4+	20.3	28.2
Total	100.0	100.0
Q517: Type of new homes most needed in local area -Bungalows		
Mentioned Bungalows	13.0	11.7
Not mentioned Bungalows	84.6	86.2
Don't know	2.4	2.0
Total	100	100

N= 3297

Source: National Centre for Social Research, *British Social Attitudes Survey, 2010* [computer file]. Colchester, Essex: UK Data Archive [distributor], February 2012. SN: 6969, http://dx.doi.org/10.5255/UKDA-SN-6969-1

The second variable shows responses to a question about types of housing needed in the local area. Here, the unweighted estimate of the percentage mentioning bungalows is higher than the weighted estimate. However, it is noticeable that the effect of weighting on this estimate is smaller than in relation to household size. Indeed, survey design can affect each variable in the dataset in a different way. The magnitude of the effect depends upon how closely a variable is associated with the factors involved in the survey design.

3.1.3. Non-response weights

Like design weights, non-response weights compensate by giving more weight to those under-represented. Non-response weights are calculated with a general formula of 1 divided by the response probability. However, survey designers do not have control over survey non-response. As a result, the response probabilities and weight adjustments need to be based on assumptions about the non-response mechanism and the limited information available about non-respondents.

For many large-scale social surveys the response probabilities are derived using information from the sample frame and data about those not responding collected by the interviewer such as the type of house and neighbourhood.

Response probabilities are estimated from the response rates calculated for population sub-groups (often termed weighting classes). The weighting classes are defined by socio-demographic factors such as gender, age and area or housing types. For each weighting class, a response rate is calculated by dividing the number of respondents by the size of the weighting class in the sample. All respondents within a weighting class receive a response probability equal to the response rate. Then, a non-response weight is calculated as the inverse of the response rate.

Alternatively, the probability of response can be estimated from a parametric model (such as the logistic regression model). There are also non-parametric techniques such as a 'Segmentation Analysis', which splits the sample into 'nodes' varying the most in terms of response rate.

When calculated appropriately, non-response weights can reduce bias in population estimates. However, non-response weights are only estimates, relying on the information available to define the weighting classes, assumptions about the non-response mechanism and the contrast between respondents and non-respondents. When calculating non-response rates, it is assumed that within a weighting class the characteristics of respondents and non-respondents are the same. When this assumption does not hold, population estimates may still be biased.

For example, the propensity to take part in activities such as volunteering or voting is likely to be positively associated with agreeing to take part in a social survey. As a result, the problem of bias may continue to occur in research on social and political engagement with surveys potentially over-representing how many people take part.

3.1.4. Calibration weights and post-stratification

Non-response alongside sampling variability (chance) can mean characteristics of a sample, for example, the proportion of men and women, vary from the population.

Calibration, or post-stratification, weights adjust sample proportions to match those in the population using information from other sources, commonly Census data. This type of weight is used to improve the accuracy and precision of population estimates and is common in surveys carried out by statistical agencies such as the Labour Force Survey (LFS), organised by the Office for National Statistics (ONS).

Like with non-response weights, calibration weights are derived by defining groups using socio-demographic variables where the population total is known such as age, gender and area.

In the simplest case of calibration (post-stratification), the response rate is calculated within each group at the population level. We first calculate the number of respondents in the 'population', based on the original design weights, and then divide by the known population total. The calibration weight is then the inverse of the 'population-level' response rate.

Typically, calibration weighting uses multiple schemes, or partitions and cross-classifications to define groups. For example, in the LFS, three partitions are defined: (1) Local Authority Districts, (2) Great Britain and Northern Ireland by male/female and age group; and (3) Male/female by Government Office Region and age groups.

The weights are calculated so that the totals for all groups are proportionate to those in the population whilst remaining close to the design weights. The calculation process is mathematically complex and different methods, and software, have been developed. For example, since 2007, the ONS has applied a Generalised Regression (GREG) approach using software developed by Statistics Canada. Previously, they used a raking ratio method (see Labour Force User Guide Volume 1 for further information). A key difference between methods is whether the weights are calculated by calibrating to all partitions simultaneously or where adjustments are made in relation to each partition separately in an iterative process. The Generalised Regression (GREG) approach is viewed as more statistically robust and efficient.

3.1.5. Scale or grossing weights

Weights can be used to scale the sample to look like a dataset of a different size. Most frequently, weights are scaled to sum either to the original sample total or to the population total.

Weights that sum to the population total are called grossing-up weights. The Labour Force Survey (LFS) uses this type of weighting by employing up-to date estimates of the population total. The effect of the weights can be seen in Table 3, which shows the unweighted and weighted frequencies for a variable, indicating those in permanent and non-permanent employment. Comparison of the columns shows how the weight dramatically increases the frequencies. These types of weights make it easier to describe the population and the prevalence of social phenomena. However, in some statistical software packages grossing-up weights cause errors when calculating the precision of estimates (see Section 4.2 for further details).

Table 3: Effect of weighting the LFS – scaling to the population

	Unweigted Frequency (%)	Weighted Frequency (%)
Permanent employment	37,128 (37.2)	23,515,226 (37.8)
Not permanent	2,485 (2.5)	1,650,350 (2.7)
Does not apply	60,288 (60.3)	37,094,742 (59.6)
Total	99,901	62,260,318

The table compares weighted and unweighted frequency (and percent) for the variable JOBTYPE.
Source: Office for National Statistics. Social Survey Division and Northern Ireland Statistics and Research Agency. Central Survey Unit, *Quarterly Labour Force Survey, July - September, 2012* [computer file]. *3rd Edition.* Colchester, Essex: UK Data Archive [distributor], March 2013. SN: 7174, http://dx.doi.org/10.5255/UKDA-SN-7174-3

4. How do I use weights?

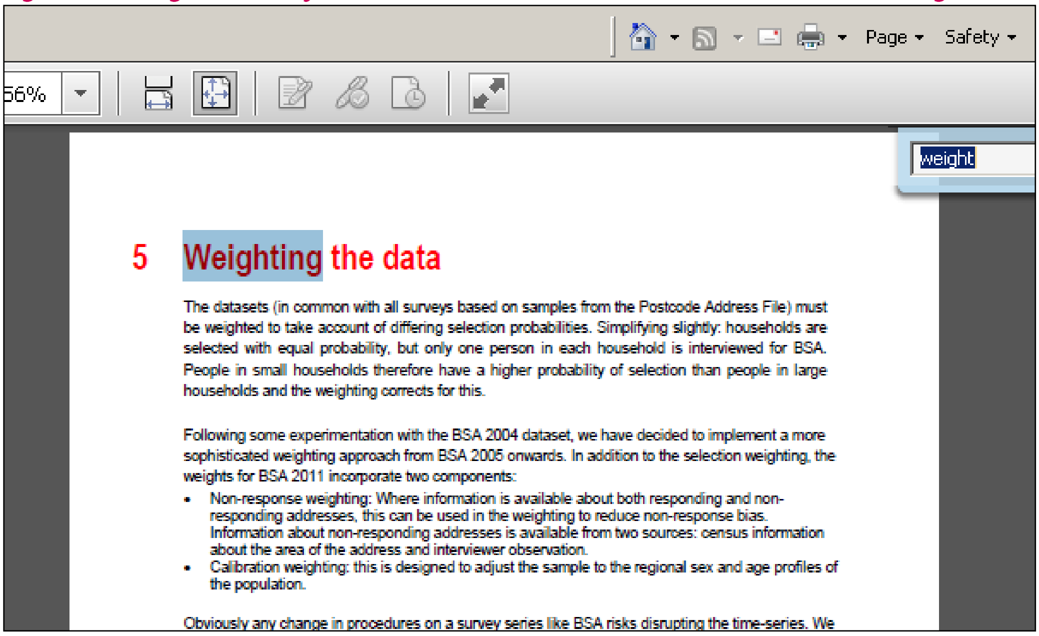
Weights appear in datasets with the other variables and can be applied using functions in statistical packages.

4.1. How do I find the weighting variables in my dataset?

Names for weighting variables are often linked to 'weight'. For example, 'wtfactor' is the "Final BSA weight" in the British Social Attitudes Survey 2010 and 'indivwgt' is the "Individual weight" in The British Crime Survey 2010-2011.

The documentation, often a 'User Guide', accompanying the dataset, will tell you which variables are the weights and how they were calculated. Figure 3 shows the relevant section in the user guide for the British Social Attitudes (BSA) 2011 (Stafford, 2011). As in this example, most user guides are in PDF form, which makes it easy to search for information using *Control-F* and terms such as "weight".

Figure 3: Using the survey documentation to find information about the weights



5 Weighting the data

The datasets (in common with all surveys based on samples from the Postcode Address File) must be weighted to take account of differing selection probabilities. Simplifying slightly: households are selected with equal probability, but only one person in each household is interviewed for BSA. People in small households therefore have a higher probability of selection than people in large households and the weighting corrects for this.

Following some experimentation with the BSA 2004 dataset, we have decided to implement a more sophisticated weighting approach from BSA 2005 onwards. In addition to the selection weighting, the weights for BSA 2011 incorporate two components:

- Non-response weighting: Where information is available about both responding and non-responding addresses, this can be used in the weighting to reduce non-response bias. Information about non-responding addresses is available from two sources: census information about the area of the address and interviewer observation.
- Calibration weighting: this is designed to adjust the sample to the regional sex and age profiles of the population.

Obviously any change in procedures on a survey series like BSA risks disrupting the time-series. We

4.1.6. There is more than one weight in my dataset, how do I choose which one to use?

Some datasets contain more than one weighting variable. The different weights may relate to different samples (such as a 'core' and 'ethnic boost' sample) or sample units (such as individuals or households). You must consult the documentation that comes with the dataset (e.g. the user guide) to find out what adjustments the weights make and the circumstances in which to use each weight.

For example, the British Crime Survey (BCS) 2010-2011 documentation shows that the dataset has three weights: 'indivwgt' for individual based analyses, 'hhdwgt' for household-based analysis and 'weighti' for incident-based analysis (Fitzpatrick & Grant, 2011).

Figure 4: Selection probabilities vary by household size

SECTION 4: BCS analysis

There are three main types of analysis that can be carried out on BCS data: individual-based analysis, household-based analysis and incident-based analysis. Both individual and household analysis can be used to produce incidence and prevalence rates for different crime types. Each of these types of analysis needs to take into account the appropriate weighting of the data.

4.1 INDIVIDUAL-BASED ANALYSIS

Individual-based analysis is carried out when the intention is to make statements about the characteristics, attitudes or experiences of adults in the sample. Analysis of attitudinal questions is individual-based, as is analysis of victims of personal crimes (such as assault). All individual-based analysis should be weighted by **indivwgt** (weighta prior to 1996 survey).

4.2 HOUSEHOLD-BASED ANALYSIS

Household based analysis is carried out when the intention is to make statements about the characteristics or experiences of households in the sample. The most common type of household-based analysis is analysis in which statements are made about households who were victims of household crimes. All household-based analysis should be weighted by **hhdwgt** (weightb prior to 1996 survey).

4.3 INCIDENT-BASED ANALYSIS (VICTIM FORM ANALYSIS)

Incident-based analysis is carried out when the intention is to make statements about characteristics of incidents of crime. It can be used to make statements about the timing, location or perceived seriousness of offences. Incident-based analysis is always carried out on the Victim Form dataset (see Section 4). All incident-based analysis should be weighted by **weighti**.

Although since 1992 incidents occurring outside of England and Wales have been given a valid offence code (see Section 2), for incident-based analysis only those incidents which occurred within England and Wales should be retained. This should be done by selecting cases based on responses to the variables '**victarea**' and '**wherhapp**'.

4.1.7. Do all survey datasets have weights?

Not all survey datasets have weights. This could be because the survey successfully achieved a representative sample of the population and no other adjustments have been made by the data provider. However, almost all large-scale surveys contain weights.

4.2. How do I apply the weights?

Statistical packages generally have functions to allow you to weight your data. In order to weight the data, you need to know the name of the weighting variable in your dataset. You also need to understand what the weight does and be sure that it is the correct weight to use for your analyses by reading the accompanying documentation.

In SPSS: Using the drop-down menus: Data > Weight Cases... You can also weight data using commands in syntax.
In Stata: Most Stata commands can deal with weighted data. Stata allows four kinds of weights. For more information, open Stata and type 'help weight'

SPSS users need to be aware that the basic method for applying a weight in SPSS can underestimate standard errors. This is because the weighted total is used to calculate the standard error rather than the actual sample size. Correct analysis of weighted data can be done using the Complex Sample module; however, the module cannot be used with all types of analysis.

The UK Data Service guides to using Stata and SPSS include sections on weighting.

4.2.8. *Do I always use weights, whatever I'm doing?*

When you are doing your initial exploratory analysis and checking that your sample size is adequate, you should look at the unweighted data. This is because some weights 'gross-up' the sample size and therefore the weighted data can give the impression of a larger sample. To decide whether you have enough cases on which to base your analyses, you must look at the unweighted data to see the sample numbers, not the grossed-up numbers. Once you are sure that you have enough data and proceed with your analyses, in most cases you will need to weight the data.

4.2.9. *Changes to weighting variables*

Once calculated, design weights are not usually changed; however, if better information becomes available to the data provider, non-response and post-stratification weights may be subsequently altered. This practice can be seen in the Labour Force Survey, when in 2007, the ONS implemented a re-weighting programme where weights were re-calculated using more up-to-date population estimates and a new method of calculating calibration weights (see discussion above).

4.3. Survey design, weighting and precision of estimates

A final point to be aware of is that survey design and weighting adjustments affect the precision of estimates.

4.3.1. *Precision*

The precision of an estimate is generally referred to by the standard error, which indicates how close an estimate such as a mean is likely to be to the population value.

Analyses of weighted data, whilst reducing bias, may be less precise than for unweighted data (this is because the method used to adjust the data and produce the estimates may be less statistically efficient). However, calibration weights and post-stratification weights can increase the precision of estimates when what is being estimated correlates with the weighting classes, or calibration groups.

4.3.2. *Design effects*

More generally, users of survey data should be aware of how features of survey design affect the precision of estimates and how we can account for this in our analysis.

As indicated above, oversampling small groups can improve precision. This feature of survey design is a form of stratification. Stratification is a process of dividing the population into groups (strata) and sampling from each with the aim of ensuring the sample is distributed across the groups. Stratifying a sample generally increases precision.

Additionally, survey designs often include multiple stages of sampling, for example, by first selecting a sample of areas (termed Primary Sampling Units, or PSUs) and then sampling addresses, households or individuals within these areas. For UK surveys, Postcode Sectors are commonly sampled first and then addresses are selected within areas. Two-stage, or 'clustered', sample designs mean those selected to take part are geographically clustered together. Hence, clustering can reduce the costs of conducting a survey by cutting interviewer travel time. However, because households in an area tend to be more similar than the population in regards to many social characteristics, clustering can reduce the precision of population estimates.

Typically, statistical packages calculate the precision of estimates, i.e. standard errors, on the assumption that data comes from a simple random sample. Hence, the estimated precision of an estimate may be incorrect. Applying weights does not adjust for these design effects on the precision of estimates. However, in some statistical software packages you can make the adjustments using other commands.

5. References and resources

References in text

Fitzpatrick A and Grant C. (2011) *The 2010/11 British Crime Survey (England and Wales)*, TNS-BMRB.

Groves, R.M., Fowler, F.J., Couper, M., Lepkowski, J.M., Singer, E. and Tourangeau, R. (2009). *Survey Methodology*, 2nd Edition. Hoboken, New Jeresy: Wiley Series in Survey.

Office for National Statistics. (2011) *Labour Force Survey User Guide Volume 1 – LFS Background and Methodology*, Version 1.0 August 2011.

Stafford, R. (2011) *British Social Attitudes 2011 User Guide,* NatCen.

Other references/resources

Lehtonen, R. and Pahkinen, E.J. (1995) Practical Methods for Design and Analysis of Complex Surveys. Chichester: John Wiley & Sons.

Lohr, S.L. (2010) Sampling: Design and Analysis, 2nd edition. Boston: Brooks/Cole

P|E|A|S (Practical Exemplars and Survey Analysis)
www.restore.ac.uk/PEAS/about.php

Other guides

- What is Stata?
- What is SPSS 20 for Windows?
- The R guide to UK Data Service key UK surveys
- What is complex sample design?

21 March 2014

T +44 (0) 1206 872143
E help@ukdataservice.ac.uk
W ukdataservice.ac.uk

The UK Data Service delivers quality social and economic data resources for researchers, teachers and policymakers.

UK Data Service

E·S·R·C ECONOMIC & SOCIAL RESEARCH COUNCIL