# Data quality

## English Housing Survey technical advice note

www.communities.gov.uk

# Data quality

## English Housing Survey technical advice note

May 2011
Department for Communities and Local Government

Department for Communities and Local Government
Eland House
Bressenden Place
London
SW1E 5DU
Telephone: 030 3444 0000
Website: www.communities.gov.uk



If you require this publication in an alternative format please email alternativeformats@communities.gsi.gov.uk

DCLG Publications
Tel: 030 0123 1124
Fax: 030 0123 1125

Email: product@communities.gsi.gov.uk
Online via the website: www.communities.gov.uk

May 2011

ISBN: 978 1 4098 2958 4

# Contents

**Introduction**

**Overview of data quality issues**

**Non-response and missing data**

**Sampling and measurement error**

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

**Surveyor variability**

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

# Introduction

This document is in one in a series of Technical Advice Notes providing background information about the methodology of the English Housing Survey to assist users in their analysis and interpretation of the survey findings.

# Overview of data quality issues

1. This note outlines the main sources of error affecting the quality of results from the English Housing Survey:

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

These are discussed separately below.

# Non-response and missing data

2. It is essential that the English Housing Survey provides a representative picture of housing in England. The sampling structure was specifically designed to provide such a picture.

3. 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.

4. 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 Grossing on the following webpage:

www.communities.gov.uk/publications/housing/ehstechnicaladvicenotes

## Sampling and measurement error

5. Like all estimates based on samples, the results of the English Housing Survey 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

6. 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.

7. 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. Since the 2008-09 English Housing Survey effectively is a simple random sample the deft arises solely from the weighting adjustments.

### Measurement error

8. Error may also arise in other ways, the main one of which is measurement error, 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.

### Systematic error

9. 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.

### Estimating sampling error

10. As outlined above, estimates of dwelling and household characteristics produced from a sample survey such as the English Housing Survey may

differ from the true population figures because they are based on a sample rather than a census. This difference is known as sampling error, and it is important to be able to estimate the size of this error when interpreting the survey results.

11. The size of the sampling error depends on the size of the sample; in general, sampling error is potentially larger in smaller samples. For example, a larger sampling error will be associated with estimates for converted flats than estimates for semi-detached or terraced houses, which are more numerous in the English Housing Survey sample.

12. 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

13. 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.

14. 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 English Housing Survey 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.

15. 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:

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

where n is the unweighted sample size

16. 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. 17,691 households in 2008-09. The sample size for the ‘2008’ (two-year) dwelling sample is 16,150.

17. The simplest approximation of a 95 per cent confidence interval for a percentage, assuming a simple random sample, may be obtained using

the look-up table in Annex 1 at the end of this note. The confidence interval can be calculated by reading off the closest figure from Table 1, where the estimated percentages are shown as columns and the unweighted sample sizes as rows, and then adding and subtracting this figure from the estimated percentage, as in the following examples:

i) Over two-thirds (67.9%) of all households were owner-occupiers in 2008-09. This figure is based on an unweighted sample of 17,691 households.

The corresponding number from the 1st row and 8th column of Table 1 is 0.7%, giving a confidence interval of 67.2% to 68.6%.

ii) Just over one-fifth (21.4%) of dwellings in 2008 were semi-detached. This figure is based on the 2-year unweighted sample of 16,150 dwellings.

The corresponding number from the 2nd row and 3rd column of Table 1 is 0.6%, giving a confidence interval of 20.8% to 22.0%.

18. Confidence intervals can be calculated more accurately by using the formula above. For example (i) the standard error is given by:

$se(p)=\sqrt{(67.9*32.1)/17,691}=0.351$

so the confidence interval is 67.9 +/- 1.96*0.351, or 67.21% to 68.59%.

19. 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 English Housing Survey 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).

20. 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 2008-09 English Housing Survey 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 subsampling by tenure and the corresponding weighting adjustments.

[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.

21. Examples of standard errors and confidence intervals calculated using the appropriate design factors are given in each of the detailed English Housing Survey annual reports, see:

www.communities.gov.uk/housing/housingresearch/housingsurveys/englishhousingsurvey/ehspublications/

22. Note that the confidence intervals for the 2008 stock data include the impact of combining the 2007-08 English House Condition Survey data, which used a clustered sample, with the 2008-09 English Housing Survey data which used an unclustered sample.

23. 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.

### Estimating measurement error

24. 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. 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.

## 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 English Housing Survey 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 region

- 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 used calibration workbooks as a means of detecting any significant shift in surveyor marking, or 'surveyor drift', between surveys. These are being introduced for the English Housing Survey from 2009-10. The workbooks are normally completed by the surveyors at the end of each year's fieldwork. The electronic workbooks contain 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 are chosen to cover a range of dwelling elements, building types and levels of severity.

28. The workbooks are 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 often 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 are 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 are then compared with those derived from calibration workbooks from previous years and statistical analysis is 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 2001-2009. Two Housing Health and Safety Rating System examples, falls on stairs and hot surfaces, were included for the third time. The 2009-10 results saw a marked improvement in the percentage of surveyors correctly assessing both hazards as category 1 hazards.

### 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. As an 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 study was conducted during the 2003-04 physical survey fieldwork for the English House Condition Survey 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 264 properties were re-surveyed by a second surveyor and the results were compared. The objectives of the study were to:

- compare the correlated surveyor variability with previous results to see whether the new survey design and contractor had had an impact and to estimate the impact of surveyor variability on standard errors for the survey and

- provide evidence for the reliability of the core survey measures, so that analysis of trends and comparisons may focus on the most reliable measures, and problematic measures can be improved through briefing or questionnaire design

33. This study was repeated using very similar methodology as part of the 2009-10 English Housing Survey physical survey fieldwork. Results from the study will be published during 2011.

34. The 2003-04 study found that:

i) overall, the levels of variability between surveyor judgements were low. However, where there is an appreciable level of error, the combined impact on the level of error surrounding the survey estimates can still be substantial

ii) in general, there was a high level of agreement between surveyors. For the 96 core survey measures tested there was, on average, 81% agreement between surveyors

iii) Kappa scores were used to measure the level of agreement after chance agreement has been excluded. 24 variables had Kappa scores that indicated ‘poor’ agreement. Ten had scores indicating ‘very good’ agreement

iv) multilevel modelling was used to calculate correlated surveyor variance. This measures the tendency of an individual surveyor to make assessments which are consistent for that surveyor but are different from the average assessment of all surveyors

v) correlated surveyor variance was found to be substantially lower on average for derived composite variables, such as whether a dwelling met the decent homes standard, than for simple variables, taken straight from the survey questionnaire. The same result was found in the previous 2001 study

vi) the most problematic variables are those with high correlated surveyor variability and a low Kappa score. For this study these variables were all concerned with surveyor assessments of problems in the area. However high correlated surveyor variability were associated with assessments of 'no' or 'some' problems rather than assessments of 'major' problems in the area

### Taking account of between-surveyor variability

35. The standard error calculations described earlier, which take account of the complex weighting of the survey, and the effect of subsampling 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 a little too narrow. Using the correlated surveyor estimates from the multilevel modelling, it will be possible to estimate the size of these downward biases in the standard error estimates and make an adjustment. Until these results are available, the results from the 2003-04 English House Condition Survey exercise can be used to give an approximate adjustment to the standard errors, as set out in Chapter 3 of the *English House Condition Survey 2007 Technical Report* at:
www.communities.gov.uk/publications/housing/ehcstechnicalreport2007

# Annex 1: Look-up table for calculating 95 per cent confidence intervals for a percentage

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

| **Sample size** | **5%** | **10%** | **20%** | **30%** | **40%** | **50%** | **60%** | **70%** | **80%** | **90%** | **95%** |
|---|---|---|---|---|---|---|---|---|---|---|---|
| ***17,691*** | *0.3* | *0.4* | *0.6* | *0.7* | *0.7* | *0.7* | *0.7* | *0.7* | *0.6* | *0.4* | *0.3* |
| ***16,150*** | *0.3* | *0.5* | *0.6* | *0.7* | *0.8* | *0.8* | *0.8* | *0.7* | *0.6* | *0.5* | *0.3* |
| **14,000** | 0.4 | 0.5 | 0.7 | 0.8 | 0.8 | 0.8 | 0.8 | 0.8 | 0.7 | 0.5 | 0.4 |
| **12,000** | 0.4 | 0.5 | 0.7 | 0.8 | 0.9 | 0.9 | 0.9 | 0.8 | 0.7 | 0.5 | 0.4 |
| **10,000** | 0.4 | 0.6 | 0.8 | 0.9 | 1.0 | 1.0 | 1.0 | 0.9 | 0.8 | 0.6 | 0.4 |
| **9,000** | 0.5 | 0.6 | 0.8 | 0.9 | 1.0 | 1.0 | 1.0 | 0.9 | 0.8 | 0.6 | 0.5 |
| **8,000** | 0.5 | 0.7 | 0.9 | 1.0 | 1.1 | 1.1 | 1.1 | 1.0 | 0.9 | 0.7 | 0.5 |
| ***7,972*** | *0.5* | *0.7* | *0.9* | *1.0* | *1.1* | *1.1* | *1.1* | *1.0* | *0.9* | *0.7* | *0.5* |
| **7,000** | 0.5 | 0.7 | 0.9 | 1.1 | 1.1 | 1.2 | 1.1 | 1.1 | 0.9 | 0.7 | 0.5 |
| **6,000** | 0.6 | 0.8 | 1.0 | 1.2 | 1.2 | 1.3 | 1.2 | 1.2 | 1.0 | 0.8 | 0.6 |
| **5,000** | 0.6 | 0.8 | 1.1 | 1.3 | 1.4 | 1.4 | 1.4 | 1.3 | 1.1 | 0.8 | 0.6 |
| **4,000** | 0.7 | 0.9 | 1.2 | 1.4 | 1.5 | 1.5 | 1.5 | 1.4 | 1.2 | 0.9 | 0.7 |
| **3,000** | 0.8 | 1.1 | 1.4 | 1.6 | 1.8 | 1.8 | 1.8 | 1.6 | 1.4 | 1.1 | 0.8 |
| **2,000** | 1.0 | 1.3 | 1.8 | 2.0 | 2.1 | 2.2 | 2.1 | 2.0 | 1.8 | 1.3 | 1.0 |

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **1,000** | 1.4 | 1.9 | 2.5 | 2.8 | 3.0 | 3.1 | 3.0 | 2.8 | 2.5 | 1.9 | 1.4 |
| **900** | 1.4 | 2.0 | 2.6 | 3.0 | 3.2 | 3.3 | 3.2 | 3.0 | 2.6 | 2.0 | 1.4 |
| **800** | 1.5 | 2.1 | 2.8 | 3.2 | 3.4 | 3.5 | 3.4 | 3.2 | 2.8 | 2.1 | 1.5 |
| **700** | 1.6 | 2.2 | 3.0 | 3.4 | 3.6 | 3.7 | 3.6 | 3.4 | 3.0 | 2.2 | 1.6 |
| **600** | 1.7 | 2.4 | 3.2 | 3.7 | 3.9 | 4.0 | 3.9 | 3.7 | 3.2 | 2.4 | 1.7 |
| **500** | 1.9 | 2.6 | 3.5 | 4.0 | 4.3 | 4.4 | 4.3 | 4.0 | 3.5 | 2.6 | 1.9 |
| **400** | 2.1 | 2.9 | 3.9 | 4.5 | 4.8 | 4.9 | 4.8 | 4.5 | 3.9 | 2.9 | 2.1 |
| **300** | 2.5 | 3.4 | 4.5 | 5.2 | 5.5 | 5.7 | 5.5 | 5.2 | 4.5 | 3.4 | 2.5 |
| **200** | 3.0 | 4.2 | 5.5 | 6.4 | 6.8 | 6.9 | 6.8 | 6.4 | 5.5 | 4.2 | 3.0 |
| **100** | 4.3 | 5.9 | 7.8 | 9.0 | 9.6 | 9.8 | 9.6 | 9.0 | 7.8 | 5.9 | 4.3 |

Department for Communities and Local Government


**ISBN: 978 1 4098 2958 4**