
Chapter 1

Sampling

The English Housing Survey (EHS) consists of two main elements: an initial interview survey of approximately 13,300 households and a follow up physical inspection of a sub-sample of around 6,200 of dwellings which includes vacant dwellings. This chapter provides information on how household and dwelling samples were selected for inclusion in the interview and physical survey elements of the 2012-13 EHS.

Overview

- 1.1 As in previous years, addresses for the initial 2012-13 EHS sample were selected using a systematic random sample design. Interviews are attempted at all of these addresses. This is the 'household sample', also referred to throughout this report as the 'interview survey sample'.
- 1.2 Unlike previous years, the design was changed in 2012-13 so that the sample was clustered, with interviewing carried out in alternate halves of England in each year. This design delivers a representative sample of households in England for any year, with the sample being unclustered for any two combined years of the survey.
- 1.3 In addition to the household sample, a sub-sample of addresses is then selected for physical inspection by a qualified surveyor. This sub-sample includes both occupied and vacant dwellings and is referred to as the 'dwelling sample' or 'physical survey sample'. To ensure a sufficient number of rented properties are included in the dwelling sample, all rented properties are selected for physical inspection while about 50% of owner occupied dwellings are randomly selected for a physical inspection (the percentage selected varied by quarter of the survey, ranging from 42% to 55%).
- 1.4 In 2012-13, 25,286 addresses were issued to interviewers. A small proportion of these were found to be ineligible, including addresses that were found to be commercial premises, second and holiday homes or demolished properties. Of the remainder, interviews were achieved at 13,652 addresses. A further 1,112 addresses were found to be vacant. After sub-sampling, 10,415 dwellings were passed to surveyors, and physical surveys were achieved at 6,304 of these.

Interview survey sample

- 1.5 The requirement for the 2012-13 household sample was to achieve 13,300 interviews across the four housing tenures (i.e. owner occupied, private rented, local authority and housing association).
- 1.6 To achieve these interviews, an initial sample of 36,000 addresses was drawn from the Postcode Address File (PAF). These addresses were drawn as a systematic two-stage random sample from the Royal Mail's Small User PAF.
- 1.7 For the first stage of sampling, 1,808 geographic 'merged areas' were generated by combining neighbouring Lower Layer Super Output Areas (LSOAs), so that each 'merged area' contained about 12,500 addresses. Half (904) of these 'merged areas' were randomly selected (as a systematic stratified sample) for the 2012-13 sample, with the remaining half then allocated for the 2013-14 sample.
- 1.8 The second stage of sampling involved selecting a stratified systematic random sample of addresses within the 'merged areas' selected for the 2012-13 sample.
- 1.9 The advantage of this two-stage approach is that it reduces the fieldwork area to half the country so interviewer and surveyor travel time and costs are reduced. The disadvantage is that for any single survey year the survey is partially clustered, which will result in a small loss¹ in statistical efficiency. However, when analysing any two years of survey data the combined sample is entirely unclustered.
- 1.10 For each sampled address, the predominant tenure within the postcode that contained that address was identified and attached to the record. Predominant tenure was identified using Experian's Residata² classifications; addresses were then grouped into strata and sub-sampled at the rates of 63% for owner occupied, 68% for private rented and 100% for social rented. As a result, 25,286 addresses were issued to interviewers, Table 1.1.

¹ NatCen estimate that the maximum design factor for standard errors would be around 1.14.

² Experian possess a database that contains information obtained from a number of sources including insurance companies, Census, etc. referred to as Residata. It is from this that information is taken on predominant tenure within a postcode as well as other information. The matching of the EHS sample to Residata is carried out by BRE.

Table 1.1: Sub-sampling of PAF addresses, 2012-13

	PAF sample	Sub-sampling rate	Issued EHS sample
	<i>addresses</i>	<i>percentages</i>	<i>addresses</i>
predominant tenure			
owner occupied	25,361	62.8%	15,919
private rented	3,935	67.7%	2,663
social rented	6,547	100.0%	6,547
unknown tenure	157	100.0%	157
Total	36,000	70.2%	25,286

1.11 Interviews were achieved at 13,652 households, Table 1.2.

Table 1.2: Number of interviews achieved, 2012-13

tenure	
owner occupiers	8,280
private rented	2,103
local authority	1,523
housing association	1,746
Total	13,652

Physical survey sample

1.12 The requirement for the 2012-13 physical survey sample was 6,200 physical surveys across the four housing tenures.

1.13 To ensure that the EHS delivers the same level of precision as its predecessor the English House Condition Survey (which was based on a clustered sample design), the EHS sample is stratified to ensure disproportionate numbers of renters are included, Table 1.3. The alternative would be to issue a very much larger random sample and carry out unnecessary owner occupier surveys in order to accumulate sufficient renting cases.

Table 1.3: Tenure distribution of achieved physical survey sample compared with the national stock

	Achieved sample		National stock
	<i>number</i>	<i>percentages</i>	<i>percentages</i>
tenure			
owner occupiers	2,640	42.0%	65.0%
private rented	1,295	21.0%	18.0%
local authority	1,120	18.0%	10.0%
housing association	1,249	20.0%	8.0%
Total	6,304	100	100

- 1.14 The issued sample for the physical survey is drawn as a stratified sub-sample of the dwellings of those households who respond to the interview survey, together with a stratified sub-sample of dwellings found to be vacant at that stage. Calculation of the size of the sample to be issued takes account of the expected physical survey response rates by tenure.
- 1.15 The sub-sampling of interview survey cases for the physical survey is carried out in the field during the interview, using the tenure of the dwelling established during the interview. Interviewers are advised via the CAPI (computer-aided personal interviewing) instrument if the household they are interviewing is eligible for the physical survey, and if so they then attempt to gain agreement from respondents to take part before passing on the address details to CADS Housing Surveys.
- 1.16 Different sub-sampling rates are applied to each tenure group to identify cases eligible for the physical survey. These rates can be varied quarterly if required in order to achieve the required annual total sample.
- 1.17 Because the rented sector is smaller than the owner occupied sector, the rented sector is over-sampled to ensure sufficient numbers for analysis. Sub-sampling rates are reviewed at the start of each year and are kept under review throughout the year in order to ensure a sufficiently large sample is achieved. The sub-sampling rates for 2012-13 are shown in Table 1.4.

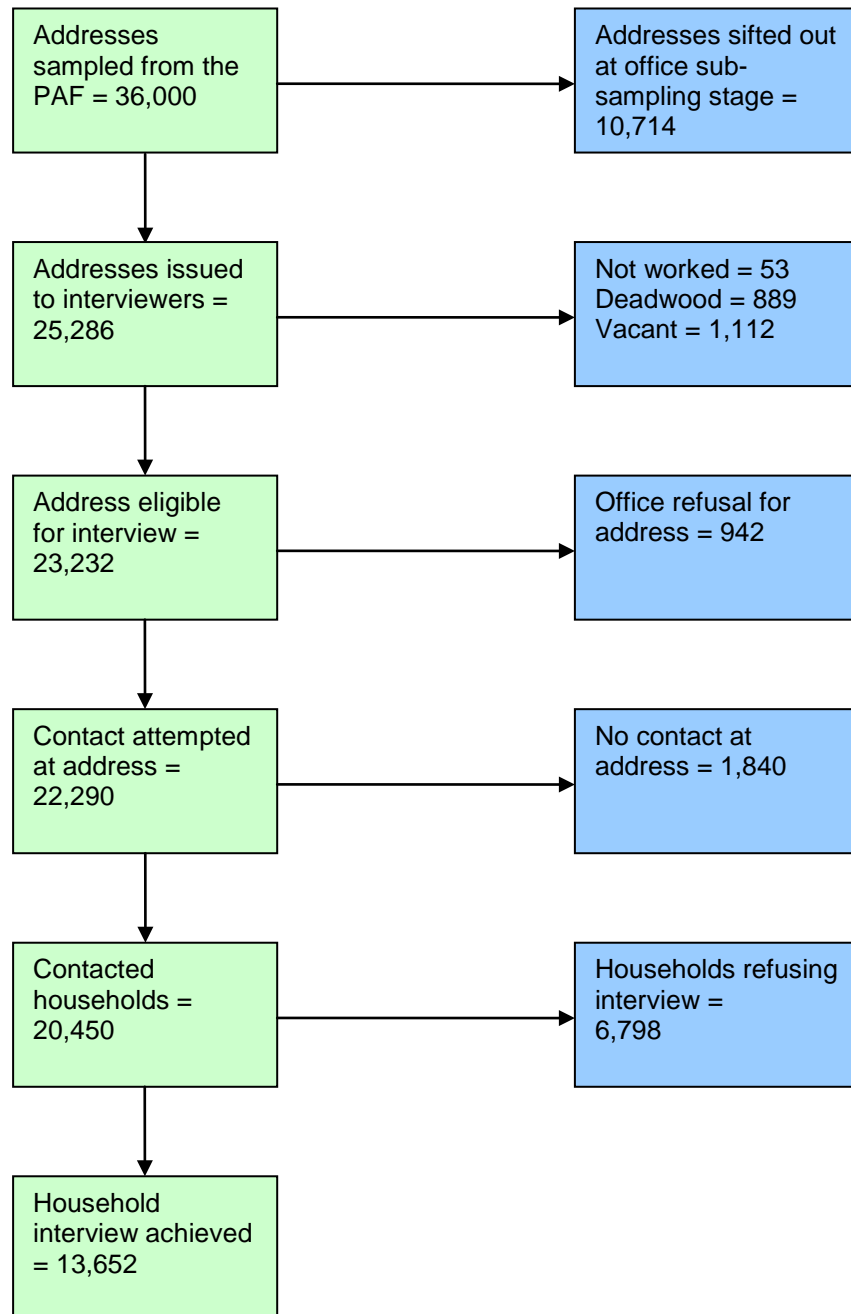
Table 1.4: Sub-sampling rates, 2012-13 physical survey

	Sub-sample rates			
	Quarter 1	Quarter 2	Quarter 3	Quarter 4
	<i>percentages</i>			
tenure				
owner occupiers	41.7%	50.0%	55.0%	55.0%
private rented	100.0%	100.0%	100.0%	100.0%
local authority	100.0%	100.0%	100.0%	100.0%
housing association	100.0%	100.0%	100.0%	100.0%

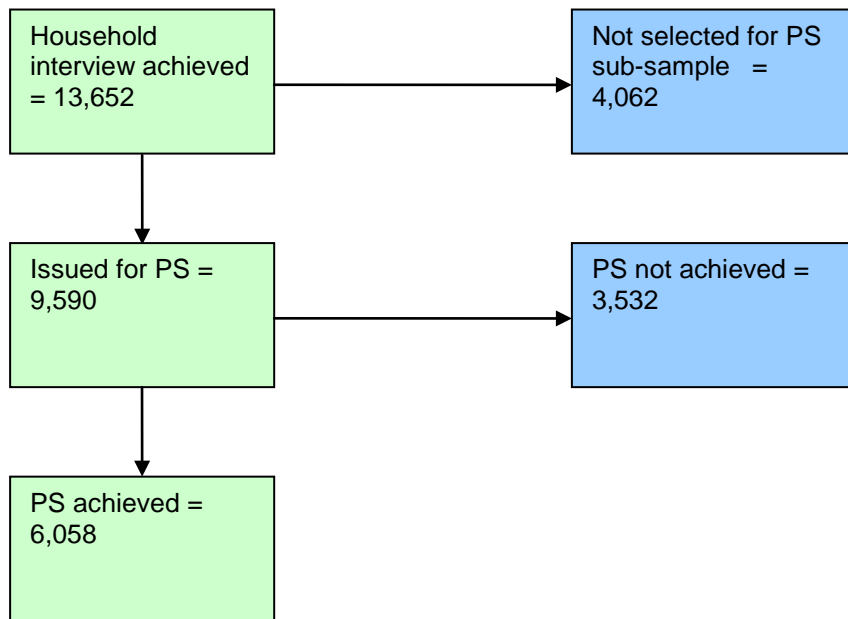
- 1.18 The aim was to keep the rate fairly constant. This was to ensure that there were roughly equal numbers of surveys across the quarters to reduce the impact of seasonality (for example, damp problems are more likely to be identified in the winter). In practice the sub-sample rate for owner occupiers was increased from 42% to 55% from quarter 1 to quarter 4. The increase in the sub-sample rate was required because of a systematic change in the accuracy of Experian’s Residata from the previous year’s database.
- 1.19 Vacant properties are sub-sampled at the same rates as occupied cases based on their last known tenure gathered by interviewers as part of their initial visit. For vacant properties, cases are selected for physical survey based on the interviewer’s best estimate of tenure using local enquiries. Permission and access for the survey is then sought by CADS Housing Surveys surveyors.
- 1.20 Not surprisingly, a lower proportion of full physical surveys are obtained in unoccupied dwellings, compared with occupied dwellings, because of the difficulty in gaining access to a property that is currently unoccupied. In 2012-13, surveyors managed to gain access and obtain full physical surveys in 30% of dwellings that were unoccupied at the time of interview. This is a decrease from 35% in 2011-12.
- 1.21 The 2012-13 sampling and response process is summarised below, Figure 1.1.

Figure 1.1: Sample structure of the EHS, 2012-13

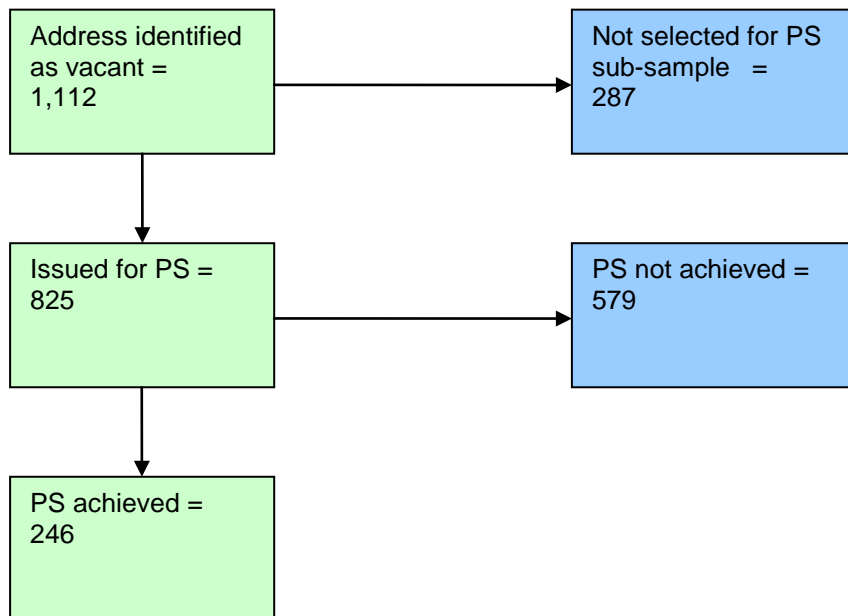
Interview survey sample



Physical survey sample: occupied dwellings



Physical survey sample: vacant dwellings



Sampling at addresses

- 1.22 Most addresses contain a single dwelling and single household. However, at a small proportion of addresses (less than 1%) this is not the case. There are standard procedures for interviewers to randomly select one dwelling and/or one household when more than one is identified.
- 1.23 Considering dwellings, the procedure is for the interviewer to list the dwellings identified at the address and then randomly select one from the list, using a pre-selected random number. The random number is obtained from a sheet (called a Kish grid) which has a column of the number of dwellings identified, and a column of which numbered dwelling to select.
- 1.24 The same procedure is used to select the household to interview when more than one is identified at a dwelling/address.