English Housing Survey

Technical Advice Note

Data quality: 2011-12 Update

This document/publication is also available on our website at www.gov.uk/dclg

Any enquiries regarding this document/publication should be sent to us at:

Department for Communities and Local Government
Eland House
Bressenden Place
London
SW1E 5DU
Telephone: 030 3444 0000

July, 2013

ISBN: 978-1-4098-3934-7

Content

Introduction

Non-response and missing data

Sampling and measurement error
- Sampling error
- Estimating sampling error
- Confidence intervals for percentages
- Measurement error
- Systematic error

Surveyor variability
- Measuring between-surveyor variability
- Taking account of between-surveyor variability

Introduction

1. This is one of a series of Technical Advice Notes about the methodology of the English Housing Survey (EHS), to assist users in their analysis and interpretation of the survey findings.

2. This note outlines the sources of error affecting the quality of EHS data. It covers:

   - the impact of non-response and missing data
   - sampling and measurement error
   - surveyor variability

   Each of these sources of error is discussed separately below.

Non-response and missing data

3. It is essential that the EHS provides a representative picture of housing in England. The sampling structure was specifically designed to provide such a picture.

4. Inevitably, not all of the addresses originally issued for the survey are retained in the final dataset. A few will prove not to be dwellings, and others will be lost due to non-response or incomplete data. In order to produce good quality, representative results from the survey, it is important to check whether valid but non-responding cases are typical of those that remain and, if not, to counter any resulting response bias in the grossed data set.

5. Where non-response biases were found at any stage of the survey, adjustments were made to the responding cases in the grossing procedures for that stage. More information about this process is given in the Technical Advice Note on Sampling and weighting, available here: https://www.gov.uk/government/publications/english-housing-survey-technical-advice

Sampling and measurement error

6. Like all estimates based on samples, the results of the EHS are subject to various possible sources of error. The total error in a survey estimate is the difference between the estimate derived from the data collected and the (unknown) true value for the population. The main sources of error are random error, measurement error and systematic error.

Sampling error

7. Sampling error is the error that arises because the estimate is based on a sample survey rather than a full census of the population. The results obtained for any single sample may, by chance, differ from the true values for the population but the difference would be expected to average to zero over a number of repeats of the survey. The amount of variation depends on the size of the sample and the sample design and weighting method.

8. A measure of the impact of the variation introduced by the sample design and the weighting is the design factor (deft). This is evaluated relative to the error that would have been produced had the survey been carried out using a simple random sample. A deft greater than one shows that the design and weighting have increased the variability of the estimate and increased the measure of the standard error relative to the reference.

Estimating sampling error

9. It is important to be able to estimate the size of the sampling error when interpreting the survey results. The size of the sampling error depends on the size of the sample; in general, sampling error is potentially larger in smaller samples.

10. A frequently used method of assessing the magnitude of sampling errors is to calculate a confidence interval for an estimate. This is an interval within which one can be fairly certain that the true value lies. The following section explains how to calculate 95% confidence intervals, using a method from standard statistical theory for large samples.

Confidence intervals for percentages

11. Although the estimates produced from a sample survey will rarely be identical to the population value, statistical theory allows us to measure the accuracy of any survey result. Standard errors can be estimated from the results obtained for the sample, and these allow calculation of confidence intervals which give an indication of the range in which the true population value is likely to fall.

12. The simplest method of estimating a standard error and the resulting confidence interval assumes that the sample in question is a simple random sample. The EHS is based on a simple random sample, so these standard confidence intervals are useful to give an approximation to the size of standard errors, particularly given that more accurate calculations are not quick to carry out.

13. The 95% confidence interval for a percentage estimate, p, is given by the formula:

$$p \pm 1.96 se(p)$$

where se(p) represents the standard error of the percentage and is calculated by:

$$se(p) = \sqrt{\frac{p(100-p)}{n}}$$

where n is the unweighted sample size.

14. Estimating confidence intervals for results in this way, based on the assumption of a simple random sample, which has no stratification, is straightforward. For percentages based on the full household sample, the sample size, n, is the unweighted sample total; i.e. 13,829 households in 2011-12. The sample size for the '2011' (2-year) component of the paired dwelling sample is 14,951.

15. The simplest approximation of a 95% confidence interval for a percentage, assuming a simple random sample, may be obtained using the lookup table in Annex 1 at the end of this note. The confidence interval can be calculated by reading off the closest figure from the table, where the estimated percentages are shown as rows and the unweighted sample sizes as columns, and then adding and subtracting this figure from the estimated percentage, as in the following examples:

- In 2011-12, around two-thirds (65.3%) of all households were owner occupied. This figure is based on an unweighted sample of 13,829 households. The corresponding number from the 35^{th} row and 10^{th} column of the lookup table in Annex 1 is 0.8%, giving a confidence interval of 64.5% to 66.1%.

- In 2011, around a quarter (26%) of dwellings were semi-detached. This figure is based on the 2-year unweighted sample of 14,951 dwellings. The corresponding number from the 26^{th} row and 7^{th} column of the lookup table in Annex 1 is 0.7%, giving a confidence interval of 25.3% to 26.7%.

16. Confidence intervals can be calculated more accurately by using the formula above. For the first example above, the standard error is given by

$$se(p) = \sqrt{\frac{65.3 * 34.7)}{13{,}829}} \quad =0.405$$

so the confidence interval is $65.3 \pm 1.96 * 0.405$ or 64.5% to 66.1%.

17. However, this method still only gives an approximation to the confidence interval, because it assumes a simple random sample. Although the sample initially drawn for the EHS is a simple random sample, the subsequent selection of cases for the dwelling sample is stratified by tenure. In addition, the bias adjustments involved in the weighting of both the household and dwelling samples also affect the size of the confidence intervals. As a result, standard errors calculated using the simple random sample method will only give a rough guide and if more accurate standard errors are required, these need to be calculated using a sample design factor (deft).

18. This design factor is a measure of the impact of the variation introduced by the sample design and the weighting. This is evaluated relative to the error that would have been produced had the survey been carried out using a simple random sample[1]. A deft greater than one shows that the design and weighting have increased the variability of the estimate and increased the measure of the standard error relative to the reference. The 2011-12 EHS household sample is effectively a simple random sample and the deft arises solely from the weighting adjustments, but for the dwelling sample the deft includes the effect of sub-sampling by tenure and the corresponding weighting adjustments.

19. Examples of standard errors and confidence intervals calculated using the appropriate design factors are given in each of the detailed EHS annual reports: https://www.gov.uk/government/organisations/department-for-communities-and-local-government/series/english-housing-survey#publications.

20. We cannot provide full confidence intervals for all survey measures, but by comparing confidence intervals derived from the lookup table in Annex 1 with those shown in the relevant annual report it is possible to gauge to what extent the deft impacts on the size of confidence intervals.

Measurement error

21. Error may also arise due to inaccuracies in individual measurements of survey variables because of the inherent difficulties of observing, identifying and recording what has been observed. Measurement error may occur randomly, or may reflect a problem experienced by most or all interviewers or surveyors. Efforts are made to minimise these effects through interviewer and surveyor training and through pilot work.

[1] Technically, the deft is the estimate of the standard error produced under the complex design divided by the standard error under an equally weighted simple random sample.

22. Although measurement error may occur at the interview stage of the survey, there are rather more practical difficulties in assessing the condition of an individual dwelling than the characteristics of a household. These difficulties mainly stem from the technical problems in the diagnosis and prognosis of any defects found in the dwelling. Difficulties are found particularly in more subjective assessments such as the state of repair.

Systematic error

23. Systematic error, or bias, covers those sources of error which will not average to zero over repeats of the survey. Bias may occur, for example, if certain sections of the population are omitted from the sampling frame, if non-respondents to the survey have different characteristics to respondents, or if interviewers unintentionally systematically influence responses in one way or another. When carrying out a survey, substantial efforts are put into the avoidance of systematic errors but it is possible that some may still occur.

Surveyor variability

24. It is quite possible that two surveyors inspecting a given dwelling may have different views on the extent and severity of disrepair and the work needed to remedy it. Assessments of the condition of the area surrounding the dwelling are also prone to subjective variation. This leads to surveyor variability.

25. Estimates of measures such as disrepair rates in the dwelling stock are based on individual surveyor assessments and are dependant on the 'average performance' of all the surveyors. However, individual surveyors will produce assessments which may vary from this average. Thus there is some uncertainty or error associated with such estimates and the greater the variability between surveyors the greater is this error. It is therefore important to control this variability as much as possible and to understand the effect that any residual variability can have on the survey results.

26. Experience has shown that surveyor variability cannot be completely eliminated or even reduced to an insignificant level, but precautions are taken during the EHS to control its impact:

- by using a large number of surveyors, and setting a limit of 60 on the number of surveys any one surveyor can complete overall, and ensuring that surveyors' workloads are spread across more than one geographical area; and
- by ensuring that the surveyors are provided with a rigorous and uniform briefing, designed to minimise subjectivity, which is backed up by survey manuals, supervision in the field, refresher briefings, and the use of calibration workbooks

27. The English House Condition Survey (EHCS) – the precursor to the physical survey element of the EHS – used calibration workbooks as a means of detecting any significant shift in surveyor marking, or 'surveyor drift', between surveys. These were introduced for the EHS during the 2009-10 and 2010-11 survey years. The workbooks were undertaken at the end of the last fieldwork period (either Q3 or Q4) by all surveyors working on the survey during that year. The electronic workbooks contained descriptions and photographs of a number of dwelling faults, and surveyors are asked to record their assessments as they would in the field. The faults were chosen to cover a range of dwelling elements, building types and levels of severity.

28. The workbooks were intended to measure the aspect of surveyor variability that arises from surveyors making different judgements about exactly the same information. Previous work has indicated that surveyors do tend to identify the same problems in a given dwelling, but that they can differ in the work that they specify to remedy these problems. For example, three surveyors looking at the same roof may agree that some slates have slipped and others are missing. However, one surveyor may say that because it is not leaking, no work is needed now but it should be replaced within 10 years; another may say that it should be repaired now and replaced within 15 years, and the third may say it should all be replaced now.

29. The surveyors' responses in the workbooks were used to devise a number of measures including: total estimated costs of all repairs required in the next 10 years specified across all examples, whether specific examples do not meet the decent homes criteria under modernisation and disrepair and the proportion of repairs marked as requiring urgent attention. These measures were then compared with those derived from calibration workbooks from previous years. Statistical analysis is then used to establish whether there have been any significant changes in these measures over time.

30. Results from the 2009-10 exercise showed no significant difference overall in the surveyors' assessments of repair costs, the overall distribution of which remained similar through 2001 to 2009-10. There was no calibration exercise undertaken in 2011-12, the next scheduled calibration exercise is due to be undertaken in 2014-15 under the next EHS contract.

Measuring between-surveyor variability

31. Despite the rigorous surveyor training program, it is natural that a degree of personal judgement and subjectivity will still affect surveyors' assessments. For example, some surveyors will be more likely, after weighing the evidence, to conclude that a particular dwelling needs a new roof, whereas others will be more likely to conclude that the roof can be adequately repaired. This between-surveyor variability is an additional source of variance in estimates from the physical survey data, and can be measured by estimating the correlated surveyor variance.

32. A Surveyor Variability Study (SVS) was conducted using data from the 2009-10 physical survey fieldwork for the EHS to analyse the effects of systematic surveyor variability on the precision of estimates from the physical survey. This involved a call-back exercise in which 300 properties were re-surveyed by a second surveyor and the results were compared. The objectives of the study were to:

- compare variability between surveyors
- highlight key survey measures on the EHS which were either subject to high variability, or low variability
- produce evidence to improve briefing of surveyors and to improve form design
- assess impact of variability on complex standard errors
- compare the levels of variability between the EHCS and EHS

33. This SVS took a similar form to the 2003-04 SVS, which was conducted on the EHCS. Information on the 2003-04 SVS can be found in the 2007 EHCS technical report: http://webarchive.nationalarchives.gov.uk/20120919132719/www.communities.gov.uk/publications/housing/ehcstechnicalreport2007

34. The 2009-10 SVS focused on 75 key survey measures spanning 5 broad topic areas:

- stock profile
- amenities and services
- external environments
- stock condition
- energy performance

35. The 2009-10 SVS found that, overall, there was a high level of agreement between surveyors' assessments of properties: 28 of the 75 survey measures included in the study had a low correlated surveyor variance, and high level of agreement when taking into account agreement by chance and 19 measures were found to have an acceptable level of agreement when taking both chance agreement and correlated surveyor variance into account. However, 28 survey measures were identified as having potentially low levels of agreement. These survey measures tended to be found in the topic areas covering external environments and stock condition and typically required a surveyor's opinion on topics with generally high degrees of variability; e.g. an opinion of the local area or the condition of a property.

36. Revised standard errors and confidence intervals were calculated taking into account any potential bias due to surveyor variability. These were calculated for 251 response options of the 75 survey measures included in this study. The majority of response options experienced insignificant changes to their standard error and confidence intervals. The largest changes to standard errors (and consequently decreasing precision surrounding the estimates) were found in the external environment measures.

37. The 2009-10 SVS found that the overall level of variability on the EHS is in line with the EHCS. We can therefore broadly conclude that variability has remained consistent across the two different survey designs.

38. More information on the 2009-10 SVS – including a more detailed discussion of the findings – can be found in the 2009-10 SVS summary report, published on the DCLG website: https://www.gov.uk/government/publications/english-housing-survey-technical-advice

Taking account of between-surveyor variability

39. The standard error calculations described earlier, which take account of the complex weighting of the survey, and the effect of sub-sampling for the dwelling sample, only partly reflect the effect of between-surveyor variability. In consequence, they are biased downwards and the confidence intervals calculated from them are likely to be too narrow. Using the correlated surveyor estimates from the SVS, it is possible to estimate the size of these downward biases in the standard error estimates and make an adjustment.

Annex 1

Lookup table for calculating 95% confidence intervals for a percentage.

This table assumes a simple random sample and no adjustment for stratification or other design factors.

p \ n	17,500	17,000	16,500	16,000	15,500	15,000	14,951	14,500	14,000	13,829	13,500	13,000	12,500	12,000	11,500	11,000	10,500	10,000
1	0.1	0.1	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.2
2	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.3	0.3	0.3	0.3	0.3
3	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3
4	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.4	0.4	0.4	0.4	0.4
5	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4
6	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.5	0.5
7	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.5	0.5	0.5	0.5	0.5
8	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5
9	0.4	0.4	0.4	0.4	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.6
10	0.4	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.6	0.6	0.6
11	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.6	0.6	0.6	0.6	0.6
12	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.6	0.6	0.6	0.6	0.6	0.6	0.6
13	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.7
14	0.5	0.5	0.5	0.5	0.5	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.7	0.7
15	0.5	0.5	0.5	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.7	0.7	0.7	0.7
16	0.5	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.7	0.7	0.7	0.7	0.7
17	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.7	0.7	0.7	0.7	0.7	0.7
18	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.7	0.7	0.7	0.7	0.7	0.7	0.8
19	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.7	0.7	0.7	0.7	0.7	0.7	0.7	0.8	0.8
20	0.6	0.6	0.6	0.6	0.6	0.6	0.6	0.7	0.7	0.7	0.7	0.7	0.7	0.7	0.7	0.7	0.8	0.8
21	0.6	0.6	0.6	0.6	0.6	0.7	0.7	0.7	0.7	0.7	0.7	0.7	0.7	0.7	0.7	0.8	0.8	0.8
22	0.6	0.6	0.6	0.6	0.7	0.7	0.7	0.7	0.7	0.7	0.7	0.7	0.7	0.7	0.8	0.8	0.8	0.8
23	0.6	0.6	0.6	0.7	0.7	0.7	0.7	0.7	0.7	0.7	0.7	0.7	0.7	0.8	0.8	0.8	0.8	0.8
24	0.6	0.6	0.7	0.7	0.7	0.7	0.7	0.7	0.7	0.7	0.7	0.7	0.7	0.8	0.8	0.8	0.8	0.8
25	0.6	0.7	0.7	0.7	0.7	0.7	0.7	0.7	0.7	0.7	0.7	0.7	0.8	0.8	0.8	0.8	0.8	0.8
26	0.6	0.7	0.7	0.7	0.7	0.7	0.7	0.7	0.7	0.7	0.7	0.8	0.8	0.8	0.8	0.8	0.8	0.9
27	0.7	0.7	0.7	0.7	0.7	0.7	0.7	0.7	0.7	0.7	0.7	0.8	0.8	0.8	0.8	0.8	0.8	0.9
28	0.7	0.7	0.7	0.7	0.7	0.7	0.7	0.7	0.7	0.7	0.8	0.8	0.8	0.8	0.8	0.8	0.9	0.9
29	0.7	0.7	0.7	0.7	0.7	0.7	0.7	0.7	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.9	0.9
30	0.7	0.7	0.7	0.7	0.7	0.7	0.7	0.7	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.9	0.9	0.9
31	0.7	0.7	0.7	0.7	0.7	0.7	0.7	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.9	0.9	0.9
32	0.7	0.7	0.7	0.7	0.7	0.7	0.7	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.9	0.9	0.9	0.9
33	0.7	0.7	0.7	0.7	0.7	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.9	0.9	0.9	0.9
34	0.7	0.7	0.7	0.7	0.7	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.9	0.9	0.9	0.9
35	0.7	0.7	0.7	0.7	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.9	0.9	0.9	0.9	0.9
36	0.7	0.7	0.7	0.7	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.9	0.9	0.9	0.9	0.9
37	0.7	0.7	0.7	0.7	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.9	0.9	0.9	0.9	0.9
38	0.7	0.7	0.7	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.9	0.9	0.9	0.9	0.9	1.0
39	0.7	0.7	0.7	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.9	0.9	0.9	0.9	0.9	1.0
40	0.7	0.7	0.7	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.9	0.9	0.9	0.9	0.9	1.0
41	0.7	0.7	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.9	0.9	0.9	0.9	0.9	1.0
42	0.7	0.7	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.9	0.9	0.9	0.9	0.9	1.0
43	0.7	0.7	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.9	0.9	0.9	0.9	0.9	0.9	1.0
44	0.7	0.7	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.9	0.9	0.9	0.9	0.9	0.9	1.0
45	0.7	0.7	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.9	0.9	0.9	0.9	0.9	1.0	1.0
46	0.7	0.7	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.9	0.9	0.9	0.9	0.9	1.0	1.0
47	0.7	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.9	0.9	0.9	0.9	0.9	1.0	1.0
48	0.7	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.9	0.9	0.9	0.9	0.9	1.0	1.0
49	0.7	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.9	0.9	0.9	0.9	0.9	1.0	1.0
50	0.7	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.8	0.9	0.9	0.9	0.9	0.9	1.0	1.0