
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 the household and dwelling samples were selected for inclusion in the interview and physical survey elements of the 2013-14 EHS.

Overview

- 1.1 As in previous years, addresses for the initial 2013-14 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 The design delivers a representative sample of households in England for 2013-14, with the sample being unclustered when combined with the 2012-13 sample.
- 1.3 In addition to the household sample, a sub-sample of addresses was selected for physical inspection. This sub-sample included both occupied and vacant dwellings and is referred to as the 'dwelling sample' or 'physical survey sample'. To ensure that a sufficient number of rented properties were included in the dwelling sample, all rented properties were selected for physical inspection while around half of owner occupied dwellings were randomly selected for a physical inspection (the percentage selected varied each quarter, ranging from 45% to 53%).
- 1.4 In 2014-15, 23,852 addresses were issued to interviewers. A small proportion of these (783 addresses) were found to be ineligible, including addresses that were found to be commercial premises, second and holiday homes or demolished properties. Productive interviews were achieved at 13,276 of the eligible addresses. Just under three-quarters of the productive addresses (9,822) were passed to surveyors, and physical surveys were achieved at 6,194 of these.

Interview survey sample

- 1.5 The requirement for the 2013-14 household sample was to achieve 13,300 interviews across the four housing tenure types (i.e. owner occupied, private rented, local authority and housing association).
- 1.6 To achieve these interviews, an initial sample of 40,000 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, the whole of England was grouped into 1,808 geographic 'merged areas'. Those 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 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 2013-14 sample.
- 1.9 The advantage of using a two-stage approach involving the 'merged areas' 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 The social rented sector (local authority and housing association) is less prevalent compared with the other sectors. To ensure that there is a sufficiently large sample of social renters for analysis, the 40,000 addresses were sub-sampled to sift out about half of the owner occupiers and private renters. The sub-sampling was carried out based on the 'predicted tenure' of the sampled addresses. The 'predicted tenure' of an address was derived from the predominant tenure within the postcode that contained that address. Predominant tenure was identified using Experian's Residata² classifications and attached to the address records. Sub-sampling was carried out by grouping the addresses into the four housing tenure types and sub-sampling at the rates of 51% for owner occupied, 52% for private rented and 96% for social rented. That process produced 23,852 addresses for issuing to interviewers, Table 1.1.

¹ NatCen estimates 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 was taken on predominant tenure within a postcode as well as other information. The matching of the EHS sample to Residata was carried out by BRE.

Table 1.1: Sub-sampling of PAF addresses, 2013-14

	PAF sample	Sub-sampling rate	Issued EHS sample
	<i>addresses</i>	<i>percentages</i>	<i>addresses</i>
predominant tenure			
owner occupied	27,904	50.5%	14,094
private rented	4,311	52.4%	2,261
social rented	7,352	96.3%	7,081
unknown tenure	433	96.1%	416
Total	40,000	59.6%	23,852

1.11 Interviews were achieved at 13,276 households, Table 1.2.

Table 1.2: Number of interviews achieved, 2013-14

tenure	
owner occupiers	7,769
private rented	2,058
local authority	1,461
housing association	1,988
Total	13,276

Physical survey sample

1.12 The requirement for the 2013-14 physical survey sample was 6,200 physical surveys across the four housing tenures.

1.13 To ensure that the EHS delivers findings on renters that have the same level of precision as its predecessor the English House Condition Survey, the EHS sample was stratified to ensure disproportionate numbers of renters are included, Table 1.3.

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	2,354	38.0%	63.2%
private rented	1,295	20.9%	19.6%
local authority	1,094	17.7%	7.2%
housing association	1,451	23.4%	10.1%
Total	6,194	100.0%	100.0%

- 1.14 The issued sample for the physical survey was drawn as a stratified sub-sample of the dwellings of those households who responded to the interview survey, together with a stratified sub-sample of dwellings found to be vacant during fieldwork. Calculation of the size of the sample to be issued took account of the expected physical survey response rates by tenure.
- 1.15 The sub-sampling of interview survey cases for the physical survey was carried out during the interview by the computer-aided personal interviewing programme. The programme used the tenure of the dwelling established at the interview to select the subsample and advised the interviewers if the household they were interviewing was eligible for the physical survey. If so they would attempt to gain agreement from respondents to take part in the physical survey and pass on the address details to CADS Housing Surveys who managed the fieldwork of the physical surveys.
- 1.16 Different sub-sampling rates were applied to each tenure group to identify cases eligible for the physical survey. As the owner occupied sector is larger than the rented sector, it was under-sampled to ensure sufficient numbers of renters for analysis. Sub-sampling rates were reviewed at the start of the year and were kept under review throughout the year in order to ensure a sufficiently large sample was achieved. The sub-sampling rates for 2013-14 are shown in Table 1.4.

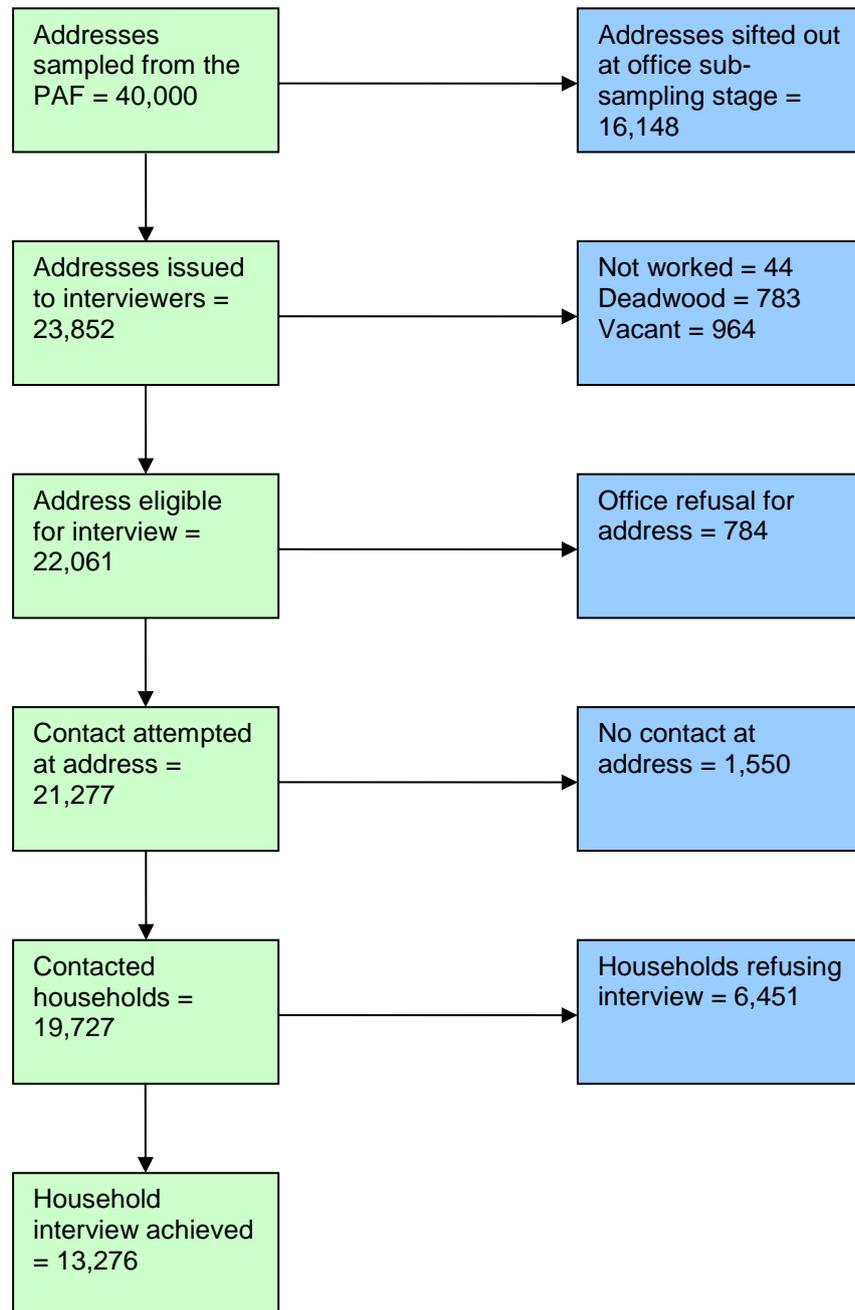
Table 1.4: Sub-sampling rates, 2013-14 physical survey

	Sub-sample rates			
	Quarter 1	Quarter 2	Quarter 3	Quarter 4
	<i>percentages</i>			
tenure				
owner occupiers	45.0%	45.0%	45.0%	53.0%
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%

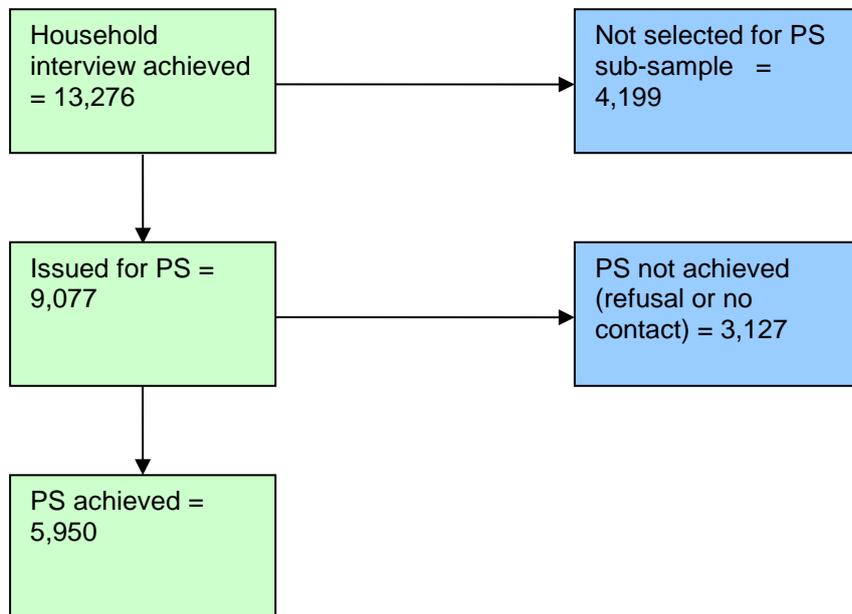
- 1.17 To ensure that there were roughly equal numbers of surveys across the quarters to reduce the impact of seasonality (for example, damp problems are more likely to be identified in the winter), the sub-sampling rate was kept fairly constant. However, the sub-sample rate for owner occupiers was increased from 45% to 53% in quarter 4 to ensure the overall target for 6,200 physical surveys across the four housing tenure groups was achieved given the response rates in quarters 1 to 3.
- 1.18 Vacant properties were sub-sampled at the same rates as occupied cases based on information about their last known tenure. This information was gathered by interviewers as part of their initial visit (from talking to the landlord or neighbours) or based on the interviewer’s best estimate of tenure derived from available evidence. Permission and access for the survey was then sought by CADS Housing Surveys surveyors.
- 1.19 Not surprisingly, a lower proportion of full physical surveys were obtained in unoccupied dwellings, compared with occupied dwellings, because of the difficulty in gaining access to a property that was unoccupied. In 2013-14, surveyors managed to gain access and obtain full physical surveys in 33% of dwellings that were unoccupied at the time fieldwork took place. This is an increase from 30% in 2012-13.
- 1.20 The 2013-14 sampling and response process is summarised in Figure 1.1.

Figure 1.1: Sample structure of the EHS, 2013-14

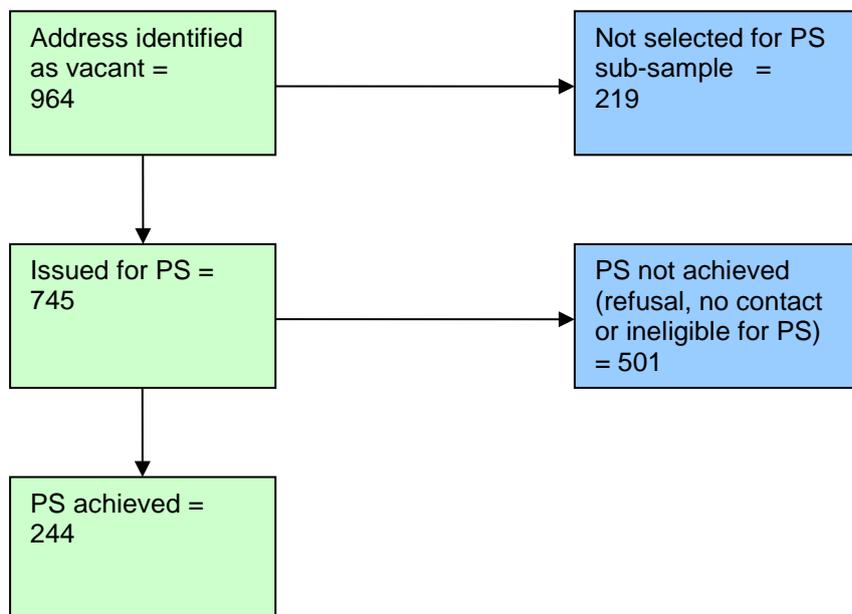
Interview survey sample



Physical survey sample: occupied dwellings



Physical survey sample: vacant dwellings



Sampling at addresses

- 1.21 Most addresses contained a single dwelling and single household. However, at a small proportion of addresses (less than 1%) this was not the case. There were standard procedures for interviewers to randomly select one dwelling and/or one household when more than one was identified.
- 1.22 For 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 was obtained from a sheet (called a Kish grid) which had a column for the number of dwellings identified, and a column for which numbered dwelling to select.
- 1.23 The same procedure was used to select the household to interview when more than one was identified at a dwelling/address.