English Housing Survey

Technical Report, 2021-22

ENGLISH
HOUSING
SURVEY

Contents

Introduction

1. The English Housing Survey (EHS) is a national survey of people's housing circumstances and the condition and energy efficiency of housing in England. In its current form, it was first run in 2008-09. Prior to then, the survey was run as two standalone surveys: the English House Condition Survey and the Survey of English Housing. It is one of the longest standing surveys in government, with 2017 marking the 50th anniversary since the first survey in 1967. The 2021-22 EHS was conducted by NatCen Social Research in partnership with the Building Research Establishment (BRE) and CADS Housing Surveys.

2. The EHS consists of two main elements: an interview survey about housing and related topics; and a physical inspection of the condition of a random sample of the housing stock.

3. This report describes the key technical features of the 2021-22 EHS. Each chapter of the report covers a specific aspect of the survey and is designed to be read as a standalone document.

Changes to the 2021-22 EHS in response to the COVID-19 pandemic

4. The survey was originally set up as a face-to-face survey. The EHS approach was amended to allow fieldwork within the social distancing restrictions introduced in response to the pandemic. EHS 2020-21 consisted of a 'push-to-telephone' interview followed by an external only physical survey: respondents were contacted by letter and invited to opt in to take part in the initial EHS telephone interview. The survey was further adapted in 2021-22 as social distancing guidelines were eased. EHS 2021-22 comprised a 'knock-to-nudge' interview followed by an external physical survey: interviewers visited issued addresses in person to introduce the survey on the doorstep and to make an appointment for the EHS telephone interview. The key differences between the EHS 2021-22 'knock-to-nudge; approach and the traditional face-to-face approach are summarised below. A similar comparison for the 2020-21 push-to-telephone approach is available in the EHS 2020-21 Technical Report.

Key changes made in 2021-22 EHS in response to the COVID-19 pandemic	Corresponding process from the 2019-20 EHS
Fieldwork was compressed into the 10-month period between mid-May 2021 and March 2022 inclusive.	Fieldwork took place between April 2019 and March 2020 inclusive.
A socially distanced 'knock-to-nudge' approach was used for the interview survey where respondents were interviewers visited each address to make an appointment for a telephone EHS interview. Interviewers did not go into respondent's home.	All the interview surveys were carried out face-to-face in the respondent's home.
The physical inspection adapted an 'external plus' approach where the visual inspection was restricted to an assessment of the exterior of the dwelling, supplemented by factual data on the dwelling's interior collected (socially distanced) at the doorstep.	The physical inspection covered both the interior and exterior of the dwelling.
The size of the issued sample was increased to 31,411 addresses as response rates from a 'knock-to-nudge' approach are typically lower than those achieved by a face to face approach.	The issued sample consisted of 28,238 addresses.
Interviewers visited issued addresses in person so were able to identify vacant properties, to collect interviewer observations and, for most properties, to find out the number of dwellings and households at each address. Physical inspections of vacant properties were included in 2021-22. The housing condition data collected by the external plus survey therefore related to the total stock.[1]	Interviewers identified vacant dwellings and collected a range of observational information when they visited the issued addresses. The physical inspections included 150 vacant dwellings, so the house condition data relate to whole housing stock (occupied as well as vacant).
The interview surveys did not use showcards; the interviewers read out the answer options over the	The interviewers used show cards at some questions.

[1] The annual reports were for occupied homes only as vacant addresses were only available for 2021-22 (not for 2020-21). However, housing quality indicators were created for occupied and vacant homes where data was available, e.g. for pass/fail for damp, HHSRS.

telephone. Some of the questions were adapted to be delivered over the telephone.	
Many data items on the condition of the interior of dwellings were not collected so detailed derived variables relating to the Housing Health and Safety Rating System, actual and standardised repairs costs are not available. Dwelling level modelled estimates on the following have been developed to fill in some key data gaps: whether the Decent Homes Standard was met; the cost of making a non-decent home meet the Standard; the presence of any HHSRS Category 1 hazards; and the presence of damp.	
The weighting scheme was amended to fit the survey approach. Interviewers visited issued addresses in person and were able to collect key information about non-contact, address-to-dwelling ratio, dwelling-to-household ratio, high- and low-rise flats for use in weighting. An additional model was included to take account of cases that were not issued (see section 6.15). The dwelling weights for the PS paired sample cases in 2020-21 and 2021-22 did not include vacant properties as information about vacant properties was not collected in 2020-21.	

Chapter 1
Sampling

1.1   The English Housing Survey (EHS) consists of two main elements: an interview survey and a follow up physical inspection of the dwelling. In previous years the interview survey included approximately 13,300 households, and the physical inspection covered around 6,000 occupied households and a further 200 vacant dwellings.

1.2   In 2021-22, face-to-face fieldwork was not possible because of restrictions introduced during the COVID pandemic. EHS was therefore carried out over a reduced fieldwork period using a 'knock-to-nudge' approach and an 'external plus' physical survey.[2] Sampling was changed to support the new approach.

1.3   This chapter provides information on how the household and dwelling samples were selected for inclusion in the interview and physical inspection elements of the 2021-22 EHS.

Overview

1.4   As in previous years, addresses for the initial 2021-22 EHS sample were selected using a systematic random sample design. Interviews were attempted at all of these addresses. This is referred to throughout this report as the 'interview survey sample'.

1.5   The design delivers a representative sample of households in England for 2021-22, with the sample being unclustered when combined with the 2020-21 sample.

1.6   In addition to the interview survey 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 'physical survey sample'.

1.7   To ensure that a sufficient number of rented properties were included in the dwelling sample, the majority of rented properties (82%) were selected for physical inspection (the percentage selected varied by tenure and quarter ranging from 49% to 100%) while around two thirds of owner-occupied

---

[2] The 'knock-to-nudge' approach was a remote approach that allowed us to carry out the EHS while COVID restrictions were in place. EHS interviewers visited the sampled address to collect respondents' contact details and then conducted the interview over the telephone. Selected respondents were invited to take part in the 'external plus' physical survey. Surveyors did not enter the household but instead asked a small number of questions on the doorstep, carried out an external only survey of the property, and used supplementary information available from EPC and other data sources.

dwellings (64%) were randomly selected for a physical inspection (the percentage selected varied each quarter, ranging from 41% to 100%).

1.8   In 2021-22, 31,411 addresses were issued. Productive interviews were achieved at 9,752 of the issued addresses. A small proportion of these (1,000 addresses) were found to be ineligible because they fall outside the scope of the EHS, they include addresses that were found to be commercial premises, second and holiday homes, demolished properties or instances where the address was not found.[3] About 77% of the productive (occupied) addresses (7,554) were passed to surveyors, and physical surveys were achieved at 5,284 of these.

Interview survey sample

1.9   The initial aim for the 2021-22 EHS was to achieve 13,300 interviews across the four housing tenure types (i.e. owner-occupied, private rented, local authority and housing association). However, sample size had to be amended post-COVID to fit the new knock-to-nudge approach and reduced fieldwork period.

Initial sample

1.10   To achieve these interviews, an initial sample of 139,997 addresses was drawn from the Postcode Address File (PAF). [4] These addresses were drawn as a systematic two-stage random sample from the Royal Mail's Small User PAF.

1.11   For the first stage of sampling, the whole of England was grouped into 1,808 geographic 'merged areas' (clusters). Those areas were generated by combining neighbouring (18 on average) Lower Layer Super Output Areas (LSOAs) within region, so that each cluster contained about 12,500 addresses (with a range from about 11,200 to 13,800).

1.12   The 1,808 clusters were stratified by region, Census estimates of the percentage of dwellings in the cluster that are owner occupied (in tertiles within region) and the percentage of owner occupied dwellings (for the tertile with the highest percentage of owner occupied dwellings) and Census estimates of the proportion of households with a Household Reference

---

[3] These categories of ineligible cases are those that do not fall within the scope of the English Housing Survey - second homes, demolished addresses, commercial addresses or instances where the address is not found. When calculating ineligible cases for the purpose of response rates we additionally include cases that fall within the scope of the EHS, i.e. cases with an identifiable primary residential address, but that are ineligible for an interview or survey, such as vacant properties and derelict properties (both ineligible for household interview but eligible for a physical survey), institutions and addresses under construction. See Annex Table 4.1. for details of cases ineligible for interview/survey.

[4] The initial sample was larger than in previous years to allow more flexibility in case a "push-to-telephone" approach (similar to the one used for EHS 2020-21) was necessary due to the COVID pandemic .

Person (HRP) who works in non-manual occupations (for the other two tertiles).

1.13   The clusters were then randomly allocated to years, by assigning them in pairs going down the list, and then randomly (and independently) allocating one of each pair to "Year 1" and the other to "Year 2" (i.e. 904 clusters per year). Thus, each year the EHS covers one half of England, giving rise to this design being referred to as the 'half-England model'. The allocation of clusters to years was fixed since 2012-13: "Year 1" clusters were used for 2012-13, 2014-15, 2016-17, 2018-19 and 2020-21 and "Year 2" clusters for 2013-14, 2015-16, 2017-18, 2019-20 and 2021-22.

1.14   At the second stage of sampling, all addresses in the PAF in the 904 clusters selected for the 2021-22 sample were stratified by cluster, LSOA ordered by Census estimates of the percentage of owner-occupied dwellings, Census output area and postcode. Then a systematic random sample of 139,997 addresses was selected.

1.15   The advantage of using a two-stage approach involving the clusters is that it reduces the fieldwork area to half of 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 results in a small loss[5] in statistical efficiency. However, when analysing any two years of survey data, the combined sample is entirely unclustered.

Issued sample

1.16   To ensure that there is a sufficiently large sample of renters for analysis, owner occupied addresses are usually sub-sampled at a higher rate compared to rented addresses before being issued to interviewers. Sub-sampling is carried out based on the 'predicted tenure' of the sampled addresses derived from the predominant tenure within the postcode that contained that address. Predominant tenure is identified using Experian's Residata[6] classifications and attached to the address records.

1.17   Sub-sampling was carried out by grouping the addresses into the four housing tenure types and sub-sampling at the rates of 15% for owner occupied, 30% for private rented and 35% for social rented (local authority and housing association renters). That process produced 31,411 issued addresses, Table 1.1.

---

[5] NatCen estimates that the maximum design factor due to clustering would be about 1.08.
[6] 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, 2021-22

	PAF sample	Sub-sampling rate	Issued EHS sample
	addresses	*percentages*	*addresses*
predominant tenure			
owner occupied	88,633	15.5%	13,703
private rented	25,351	30.1%	7,636
social rented	24,411	34.7%	8,470
unknown tenure	1,602	100.0%	1,602
Total	**139,997**	**22.4%**	**31,411**

1.18   Interviews were achieved at 9,752 households, Table 1.2.

Table 1.2: Number of interviews achieved, 2021-22

tenure	
owner occupiers	5,617
private rented	1,735
local authority	971
housing association	1,429
Total	**9,752**

Physical survey sample

1.19   The initial aim for EHS 2021-22 was to achieve 6,200 physical surveys across the four housing tenures. As with the interview survey, the target was reduced due to the change in fieldwork approach and shorter fieldwork period.

1.20   To ensure that the EHS delivers findings on renters that have the same level of precision as its predecessor, the English House Condition Survey, a disproportionate number of renters is included by under-sampling owner-occupied dwellings, Table 1.3.[7]

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

		Achieved sample	National stock
	number	*percentages*	*percentages*
tenure			
owner occupiers	2,663	48.6%	63.9%
private rented	1,032	18.8%	19.6%
local authority	691	12.6%	6.3%
housing association	1,092	19.9%	10.1%
Total	**5,478**	**100.0%**	**100.0%**

[7] The "national stock" figures are based on DLUHC counts of all dwellings (occupied + vacant) by tenure for each region adjusted to reflect the reference date of the EHS 2021-22 (1 October 2021).

1.21   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.

1.22   The sub-sampling of interview survey cases for the physical survey was carried out during the interview by the Computer-Assisted Personal Interviewing (CAPI) program. The program 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.

1.23   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 2021-22 are shown in Table 1.4.

Table 1.4: Sub-sampling rates, 2021-22 physical survey

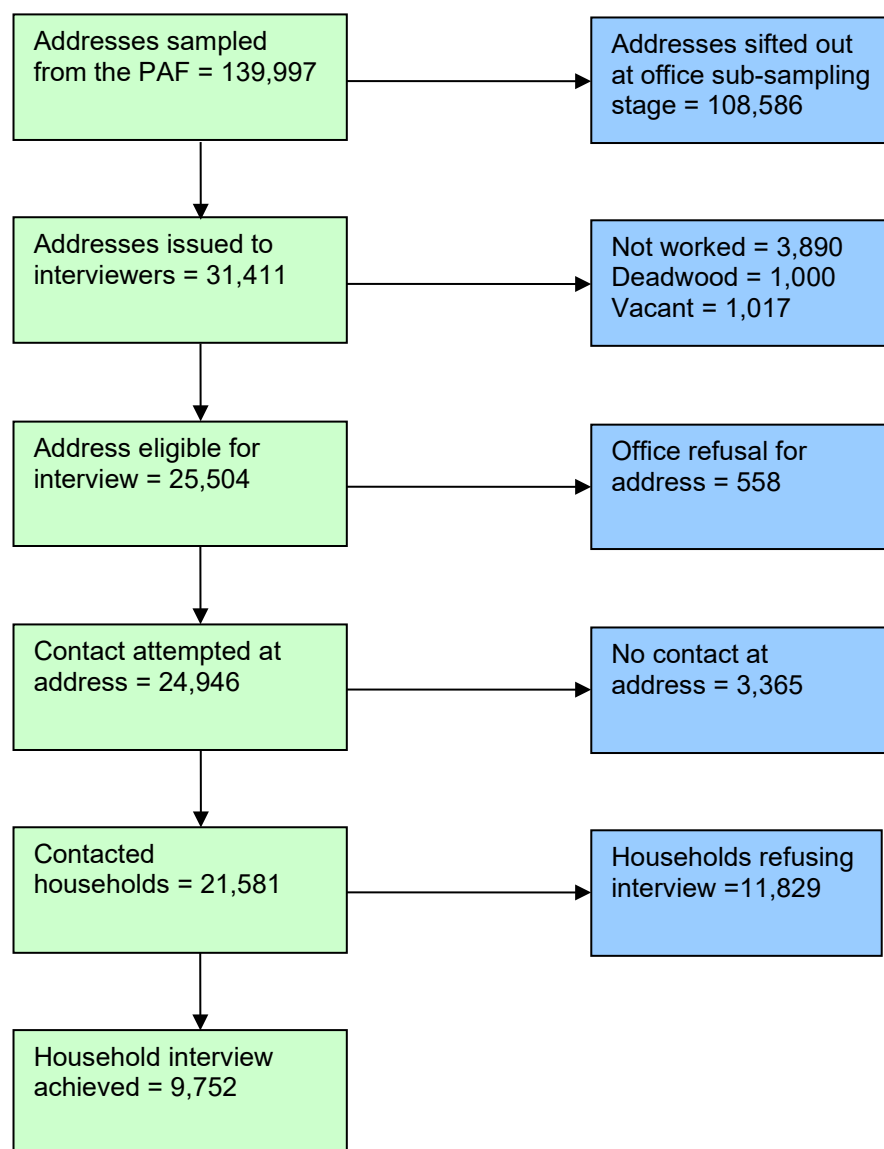| | Sub-sample rates | | | |
	Quarter 1	Quarter 2	Quarter 3	Quarter 4
				percentages
tenure				
owner occupiers	62.5%	40.9%	68.0%	100.0%
private rented	100.0%	48.7%	86.0%	100.0%
local authority	100.0%	86.4%	82.2%	100.0%
housing association	100.0%	100.0%	100.0%	100.0%

1.24   To ensure there were roughly equal number 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 adjusted accordingly after taking account of differential response by tenure. [8]

---

[8] The original issued sample allocation was not always reflected in the actual fieldwork allocation because of the difficulty of allocating points to interviewers during covid. The sub-sampling rates shown in Table 1.4 reflect the original sample allocation by quarter.

1.25   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 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 the surveyors managed by CADS Housing Surveys.

1.26   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 2021-22, surveyors managed to gain access and obtain 'external plus' physical surveys in 25% of dwellings that were unoccupied at the time fieldwork took place.

1.27   The 2021-22 sampling and response process is summarised in Figure 1.1.

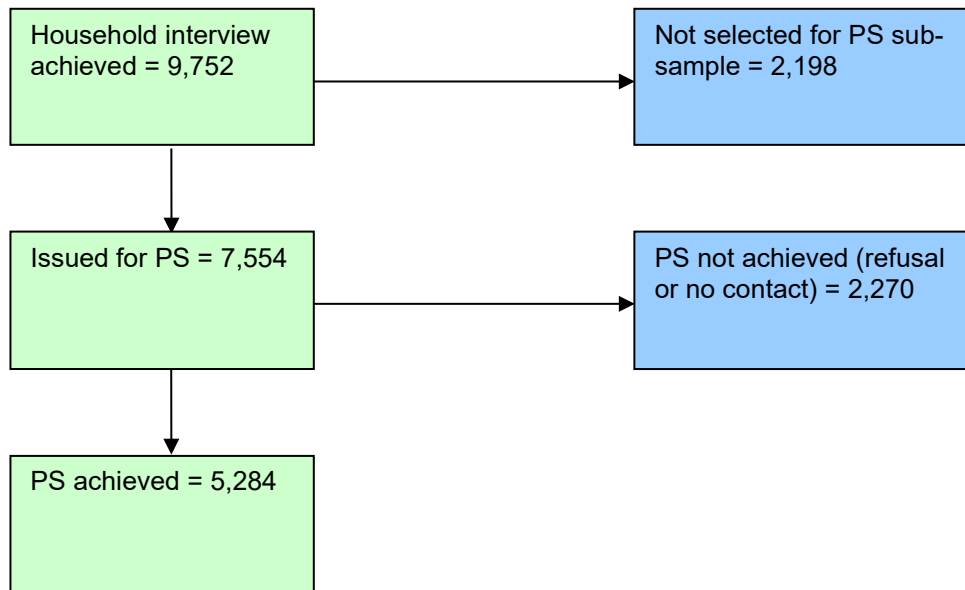## Figure 1.1: Sample structure of the EHS, 2021-22[9]

Interview survey sample

```
┌─────────────────────┐          ┌─────────────────────┐
│ Addresses sampled   │────────▶ │ Addresses sifted out │
│ from the PAF =      │          │ at office sub-sampling│
│ 139,997             │          │ stage = 108,586      │
└─────────────────────┘          └─────────────────────┘
          │
          ▼
┌─────────────────────┐          ┌─────────────────────┐
│ Addresses issued to │────────▶ │ Not worked = 3,890   │
│ interviewers =      │          │ Deadwood = 1,000     │
│ 31,411              │          │ Vacant = 1,017       │
└─────────────────────┘          └─────────────────────┘
          │
          ▼
┌─────────────────────┐          ┌─────────────────────┐
│ Address eligible for│────────▶ │ Office refusal for   │
│ interview = 25,504  │          │ address = 558        │
└─────────────────────┘          └─────────────────────┘
          │
          ▼
┌─────────────────────┐          ┌─────────────────────┐
│ Contact attempted at│────────▶ │ No contact at        │
│ address = 24,946    │          │ address = 3,365      │
└─────────────────────┘          └─────────────────────┘
          │
          ▼
┌─────────────────────┐          ┌─────────────────────┐
│ Contacted           │────────▶ │ Households refusing  │
│ households = 21,581 │          │ interview =11,829    │
└─────────────────────┘          └─────────────────────┘
          │
          ▼
┌─────────────────────┐
│ Household interview │
│ achieved = 9,752    │
└─────────────────────┘
```
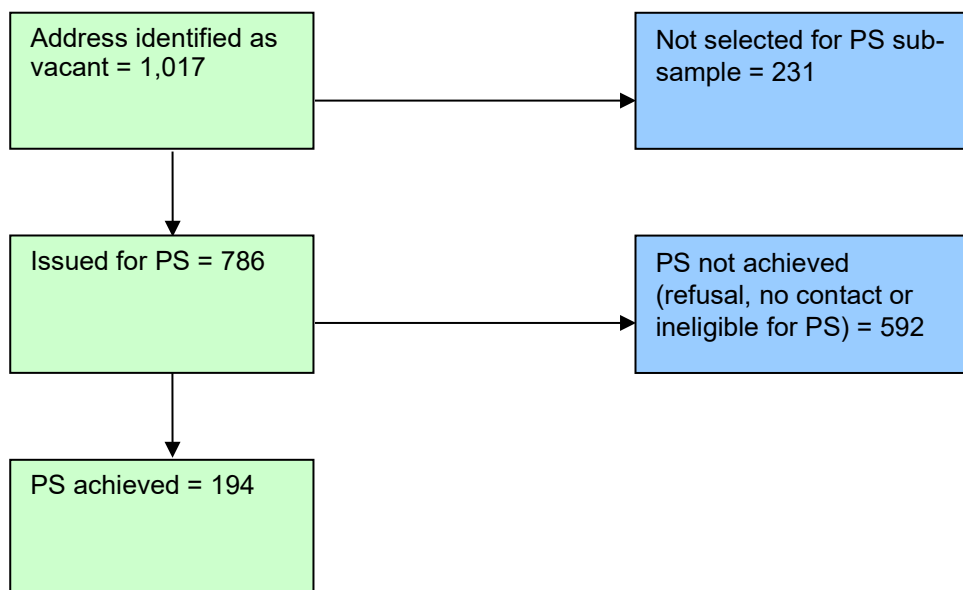
---

[9] The way case outcomes are grouped for the purposes of calculating response rates is slightly different to the sampling overview presented here. See Annex Table 4.1 for details of household interview outcomes. Note:
1) Categories 'unable to locate address' and 'not worked' are included under in-scope addresses for the purposes of response rate calculations but grouped within ineligible in the above.
2) In the diagram above 'no contact' includes categories 'unknown whether residential: no contact', 'residential but unknown eligibility: no contact', and 'other unknown eligibility'. These are treated as a separate category 'unknown eligibility' for the purposes of response rate calculations.
3) 'Households refusing interview' above includes other unproductive categories including 'language difficulties', 'lost productive' 'away in hospital during fieldwork period' and 'respondent requested deletion'. These are treated as a separate category 'other unproductive' for the purposes of calculating response rates.
4) The 1,000 cases assigned as 'deadwood' include those referred to in Footnotes 4, 5 and 6 of annex table AT4.4 of Chapter 4.

Physical survey sample: occupied dwellings

```
┌─────────────────────────┐              ┌─────────────────────────┐
│ Household interview      │─────────────▶│ Not selected for PS sub- │
│ achieved = 9,752         │              │ sample = 2,198           │
└─────────────────────────┘              └─────────────────────────┘
            │
            ▼
┌─────────────────────────┐              ┌─────────────────────────┐
│ Issued for PS = 7,554    │─────────────▶│ PS not achieved (refusal │
│                          │              │ or no contact) = 2,270   │
└─────────────────────────┘              └─────────────────────────┘
            │
            ▼
┌─────────────────────────┐
│ PS achieved = 5,284      │
│                          │
└─────────────────────────┘
```

Physical survey sample: vacant dwellings

```
┌─────────────────────────┐              ┌─────────────────────────┐
│ Address identified as    │─────────────▶│ Not selected for PS sub- │
│ vacant = 1,017           │              │ sample = 231             │
└─────────────────────────┘              └─────────────────────────┘
            │
            ▼
┌─────────────────────────┐              ┌─────────────────────────┐
│ Issued for PS = 786      │─────────────▶│ PS not achieved          │
│                          │              │ (refusal, no contact or  │
│                          │              │ ineligible for PS) = 592 │
└─────────────────────────┘              └─────────────────────────┘
            │
            ▼
┌─────────────────────────┐
│ PS achieved = 194        │
│                          │
└─────────────────────────┘
```

Sampling at addresses

1.28 Most addresses contained a single dwelling and a single household. However, at a small proportion of addresses (around 1%) this was not the case. There were standard procedures for interviewers to select one dwelling and/or one household at random when more than one was identified.

1.29 The interviewer listed the dwellings identified at multi-dwelling address and then randomly selected 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.

1.30 The same procedure was used to select the household to interview when more than one was identified at a dwelling.

Chapter 2
Questionnaires

2.1   The English Housing Survey (EHS) collects data in two separate phases. In the first phase, an interviewer conducts an interview using Computer-Assisted Interviewing. Following the interview, a second phase of data collection occurs when, for a sub-sample of cases, a qualified surveyor visits the address to make an assessment of physical aspects of the property. The surveyor completes a detailed survey form using a digital pen, an input device which captures the surveyor's handwriting and brush strokes and converts them into digital data.

Overview

2.2   The Department for Levelling Up, Housing and Communities (DLUHC), in consultation with the Department for Energy Security and Net Zero (DESNZ) and other users of the survey, review the content of the interview and physical survey annually. The review revises or adds questions in order to cover new salient topics, to reflect technical developments in the construction and energy efficiency of homes or to improve data quality. If required, the review also removes questions if the information they elicit is no longer relevant to users' needs

2.3   Compared with the physical survey, the interview questionnaire has changed more since the launch of the EHS in 2008-09. The most radical overhaul of the interview questionnaire and physical survey form took place during the development of the 2011-12 EHS. That was driven by a significant reduction in funding, and it resulted in a shortening of the average length of the interview from 50 to 30 minutes and a reduction in the data collected about Housing Health and Safety hazards on the physical survey form.[10] By comparison, changes to the interview questionnaire and physical survey form since that time have been relatively minor.

---

[10] As a result of the review, the number of Housing Health and Safety Rating System (HHSRS) hazards that were directly measured by the surveyor was reduced from ten to six; the others were flagged only when an extreme risk (equating to a Category 1 Hazard) was found as part of the physical survey.

Interview questionnaire

2.4   A core set of questions is asked of respondents every year. They cover:

- household composition, ethnicity, nationality, economic status, education and health;
- household accommodation and length of residence;
- housing history and aspirations;
- rent and mortgage payments;
- satisfaction with landlord/attitudes to neighbourhood; and
- income.

These topics cover the key attributes of a household and the dwelling it occupies. The permanent inclusion of questions on these topics (with minimal change to the phrasing of the questions) ensures that a consistent picture is provided over time. A number of topics (such as marital status and ethnic group) use the questions developed by the Government Statistical Service Harmonisation Programme to allow comparison of findings across different government surveys: https://gss.civilservice.gov.uk/guidances/0-harmonised-principles/#demographics

2.5   The questionnaire also contains a number of rotating question sets which come in and out of the survey on an annual, biennial or, in some cases, less frequent basis. Rotating modules ensure the interview length is manageable to reduce respondent burden, while also meeting policy requirements as it allows key topics to be covered which are not required every year. Topics covered in these modules include:

- second homes;
- fire and fire safety;
- satisfaction with the neighbourhood;
- work undertaken to improve energy efficiency of the home;
- adaptations made to the home to improve accessibility; and
- tenancy deposits.

2.6   Table 2.1 outlines the topics included in the 2021-22 questionnaire. The reintroduced rotating modules were tenancy deposits and second homes. The modules on energy performance certificates and adaptations for disability were removed. New ARF questions on the HRP's dwelling, household, main/second home were included in the CAPI. Additionally, checks have been added to confirm who the partner of the HRP is and the definition of smart meters has been updated.

Table 2.1 Household interview questionnaire topics, 2021-22

demographics including age, sex, marital status, household reference person and household relationships
type of property
tenure
individual characteristics and perceptions including: nationality; country of birth; ethnicity; time at address; health; disability; education; wellbeing; and direct measure of loneliness
age of accommodation
housing history
subletting
waiting lists (for social housing)
rooms available to the household and shared facilities
type of dwelling and household
satisfaction with accommodation and neighbourhood (reduced version)
institutional discrimination
council tax and utilities
energy efficiency (reduced version)
ownership type (i.e. leasehold or freehold)
satisfaction with repairs and maintenance
ownership details including when bought property, how financed purchase, who bought it from, how much it cost
mortgages, including type, payments and arrears
tenancy type
tenancy deposit scheme
tenancy deposit
social renting details
experience of homelessness
rent and housing benefit including payments and arrears
number of tenancy agreements
fire safety (reduced version)
second homes
buying aspirations
working status and job details
economic status
income and earnings
receipt of benefits
income support and mortgage interest
savings and investments
formal complaints renters have made to the landlord or letting agent
satisfaction with landlord and/or letting agent (asked of private renters only)
casual lettings

2.7   The 2021-22 EHS used the 'knock-to-nudge' approach to collect the interview data because social distancing restrictions introduced during the COVID-19 pandemic had prevented interviewers from completing the interview in the

respondent's homes. This approach involved visiting a respondent at their home and, while socially distanced, encouraging them to respond to the survey by telephone. The use of telephone interviews has meant that the 2021-22 EHS did not use showcards during the interview; instead, interviewers read out the answer options of the questions requiring showcards to the respondent. Most questions have had minor wording changes to be adapted to the telephone data collection method. Details of the changes are recorded on the 2021-22 EHS questionnaire documentation[11].

2.8 In 2021-22, the median interview length including recruitment to the physical survey was 41 minutes. This length is calculated on full interviews only; partial interviews are excluded.[12]

Physical survey

2.9 The physical survey form is designed to collect information about the dwelling and its surroundings, and about the condition of the property and what would need to be done to remedy defects.

2.10 The content of the physical survey has remained largely unaltered from the former EHCS. In a typical year, surveyors record the following:

- details of the nature and type of each dwelling;
- the presence and condition of facilities and services;
- the condition of the internal and external building materials;
- the presence and condition of shared facilities and services in blocks of flats or on estates;
- an assessment of the environment in which the dwelling is located; and
- assessments of the health and safety risks associated with the dwelling.

2.11 It was not possible to carry out an internal inspection of the property in 2021-22 due to social distancing restrictions. Surveyors therefore used the 'external plus' approach instead, which was used for 2020-21. The approach involved the surveyor carrying out an external inspection of the property, asking a small number of questions about the interior of the dwelling on the doorstep. Despite this, it was not possible for the 2021-22 EHS to collect the 'usual' levels of data about the interior of the inspected dwellings. The section on the physical survey form covering the following topics were, therefore, not collected in full due to the use of the 'external plus' methodology:

---

[11]

https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/1124145/EHS_Questionnaire_documentation_2021-22.pdf

[12] A partial interview is when a substantial part of the interview is carried out (up to the question PlanTen which asks about the type of housing the respondent expects occupy in the longer term) but the interview is stopped before the end is reached.

- information on amenities
- internal defects
- housing health and safety rating system
- common parts
- shared facilities.

Data items collected by the 'external plus' approach are highlighted with a red border around them on the 2021-22 EHS physical survey form.[13]

2.12　Table 2.2 outlines the topics covered in the 2021-22 physical survey. The full EHS 2021-22 physical survey form is published on Gov.uk.[13]

Table 2.2 Physical survey topics, 2021-22

Services, heating and energy
Construction
Measurement
Exterior and plot
Age of external building elements (or components)
External defects
Structural faults
Pests
Drains
Flat construction and faults
Local area and environment

2.13　New data was collected on electric vehicle charge points for the physical survey in 2021-22:

- On page 8, households were asked whether access to an electric vehicle charge point existed and if so, in what capacity (exclusive use, communal or on-street point within 100 metres of the home).

13

https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/1124155/EHS _2021-22_Physical_Survey_Form.pdf

Chapter 3
Fieldwork

3.1   Fieldwork for the 2021-22 English Housing Survey (EHS) was managed on behalf of the Department for Levelling Up, Housing and Communities (DLUHC) by NatCen Social Research. There are two parts to the EHS: the interview survey and the physical survey. This chapter provides further details of fieldwork procedures, including interviewer and surveyor training.

Overview

3.2   The 2021-22 survey was conducted by NatCen Social Research in consortium with CADS Housing Surveys and Building Research Establishment (BRE). NatCen carried out the interview surveys while CADS was responsible for the field force of professional surveyors who undertook the visual inspection of properties occupied by a sub-sample of the households who had taken part in the interview survey. BRE developed the physical survey questionnaire as well as the surveyor training manuals and delivered the surveyor training sessions.

3.3   To conform with the social distancing restrictions introduced during the COVID-19 pandemic, the 2021-22 EHS carried out the interview and physical survey using approaches which minimise social contact with respondents. The interview survey used a 'knock-to-nudge' approach where a letter was sent out to sampled addresses explaining the purpose of the survey and informing respondents that an interviewer would call at their address. Interviewers visited the sampled addresses in person to arrange an appointment for the interview and to collect the respondents contact details. The interview was carried out by telephone. The physical survey adopted an 'external plus' approach where the visual inspection was restricted to an assessment of the exterior of the dwelling, supplemented by factual data on the interior of the dwelling the surveyor collected (socially distanced) at the doorstep.

Fieldwork period

3.4   Fieldwork took place over a slightly reduced ten-month period starting from May. Fieldwork was carried out in eight separate waves. Each wave consisted of an eight-week interviewer and surveyor fieldwork period, with surveyor fieldwork starting one week after interviewer fieldwork began. Each wave started approximately one month after the previous wave, with a longer gap between waves four and five.

3.5   Interviewers began fieldwork on the 17th May 2021 and surveyors on the 12th June. Fieldwork dates for 2021-22 were as follows:

Fieldwork Quarters	Interviewer		Surveyor	
	Start	finish	Start	finish
Quarter 1				
Wave 1	17th May 2021	27th June 2021	12th June 2021	4th July 2021
Wave 2	21st June 2021	1st August 2021	17th July 2021	8th August 2021
Quarter 2				
Wave 3	26th July 2021	5th September 2021	21st August 2021	12th September 2021
Wave 4	30th August 2021	10th October 2021	25th September 2021	17th October 2021
Quarter 3				
Wave 5	4th October 2021	14th November 2021	30th October 2021	21st November 2021
Wave 6	8th November 2021	17th December 2021	4th December 2021	17th December 2021
Quarter 4				
Wave 7	4th January 2022	13th February 2022	29th January 2022	20th February 2022
Wave 8	31st January 2022	13th March 2022	26th February 2022	20th March 2022

Training and project briefings

Interviewers

3.6   Interviewers working on the EHS are drawn from the NatCen Social Research pool of interviewers. Prior to starting work at NatCen, all interviewers receive three days of intensive training, which includes training in Computer-Assisted Personal Interviewing (CAPI) and workshops on doorstep technique. After this, interviewers are supervised for up to three days in the field. Interviewers are also encouraged to 'shadow' more experienced colleagues to develop their skills further and to share experiences.

3.7   Before starting work on the EHS, all interviewers are required to complete a briefing on the survey.

3.8   Experienced EHS interviewers attended a two-hour refresher briefing delivered remotely via Teams by NatCen Social Research. Interviewers who were new to the EHS attended a 6-hour full briefing. In advance of the briefings, interviewers were required to undertake a pre-briefing exercise. This involved reading the project instructions, a guide to changes and completing a practice questionnaire covering the main survey procedures. In total, 208 interviewers were briefed at such briefings. The briefings covered:

- the purpose of the study;
- the new fieldwork processes including how respondents were invited to take part, how they could opt in, how they would then be allocated to interviewers and how interviewers would be able to make contact with them;
- interview procedures, including a summary of the key changes made to the questionnaire to adapt it to push-to-telephone;
- survey documents;
- booking appointments for the physical surveyor visit;
- field admin;
- various exercises to test their understanding of the EHS; and
- going through a practice interview with a scenario provided by research.

Surveyors

3.9   EHS surveyors are contracted by CADS Housing Surveys and trained by BRE on the technical content of the survey. Surveyors working on the EHS come from a range of professional backgrounds with approximately half being qualified Environmental Health Officers and the remaining from a variety of professional groups including Architects and Building Surveyors. New surveyors undertook a rigorous five-day residential briefing before working on their first survey. Chapter 3 of the 2019-20 EHS Technical Report contains full details on the training of new surveyors.[14]

3.10   Annual turnover of surveyors is low. In 2021-22,156 surveyors (including Regional Managers) worked on the EHS, 136 of which had worked on the previous year's survey. 20 new surveyors were recruited during 2021-22.

3.11   Existing surveyors receive a refresher briefing every year to ensure that they are up to date with any changes in the physical survey form and survey methodology. The format of the refresher briefing varies depending on the amount of changes to the physical survey form. In 2021-22 surveyors undertook distance learning on the updates to the BRE EHS website (where data from the physical inspections are downloaded) and the findings of the

---

[14] https://www.gov.uk/government/publications/english-housing-survey-2019-to-2020-technical-report

Surveyor Quality Report[15], with briefing material focused on data recorded on conservatories (and their thermal properties) All surveyors that continued to work on the EHS during 2021-22 took part in the training.

Making contact with respondents

Letters inviting respondents to take part

3.12   All addresses sampled for the EHS receive a letter, in advance of the interviewer's first visit, explaining the purpose of the study and stating that an interviewer will be visiting shortly. The letter also includes details of how the address has been selected and gives reassurances about confidentiality by stating:

We will treat information you give in strictest confidence under the EU General Data Protection Regulation. Your answers will be combined with those of others who take part and presented anonymously in research reports. For details of our purpose in collecting and processing the data of this survey, who is going to be able to access the data and all other privacy-related information, please visit https://www.gov.uk/government/publications/english-housing-survey-privacy-notice-for-survey-participants. We undertake to use and store the information you give us in accordance with the purposes and processes outlined in the privacy statement available on this website. You will not receive any junk mail or marketing as a result of doing the survey.

3.13   There are three different versions of the letter, one targeted at owner occupiers, one at private renters and the third at social renters. These letters are sent out centrally by NatCen Social Research, not by the interviewers. The interviewers carry a laminated copy of the generic version of the letter with them to aid respondents' recall of the advance letter on the doorstep. Copies of the various versions of the advance letter are available at Annex 3.2.

Leaflets

3.14   Respondents are also sent a leaflet about the study with their advance letter. This provides additional information about the survey including examples of how previous EHS data have been used. It also addresses potential concerns about data protection, provides contact details for NatCen Social Research and web addresses for DLUHC and NatCen Social Research.

---

[15] The Surveyor Quality Study (SQS) explored measurement errors in the physical survey component of the EHS. The SQS required all surveyors to survey two test properties for which a model answer had been produced by BRE and CADS. The test properties contained dwelling condition features of particular interests to the EHS https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/1135176/2021-22_EHS_Quality_Report.pdf

3.15 There are three different versions of the leaflet, one targeted at owner occupiers, one at private renters and the third at social renters. Each version includes pictures and study findings considered to be particularly relevant for that tenure group. Copies of the 2021-22 interviewer leaflets are available at Annex 3.3.

Interview survey data collection

3.16 Prior to seeking an interview with a respondent at a sampled address, the interviewer undertakes a series of contact procedures using the EHS 'doorstep form'. These include:

- Collecting 'first impression' data about the sampled address for use in non-response analysis.
- Identifying the sampled dwelling. A dwelling is defined as a unit of accommodation where all the rooms and amenities are for the exclusive use of the household(s) occupying them.
- In cases where the sample address has more than one dwelling, for example, if a house has been split into two self-contained flats, randomly selecting a dwelling at the sampled address for the interview.
- Identifying the households living at the dwelling and, in cases where a dwelling contains more than one household, randomly selecting a household for interview.
- Collecting information from neighbours about addresses that are vacant or where the interviewer has been unable to make contact with the occupants.

3.17 The data collection process for the interview survey is as follows:

- interviewers conduct a Computer-Assisted Personal Interview (CAPI) with the household reference person (HRP) or their partner over the telephone and record respondents' answers onto a laptop computer;
- interviewers upload the survey data to NatCen; and
- interviews are checked and edited where necessary by a team of editors. For example, if a respondent has an 'other' method of payment for electricity or gas not covered in our responses, the interviewer will record their response in a follow up questions and the edit team will attempt to code this to one of the existing response categories, or suggest new response category if needed.

3.18 Every year, the EHS physical survey inspects around 200 vacant dwellings. Properties that are vacant at the time of the interviewer's first call are passed to CADS Housing Surveys to approach for a physical survey. These are sampled on the same basis as occupied dwellings based on the last known tenure of the property.

3.19 The contents of the interview survey are reviewed annually (see Chapter 2 for more detail). NatCen and DLUHC thoroughly check all new and revised questions in the CAPI program before the beginning of fieldwork, including:

- question and answer texts;
- all routing alternatives;
- textfills inserted into the CAPI program to customise the questions; and
- checks built into the CAPI program to improve the accuracy of data collected during the interview.

3.20 The data for the first wave are also checked to ensure that the new/revised questions are working properly.

Interviewing non-English speakers

3.21 NatCen employs some interviewers who can conduct the interview in a language other than English. This is not always possible, and interviewers tend to rely on household interpreters to complete interviews with non-English speakers. Household interpreters must be aged 13 or over and willing to interpret on behalf of the interviewer. None of the survey documents are translated.

3.22 In 2021-22, 12 interviews were carried out in a language other than English; of these, 9 were carried out by NatCen interviewers in a language other than English, 1 was carried out with a member of the respondent's household acting as an interpreter, and 2 with an interpreter who was a relative or friend outside of the household.

Maximising response

3.23 In 2021-22, interviews were achieved for 9,752 households. This represents 33% of sampled addresses eligible for interview. Interviewers used a range of established procedures to maximise response rates.

3.24 Interviewers are encouraged to make a minimum of nine calls at each address before classifying the address as a non-contact. The calls have to be at different times of the day and spread across the fieldwork period. At least two of the calls need to be in the evening from Monday to Thursday and two have to be at the weekend.

3.25 In an attempt to maximise response, some survey agencies have procedures to reissue/reassign addresses that have not yielded a productive interview to another interviewer to attempt. EHS addresses are not generally reissued because doing so would delay the completion of the interview survey component and would reduce the amount of time the physical surveyors have to complete the fieldwork of the physical surveys. Instead of reissuing unproductive addresses, interviewers are required to make a comparatively

high number of calls at each address to contact the occupants and attempt to secure an interview.

3.26   In 2021-22, all households invited to take part in the EHS were given an unconditional £5 Post Office voucher, and interviewers had up to two £25 discretionary gift cards they could issue to encourage reluctant households. This was a return to the previous incentive strategy, based on the results from the full split sample incentive experiment run in 2017-18. More information about this experiment can be found with the other technical documents on the EHS on gov.uk: https://www.gov.uk/government/publications/english-housing-survey-findings-from-the-2017-to-2018-incentives-experiment.

Surveyor appointments

3.27   Interviewers are also responsible for gaining consent from those households eligible for the physical survey. Interviewers are asked to make a firm, timed appointment wherever possible using information about the surveyor's availability. The interviewers are provided with a list of dates and 59-minute 'slots' when their surveyor has indicated they are available. This information is built into the CAPI interview programme. Surveyors are also able to provide daily notes about each of their available days and overarching notes about their general availability and preferences for working. Those notes are also displayed on the interviewers' laptop.

3.28   The surveyor's availability is updated each night during the fieldwork period. Interviewers obtain current availability, plus details of any existing appointments allocated to their surveyor, each time they connect with NatCen Social Research's servers, which they are instructed to do before starting work and when finishing work each day.

3.29   The approach to making appointments is kept under close review in order to maximise the consent rate. After the information about respondents' appointment times/preferences is transferred from NatCen Social Research to the secure CADS Housing Surveys website, CADS Helpline staff assist in changing and confirming arrangements between the surveyors and respondents for the physical inspection.

3.30   Interviewers and surveyors are encouraged to speak to each other before interviewer fieldwork commenced to discuss any issues there may be on availability and booking appointments for the physical survey. Interviewers and surveyors are also provided with reciprocal contact details so that they can discuss work patterns and scheduling preferences if necessary.

3.31   As part of their training, interviewers are briefed on how to explain the physical survey to respondents and arrange the appointment for the surveyor to visit. Interviewer training includes:

- the surveyor fieldwork periods;
- transfer of information about appointments between CADS and NatCen;
- how to make appointments and what to take into account when making them;
- how to gain agreement from respondents for the physical survey; and
- a practice of the booking appointment section of the CAPI program.

Physical survey data collection

3.32   To conform with social distancing restrictions introduced during the COVID-19 pandemic, the 2021-22 physical survey data were collected using an 'external plus' approach where the inspection was restricted to an assessment of the exterior of the dwelling and supplemented by data on the interior of the dwelling the surveyor collected (socially distanced) at the doorstep. This differed from the previously standard approach, last carried out in 2019-20 before the COVID-19 pandemic, which involved a full assessment of both the interior and exterior of the dwelling by the surveyor.

3.33   As surveyors did not have access to the interior of the dwelling and were only able to collect data about the interior features by asking questions of the respondent at the doorstep, they could not complete all the sections of the physical survey form in 2021-22. The data items collected by the 'external plus' approach are highlighted with a red border on the 2021-22 EHS physical survey form available on Gov.uk.[16] After introducing him/herself at the time of the appointment, the surveyor began 'the external plus' survey by covering the sections on Pages 1 to 8 of the Physical Survey form that were applicable to the 2021-22 EHS. To help ease this process households were asked to note down, in advance of the appointment, the boiler make and model number along with type and thickness of the loft insulation. Following the completion of the internal questions, the surveyor would continue with the external survey, taking measurements and making judgements on all parts of the exterior of the dwelling. When the external survey was complete, the surveyor may ask further clarification questions of the household, for example, the age of the external elements that may not be original, such as roof cover, windows or doors, the age of extensions and alterations. After that, the surveyor would complete the survey by carrying out a visual assessment of the area in which the dwelling was located.

3.34   Surveyors tended to carry out more pre-survey preparation with the 'external plus approach than with the full survey approach last carried out in 2019-20. They were more likely to refer to sources such as Google Earth, Rightmove

---

[16]
https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/1039253/English_Housing_Survey_2021_to_2022_physical_survey.pdf

and Energy Performance Certificates data prior to the visit and indeed may review these again when validating the survey.

3.35 Other than the use of the 'external plus' approach and the external reference sources, the actual process of collecting the Physical Survey data collection has remained the same. The key elements of that process are as follows:

- Surveyors conduct the survey using a paper form designed to be used with a digital pen. They also take digital photographs of the outside of the property.
- The paper form features a series of small black dots arranged in a special non-repeating pattern (the Anoto pattern) which can be read by a digital pen which features an inbuilt digital camera. Every page has a unique Anoto pattern. As the pattern on each page is unique and non-repeating, the camera in the pen can determine the page the pen is on as well as the exact position of the pen as it writes across each page. During a survey the digital pen's inbuilt camera collects data by recording the position of the pen and captures and stores digital snapshots of the pen marks across the Anoto pattern.
- After the survey, surveyors upload the survey data to a dedicated website which displays the data as a facsimile of each page for the surveyor to check and correct electronically. An extensive series of validation checks on range, plausibility and consistency have been developed which enable surveyors to validate their own survey before submitting it to their regional manager.
- Photographs of the properties surveyed are added to the website.
- Corrected data are submitted to the surveyors' regional manager for final checking and, if necessary, returned to the surveyor for correction or clarification via the website.
- The checked data are forwarded to BRE by regional managers for consolidation with the rest of the EHS results.
- All cases showing errors on the validation program are checked by a single appointed regional manager and, if necessary, discussed with BRE. More information on data validation can be found in section 5.4 of Chapter 5 'Data Processing'.

3.36 The subjective nature of some assessments required for the physical survey means that a degree of variability is inevitable between surveyors in some of their judgements. To minimise the impact that any one surveyor can have on the results of any one area or type of property an annual upper target of 65 is set on the number of surveys any one surveyor can complete. An additional restriction is set relating to work within any one region, such that no one surveyor should complete more than 45 full surveys in any region (a limit of 36 exists for the North East as fewer physical surveys are conducted in this area compared with other regions). These rules help to minimise bias within the survey and to provide more robust measures of housing condition below the

national level. In 2021-22, 79% of surveyors did not exceed any regional target and 92% of surveyors did not exceed the annual upper target of 65 full surveys.

3.37   In 2014-15, a calibration exercise was undertaken to detect any variability that arises from surveyors making different judgements given the same information. Results from the 2014-15 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. More detail on the rationale, process and how those exercises have informed previous years of the survey can be found in the 2011-12 technical advice note on data quality, published on Gov.uk: https://www.gov.uk/government/publications/english-housing-survey-technical-advice.

3.38   To provide further insight into the effects of systematic surveyor variability on the precision of estimates from the physical survey, a Surveyor Variability Study (SVS) is routinely undertaken. This exercise which was last conducted in 2014-15 involved a call-back methodology in which 300 properties were surveyed again by a second surveyor and the results were compared. More information on the SVS can be found on gov.uk: https://www.gov.uk/government/collections/english-housing-survey-technical-advice.

3.39   In March 2019, a Surveyor Quality Study (SQS) was carried out to gauge the accuracy of the work of the surveyors against model answers compiled by CADS and BRE. The study involved all the surveyors and took place during the residential refresher briefing. Each surveyor was asked to complete a full survey on two of four test properties purposefully selected because they contained a range of house conditions and energy efficiency issues. Information on the SQS is available here: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/1135176/2021-22_EHS_Quality_Report.pdf

Annex 3.1: The briefing programme for surveyors

3.40   20 new surveyors were recruited in 2021-22 therefore a new surveyor briefing held in August/September 2021. When new surveyors are recruited, they undergo an intensive five day briefing on what is a highly specialised survey method. The briefing includes practical surveying exercises. The annual refresher briefing keeps surveyors up to date with any changes in the physical survey form and survey methodology.

EHS briefing programme for new surveyors

Day 0	
19.00-20.15	Dinner
20.15-20.45	Welcome/Introduction
20.45-21.30	EHS overview. S1-4 of form, HMO's
21.30-22.00	Dwelling definition exercise

Day 1	
08.30-10.00	Interior, incl. selected faults & treatments
10.00-10.45	Amenities, incl. selected faults & treatments
10.45-11.00	Coffee
11.00-13.00	Services, heating, energy
13.00-14.00	Lunch
14.00-15.00	Energy exercise
15.00-18.00	Fieldwork interior of Simple larger house
	Debriefing on site
18.00-19.00	Free
19.00-20.15	Dinner
20.15-21.00	Construction, materials, improvements
21.00-	Free

Day 2	
08.30-09.45	Shape/measurement, views
09.45-10.15	Shapes and Views exercise
10.15-10.45	Principles of disrepair
10.45-11.00	Coffee
11.00 -12.30	Exterior/plot, incl selected faults & treatments
12.30-13.00	Structural faults
13.00-14.00	Lunch
14.00-18.00	Fieldwork/debriefing
	Exterior of 2 dwellings
18.00-19.00	Free
19.00-20.15	Dinner
20.15-	Free

Day 3	
08.30-09.30	HHSRS and the EHS
09.30-10.45	HHSRS worked examples
10.45-11.00	Coffee
11.00-11.20	Introduction to digital pen
11.20-13.00	Digital pen practice.
13.00-14.00	Lunch
14.00-18.30	Fieldwork -
	1 full survey including HHSRS using digital pen and 1 HHSRS debrief in field
18.30-18.45	Downloading survey
19.00-20.15	Dinner
20.15-21.30	Validating data from this afternoons survey
21.30-	Free

Day 4	
08.30-09.15	Module/flat surveys
09.15-10.30	Common parts, shared facilities, Incl. selected faults & treatments
10.30-10.45	Coffee
10.45-11.15	Planning flat surveys
11.15-12.00	Local area
12.00-13.00	Module definition Exercise
13.00-14.00	Lunch
14.00-16.30	Fieldwork : full survey of flat in large block
17.00-18.00	Debriefing in classroom
18.00-19.00	Free
19.00-20.15	Dinner
20.15-	Free

Day 5	
08.30-09.00	Check out of rooms
09.00-10.30	Full survey debriefing
10.30-10.45	Coffee
10.45-11.15	Validation hints and tips
11.15-12.15	Using your data
12.15-13.00	Interacting with households, HH interview Surveying vacant dwellings
13.00-14.00	Lunch
14.00-15.00	Survey admin/website
	Coffee, Pick up packs and leave

Annex 3.2: Advance letters

Advance letter (generic version, front page)

Ministry of Housing,
Communities &
Local Government

ENGLISH
HOUSING
SURVEY

The Householder

Reference no.: P15602

Help shape the future of housing

Dear Sir or Madam,

You've been selected to take part in the English Housing Survey—we hope we can count on your help in this important survey.

Each year a number of people are chosen at random from across the country to be interviewed, so this is your unique chance to take part in this research.

The survey will help to shape the future of housing—improving the experiences of home owners, social and private renters, reducing our carbon emissions, and helping young people get a foot on the housing ladder.

What's next?
One of our interviewers will call at your house in the next week or so. So you know who they are, they all carry a photo ID.

Any questions?
If you have any questions, please call us on freephone **0800 652 4572** or visit **www.natcen.ac.uk/taking-part/ehs**

Thanks for your help
Your help will really give us a better understanding of the housing situation in England today, so thanks again for playing such an important part. To say thank you for taking part, we have enclosed a £5 **post office voucher** for your household.

Yours Sincerely,

Reannan Rottier
Head of English Housing Survey
ehs@communities.gov.uk

NatCen
Social Research that works for society

Gen_May2021

Advance letter (owner occupied sector, front page)

Ministry of Housing,
Communities &
Local Government

ENGLISH HOUSING SURVEY

The Householder
<Address1>
<Address2>
<Address3>
<Address4>
<Address5>
<postcode>

<Post_Date>
Reference no.: P15602/<SerialNumber><CL1>

Help shape the future of housing

Dear Sir or Madam,

You've been selected to take part in the English Housing Survey—we hope we can count on your help in this important survey.

Each year a number of people are chosen at random from across the country to be interviewed, so this is your unique chance to take part in this research.

The survey will help to shape the future of housing—improving the experiences of home owners, social and private renters, reducing our carbon emissions, and helping young people get a foot on the housing ladder.

What's next?

One of our interviewers will call at your house in the next week or so.
So you know who they are, they all carry a photo ID.

Any questions?

If you have any questions, please call us on freephone 0800 652 4572
or visit www.natcen.ac.uk/taking-part/ehs

Thanks for your help

Your help will really give us a better understanding of the housing situation in England today, so thanks again for playing such an important part. To say thank you for taking part, we have enclosed a £5 post office voucher for your household.

Yours Sincerely,

Reannan Rottier
Head of English Housing Survey
ehs@communities.gov.uk

NatCen
Social Research that works for society

Own_May2021

Advance letter (social rented, front page)

The Householder
<Address1>
<Address2>
<Address3>
<Address4>
<Address5>
<postcode>

<Post_Date>
Reference no.: P15602/<SerialNumber><CL1>

One chance to help improve homes and housing

Dear Sir or Madam,

You've been selected to take part in the English Housing Survey—we hope we can count on your help in this important survey.

Each year a number of people are chosen at random from across the country to be interviewed, so this is your unique chance to take part in this research.

From giving a helping hand with fuel bills to helping improve poor housing conditions—the survey will help shape the future of housing. So you could end up helping your neighbours, family and friends.

What's next?
One of our interviewers will call at your house in the next week or so.
So you know who they are, they all carry a photo ID.

Any questions?
If you have any questions, please call us on freephone **0800 652 4572**
or visit **www.natcen.ac.uk/taking-part/ehs**

Thanks for your help
Your help will really give us a better understanding of the housing situation in England today, so thanks again for playing such an important part. To say thank you for taking part, we have enclosed a **£5 post office voucher** for your household.

Yours Sincerely,

Reannan Rottier
Head of English Housing Survey
ehs@communities.gov.uk

NatCen
Social Research that works for society

Soc_Mar2020

Advance letter (private rented, front page)

Ministry of Housing, Communities & Local Government

ENGLISH HOUSING SURVEY

The Householder
<Address1>
<Address2>
<Address3>
<Address4>
<Address5>
<postcode>

<Post_Date>
Reference no.: P15602/<SerialNumber><CL1>

Help improve homes for your family, friends and neighbours

Dear Sir or Madam,

You've been selected to take part in the English Housing Survey—we hope we can count on your help in this important survey.

Each year a number of people are chosen at random from across the country to be interviewed, so this is your unique chance to take part in this research.

The survey will help to shape the future of housing—improving the experiences of home owners, social and private renters, reducing our carbon emissions, and helping young people get a foot on the housing ladder.

What's next?

One of our interviewers will call at your house in the next week or so. So you know who they are, they all carry a photo ID.

Any questions?

If you have any questions, please call us on freephone 0800 652 4572 or visit www.natcen.ac.uk/taking-part/ehs

Thanks for your help

Your help will really give us a better understanding of the housing situation in England today, so thanks again for playing such an important part. To say thank you for taking part, we have enclosed a £5 post office voucher for your household.

Yours Sincerely,

Reannan Rottier
Head of English Housing Survey
ehs@communities.gov.uk

NatCen
Social Research that works for society

Priv_May2021

Advance letter (back page of all versions)

Frequently Asked Questions

How did you choose my address?

We chose your address at random from the Postcode Address File. This file is held by the Post Office and is available to the public.

What is the interview about?

The interview covers a range of topics including your housing circumstances, energy efficiency in your home and satisfaction with your accommodation and neighbourhood. The information you give helps to provide a clearer picture of homes and housing in England. This means the Government can make more informed decisions to shape the future of housing for everyone.

Who is carrying out the study?

NatCen Social Research carries out the study which is funded by the Ministry of Housing, Communities and Local Government (MHCLG). NatCen is independent of all government departments and political parties. For more information about NatCen visit **www.natcen.ac.uk**.

What will happen to any information I give?

We will treat information you give in strictest confidence under the EU General Data Protection regulation. Your answers will be combined with those of others who take part and presented anonymously in research reports. For details of our purpose in collecting and processing the data of this survey, who is going to be able to access the data and all other privacy-related information, please visit **https://www.gov.uk/ government/publications/english-housing-survey-privacy-notice-for-survey-participants**. We undertake to use and store the information you give us in accordance with the purposes and processes outlined in the privacy statement available on this website. You will not receive any junk mail or marketing as a result of doing the survey.

What are my rights?

Participation in the research is voluntary. We hope that you will choose to take part as we rely on the voluntary cooperation of people invited to take part to get a clear picture of housing in England.

When we ask you for information, we will keep to the law, including the 2018 Data Protection Act and General Data Protection Regulation. You have the right to request that your data be amended if it is not accurate.

MHCLG is the data controller. You can contact the Data Protection Officer at **dataprotection@communities.gov.uk**. If you wish to lodge a complaint, you can contact NatCen at **info@natcen.ac.uk** or **0800 652 4572**, quoting your reference number printed on the other side of this letter. If we cannot resolve your complaint, you can contact the Information Commissioner's Office – **www.ico.org.uk** or phone **0303 123 1113**.

Where can I find out more?

Visit **www.natcen.ac.uk/taking-part/ehs** or phone us free on **0800 652 4572**.
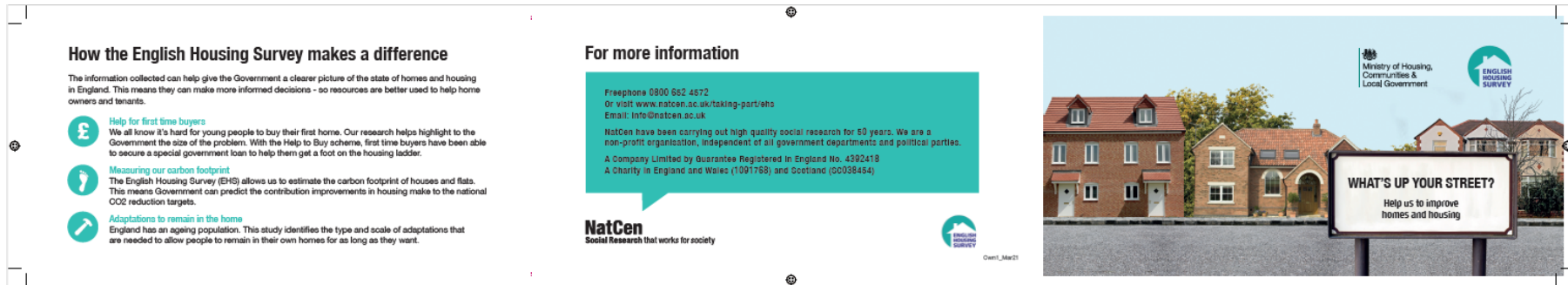
NatCen
Social Research

NatCen Social Research, Kings House, 101-135 Kings Road, Brentwood, Essex CM14 4LX Tel. 0800 526 397. Company limited by guarantee. Reg No. 4392418. A Charity registered in England and Wales (1091768) and in Scotland (SC038454).

Annex 3.3: Leaflets

Interviewer (interview survey) leaflet

Leaflet for owner occupiers

Leaflet for social renters

Leaflet for private renters

How the English Housing Survey makes a difference

The information collected can help give the Government a clearer picture of the state of homes and housing in England. This means they can make more informed decisions - so resources are better used to help tenants and home owners.

Deposits in safe hands
The Tenancy Deposit Scheme helps private tenants protect their deposit from unscrupulous landlords. This study provided information to help set up the scheme.

Health spending
The English Housing Survey (EHS) has been used to estimate that the NHS spends around £1.4 billion per year on treating health problems that are directly linked to poor housing conditions.

Help for first time buyers
We all know it's hard for young people to buy their first home. Our research helps highlight to the Government the size of the problem. With the Help to Buy scheme, first time buyers have been able to secure a special government loan to help them get a foot on the housing ladder.

For more information

Freephone 0800 652 4572
Or visit www.natcen.ac.uk/taking-part/ehs
Email: info@natcen.ac.uk

NatCen have been carrying out high quality social research for 50 years. We are a non-profit organisation, independent of all government departments and political parties.

A Company Limited by Guarantee Registered in England No. 4392418
A Charity in England and Wales (1091768) and Scotland (SC038454)

NatCen
Social Research that works for society
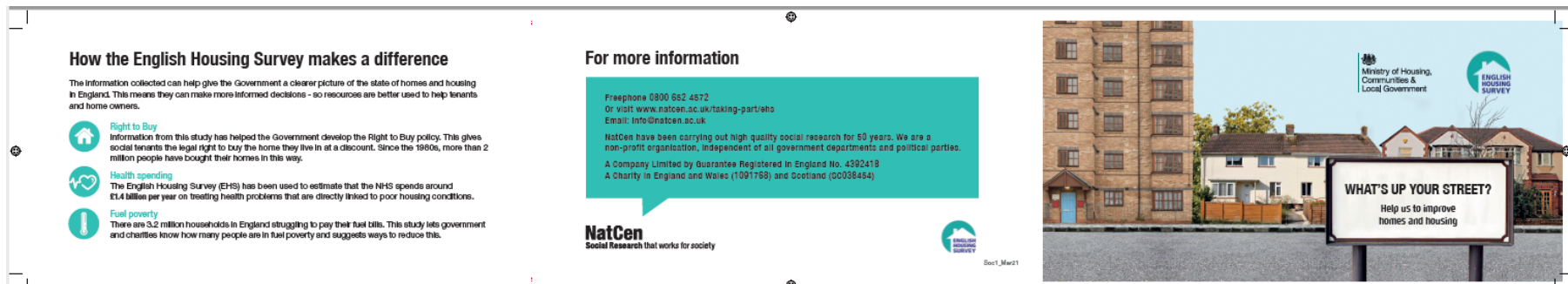
Priv1_Mar21

Ministry of Housing, Communities & Local Government

ENGLISH HOUSING SURVEY

WHAT'S UP YOUR STREET?
Help us to improve homes and housing

Chapter 4
Response rates

4.1   This chapter provides details of the final response rates and information on action taken to maximise response.

Overview

4.2   In 2021-22, the EHS amended its fieldwork approach because of the COVID-19 pandemic and that had implications for response achieved. A socially distanced 'knock-to-nudge' approach was used for the interview. Interviewers visited each issued address in person to introduce the survey and to make an appointment for the interview which was carried out by telephone. The physical survey was amended to an 'external plus' approach where the inspection was restricted to an assessment of the exterior of the dwelling and supplemented by factual data on the interior of the dwelling the surveyor collected (socially distanced) at the doorstep. Fieldwork took place over a compressed fieldwork period, between May, 2021 and April, 2022.

4.3   The Department for Levelling Up, Housing and Communities (DLUHC) and NatCen monitored response on the survey closely through analysis of electronic data reports and direct contact between interviewers and their team leader. NatCen also set and monitored targets for coverage within each wave of fieldwork, monitored response at an area and interviewer level and adapted the approach to optimise response.

4.4   Interviewers collected similar information to standard EHS years and used similar response categories. Vacant and derelict addresses were identified and passed on to the surveyor for the physical survey. Interviewers identified Houses of Multiple Occupation, enumerated the number of dwellings and households and carried out a dwelling or household select where needed. Non-contact and refusals were included as response categories.

Interview survey

4.5   In 2021-22, interviews were achieved at 9,752 households. This represents a response rate of 33%. More details on the response rate to the EHS interview survey in 2021-22 are provided in Table 4.1.

Table 4.1: Interview survey response rate, 2021-22

all issued households

	number (N)	issued cases (%)	in-scope cases (%)
total issued addresses	31411		
not yet built/under construction[1]	49	0.2	
demolished/derelict[1]	58	0.2	
vacant/empty or derelict housing unit[2]	1033	3.3	
non-residential address[1]	395	1.3	
address occupied - no resident household[1,5]	127	0.4	
communal establishment/institution[1]	52	0.2	
other ineligible[1]	110	0.4	
total ineligible addresses	**1824**	**5.8**	
total in-scope addresses	**29,587**	**94.2**	
not issued[3]	0	0.0	0.0
issued, but not attempted[3]	3893	12.4	13.2
inaccessible[3]	44	0.1	0.1
unable to locate address[1]	188	0.6	0.6
unknown whether residential: Info refused	7	0.0	0.0
unknown whether residential: no contact	90	0.3	0.3
residential but unknown eligibility : info refused	24	0.1	0.1
residential but unknown eligibility : no contact	111	0.4	0.4
info refused about whether address is residential	4	0.0	0.0
info refused whether resident(s) are eligible	4	0.0	0.0
eligibility not confirmed: language barrier	23	0.1	0.1
other unknown eligibility	163	0.5	0.6
total unknown eligibility	**4,551**	**14.5**	**15.4**
office refusal	555	1.8	1.9
information refused on no. of dwellings	138	0.4	0.5
information refused on no. of households	315	1.0	1.1
can't identify target respondent(s): info refused	454	1.4	1.5
refusal before interview: by selected respondent	6915	22.0	23.4
proxy refusal	159	0.5	0.5
refusal during interview (unproductive partial)	202	0.6	0.7
broken appointment, no re-contact	2024	6.4	6.8
total refusals[4]	**10,762**	**34.3**	**36.4**
no contact with anyone at address	2761	8.8	9.3
multi dwellings - No contact made with selected dwelling	9	0.0	0.0
no contact with responsible adult	146	0.5	0.5
contact made at dwelling, but not from selected household	115	0.4	0.4
total non-contact[5]	**3,031**	**9.6**	**10.2**
ill at home during survey period: Head Office	64	0.2	0.2
ill at home during survey period: Interviewer	272	0.9	0.9
away or in hospital all survey period: Head Office	20	0.1	0.1
away or in hospital all survey period: Interviewer	216	0.7	0.7
physically/mentally unable/incomp: Head Office	43	0.1	0.1
physically/mentally unable/incomp: Interviewer	444	1.4	1.5
language difficulties: Head Office	17	0.1	0.1
language difficulties: Interviewer	289	0.9	1.0
lost productive	0	0.0	0.0
interview achieved but resp requested data deleted	0	0.0	0.0
interview achieved but research requested data deleted	5	0.0	0.0
other unproductive	121	0.4	0.4
total other unproductive	**1,491**	**4.7**	**5.0**
full interview	9725	31.0	32.9
partial interview	27	0.1	0.1
total interviews	**9,752**	**31.0**	**33.0**

Note: The text in superscript marks the instances where the response rate calculations and sampling processes treat some of the survey outcomes differently. Please see Annex Table 4.1, published alongside this report, for full notes on these differences

4.6   Of these 9,752 interviews, 80% were conducted with the household reference person (HRP), 18% with the HRP's partner and 1% with a proxy respondent answering on behalf of the HRP or partner, Table 4.2.

Table 4.2: Interview respondents, 2021-22

all productive cases

	number (N)	percentage (%)
HRP	7,832	80.3
HRP's partner	1,791	18.4
total non-proxies	**9,623**	**98.7**
proxy for the HRP	124	1.3
proxy for the Partner	5	0.1
total proxies	**129**	**1.3**

Physical survey

4.7   Not every household that takes part in the interview survey is eligible to take part in the physical survey. Cases eligible[17] for the physical survey are identified by the computer-assisted personal interviewing program during the interview survey. Interviewers are then responsible for securing the consent of the householder to a physical survey. The interviewer explains the purpose of the survey and describes briefly what it will involve. If the respondent is willing, the interviewer arranges a fixed appointment for the physical survey between 10 and 12 days after the interview. The CADS Helpline then allocated these appointments to surveyors.

4.8   Of the 7,554 interviewed households eligible for the physical survey, 74% agreed to have a physical survey, which is just below the target of 78%.

4.9   In 2021-22, 5,478 physical surveys were achieved. Of which 5,284 were surveys in occupied properties and 194 of vacant dwellings. This represents 95% of households which agreed to a physical survey at the interview and 32% of identified vacant dwellings.

---

[17] See Chapter 1 for more information on the sub-sampling process that determines how addresses are deemed eligible for the physical survey.

Table 4.3: Physical survey response rate, 2021-22

all addresses eligible for physical survey

	number (N)	cases eligible for PS (%)	cases agreed to PS at IS (%)
occupied addresses			
total occupied addresses eligible for physical survey	**7,554**		
eligible but refused appointment at interview	1,974	26.1	
eligible and agreed appointment at interview	5,580	73.9	
total unproductive			
incomplete survey	0	0.0	.0
refusal on doorstep	146	1.9	2.6
refusal to HQ	19	0.3	.3
partial survey then refusal	3	0.0	.1
household missed appointment	21	0.3	.4
surveyor missed appointment - no contact	1	0.0	.0
surveyor missed appointment - rescheduled	0	0.0	.0
speculative call - no contact	97	1.3	1.7
address untraceable	0	0.0	.0
other reason for non-survey	9	0.1	.2
full survey (paired cases)	**5,284**	**69.9**	**94.7**
vacant/derelict addresses			
total vacant/derelict addresses eligible for physical survey	**786**		
eligible but refused appointment at interview	188	23.9	
total unproductive	**592**	**75.3**	**99.0**
partial survey then refusal	0	0.0	.0
refusal on doorstep	56	7.1	9.4
refusal to HQ	4	0.5	.7
household missed appointment	5	0.6	.8
speculative call no contact	304	38.7	50.8
address untraceable	2	0.3	.3
dwelling derelict	6	0.8	1.0
dwelling demolished	2	0.3	.3
no longer usable as dwelling	1	0.1	.2
other reason for non-survey	24	3.1	4.0
survey achieved (vacant)	189	24.0	31.6
survey achieved (derelict)	5	0.6	.8
total vacant/derelict physical surveys achieved	**194**	**24.7**	**32.4**
total physical surveys achieved	**5,478**	**65.7**	**88.7**

Notes:
  1) for the occupied addresses the cases eligible for a physical survey is the total occupied addresses eligible for the physical survey; and the cases agreed to a physical survey at interview survey is the total occupied addresses who agreed to have a physical survey at the interview.
  2) for the vacant addresses the cases eligible for a physical survey is the total vacant addresses eligible for the physical survey; and the final column is the percentage out of all the eligible occupied addresses excluding those who refused an appointment at the interview stage (not all the vacant addresses were contacted at the interview stage so did not have the chance to refuse).
  3) the cases eligible for a physical survey is all the occupied and vacant addresses eligible for a physical survey: and the cases agreed to a physical survey at the interview is all the occupied and vacant addresses who did not refuse to have a physical survey at the interview stage
  4) a small number of cases with outcome 'address untraceable' and 'dwelling demolished' have been added to other reasons for non-survey for both occupied and vacant dwellings
  5) vacant and appointment led cases not covered due to COVID-19 are included in 'other reason for non-survey

Chapter 5
Data processing

5.1   The data of the English Housing Survey (EHS) undergo a series of checks and quality assurance after they have been collected. The process begins at the point of data collection, both through the computer-aided personal interviewing (CAPI) system and through surveyors validating their forms using the online system developed by the Building Research Establishment (BRE). This chapter outlines those procedures and also gives information about the main derived variables and data outputs.

5.2   To conform with social distancing restrictions introduced during the COVID-19 pandemic, the 2021-22 EHS interview surveys were carried out on the telephone. Prior to the telephone survey, a 'knock to nudge' approach was undertaken which involved visiting a respondent at their home and, while socially distanced, encouraging them to respond to the survey by telephone. The physical survey data were collected using an 'external plus' approach where the inspection was restricted to an assessment of the exterior of the dwelling and supplemented by information about the interior of the dwelling the surveyor collected (socially distanced) at the doorstep. Some data items could not be collected at all. Data items collected by the 'external plus' approach are highlighted by a red border on the version of the 2021-22 EHS physical survey form available using this link
https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/1124155/EHS_2021-22_Physical_Survey_Form.pdf

5.3   In some instances, modelled estimates were produced for the 2021-22 EHS annual reports. This Chapter outlines the approaches used to process the 2021-22 data. Some of those have been used on previous surveys while others are new approaches required to take account of the changes to the data collection process brought about by the COVID-19 pandemic. The impact of COVID-19 on data quality and analysis is summarised in Chapter 8 of this report.

Data editing

Interview data

5.4   The Computer Assisted Personal Interview (CAPI) program has numerous built-in checks for identifying obvious discrepancies so that they can be resolved by the interviewer during the interview. The discrepancies are

resolved by either correcting a data entry error or by clarifying a response directly with the respondent. The CAPI checks include:

- range checks to identify where the answer falls outside a pre-specified range of responses, for example, an unusually high/low weekly rent is entered; and
- conflicting answers to different questions, for example, if the number of years living in the current accommodation is greater than the respondent's age.

5.5   There are two types of checks.

- Hard checks – where the interviewer cannot continue with the interview until they have changed the data entered in some way to remove the inconsistency. Hard checks are used when the inconsistency is impossible as with the example of the number of years living in current accommodation being greater than the respondent's age.
- Soft checks (signals) – where the interviewer is told about the error, but they can ignore it and move on to the next question. Soft checks are used when an answer is unlikely but not impossible, e.g. if a respondent says they have more than 5 bathrooms. These checks are used to get the interviewer to confirm that the answer is correct and is not a data entry error, checking the answer with the respondent if appropriate.

Physical survey data

5.6   For the physical survey, a system of automatic data validation was introduced in 2008 as part of the move to using digital pens to collect the data. The process is subject to continuous development and operates in three stages.

5.7   First, a large number of checks are built into the EHS surveyors' website as surveys are uploaded. These include:

- range checks – to identify where the entered answer falls outside a pre-specified range of responses;
- logic checks – where a combination of responses to certain questions are not logically consistent (e.g. to check that the sum of 'tenths of area' across rows added up to ten);
- consistency checks – to determine whether linked responses in different parts of the form are consistent with each other (e.g. that detailed room data are only entered where a room coded as existing); and
- plausibility checks – to determine whether a response is reasonable given that there is not a well-defined range of possible answers (e.g. ceiling height of a room entered as 24 metres instead of 2.4 metres).

5.8   Surveyors also visually check all pages to ensure that the digital pen entries mirror those on the paper form, i.e. that handwritten numbers have not been misinterpreted by the software.

5.9   Second, the CADS Housing Surveys regional managers check the data and where necessary discuss with surveyors to agree on a final set of responses.

5.10   Once all EHS physical surveys have been submitted by the surveyors for the survey year, BRE undertakes further consistency and plausibility checks on the raw physical survey data. The purpose of these checks is, firstly, to detect and eliminate certain logical inconsistencies that would cause problems for data modelling and, secondly, to identify highly implausible answers, which if deemed necessary after investigation, are corrected. In some cases, the raw EHS physical survey data are altered following these consistency and plausibility checks as outlined below.

- Levels checks – data may be inconsistent regarding the number of storeys in the building, and the floor occupied by the dwelling. The BRE checks test for the following possible errors:
    o a room on a floor that does not exist (e.g. 3rd floor of a three storey block, the 3 floors being recorded as ground, first and second);

    o a room on a level that is not part of the flat (e.g. room on the 3rd floor but flat on the 2nd floor);

    o a measured floor that is not part of the block (e.g. dimensions for 3rd floor when the dwelling only has three storeys);

    o a flat on a level that does not exist (e.g. flat on the 3rd floor when the module only has three storeys); and

    o presence of a habitable attic/basement is inconsistent with the number of floors.

- Plausible dimensions – checks are carried out on the dimensions, to identify any floor area that seems too large or too small. Where a reliable measurement is missing, BRE will attempt to work out the data from any measurements thought to be correct, or failing this, by estimating the dimensions as best as possible from the photographs.

- Non-permissible values – on rare occasions a surveyor response may happen to be equal to a value that is reserved for special purposes. The numbers 77, 88 and 99 are reserved to indicate that the section is not applicable, question is not applicable, or the information is unknown. When these figures occur as real measurements or counts, they are reduced by one.

- Incorrect number of flats – the dimensions of the surveyed flat are checked against the total floor area of the survey module to identify if the number of flats per module seems realistic.

- Incorrect roof type – certain roof types (chalet and mansard) can only occur where the dwelling has an attic. On occasions, surveyors may mistake steep pitched roofs for chalet roofs. In this situation, the data for pitched and chalet roofs are swapped over.
- Implausible wall and window areas/fenestration ratio – where a dwelling seems to have a wall or window area/fenestration ratio that is either too high or too low, the data are checked. The surveyor's judgement is deemed correct unless there is clear evidence (e.g. from photographs) to amend the data.
- Wall thickness – cases are identified where the wall thickness as measured by the EHS surveyor is not typical of the wall selected, i.e. cases where the EHS surveyors' website has triggered a wall thickness range check. Each case is checked by viewing the EHS surveyors' website and looking at the details recorded on the physical survey form in conjunction with the photos/EHS surveyor comments. Based upon the information gained, the action is decided upon for each case. This could be no action required or it could be that the physical survey data looks incorrect, either the wall thickness value or the way the surveyor has coded something as wall that should not be counted as wall. Where required, the appropriate modifications are applied to the physical survey data.
- Heating system consistency checks – cases which contain inconsistent heating system data on the physical survey form are flagged in the validation process at BRE. Each case is checked by returning to the raw data; in cases where alterations can confidently be made, the data are modified accordingly.

Comparison edits

5.11 A further important quality check involves comparing interview survey data with the corresponding physical survey data for each case. The first step is a series of global edits to resolve discrepancies in the data. For example:

- If tenure in the interview survey (IS) was owner occupied AND tenure in the physical survey (PS) was another tenure, the PS tenure would be changed to owner occupied.
- If tenure in the interview survey (IS) was renting from local authority AND tenure in the physical survey (PS) was another tenure, the PS tenure would be changed to renting from local authority.

5.12 The remaining discrepancies between the two parts of the survey are flagged, investigated and recoded where applicable. This process is carried out to:

- check that the correct sampled dwelling was visited at both the interview survey and the physical survey; and
- correct any inconsistencies in key variables (e.g. tenure or property type) between the two different parts of the survey. Where possible

other information from the survey (e.g. number of bedrooms, interviewer's and surveyor's comments, photo of the property) is checked to help decide what information is correct.

Houses in multiple occupation (HMO) edits

5.13   An HMO is a property rented by more than one person who are not from one 'household' (e.g. a family) but share facilities like the bathroom or kitchen. These differ from a shared house in that the residents in an HMO generally have separate tenancy agreements and usually have begun their tenancies independently of each other. The identification of HMOs is critical in order to help ensure the accuracy of the weighting for the sample dwelling. The procedure for monitoring, reconciliation and validation of cases which have been flagged as HMOs by NatCen Social Research interviewers and/or CADS Housing Surveys surveyors is described below.

5.14   Whether a case is flagged as an HMO or not is dependent upon responses to certain key questions in the household questionnaire. Interviewers are trained in applying the EHS household definition and assessing the type of occupancy in complex situations, particularly in making the distinction between a group of sharers forming one household and separate households sharing facilities. Where necessary, reference is made to a check list of supplementary questions on the HMO Rules Card issued to interviewers (Annex 5.1) to help determine whether an address should be classified as an HMO.

5.15   Where the responses to the interview questions lead to the dwelling being flagged as an HMO or possible HMO, and the dwelling is eligible for a physical survey, the CADS Housing Surveys Regional Manager is notified. The Regional Manager will contact the interviewer to discuss the layout and occupation of the premises. The purpose of this contact is twofold:

- to confirm, as far as possible, that the address is an HMO for EHS purposes; and
- to determine whether the case is one that should be visited by the Regional Manager personally, as a complex HMO, or whether it should be allocated to a surveyor.

5.16   There are occasions when a physical surveyor considers that a referred address appears to be an HMO despite not being flagged as such by the NatCen Social Research interviewer. In such cases, the surveyor will treat the case as an HMO, and a reconciliation process is applied to the interview and physical data during the final data validation stage, with confirmation on the exact nature of the case sought from the interviewer or surveyor where necessary.

5.17   CADS Housing Surveys Regional Managers compile and maintain a database of all cases they know to be HMOs. These cases, along with cases flagged as HMOs at the interview survey but which did not have a subsequent physical survey, are reviewed by BRE for data validation as part of the comparison edits process. The HMO checking process also includes cases that were not identified as HMOs at the interview survey but which the data suggest could potentially be HMOs. BRE checks relevant interview and physical survey data such as number of households (NumHhld) and number of accommodation units (AcNumber). Where there are inconsistencies, further investigation is undertaken, and the data altered to the correct values.

5.18   During the HMO comparison process, BRE also derives the ratios of addresses to dwellings and dwellings to households. This information is required to ensure the correct numbers of dwellings and households are used in the production of weights. As part of the QA process, DLUHC conducts spot checks on these ratios as well as the HMO edits resulting from the process above.

5.19   A record of all address changes is kept by interviewers and/or Regional Managers for HMO cases as part of a comprehensive system for recording address changes for all issued cases. This feeds into the address file supplied to DLUHC at the end of fieldwork.

5.20   Due to the change in survey methodology for the interview and physical surveys required to conform with the social distancing restrictions, it was not possible for many HMOs to be identified in comparison with previous EHS surveys. Consequently, the weighting methodology was adjusted to account for this (see Chapter 6).

Coding

5.21   After the interview, the data are coded and edited by trained coders and editors at NatCen Social Research. An edit program is used to code open answers and back-code responses as appropriate. For example, at the interview, respondents are asked how they pay for their electricity (question HmpyElec3), and the respondent is shown eight possible answers on a card. If their payment method is not on the list the interviewer will code 'other' and is asked to enter the details of the payment method at a follow up question (Hmelothr). After the interview, the coder will look at the details given at Hmelothr and check it against the eight answer codes to see whether it could be classified as one of these payment methods and if it can they will change the answer as appropriate (i.e. backcode the answer). Job details are coded to the Standard Occupation Classification (SOC) and the Standard Industry Classification (SIC).

5.22   Errors detected by the edit program are resolved by referring back to the original questionnaire documents by experienced editors. Individual corrections are made to the data and the corrected data are rerun through the edit programme until it confirms that the data have passed all the checks. Queries arising from the coding and editing process are recorded in a standardised way and these are examined by the supervision team on completion of each batch of work to ensure that they have been carried out correctly.

5.23   After the coding and editing stage further internal consistency checks on the data are carried out by a data manager and the data are corrected where appropriate.

Derived variables

5.24   Derived variables are created either by simply recoding a particular survey question or by combining the information collected from a number of questions, which can involve complex modelling. Examples of basic derived variables include dwelling age and dwelling type and examples of complex derived variables include basic repair costs (which could not be modelled for the EHS in 2020 and 2021 using the external plus survey approach), usable floor area and energy efficiency rating. The derived variables and geo-demographic variables, such as region, rurality and Index of Multiple Deprivation, included in the key EHS derived datasets interview.sav, physical.sav and general.sav can be found in Annex 5.2.

5.25   In addition to the three key EHS derived datasets, further detailed derived files such as dimensions 20+21.sav, energy performance 20+21.sav and 20+21 paired sample equivalised income.sav are available via the UK Data Service, as listed in Table 5.1. For the 2021-22 EHS, however, some of the physical derived variables, for example those relating to the Housing Heath and Safety Rating System (HHSRS), Decent Homes and repair costs, are not available because the raw data required for creating/modelling them could not be collected. Nevertheless, it was possible to produce modelled estimates at dwelling level on a limited number of housing quality variables. They include: whether the Decent Homes Standard was met; the cost of making a non-decent home meet the Standard; the presence of any HHSRS Category 1 hazards and the presence of damp. These variables are included in an additional file called modelled housing quality variables 20+21.sav (see Annex Table 5.2.4 and Annex 5.5).

5.26   Further details on the derivation of these derived and detailed variables are available in the EHS Data Dictionary, made publicly available via the UK Data Service (http://ukdataservice.ac.uk/).

5.27   The EHS derived variables are included in the datasets made available from UK Data Service. To comply with the data disclosure control guidance issued by the Government Statistical Service, some of the variables are released under the more restricted Special Licence rather than through the End User Licence. In addition, the very disclosive geo-demographic variables (local authority and postcode) are available only through the Service's Secure Data Service.

Modelling

5.28   The derivation of some of the derived variables involves complex data modelling. A detailed description of how the more complex derived variables are defined and modelled is covered in Annexes to this chapter:

- Annex 5.3: Accessibility indicators
- Annex 5.4: Household derived indicators
- Annex 5.5: Housing conditions
- Annex 5.6: Energy efficiency
- Annex 5.7: Dimensions
- Annex 5.8: Poor quality environments

Imputation

5.29   As part of the modelling processes, it is sometimes necessary for any missing data to be substituted with imputed values. The imputation of missing data is more prevalent with the interview survey data than the physical survey data. This is because the interview survey data are based on information provided by the householder who can choose to refuse questions or who may not know the answer to particular questions. The physical survey data are based upon a physical inspection of the property and there are only a few sections of the physical survey form where the trained surveyor can select 'information unknown' as an option; the most notable is the loft inspection, where surveyors cannot always obtain access.

5.30   Imputation of data also takes place in the modelling of derived variables where a value provided in the raw data falls outside the pre-specified range of the consistency/plausibility checks. Such values are interrogated and only changed when it is almost certain that the data are incorrect. See Annexes 5.3, 5.4, 5.5, 5.6, 5.7 and 5.8 for further details.

5.31   Examples of imputation that occurred in the modelling of EHS 2021-22 derived variables are as follows (figures are based upon weighted data):

- In the modelling of the derived variables from the EHS 2021-22 interview survey, 50% of the weighted full sample had some form of

income imputation (the highest imputation rate of all of the derived variables due to the sensitive nature of the questions), 11% of renters had weekly rents imputed and 20% of households with a mortgage had their weekly mortgage payments imputed. These imputations were due to a combination of missing raw data and implausible values. The 50% figure for imputation of income includes any change to any component of household income. This may only be to change the amount received from a particular benefit by a very small amount, which would not significantly affect the total household income.

- In the modelling of derived variables from the EHS 2021-22 physical survey on the dimensions of the property e.g. derivation of floor area, external wall area etc., a total of 250 cases in the paired single year dataset had some form of alteration to the raw physical survey data following consistency and plausibility checks on the raw physical survey data.

5.32   Where appropriate, the EHS Annual Reports contain details on the approach used to handle the cases that are missing from the raw physical and interview data during analysis.

Data outputs

5.33   A range of EHS data files are produced annually and released via the UK Data Service under the End User Licence and the Special Licence, Table 5.1.

5.34   For the EHS 2021-22 some of the physical derived variables and detailed derived data files are not available because the raw data required for the modelling could not be collected due to the change in physical survey methodology required to conform with social distancing restrictions. While detailed derived variables relating to HHSRS, actual and standardised repairs costs are not available, it was possible to produce modelled estimates at dwelling level on the following: whether the Decent Homes Standard was met; the cost of making a non-decent home meet the Standard; the presence of any HHSRS Category 1 hazards (see Annex 5.5) and the presence of damp. These variables are included in an additional file called modelled housing quality variables 20+21.sav.

Table 5.1 List of annual datasets

Physical data files	Interview data files	Detailed derived data files	Derived data files (paired sample)	Derived data files (full household sample)
Available via the Special Licence only			**Available under both End User Licence and Special Licence**	
Amenity.sav			General 20+21.sav	Generalfs21.sav
Around.sav		Dimensions 20+21.sav	Physical 20+21.sav	Interviewfs21.sav
Chimney.sav	Attitudes.sav	Energy performance 20+21.sav	Interview 20+21.sav	
Commac.sav	Contact.sav	2021 full sample equivalised income.sav 20+21 paired sample equivalised income.sav	Modelled housing quality variables 20+21.sav	
Common.sav	Damp.sav			
Damppc.sav	Disability.sav			
Doors.sav	Dwelling.sav			
Dormers.sav	Employment.sav			
Elevate.sav	Energy.sav			
Firstimp_PS.sav	Fire.sav			
Flatdets.sav	Firstimp.sav			
Hhsrs.sav				
HQ.sav	HhldType.sav			
Interior.sav	Identity.sav			
Introoms.sav	Income.sav			
Numflats.sav	Owner.sav			
Plotlvl.sav	People.sav			
Roofcov.sav	Renter.sav			
Rooffeat.sav	Rooms.sav			
Roofstru.sav	Vacant.sav			
Services.sav	WaitList.sav			
Shape.sav				
Shared.sav				
Structure.sav				
Wallfin.sav				
Wallstru.sav				
Windows.sav				

5.35 The data, user guides and supporting documentation are publicly available from the UK Data Service (https://ukdataservice.ac.uk/). Datasets can be downloaded in SPSS and SAS format.

5.36 Prior to releasing the data to the UK Data Service, all disclosive variables are removed to maintain the confidentiality of respondents. Some response categories are also condensed, several variables are top coded, and, in a few rare situations, data swapping between cases takes place for disclosure control reasons.

Annex 5.1: HMO rules card

Determining Houses in Multiple Occupation ('HMOs')

The following supplementary questions will help determine whether the dwelling is an HMO. Count the number of "yes" responses to the **key questions** and **additional questions** then use the table below to determine whether the pattern of responses indicates that the dwelling is an HMO or instead identifies the dwelling as a single household/shared household (i.e. NOT an HMO).

<u>Key questions</u> (ranked in order):
1. Does the landlord find new tenants to fill any rooms that become vacant (as opposed to the remaining tenants fulfilling this function)?
2. Does the landlord bear the cost of any shortfall in rent if one or more tenants defaults or vacates (as opposed to the remaining tenants having to pay)?
3. Did the occupiers come to rent the house separately (as opposed to a single group of renters)?

<u>Additional questions</u> (not in any ranking order):
4. Do the tenants identify themselves as multiple households?
5. Is there a large group of occupiers (i.e. more than 5 persons)?
6. Is there a separate tenancy agreement for each occupant (as opposed to a joint tenancy)?
7. Do individual tenants keep their rooms locked, excluding other tenants from their accommodation?
8. Has there been a large turnover of occupiers since the commencement of the tenancy? (Say more than 40% over last 18 months.)
9. Is there a mix of different types of occupants at the premises? (e.g. they are not all students)

Pattern of responses	Meaning
6+ questions answered 'Yes'	= bed-sits (which collectively would form an HMO)
4-5 questions answered 'Yes' including at least 1 'key question'	= bed-sits (which collectively would form an HMO)
4-5 questions answered 'Yes' but not any 'key questions'	= single household/shared house (i.e. NOT an HMO)
Less than 4 questions answered 'Yes'	= single household/shared house (i.e. NOT an HMO)

Annex 5.2: List of derived variables

5.37   This Annex includes a full list of derived and geo-demographic variables found on the End User Licence (EUL) and Special Licence (SL) versions of interview.sav, physical.sav, general.sav and modelled housing quality estimates files, Annex Tables 5.2.1, 5.2.2, 5.2.3 and 5.2.4 respectively. Further details on the derivation of these variables are available in the EHS Data Dictionary, publicly available on the UK Data Service (http://ukdataservice.ac.uk/).

5.38   Due to COVID-19 restrictions in 2021-22 it was not possible to collect data about the internal condition of the home. Instead, EHS interview surveys were carried out on the telephone and the physical survey data were collected from an external inspection of properties by a surveyor and supplemented by information about the interior of the dwelling the surveyor collected (socially distanced) at the doorstep. Data items collected by the 'external plus' approach are highlighted by a red border on the version of the 2021-22 EHS physical survey form. Although many physical derived variables are not available for the 2021-22 EHS, it was possible to produce modelled estimates of some housing quality measures (non-decency, HHSRS Category 1 hazards and damp) to support users interested in time series findings. Variables representing these estimates can be found in modelled housing quality variables 20+21.sav, Annex Table 5.2.4. For further details on how these were modelled, please see Annex 5.5.

5.39   In Table 5.2.1 the Household Reference Person (HRP) is the person in whose name the dwelling is owned or rented or who is otherwise responsible for the accommodation. In the case of joint owners and tenants, the person with the highest income is taken as the HRP. Where incomes are equal, the older is taken as the HRP. This procedure increases the likelihood that the HRP better characterises the household's social and economic position.

Annex Table 5.2.1: interview 20+21.sav and interviewfs21.sav

accomhh	Type of accommodation for household	SL only
accomhh1	Type of accommodation for household & if not self-contained	EUL and SL
agehrp2x	Age of the Household Reference Person (HRP) - 2 band	EUL and SL
agehrp4x	Age of HRP - 4 band	EUL and SL
agehrp6x	Age of HRP - 6 band	EUL and SL
agehrpx	Age of HRP - continuous	SL only

agen16	Number of persons under 16 in household	SL only
ageoldbx	Age of oldest person in household - banded	EUL and SL
ageoldx	Age of oldest person in household	SL only
agepart6x	Age of partner - 6 band	EUL and SL
agepartx	Age of partner - continuous	SL only
ager	Report age categories	EUL and SL
AHCinceq	AHC equivalised weekly income (modified OECD scale)	EUL and SL
AHCinceqv5	AHC equivalised income quintiles (weighted by peoplegross)	EUL and SL
AHCinceqv60h	AHC: below 60% of median income (weighted by peoplegross)	EUL and SL
ALLincx	Annual gross income of the HRP and partner inc. income from housing benefit and LHA	EUL and SL
amthbenx	Weekly housing benefit	EUL and SL
atspaany	Is any household member at state pension age?	EUL and SL
atspahrp	Is HRP at state pension age?	EUL and SL
atspaprt	Is partner at state pension age?	EUL and SL
bedrqx	Number of bedrooms required by the household (2006 definition)	EUL and SL
bedstdx	Bedroom standard (2006 definition)	EUL and SL
BHCinceq	BHC equivalised weekly income (modified OECD scale)	EUL and SL
BHCinceqv5	BHC equivalised income quintiles (weighted by peoplegross)	EUL and SL
BHCinceqv60h	BHC: below 60% of median income (weighted by peoplegross)	EUL and SL
buypresh	Year HRP bought present accommodation	SL only
cohabhrp	If HRP is cohabiting	SL only
cohabprt	If HRP's partner is cohabiting	SL only
emphrp3x	Working status of HRP (primary) - 3 categories	EUL and SL
emphrpx	Employment status (primary) of HRP	EUL and SL
empprt3x	Working status of partner (primary) - 3 categories	EUL and SL
empprtx	Employment status (primary) of partner	EUL and SL
equityr	Equity in home (based on respondent valuation only)	EUL and SL
equityr5	Equity in home (based on respondent valuation only)	EUL and SL
ethhrp2x	Ethnic origin of HRP - 2 categories	EUL and SL
ethhrp4y	Ethnic origin of HRP - 4 categories	SL only
ethhrp8y	Ethnic origin of HRP - 8 categories	SL only
ethprt2x	Ethnic group of HRP's partner - 2 categories	EUL and SL
ethprt8y	Ethnic group of HRP's partner - 8 categories	SL only
famnumx	Number of family units in household	SL only
freeLeas	Freehold or leasehold	EUL and SL
ftbuyer	If first-time buyer	EUL and SL
grossA	Age group & sex of youngest person in household	SL only
hatentp3	Type of housing association tenancy	EUL and SL
hhbensx	Household on means tested bens or tax credits with a relevant income below the threshold	EUL and SL
hhcomp1	Household composition	EUL and SL
hhcompx	Household composition	EUL and SL
hhempx	Employment status of HRP and partner combined	EUL and SL

hhinc5x	All households - income in 5 bands	EUL and SL
hhincflg	Imputations used to create net total household income	SL only
hhincx	EHS Basic Income (annual net household income (HRP + Partner) including savings)	EUL and SL
hhltsick	Anyone in household have long term illness or disability?	EUL and SL
hhsizex	Number of persons in the household	EUL and SL
hhtype6	Household type - 6 categories	EUL and SL
hhtype7	Household type - 7 categories	EUL and SL
hhtype11	Household type - full 11 categories	EUL and SL
hhvulx	Household vulnerable - on means tested or certain disability related benefits?	EUL and SL
hhwhch	Anyone in household uses a wheelchair	EUL and SL
housbenx	Household (HRP + partner) receives any housing benefit?	EUL and SL
HYEARGRx	Household gross annual income (inc. income from all adult household members)	EUL and SL
JOINTINCx	Annual gross income of the HRP and partner	EUL and SL
lenown	Length of ownership (years)	SL only
lenownb	Length of ownership to date of survey	EUL and SL
lenres	Length of residence (years)	SL only
lenresb	Length of residence	EUL and SL
lharqx	Number of bedrooms required by the household (2011 definition)	EUL and SL
lhastdx	Bedroom standard (2011 definition)	EUL and SL
loncoupx	Single householder or with partner	SL only
market_rent	Tenancy types of renters	EUL and SL
mortwkx	Weekly mortgage payments	EUL and SL
nbatha	If shares a bathroom, shower room or WC	SL only
nbedsx	Total number of bedrooms household actually has	EUL and SL
ncouple	Number of couples in household	SL only
ndepchild	Number of dependent children in household	EUL and SL
nEmp	Number of employed persons in household	SL only
nInac	Number of economically inactive persons in household	SL only
nkita	If shares a kitchen	SL only
nliving	If shares other room (living room)	SL only
nlpar	Number of lone parent families in household	SL only
noUnits1	Banded number of family units in household	EUL and SL
nrooms1a	Number of rooms available to household (grouped)	EUL and SL
nroomsa	Number of rooms available to household	SL only
nshare	If shares any part of accommodation	SL only
nsing	Number of one-person family units in household	SL only
nssech9	NS-SEC Socio-economic Classification - HRP	EUL and SL
nssecp9	NS-SEC Socio-economic Classification - HRP's partner	EUL and SL
nStud	Number of students in household	SL only
nUnemp	Number of unemployed persons in household	SL only
nxdepch	Number of non-dependent children in household	SL only
olderx	Number of people aged 60 plus who are HRP or partner	SL only

otherfam	Additional families present in household	SL only
othfamlp	Type of additional families in household	SL only
owntype	Type of ownership	EUL and SL
prevten	Previous tenure	EUL and SL
pyngbx	Age band of youngest person in household	EUL and SL
pyngx	Age of youngest person in household	SL only
rentExS	Total weekly rent excluding the cost of services	EUL and SL
rentflg	Rent/housing benefit changed/imputed	SL only
rentwkx	Total weekly rent payable (rent plus housing benefit)	EUL and SL
sexhrp	Sex of HRP	EUL and SL
sexprt	Sex of HRP's partner	SL only
sft	Number of full-time workers in household	EUL and SL
sharer	If shares accommodation with other household	EUL and SL
srtentype2	Type of social sector tenancy	EUL and SL
studhrp	If HRP is a full-time student	SL only
studprt	If HRP's partner is a full-time student	SL only
tenex	Extended tenure of household	EUL and SL
tenure1	Tenure group 1	SL only
tenure2	Tenure group 2	EUL and SL
tenure3	Tenure group 3	SL only
tenure4	Tenure group 4	EUL and SL
totkitsa	Number of unshared kitchens	SL only
workless	Household with no one of working age employed - ILO defn	EUL and SL

5.40   The derived physical variables that are not available for the 2021-22 EHS, are repair cost variables, reasons for failing the Decent Homes Standard, types of damp and presence of secure windows and doors. Twenty five variables are not available in physical 20+21.sav. They are highlighted in Annex Table 5.2.2 using strikethrough text.

Annex Table 5.2.2: physical20+21.sav

alltypex	Dwelling age and type	EUL and SL
area3x	Type of area	SL only
arnatx	Nature of area	SL only
attic	Attic present in dwelling	EUL and SL
basement	Basement present in dwelling	SL only
boiler	Type of boiler	EUL and SL
constx	Construction type	SL only
~~Cstactbx~~	~~Basic repair costs (actual)~~	~~EUL and SL~~
~~cstactcx~~	~~Comprehensive repair costs (actual)~~	~~EUL and SL~~
~~cstactux~~	~~Urgent repair costs (actual)~~	~~EUL and SL~~
~~cststdbx~~	~~Basic repair costs (per square metre)~~	~~EUL and SL~~
~~cststdcx~~	~~Comprehensive repair costs (per square metre)~~	~~EUL and SL~~

~~cststdux~~	~~Urgent repair costs (per square metre)~~	~~EUL and SL~~
~~dampalf~~	~~Dampness problems in one or more rooms~~	~~EUL and SL~~
~~dampcdf~~	~~Serious condensation in one or more rooms~~	~~EUL and SL~~
~~damppnf~~	~~Penetrating damp in one or more rooms~~	~~EUL and SL~~
~~damprsf~~	~~Rising damp in one or more rooms~~	~~EUL and SL~~
dblglaz2	Extent of double glazing	EUL and SL
dblglaz4	Extent of double glazing	EUL and SL
~~dhcosty~~	~~Cost to make decent (15 hazard HHSRS model)~~	~~EUL and SL~~
~~dhdisrx~~	~~Decent homes repair criterion~~	~~EUL and SL~~
~~dhhhsrsx~~	~~Decent homes HHSRS (15 hazard) criterion~~	~~EUL and SL~~
~~dhhhsrsy~~	~~Decent homes HHSRS (26 hazard) criterion~~	~~EUL and SL~~
~~dhmodx~~	~~Decent homes modern facilities criterion~~	~~EUL and SL~~
~~dhnumy~~	~~Decent homes: number of criteria failed (15 hazard HHSRS model)~~	~~EUL and SL~~
~~dhnumz~~	~~Decent homes: number of criteria failed (26 hazard HHSRS model)~~	~~EUL and SL~~
~~dhomesy~~	~~Decent homes - overall standard (15 hazard HHSRS model)~~	~~EUL and SL~~
~~dhomesz~~	~~Decent homes - overall standard (26 hazard HHSRS model)~~	~~EUL and SL~~
~~dhreasny~~	~~Decent homes criterion not met (15 hazard HHSRS model)~~	~~EUL and SL~~
~~dhreasnz~~	~~Decent homes criterion not met (26 hazard HHSRS model)~~	~~EUL and SL~~
~~dhtcacty~~	~~Requirement to pass decent homes thermal comfort criterion~~	~~SL only~~
~~dhtcreasy~~	~~Reason for failing decent homes on thermal comfort?~~	~~EUL and SL~~
~~dhthermy~~	~~Decent homes thermal comfort criterion~~	~~EUL and SL~~
dwage4x	Dwelling age	EUL and SL
dwage5x	Dwelling age	EUL and SL
dwage6x	Dwelling age	EUL and SL
dwage7x	Dwelling age	EUL and SL
dwage10x	Dwelling age	SL only
dwtype3x	Dwelling type	EUL and SL
dwtype7x	Dwelling type	SL only
dwtype8x	Dwelling type	EUL and SL
dwtypenx	Dwelling type	EUL and SL
EPceeb12e	Energy efficiency rating band (SAP 2012)	EUL and SL
EPceib12e	Environmental impact rating band (SAP 2012)	EUL and SL
EPceir12e	Environmental impact rating (SAP 2012)	EUL and SL
floor5x	Useable floor area - original EHS definition	EUL and SL
floor5y	Useable floor area - building regulations definition	EUL and SL
floorx	Useable floor area (sqm) - original EHS definition	EUL and SL
floory	Useable floor area (sqm) - building regulations definition	EUL and SL
fuelx	Main fuel type	EUL and SL
heat4x	Main heating system	EUL and SL
heat7x	Main heating system	EUL and SL
heatsec	Secondary heating type	SL only
housex	Dwelling type	EUL and SL

loftins4	Loft insulation thickness	EUL and SL
loftins6	Loft insulation thickness	EUL and SL
loftinsu	Loft insulation thickness with unknowns	EUL and SL
loftinsx	Loft insulation thickness	EUL and SL
loftu4	Loft insulation thickness with unknowns	EUL and SL
lv1upkpx	Poor quality environment - upkeep problems	EUL and SL
lv2trafx	Poor quality environment - traffic problems	EUL and SL
lv3utilx	Poor quality environment - utilisation problems	EUL and SL
lvanyx	Poor quality environment	EUL and SL
lvnumx	Number of liveability problems present	EUL and SL
mainfuel	Main heating fuel	SL only
mainsgas	Mains gas supply present	EUL and SL
neivisx	Appearance of area	EUL and SL
parking	Parking provision of survey dwelling	EUL and SL
pcavwallx	Percentage of external wall area that is cavity masonry	EUL and SL
rdsap09	Reduced data energy efficiency rating (SAP 2009 based)	EUL and SL
sap12	Energy efficiency (SAP 2012) rating	EUL and SL
sap412	Energy efficiency (SAP 2012) rating	EUL and SL
~~secure~~	~~Secure windows and doors~~	~~EUL and SL~~
storeyx	Number of floors above ground	EUL and SL
sysage	Age of heating system	EUL and SL
typercov	Predominant type of roof covering	SL only
typerstr	Predominant type of roof structure	SL only
typewfin	Predominant type of wall finish	SL only
typewin	Predominant type of window	SL only
typewstr2	Predominant type of wall structure	SL only
wallcavy	Type of wall	EUL and SL
wallinsz	Type of wall and insulation	EUL and SL
watercyl	Hot water cylinder present	EUL and SL
watersys	Water heating system	EUL and SL
wins95x	Type of wall - post 1995 assumption	EUL and SL

Annex Table 5.2.3: general20+21.sav and generalfs21.sav

Aagfh20	household weight (2021) (generalfs only)	EUL and SL
aagpd1920	Rounded dwelling weight for 2 year physical survey sample (paired general only)	EUL and SL
Aagph1920	Rounded household weight for 2 year paired sample (paired general only)	EUL and SL
fimonth	Interview survey month (interviewed)	SL only
fiqtr	Interview survey quarter (interviewed)	SL only
fiyear	Interview survey year (interviewed)	SL only
fmonth	Interview survey month (issued)	SL only
fqtr	Interview survey quarter (issued)	SL only
fyear	Interview survey year (issued)	SL only
gorehs	Government office region EHS version	EUL and SL

govreg1	Government office Region, grouped	EUL and SL
imd1510	IMD 2015 decile ranking of areas (lower layer SOA)	SL only
imd1910	IMD 2019 decile ranking of areas (lower layer SOA)	SL only
paired	Whether paired sample case (generalfs only)	EUL and SL
pmonth	Physical survey fieldwork month (completed)	SL only
pqtr	Physical survey fieldwork quarter (completed)	SL only
pyear	Physical survey fieldwork year (completed)	SL only
region3x	Overall region of England	EUL and SL
ru11combin	Rurality classification - combined (2011 COA)	SL only
ru11contxt	Rurality classification - context (2011 COA)	SL only
ru11morph	Rurality classification - morphology (2011 COA)	SL only
tenure2x	Tenure	EUL and SL
tenure4x	Tenure	EUL and SL
tenure8x	Tenure with vacancy	EUL and SL
~~vacantx~~	~~Whether occupied or vacant (paired general only)~~	~~EUL and SL~~
~~vaclngth~~	~~Length of vacancy (paired general only)~~	~~SL only~~

5.41   The variables vacantx and vaclngth are not available in 2021-22 as no vacant dwellings were surveyed due to COVID-19 restrictions.

Annex Table 5.2.4: modelled housing quality variables 20+21.sav

dampalfm	Modelled dampness problems in one or more rooms	EUL and SL
dhhhsrsym	Modelled decent homes HHSRS (26 hazard) criterion	EUL and SL
dhomeszm	Modelled decent homes - overall standard (26 hazard HHSRS model)	EUL and SL
dhcostzm	Modelled cost to make decent (26 hazard HHSRS model)	EUL and SL

5.42   As Covid restrictions prevented data being collected on the internal condition of the home for both the 2020-21 and 2021-22 survey years only modelled housing quality data was available for both these years. The variable names in 'modelled housing quality variables 20+21.sav' have the suffix 'm' to reflect this. For the previous year's datafile (modelled housing quality variables 19+20.sav) the variables were a hybrid of measured (2019-20) and modelled (2020-21) data and so variable names had the suffix 'h' to reflect this.

Annex 5.3: Accessibility indicators

5.43   The EHS collects a good deal of information on whether dwellings possess certain features or attributes to make them more accessible and useable for people with disabilities. In 2021-22, however, as it was not possible to undertake any internal inspections, most data on the accessibility of the dwelling could not be collected. Despite this, some data items on accessibility features of the exterior of the dwelling were collected. They included the existence of level access and lighting at the main entrance of the dwelling where there was a private plot. In addition, the surveyors collected data on the floor level of the dwelling and information about the types of rooms in the dwelling and the level of those rooms from the questions they asked of households at the doorstep before the start of the physical survey. Data items collected by the 'external plus' approach are highlighted by a red border on the version of the 2021-22 EHS physical survey form - https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/1039253/English_Housing_Survey_2021_to_2022_physical_survey.pdf

Annex 5.4: Household derived indicator

5.44 This Annex focuses on the more complex derived household variables created using EHS interview survey data where the home is occupied and covers the following variables:

- income
- equivalised income
- rents and housing related benefit
- modelling of mortgage repayments
- equity
- household composition

5.45 These complex derived household variables, along with more straight forward household variables e.g. age of the household reference person (HRP), ethnicity of the HRP etc, are used throughout the 2021-22 EHS annual reports. The HRP is the person in whose name the dwelling is owned or rented or who is otherwise responsible for the accommodation. In the case of joint owners and tenants, the person with the highest income is taken as the HRP. Where incomes are equal, the older is taken as the HRP. This procedure increases the likelihood that the HRP better characterises the household's social and economic position.

5.46 The following information details the data processing methods and modelling assumptions for these derived variables, while Chapter 8 of the technical report will focus on the impact of COVID-19 on data quality and analysis.

5.47 Checks are made on the derived variables to ensure as far as possible that the data values are reasonable, and that missing data have been assigned correctly. Implausible values are investigated and only when it is as certain as possible that the data are incorrect is a change made. To assist in analysis, changes made to the data are flagged in the derived EHS interview variable dataset indicating the nature and extent of any imputation.

5.48 All interview based variables are derived from the full annual EHS sample. Most of the EHS annual reports published in 2020-21 are based mainly on one full annual EHS sample. The exceptions are the reports which predominantly use physical survey data: Energy efficiency, housing quality and the stock condition chapters of the main tenure reports which include analysis based on two or more years of data. The EHS datasets containing these household variables are available to users via the UK Data Service (http://ukdataservice.ac.uk/)

Income

5.49　The EHS reports present household/housing related characteristics in relation to various income indicators, to examine housing costs and affordability.

5.50　The interview survey collects information on the main components of income for the HRP and their partner as well as the income of other additional household members aged 16 or over. The components of income collected on the HRP and partner include:

- earnings from regular employment (including government training scheme income) or as self-employed;
- income from occupational and private pensions;
- income from other private sources such as rent from lodgers, student loans, maintenance payments, interest from investments etc;
- state benefits including state pensions; and
- income from savings.

5.51　The income data are thoroughly checked for inconsistencies and errors to make sure, as far as possible, that the data are reasonable and that missing data have been imputed correctly. Implausible values are interrogated and only changed when it is almost certain that the data are incorrect. Any changes made are flagged to indicate the nature and extent of any imputation (variable hhincflg).

5.52　Where respondents report receipt of private income sources, e.g. employment, self-employment income etc., but are unable/refuse to specify an amount, then an estimated amount is assigned according to the methods outlined in Annex Table 5.4.1. All private incomes are modelled as gross amounts and converted (if applicable) to net amounts later in the income modelling process according to Income Tax and National Insurance rates and allowances.

- From 2015 modelling onwards, where the HRP/partner selected working in terms of their employment status, stated as being in receipt of paid work, but did not select any employment or self-employment income under their income sources (nor responded refused to their income sources) and is not in receipt of any 'other' private income, then the HRP/partner is assigned an employment income and an employment income amount is imputed based upon the method outlined in Annex Table 5.4.1. In this situation it is assumed that the HRP/partner unintentionally omitted employment income as an income source in the EHS interview survey.

- From 2017 modelling onwards, in an attempt to address the issue of low student household incomes, several minor modelling changes surrounding the income of student households were incorporated.

1. If an additional adult is a full-time student (likely to be a university student), with no reported income, living in a house where the HRP is a full-time student in further education, then the additional adult income is set to missing (highly likely that they have misreported presence of an income source) and an income amount is imputed.

2. If the HRP/partner is in full-time education (and not studying for A-Levels), with no reported income, living in what appears to be a university household, then the HRP/partner income is set to in receipt of 'other income', rather than 'no income'.

Annex Table 5.4.1: Imputation procedure for private sources

	Type of missing data	Method of imputation
Self-employed	Amount missing	Uses data from the Annual Survey of Hours and Earnings (ASHE) based on age, sex, part-time/full-time, social economic group and geographical location
Regular employment	Amount missing	
Occupational pension	Amount missing	Sample median based on sex and social economic group
Private pension	Amount missing	
Other private sources	Amount missing	Sample median based on working status

5.53 Average values are based on the sample median rather than the sample mean as use of median values better reflects the characteristics of skewed distributions such as are common with income data.

5.54 Where respondents state receipt of particular types of benefits but are unable/refuse to specify an amount, an estimate is inserted based on their theoretical entitlement to the particular benefit. The EHS interview survey incorporates checks on the missing benefit amounts to ascertain whether this were due to the inclusion of the missing amounts with other specified benefits. Use of this data is incorporated into the assessment of missing benefit incomes to avoid double counting of benefit income where this looks probable. Only households that state they are in receipt of benefits are allocated income from benefits. If they are entitled to other benefits but are not claiming them, then estimates for these are not included. All state benefit incomes are modelled as gross amounts and converted (if applicable) to net amounts later in the income modelling process according to Income Tax rates and allowances.

5.55 Since its introduction in April 2013, attention is paid to cases where the HRP or partner selects receipt of Universal Credit in the benefit module of the EHS interview survey. To better capture the data on Universal Credit amounts,

questions regarding Universal Credit were altered in the EHS 2018-19 interview survey. The modelling of Universal Credit has been updated to reflect the EHS interview survey revisions. From 2018 modelling, Universal Credit cases are assigned a Universal Credit scenario (between 1 to 20) that is defined based upon how the suite of Universal Credit/housing related benefit questions are answered in the EHS interview survey, and for some scenarios, also in combination with the household theoretically modelled Universal Credit award. Depending on the scenario assigned, the Universal Credit amount collected in the benefit section of the EHS interview survey is either treated as the total Universal Credit amount that includes help towards housing costs i.e. an 'income' element and a 'housing' element combined, or as an 'income' only amount i.e. an amount that excludes the housing element. In the former situation, the income only amount is derived by deducting the housing element (as collected in the housing benefit module of the EHS interview survey) from the total amount. The housing element of Universal Credit is not incorporated in the derived annual income of the HRP and any partner variables (hhincx and JOINTINCx), but the housing element of Universal Credit is included in the derived annual income of the HRP and any partner variable that includes housing related benefit (variable ALLincx).

5.56   Information is also collected on saving amounts held by the HRP and partner. Where the amount of savings has not been provided, a method based on CHAID analysis (Chi-square automatic interaction detection) is used to estimate the combined savings of the HRP and any partner using predictor variables such as tenure, age/sex of HRP, number of jobs (HRP plus any partner) etc. Once all the cases with savings have an assigned savings amount, income received from these savings is calculated using a gross interest rate of 3.75%. No information is collected on the EHS interview survey that relates to the type of savings account, the term of the account or the rate of interest. The rate of 3.75% (gross) is fixed for all householders and has been used for all years of the EHS income calculation regardless of whether interest rates are higher or lower than this. Income from savings is modelled as a gross amount and converted (if applicable) to a net amount later in the income modelling process according to Income Tax rates and allowances.

5.57   Up to this point in the income modelling, the components of income are calculated as gross amounts. When deriving the annual net income measure, Income Tax and National Insurance payable for the HRP and their partner are calculated where applicable according to Income Tax and National Insurance rates and allowances (incorporating the personal savings allowance where applicable from 2016 modelling), and deducted from the gross income to give the total net annual income of the HRP and partner. From EHS 2017 modelling, if the HRP/partner selects receipt of income from rent, the income modelling assumes that the HRP/partner benefits in tax relief from either the £1,000 property trading allowance scheme (a new policy from April 2017) or, if

a lodger is present, the Rent-a-room relief legislation (existing policy but not previously modelled). Recipients cannot get both allowances, therefore the greatest applicable tax allowance figure is derived based on the HRP/partner modelled rental income and circumstances and assigned to the individual.

5.58   Low incomes in the dataset are uplifted, the justification being that it is likely the respondent under reported their income, either deliberately or by mistake. Where the calculated income of the HRP and any partner is lower than the household's calculated basic theoretical Universal Credit (excluding the household element) /pension credit entitlement, the income amount is changed as follows. Households in receipt of one or more of the main income based benefits (excluding child benefit) and with an income below their theoretical Universal Credit (excluding the housing element)/pension credit entitlement are allocated their Universal Credit (excluding the housing element)/pension credit level plus any disability premiums that they might qualify for. Households that are not in receipt of any of the main benefits and with an income below their theoretical Universal Credit (excluding the housing element)/pension credit entitlement have their income initially set to missing as it was assumed key components of income had been missed or seriously under-reported. An imputed household income value is then derived as further noted below.

5.59   Prior to the EHS 2020-21 income modelling process the low income assessment for the working age group (and associated uplift of income) was based on theoretical income support entitlement. The modelling converted to using theoretical Universal Credit entitlement in the low income assessment as Universal Credit replaced income support as part of the 2013 Welfare Reform Act and Universal Credit overtook income support as the predominant working age means tested benefit.

5.60   Households where the total HRP and partner income is missing have this estimated using the median income for households as defined by working status, social economic group and whether a partner of the HRP is present in the household.

5.61   There are two versions of the variable for the annual income of the HRP and any partner (known as the primary benefit unit income). One variable is in terms of gross income i.e. income before tax and National Insurance deductions (labelled JOINTINCx) and the other in terms of net income i.e. with the deduction of Tax and National Insurance where applicable (labelled hhincx). It should be noted that these two income variables do not include any housing related benefit/allowances.

5.62   In addition, a variable giving the gross income of the HRP and partner has been created that includes housing related benefit (i.e. housing benefit, Local Housing Allowance (LHA) and Universal Credit payments towards rent) as income (labelled 'ALLincx'). This variable is derived by simply adding together

the annual gross income of the HRP and partner (JOINTINCx) and an annualised housing related benefit amount (amthbenx * 52). See the section on 'Rents and housing related benefit' for the calculation of housing related benefit using EHS data.

5.63   There is also a further gross income variable available, labelled HYEARGRx which is an extension of the gross income of the HRP and any partner. This variable represents the household gross income of the HRP and any partner but also includes the gross income of other additional adults living within the household that are not part of the primary benefit unit e.g. a grown-up child living with their parents or two or more unrelated individuals sharing a house (see below for the derivation of additional adult income). Note that this income variable does not include any housing related benefit/allowances.

5.64   Data on the total gross income of other additional household members aged 16 or over (who are not the HRP or partner) are collected at the person level in the EHS interview survey. Household members aged 16 or over that are not in the same benefit unit as the HRP are considered as additional adults and form Other Benefit Units. If the household member is a child of the HRP/partner, aged between 16 and 18 in full-time education (an age threshold of 19 is used if the dependent child is studying for A-Levels or equivalent), then they will be included in the same benefit unit as the HRP and therefore not considered as an additional adult. Also, additional adult household members reported during the EHS interview survey to be living in halls of residence are excluded from the analysis and their income is not considered to be part of the household income.

5.65   If a gross income amount is provided for the additional adult then this value is used for the income of the household member. Also, from 2015 modelling onwards, where an additional adult household member selected working in terms of their employment status, stated as being in receipt of paid work, but provided a gross income value of zero, then their gross income is overwritten and set to missing, and an income amount imputed based upon the standard imputation method for additional adult income as outlined below.

5.66   If an income has not been provided for the additional adult, an amount for their income is then imputed based on a 'hot-decking approach'. The process of hot-decking involves finding cases in the data set that provided an income amount, which are similar in other parts of their responses to the cases with the missing value. For imputing missing additional adult income values, a specification to find similar cases is created for each case based on age (banded), sex, working status, socio-economic group (where applicable), grouped geographical location for those in work, and (from 2015 modelling onwards), presence of an income source. The case with the missing value has a precise specification and it is matched at random to a case with an

income value with the same specification, this income value is then used for the missing case.

5.67   Not all missing cases are matched and this occurs in two situations:

1. if a case with a missing value has a specification which is not matched by a case with a non-missing value; or

2. when there are more cases with missing values than with non-missing values of the same specification.

5.68   Missing cases that are not matched during the hot-decking process are imputed to a sample median based on working status, and for some working status categories with large samples, age (banded) and sex.

Equivalised income

5.69   A measure known as 'in poverty' is sometimes reported on in the EHS Reports and associated tables e.g. the English Housing Survey Housing Costs and Affordability Report, 2015-16. This is based upon an income measure that is 'equivalised' before housing costs are considered. All analysis makes clear the precise measures being used.

5.70   The purpose of income equivalisation is based on the concept that the cost of living varies according to size and type of household. It recognises that, for example, a household of three people requires a higher income than a one person household to achieve the same standard of living. Accordingly, the modified Organisation for Economic Co-operation and Development OECD equivalence scale is employed, taking each household's size and composition into account, to make sensible comparisons.

5.71   Equivalised income measures have been constructed for the EHS in order to assess the relationship between relative poverty and housing conditions and amenities, not to provide estimates of poverty as such. Across Government, poverty is assessed principally through the DWP Households Below Average Income (HBAI) series. Information on household incomes is not collected in as much detail by the EHS as it is by the Family Resources Survey (the data source for the HBAI series). Therefore, there are some limitations to which components can be included in the income measures produced for the EHS. Annex Tables 5.4.2, 5.4.3, 5.4.4 list the HBAI components of the Before Housing Cost (BHC) and After Housing Cost (AHC) measures and describe how the information is addressed through the EHS.

Annex Table 5.4.2: BHC Income components in HBAI and EHS

HBAI Income component	Treatment in the EHS
Income from all household members	The EHS collects income data for the Primary Benefit Unit from the respondent (HRP or any partner). Income data for any additional adult household members is also collected.
Net earnings from employment	Collected
Profit or loss from self-employment	Self-employment income is collected in the EHS and it is included as an income component. The EHS does not collect information on negative self-employment income amounts (i.e. if the respondent experienced a self-employment loss) and thus income losses are not included.
Social security benefits and Tax Credits	Collected for the HRP and any partner
Income from occupational and private pensions	Asked about explicitly in EHS interview
Investment income	Collected
Maintenance payments	Would only be picked up as an 'other' source of income
Income from educational grants and scholarships	Would only be picked up as an 'other' source of income
Cash value of certain forms of income in kind	Not collected

Annex Table 5.4.3: BHC deductions in HBAI and EHS

HBAI BHC Income deduction	Treatment in the EHS
Income Tax payments	Deducted using standard rules
National Insurance contributions	Deducted using standard rules
Council tax	Deducted based on information from the council tax band for the property and the council tax rate from the local authority
Contributions to occupational pension	Not collected
Insurance premium payments made in case of sudden loss of earnings	Not collected
Maintenance and child support payments	Not included
Parental contributions to students living away from home	Not collected
Student loan repayments	Not collected

Annex Table 5.4.4: AHC deductions in HBAI and EHS

HBAI AHC Income deduction	Treatment in the EHS
Rent payments	Collected
Water rates, community or council water charges	Not collected
Mortgage interest payments	Total mortgage repayments used as a proxy (i.e. interest and capital repayments).
Structural insurance premiums	Not collected
Ground rent and service charges	Not included

5.72   The HBAI report uses two different equivalised income measures: Before Housing Costs (BHC) and After Housing Costs (AHC). The income components that make up the measured EHS BHC income variable include: net income of the HRP and any partner, net income from additional adults in the household, modelled winter fuel payment and the addition of council tax support/reduction (previously referred to as council tax benefit) and housing related benefit. For each household the BHC income measure adds up the income from these specified sources and then deducts the amount of council tax payable. These income sources are outlined in Annex Table 5.4.5 together with the method of calculating each income component. The AHC

income measure is derived by deducting rent and mortgage payments from the BHC measure, as outlined in Annex Table 5.4.6.

Annex Table 5.4.5: Income components of the BHC income

Components of the BHC income measure	Method of calculating the income component
Net income of the HRP and any partner	The income variable hhincx is used. See the section above on 'Income'.
Net income from additional adults in the household	The EHS interview survey collects gross income data at a basic level for all other additional household members that are 16 or over which is used in the derivation of the income variable HYEARGRx (see the section above on 'Income'). The gross income amount (as used for HYEARGRx) is converted to a net amount by deducting the applicable Income Tax and National Insurance.
Winter Fuel Payment	The applicable amount of WFP for the household is modelled based on the number of household members that have reached the qualifying age for WFP.
Council tax support (also known as Council tax reduction)	The BHC income measure includes income from council tax support. The EHS interview survey collects information on council tax support receipt of the HRP and any partner. Council tax support is assigned based on this information combined with knowledge of theoretical eligibility (for partial payments) and the amount of council tax due.
Housing related benefit i.e. housing benefit, Local Housing Allowance (LHA) and Universal Credit payments towards rent	The BHC income measure includes income from housing related benefit. See the section on 'Rents and housing related benefit' for the calculation of housing related benefit using EHS data. The derived variable amthbenx is used.
Deduction of council tax payable	The BHC income measure deducts council tax paid by the household. The amount of council tax paid by the household is modelled using the council tax band of the dwelling and information about charges in the relevant local authority area. Single person discount is applied where appropriate which is based upon information collected in the EHS interview survey. A final adjustment is made where the householder states that they are not in receipt of council tax support, where council tax is not included in their rent and that they do not pay/not liable for council tax. In this situation the council tax amount due is set to £0. For a small number of HMO properties, where the council tax band is deemed to be for the whole dwelling, the council tax due for the household is derived by dividing the council tax amount

	that corresponds to the band by the total number of households that share the accommodation.
	The council tax band for each dwelling is collected via a data matching exercise undertaken by the Valuation Office Agency. The council tax band information is only for use in the production of statistics.

Annex Table 5.4.6: Income components of the AHC income

Components of the AHC income measure	Method of calculating the income component
Deduction of rent	The AHC income measure deducts the amount of net rent paid by the household (if applicable). See the section on 'Rents and housing related benefit' for the calculation of rent using EHS data. The derived variable RentExS is used.
Deduction of mortgage payment	The AHC income measure deducts the mortgage payments paid by the household (if applicable). See the section on 'Mortgage Payment' for the calculation of mortgage payments using EHS data. The derived variable mortwkx is used.

5.73 The HBAI series and poverty estimates are based on a modified OECD scale (modified so that a couple with no children is considered the reference point and has an equivalence factor equal to unity). The EHS follows this approach and determines the number of 'first' adults (i.e. HRP), other adults, children aged 14 years and over and children under 14 for each sample case in the dataset. This provides the necessary information to be able to apply the OECD equivalisation factors to produce an equivalised BHC and AHC income from the measured EHS BHC and AHC income variables, Annex Table 5.4.7.

Annex Table 5.4.7: OECD equivalisation factors

Equivalence scales	Modified OECD scaled to couple without children = 1	
	BHC	AHC
First adult	0.67	0.58
Spouse	0.33	0.42
Subsequent adults	0.33	0.42
Children aged under 14 years	0.20	0.20
Children aged over 14 years	0.33	0.42

Note: additional adult household members reported during the EHS interview survey to be living in Halls of Residence are excluded from the analysis.

5.74   The measured BHC and AHC household income constructed as outlined above are divided by the appropriate equivalisation factor. Any household with an equivalisation factor of less than one (e.g. a single person household) will have their BHC/AHC income inflated, reflecting the fact that they are relatively better off than a larger household with the same income. Households with an equivalisation factor greater than one have their BHC/AHC incomes reduced, reflecting the fact that they are relatively worse off than a smaller household. The BHC and AHC incomes of households containing two adults without children will not change. This calculation creates the equivalised BHC and AHC income variables included in the main derived interview file.

5.75   Note, no cases have a measured AHC income greater than the measured BHC income. However, as outlined in Annex Table 5.4.7, different equivalisation factors are applied to the BHC income measure than the AHC income measure. When the equivalisation factors are applied to the measured BHC income and AHC income, this can result is an equivalised AHC income, greater than the equivalised BHC income.

5.76   On occasions, the EHS Reports use the BHC/AHC equivalised weekly income measures ranked and grouped into five equal sized categories. Thus the first quintile relates to the households with the lowest 20% of BHC/AHC equivalised weekly incomes and the fifth quintile relates to the households with the highest 20% of BHC/AHC equivalised weekly incomes. As previously mentioned, a term referred to as 'in poverty' is also used in the EHS Reports. Households are defined to be 'in poverty' if their equivalised income is below 60% of the median household income before housing costs are considered.

Rents and housing related benefit

5.77   Information on rents and housing related benefit are presented in detail in EHS reports in analysis of social and private renters. The amount of rent and housing related benefit also feeds into the calculation of Equivalised income (see above).

5.78   Household rents and housing related benefit receipt are collected in the EHS interview survey and apply only to households that rent their own home or households in a shared ownership scheme. Renting households that live rent-free are not asked the series of rent and housing related benefit questions in the interview survey. For rent-free cases the rent and housing related benefit amounts are set to zero.

5.79   The total weekly rent payable for the property (variables rentwkx) includes the rent paid by the householder plus any housing related benefit received towards the rent payment i.e. housing benefit, Local Housing Allowance (LHA) and Universal Credit payment intended towards the rent (variable amthbenx). These variables are calculated based on the householder's response to the set of detailed rent and housing related benefit questions asked in the EHS interview. For households that report a rent holiday, an adjustment is made so that the actual total amount of rent/housing related benefit paid over the course of the year (over n weeks) is averaged out over the full year (as if paid over 52 weeks). From 2019 modelling, for Universal Credit cases that report a rent holiday and provide an amount of Universal Credit housing element that helps towards rent during the interview survey, only the rent is adjusted to incorporate the rent free weeks, and not the Universal Credit housing element amount. This aligns the modelling to how Universal Credit housing element payments are issued to claimants.

5.80   Households that pay rent but do not provide an amount for their rent/housing related benefit, because the amount was either unknown or refused, are assigned an estimated total weekly guide rent amount. This is based on tenure, number of bedrooms and area where they live, (for private renters variables such as the type of landlord and level of furnishing are also used in the imputation), using the sources of rent data outlined in Annex Table 5.4.8. For these cases, the estimated total weekly rent payable (rentwkx) is calculated by adding the total weekly guide rent to the estimated modelled amount for services e.g. heating and regular meals etc. that are included in the rent (where applicable as specified by the householder). More information on services is detailed in the paragraph below.

Annex Table 5.4.8: Data sources used for missing rent amounts

Tenure of the renting household	Type of missing data	Rent data source for missing values
Local authority	Rent amount missing	Local Authority average general need net rents based on rent data from the Local Authority Data Return (LADR) statistics submitted by Local Authority Registered Providers, supplied by the Regulator of Social Housing (RSH). The data provided by the RSH represents rent levels for the year prior to the EHS data collection period, therefore, the guide rents supplied by the RSH are adjusted using the social sector rent uplift mechanics as outlined in the 'Policy statement on rents for social housing' (i.e. by a factor based on CPI for September of the previous year + 1%).
Housing association and shared owners	Rent amount missing	Housing Association average general need net rents based on rent data from the Statistical Data Return submitted by Private Registered Providers, (supplied by the RSH. The data provided by the RSH represents rent levels for the year prior to the EHS data collection period, therefore, the guide rents supplied by the RSH are adjusted using the social sector rent uplift mechanics as outlined in the 'Policy statement on rents for social housing' (i.e. by a factor based on CPI for September of the previous year + 1%).
Private renters	Rent amount missing	EHS data using the year in question and modelled accordingly

5.81 Households that receive housing related benefit towards their rent payment but do not provide an amount or households that do not know if they receive housing related benefit are assigned an estimated benefit amount as follows:

- If the household states that they are in receipt of full housing related benefit then the weekly housing benefit is set to their theoretical guide net rent amount (as opposed to the assigned total weekly rent payable, since housing related benefit does not cover the cost of services such as heating and regular meals).
- If the household states that they are in receipt of partial housing related benefit or if they do not know if their housing related benefit covers all or some of their rent, then an amount of housing related benefit is imputed based on their total rent amount payable and their theoretical entitlement to housing benefit, incorporating the social housing under occupation charge for social tenants.

5.82 There are a selected number of Universal Credit cases where the missing weekly housing related benefit amount received by the household is

calculated earlier on in the income modelling process, rather than following the above rules. Where this applies, the housing related benefit value is brought into the rent and housing benefit modelling to calculate the final amounts.

5.83   An extension of the derived variable for total weekly rent payable for the property (labelled rentwkx) is the total weekly rent payable for the property excluding the cost of services e.g. heating, council tax etc (labelled rentExS). The EHS interview survey asks the householder if the rent amount included any of the following services:

- Council tax;
- Heating;
- water and sewerage;
- lighting;
- hot water;
- fuel for cooking;
- regular meals; and
- TV licence.

5.84   If the amount of rent the householder provided does not include any of those services, then the total weekly rent payable for the property is the same as the total weekly rent payable for the property excluding the cost of services i.e. rentwkx equals rentExS.

5.85   From 2018 modelling, if the rent amount provided by the householder does include one or more of the services stated above then the net rent is modelled consistently for all cases by subtracting an estimated service charge from the total weekly rent payable for the property. An estimated amount for the selected services is modelled according to the methods shown in Annex Table 5.4.9 in order to derive a rent amount excluding the cost of services.

Annex Table 5.4.9: Imputation procedure for service amounts

Service	Method of imputation
Council tax	The amount is derived using the council tax band information for the property and the council tax rate from the local authority. For any cases where the respondent selects that council tax is included in their rent but also selected full housing related benefit then the council tax service amount is set to £0.
Water and sewerage	This is calculated using the mean average water and sewerage rate for a household in England for the given year multiplied by a dwelling factor (that reflects the size/type of property) multiplied by a factor for that geographical area
Heating	
Lighting	
Hot-water	Amounts are calculated based on DWP deductions from rent (as used in the assessment of housing benefit)
Fuel for cooking	
Regular meals	
TV licence	Based on the cost of a colour TV licence

5.86   EHS reports mainly focus on the weekly rent amount payable for the property, excluding the cost of services, e.g. heating, council tax etc. Between the 2007-08 and 2008-9 data there were some methodological changes to the way in which rent data were processed in the EHS compared with the previous Survey of English Housing (SEH). These changes are detailed in the EHS Households Report 2008-09 where an assessment of the impact of the methodological changes between 2007-08 and 2008-09 on the private rent estimates and further details of the changes to the calculation of social rents are provided.

5.87   The rent and housing related benefit amounts are thoroughly checked for inconsistencies and errors to make sure as far as possible that the data are reasonable and that missing data have been imputed correctly. Implausible values are interrogated and only changed when almost certain that the data are incorrect. Any changes made are flagged indicating the nature and extent of any imputation (variable rentflg).

Modelling of mortgage repayments

5.88   Mortgage repayments are calculated from raw data collected from respondents on payments for all mortgages/loans secured on the dwelling,

deducting any notional amounts for building and contents insurance, mortgage protection, and other insurance payments where they have been accidentally included. Endowment policy premiums are included in repayments. Information collected is converted to weekly amounts (variable mortwkx).

5.89   Where repayment amounts are unknown or missing, data is imputed. Where alternative data is available, using the formula below for monthly mortgage repayment:

$$\frac{\text{(value of current main mortgage x annual interest rate/12)}}{(1-1/((1+\text{annual interest rate}/12)^{(\text{length of main mortgage in months})}))}$$

5.90   The 'current main mortgage' refers to the highest value mortgage if there is more than one loan secured on the property. Where the value of the current main mortgage is not provided, it is derived from original purchase price minus deposit paid, if these data are available. The purchase price is imputed if unknown, based on current market value, year dwelling bought (if available) and DLUHC annual house price inflation indices for each area. From 2010, the EHS ceased collecting current market valuations on a regular basis from the Valuation Office Agency (VOA) for all households participating in the interview survey who had a physical inspection of their dwelling. In the interview survey, households are now asked to provide a recent (in last 12 months) valuation of their dwelling, and, if not available, their own estimate of the market value. This valuation/estimate is now used in the modelling as a proxy for current market value. Any missing values are imputed based on dwelling type and geographical area. Valuations are checked for outliers, which are reviewed and corrected if necessary and possible, based on other information on the property. Payments for other loans secured on the property are not included in this imputation process.

5.91   If the length of the main mortgage is unknown, it is modelled where data are available, based on the age of HRP and when the main mortgage was taken out (or if missing, when the dwelling was bought), assuming a maximum mortgage length of 25 years and that it will be paid off when the HRP is 60 years old. A minimum mortgage length is input for households that are aged 60 or older and have owned their home less than 5 years.

5.92   The annual interest rate is taken from monthly data provided by UK Finance. An average is calculated for the period covered by the EHS survey.

5.93   Monthly mortgage payments are set to not applicable (-9) for tenancies. If the owner owns the dwelling outright, payments are zero. All shared owners are treated as owners and mortgage payments are derived as above and where necessary adjusted for the proportion of ownership if data imputed. For shared owners that stated they had no outstanding mortgage (onoutmrg1=3) and validated as correct, these were set to a nominal value (from 2016-17).

Flexible/all in one/offset mortgages are also derived as above. Payments for equity release mortgages are imputed and are modelled as above.

5.94   Monthly repayment amounts are then converted to weekly payments. The data are thoroughly checked for inconsistencies, outliers and errors although data are only corrected where deemed totally implausible and it is possible to determine an alternative more reliable imputed value.

5.95   It should be noted that the weekly mortgage variable contains amounts for a mixture of mortgage types for example repayment and interest only cases, when the information was provided by the household. However, all imputed cases, irrespective of the mortgage type (for example, interest only mortgages) are calculated as repayment mortgages (interest and capital).

Equity

5.96   The value of a household's equity in their property is calculated for all owner occupied (including shared owner) households who participate in the interview survey. It is based on the current market value of the property minus the amount of mortgage outstanding (for shared owners this is checked and adjusted where necessary to ensure their equity is calculated in proportion to their ownership).

i.e. Equity = current market value – amount of mortgage outstanding including other loans on the property

5.97   All owner occupiers are asked in the interview for an estimated current valuation of their property. This is used to derive an approximate equity value as detailed above (variables 'equityr' and 'equityr5').

5.98   Research conducted on the EHS 2008 data demonstrated that owner occupiers are more likely to overestimate the value of their homes than underestimate it, and only 40% of household estimates are within 10% of the VOA valuation. Those who moved in just over a year ago, who are on the highest incomes or in the highest value properties are most likely to overestimate value. Those homes most likely to be underestimated in value contain singletons and couples aged over 60 and those resident for at least 20 years. As a result of respondents tending to overestimate property value, equity based on this source of market value ('equityr' and 'equityr5') is likely to be an overestimate.

5.99   Missing equity valuations are imputed using data from the regulated mortgage survey based on dwelling type and area. Valuations are checked for outliers and corrected if it is possible to determine a more reliable imputed value.

5.100 Where data used in the above formula is not available, equity is set to unknown (-88888888) or not applicable (-99999999) for tenancies.

Household composition

5.101 Each year, English Housing Survey (EHS) data are used to create a number of household typologies. These descriptors are a way of classifying households according to the relationships between the household members and are used throughout the EHS reports to provide context on household circumstances and characteristics.

5.102 The current EHS household survey is the result of the 2008 merger of the Survey of English Housing (SEH) and the English House Condition Survey (EHCS). Both these surveys collected similar information and modelled some similar household derived descriptors, however their methods and assumptions were not always the same for deriving these similar classifications. A key difference was that SEH modelled its descriptors based on the eldest person in the household whereas the EHCS used the Household Reference Person (HRP). For the majority of the simple derived descriptors this made no difference or only minor changes were required to harmonise them, however some descriptors are still not derived from the same assumptions. This has the potential to cause confusion when they are combined in analysis. This report aims to clarify the modelling assumptions of these conflicting cases. It will outline how the simpler derived variables (Annex Table 5.4.10) feed into the more complex ones (Annex Table 5.4.11 and Annex Table 5.4.12).

5.103 The flow chart below (Figure 5.4.1) outlines the modelling process of the household descriptors. As part of the modelling, the raw data are cleaned to correct any inconsistent relationship data mainly relating to the family unit. This is done within the modelling only and the raw data files remain unchanged. For this reason, although correct, there may be inconsistency between some derived descriptors and the raw data. The derived indicators are considered to be correct and in line with all guidance for classifying individuals and family units. It is these variables that are used for reporting.

5.104 Many of the simpler derived variables are either directly related to the HRP, for example Cohabhrp or Cohabprt, or derive the quantity of a particular type of household member or household type within the household, for example the number of couples in the household (ncouple). These simple derived variables are listed in Annex Table 5.4.10, which highlights both the modelling key issues and recoding and the potential data conflicts which may cause a descriptor to be at odds with another descriptor.

5.105 Some of these simpler derived variables and raw data are used to model the more complex final household composition descriptors, hhtype11, hhtype7,

hhtype6, hhcompx (originating from the EHCS) and Hhcomp1 (originating from the SEH). This can be seen in the flowchart (Figure 5.4.1) as well Annex Table 5.4.11 and Annex Table 5.4.12.

Figure 5.4.1: Simple flowchart of the modelling of the derived household variables
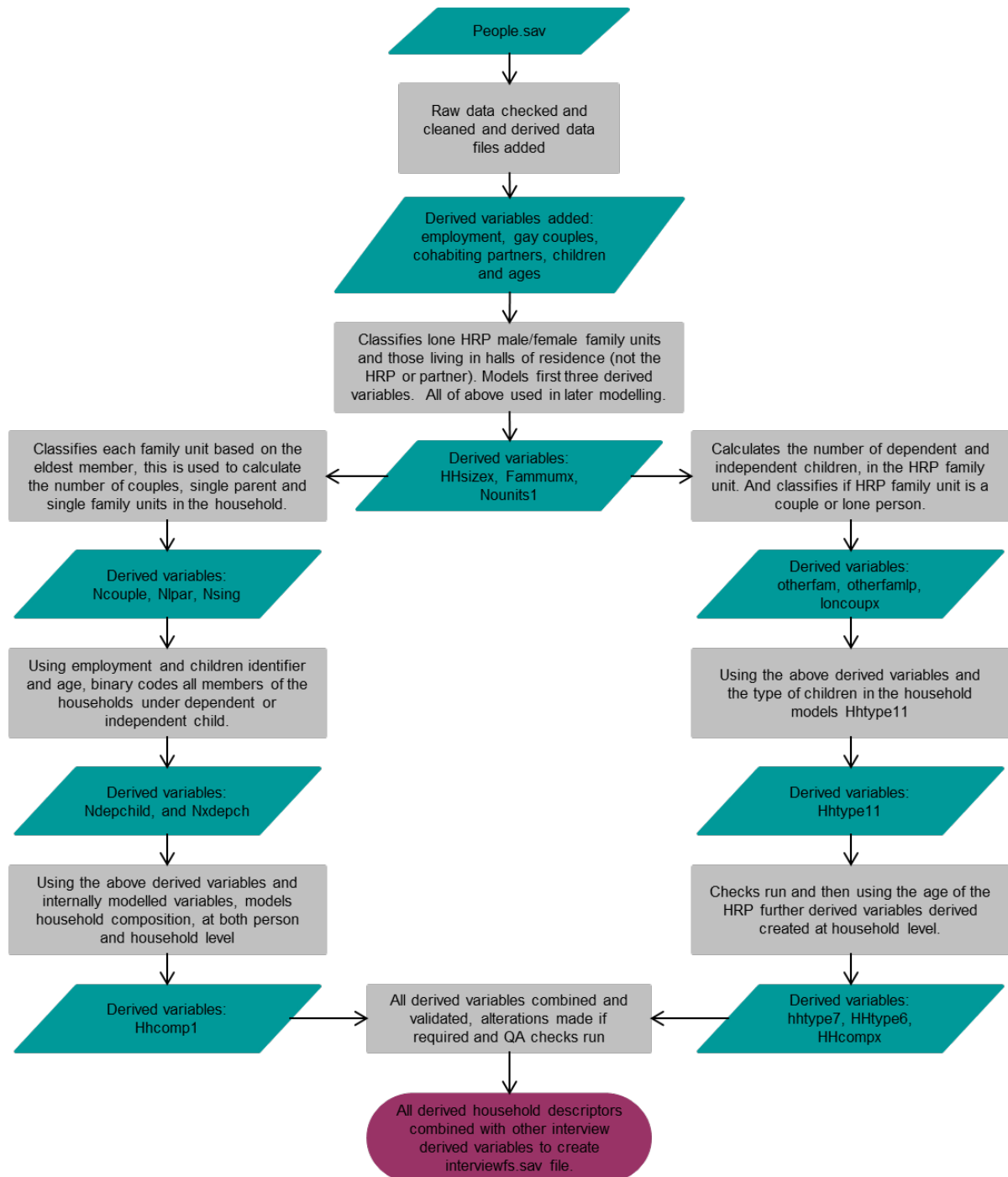
Annex Table 5.4.10: Simple household descriptors

Derived descriptor	Descriptor variable label	Modelling key issues	Data conflict issues
Cohabhrp	If HRP is cohabiting	Includes same sex couples, even if they are not in the same family unit as the HRP	May conflict with hhcomp1 which is not based on HRP. Also may conflict with other descriptors such as famnumx, NoUnits1, otherfam and othfamlp if it is a same sex couple
Cohabprt	If HRP partner is cohabiting	Includes same sex couples, even if they are not in the same family unit as the HRP	May conflict with hhcomp1 which is not based on HRP. Also may conflict with other descriptors such as famnumx, NoUnits1, otherfam and othfamlp if it is a same sex couple
Hhsizex	Number of persons in the household	Excluding those aged 16+ living away in halls of residence/boarding school, who are not the HRP or partner.	-
Famnumx	Number of family units in household	The number of family units in the household, excludes those aged 16+ living in halls of residence who are not the HRP or partner (unlike old SEH definition). Also, same sex cohabiting couples are counted as separate family units, despite being treated as a couple living together in other variables (e.g. 'loncoupx', 'hhcompx'). Originates from EHCS.	May conflict with other descriptors due to some same sex couples being recoded as a couple, but in different family units.
NoUnits1	Banded number of family units in household	Famnumx banded, see above for modelling issues.	See above
Nlpar	Number of lone parent families in household	For some households that state they are married or cohabiting (Xmarsta2), but there is no partner in the household, they are recoded from ncouple to nlpar if there are children in the same family unit in the household. For consistency with EHCS. Excludes those aged 16+ that live in halls of residence who are not the HRP or partner.	-

Derived descriptor	Descriptor variable label	Modelling key issues	Data conflict issues
Ncouple	Number of couples in household	Some cases in the data set, that state they are married (under Xmarsta2) and have children but there is no partner in the household, these are recoded to lone parents. Excludes those aged 16+ that live in halls of residence who are not the HRP or partner.	Same sex couples not in a civil partnership/married may conflict with other descriptors such as famnumx, NoUnits1, as they are in separate family units.
Nsing	Number of one-person family units in household	Same sex couples are recoded from nsing to couples, this is a model issue. Excludes those aged 16+ that live in halls of residence who are not the HRP or partner.	It is possible to be a one-person family household under nsing and a dependent or independent child, for example dependent foster children.
Ndepchild	Number of dependent children in household	All dependent children are those under 16 or those aged 16 to 18 that are in full time further education (i.e. up to A-level or equivalent). This is based on all children in the household, irrespective of which family unit they are in, but not those living in halls of residence.	All children in the home, not just the main family unit, if other family units are present in the household. It is possible to be a dependent child and a one-person family household under nsing for example in the case of non-relatives and foster children. This variable will not always match with EHCS household composition descriptors as they are modelled based on the children in the HRP family unit only. The modelling does not consider other children in the household.
Nxdepch	Number of non-dependent children in household	All children aged 16+ that are not in full time education or children that are aged 19 or older. (Children aged 16-18 in full time further education are considered dependent). Excludes those living in halls of residence.	All children in the home, not just the main family unit, if other family units are present in the household. It is possible to be an independent child and be coded against nlpar, nsing and ncouple, if the parents of the adult child live in the same household. This variable will not always match with EHCS household composition descriptors as they are modelled based on the children in the HRP family unit only. The modelling does not consider other children in the household.
Otherfam	Additional families present in household	Based on the recoding on famnumx, the number of separate family units. Excludes those aged 16+ that live in halls of residence who are not the HRP or partner.	May conflict with other ncouple due to some same sex couples being recoded under couples.

Derived descriptor	Descriptor variable label	Modelling key issues	Data conflict issues
Othfamlp	Type of additional families in household	Based on otherfam and the size of each family unit. Excludes those aged 16+ that live in halls of residence who are not the HRP or partner.	May conflict with other ncouple due to some same sex couples being recoded under couples.
Loncoupx	Single householder or with partner	Based on HRP household only, will include same sex couples even if they are not in the same family unit as the HRP	May conflict with other famnumx, NoUnits1 otherfam and othfamlp due to HRP same sex couples.

5.106   Although the modelling of the household composition descriptors is, by and large, not complicated, the difference in assumptions is important; and not always clear. For the majority of households, where the household consists of just one family unit, the derived descriptors are consistent. Complications arise when there is more than one family unit in the household or where the HRP is an independent child within the sole family unit.

5.107   Annex Table 5.4.11 shows the household composition descriptors that originate from the EHCS. It illustrates how the descriptors are derived, the key modelling assumptions and any recoding that takes place to deal with modelling difficulties and to harmonise the descriptors. The EHCS household composition descriptors are modelled based on the HRP family unit by firstly, defining the HRP family unit and then, if present, any additional family units in the household. For this reason, dependent children that are not part of the main family unit are not considered children in the household, but as separate family units coded as an independent single person. Hhtype11 is the first derivation of these four derived household composition descriptors. The other three, hhtype7, hhtype6 and hhcompx, are all derived from hhtype11 and the age of the HRP for the more detailed classifications. Understanding the derivation of hhtype11 should explain the coding of the other variables.

5.108   This model assumes:

- Same sex couples are treated as couples.
- Those in halls of residences that are not the HRP or partner are not included in the modelling.
- Dependent children are all those under 16 or aged between 16 and 18 and in full time education and part of the main family unit i.e. the family unit of the HRP and partner.
- Couple households with only independent children are classed as couples, and single parent households with independent children only are classed as other multi-person households.

Annex Table 5.4.11: Household composition descriptors that originate from the EHCS

Hhtype11 - Household type - full 11 categories			Hhtype7 (recoded from hhtype11) - Household type - 7 categories	Hhtype6 (recoded from hhtype7 using age of HRP) - Household type - 6 categories	Hhcompx (recoded from Hhtype6 using age of HRP) - Household composition
Variable code and value label	**Modelling key issues**	**Recoding or possible data conflict issues**	**Variable code and value label**	**Variable code and value label**	**Variable code and value label**
1. Couple no child(ren)	All in the same family unit, but possible to have a lone person single family unit in the household as well	Possible to have an additional single-family unit if not HRP	1. Couple no dependent child(ren)	1. Couple no dependent child(ren)	Becomes: 1. couple, no dependent child(ren) under 60 or 2. couple, no dependent child(ren) aged 60 or over
2. Couple dependent child(ren) only	All in the same family unit, but possible to have a lone person single family unit in the household as well.	Possible to have an additional single-family unit if not HRP	2. Couple with dependent child(ren)	2. Couple with dependent child(ren)	3. Couple with dependent child(ren)
3. Couple with dependent and independent children	All in the same family unit, but possible to have a lone person single family unit in the household as well.	Possible to have an additional single-family unit if not HRP	2. Couple with dependent child(ren)	2. Couple with dependent child(ren)	3. Couple with dependent child(ren)

Hhtype11 - Household type - full 11 categories			Hhtype7 (recoded from hhtype11) - Household type - 7 categories	Hhtype6 (recoded from hhtype7 using age of HRP) - Household type - 6 categories	Hhcompx (recoded from Hhtype6 using age of HRP) - Household composition
Variable code and value label	**Modelling key issues**	**Recoding or possible data conflict issues**	**Variable code and value label**	**Variable code and value label**	**Variable code and value label**
4. Couple with independent child(ren) only	All in the same family unit, but possible to have a lone person single family unit in the household as well.	Possible to have an additional single-family unit if not HRP	1. Couple no dependent child(ren)	1. Couple no dependent child(ren)	Becomes: 1. couple, no dependent child(ren) under 60 or 2. couple, no dependent child(ren) aged 60 or over
5. Lone parent with dependent child(ren) only	All in the same family unit, but possible to have a lone person single family unit in the household as well.	Possible to have an additional single-family unit if not HRP	3. Lone parent with dependent child(ren)	3. lone parent with dependent child(ren)	4. Lone parent with dependent child(ren)
6. Lone parent with dependent and independent children only	All in the same family unit, but possible to have a lone person single family unit in the household as well.	Possible to have an additional single family unit if not HRP	3. Lone parent with dependent child(ren)	3. Lone parent with dependent child(ren)	4. Lone parent with dependent child(ren)
7. Lone parent with independent child(ren) only	All in the same family unit, but possible to have a lone person single family unit in the household as well.	Possible to have an additional single family unit if not HRP	4. Other multi-person households	4. Other multi-person households	5. Other multi-person households

Hhtype11 - Household type - full 11 categories			Hhtype7 (recoded from hhtype11) - Household type - 7 categories	Hhtype6 (recoded from hhtype7 using age of HRP) - Household type - 6 categories	Hhcompx (recoded from Hhtype6 using age of HRP) - Household composition
Variable code and value label	**Modelling key issues**	**Recoding or possible data conflict issues**	**Variable code and value label**	**Variable code and value label**	**Variable code and value label**
8. Two or more families	Other additional families, the second family unit needs to contain 2 or more people in it; the HRP family unit can be a single person	If the HRP is a single person and lives with a couple or another family unit of 2+ people, it will be coded 8 however this will not be the case for HHcomp1 (coded 1 see modelling explanation below).	4. Other multi-person households	4. Other multi-person households	5. Other multi-person households
9. Lone person sharing with other lone persons	More than one family unit in the household, but with only one person in each of them	-	4. Other multi-person households	4. Other multi-person households	5. Other multi-person households
10. One male	One-person family households	-	5. One male	Becomes: 5. one person under 60 or 6. one person aged 60 or over	Becomes: 6. one person under 60 or 7. one person aged 60 or over
11. One female	One-person family households	-	6. One female	Becomes: 5. one person under 60 or 6. one person aged 60 or over	Becomes: 6. one person under 60 or 7. one person aged 60 or over

Note: The variable hhtype7 does have a seventh category 'one person (sex unknown)' but this category is currently redundant due to no missing EHS data on the sex of individuals.

5.109 It is important to note that there are a few households that are not derived based on the HRP, to provide better consistency with the SEH. These households are where the HRP is a young unmarried adult child living with their parents in the household. Under EHCS rules, if an adult child is single and they don't have children of their own they should be part of the main family unit, with their parent/s (not in a separate family unit as the raw data suggests). So, where there is a 'normal' family set up for example a couple with two children where one is the HRP, the household composition of these cases is recoded based on the overall composition of the family unit. If the data are not recoded, the household is coded 8 'as two or more families' which is not in accordance with the guidance. Usually this just affects a few households in the data set each year.

5.110 Annex Table 5.4.12 shows the SEH household composition descriptor (hhcomp1), its coding, key modelling issues and recoding or potential data conflicts. This descriptor is derived firstly based on the eldest person in the household, however in the modelling, additional family units of couples and single parents will over-ride the eldest person in the household if they are a single person family unit. This means that households that may have been coded two or more families under EHCS method, can be coded as married/cohabiting couple or a lone parent (male or female) under the SEH method.

Annex Table 5.4.12: Household composition descriptors that originate from the SEH

Hhcomp1 - Household composition		
Variable code and value label	**Modelling key issues**	**Recoding or possible data conflict issues**
1. Married / cohabiting couple	A married/cohabiting couple with or without children (independent or dependent). There can be additional single person family units in the household. Note - couple households over-ride the eldest household if the eldest household is a single household.	Some cases are recoded to lone parent, because modelled on married status, and there are some households that state they are married but there is no partner in the household, so these are recoded to match hhtype11 that codes these cases as single parents. (Also ncouple and nsing are recoded see above). This may not fit with the derived variables based on an HRP profile, for example Cohabhrp.

Hhcomp1 - Household composition		
Variable code and value label	Modelling key issues	Recoding or possible data conflict issues
2. Lone parent, male HRP	A lone parent with child(ren) dependent or independent or both with a male HRP. There can be additional single person family units in the household. Single parent households over-ride single households if they are the eldest household.	Could be coded as two or more families under hhtype11, if there is an additional single-family unit who is the HRP. If this is the case the HRP might not be the lone parent.
3. Lone parent, female HRP	A lone parent with child(ren) dependent or independent or both with a female HRP. There can be additional single person family units in the household. Single parent households over-ride single households if they are the eldest household.	Could be coded as two or more families under hhtype11, if there is an additional single-family unit who is the HRP. If this is the case the HRP might not be the lone parent.
4. Multi-family household, male HRP	Has at least two family units and it is either made up of: at least a couple family unit and single parent family unit or 2 couples or 2 single parents family units or 2 or more single people in their own family unit. It is not possible to have a single family unit with a family unit of 2 or more people in this derivation; these will be coded 1 to 3 above.	-
5. Multi-family household, female HRP	Has at least two family units and it is either made up of: at least a couple family unit and single parent family unit or 2 couples or 2 single parent family units or 2 or more single people in their own family unit. It is not possible to have a single person family unit with a family unit of 2 or more people in this derivation; these will be coded 1 to 3 above.	-
6. One male	One person households - single family unit	-
7. One female	One person households - single family unit	-

5.111 Although, the majority of household derived descriptors are consistent with the raw data and one another, there are circumstances where these may conflict. In the first instance, any conflict between the derived household descriptors and the raw data, the derived descriptors would be correct (following the

guidance for classifying individuals and households) and the household raw data inconsistent. Once the descriptors are modelled, a few households are recoded which means they are not consistent with the original modelling rules, but this has to be done to reduce data conflict between key descriptors. Even so, there are still some household descriptors which will still conflict with others. The key areas of conflict are:

- Number of dependent children in the household (ndepchild) with hhtype11, hhtype7 hhtype6 and hhcompx, where it's possible to be coded as a household with no dependent children in it, when in fact the household contains dependents in a separate family unit to the HRP.
- Hhcomp1 can conflict with hhtype11, hhtype7, hhtype6 and hhcompx; these are due to the treatment of additional single person households, depending on modelling assumptions of either an EHCS or SEH based descriptor.
- Hhcomp1 may conflict with Cohabhrp and Cohabprt as it is not modelled based on the HRP.
- Ncouple, loncoupx, Cohabhrp and Cohabprt may conflict with otherfam, othfamlp famnumx and NoUnits1 due to the modelling of cohabiting same sex couples; although these conflicts are not easy to identify through top level analysis.

Annex 5.5: Derived indicators on housing conditions

5.112 Due to COVID-19 restrictions in 2021-22 it was not possible to collect the detailed data required to derive key housing quality variables. The list of derived variables for the EHS 2021-22 is provided in Annex 5.1. To support users interested in time series findings, DLUHC first released selected national level extrapolated estimates on the prevalence of carbon monoxide detectors in the 2021-22 EHS Headline Report. https://www.gov.uk/government/statistics/english-housing-survey-2021-to-2022-headline-report

5.113 Binary classification models, first developed for the 2020-21 EHS Housing Quality and Condition Report, were also used for the 2021-22 Headline Report and the subsequent annual reports, to produce predictive modelled estimates at dwelling level to indicate whether or not a dwelling: had damp problems; had any Category 1 hazards assessed through the Housing Health and Safety Rating System (HHSRS); or met the Decent Homes Standard. It should be noted that it was possible to model Category 1 hazards related to excess cold, overcrowding and radon in the normal way and, where possible, surveyors also completed HHSRS assessments for the exterior of the dwelling. Data items from those sources were used to inform the predictive modelling.

Approach for producing the national level extrapolated estimates released in the 2021-22 EHS Headline Report

5.114 The 2021-22 EHS Headline Report estimates the prevalence of homes with carbon monoxide detectors (Annex Table 2.19) for the total housing stock, including vacant homes. The data were extrapolated from previous EHS trends using EHS findings on stock condition for the years '2019' and the previous extrapolated estimates for '2020'. Stock condition findings from the EHS are produced by combining two years of data. So stock condition findings in the 2021-22 EHS Headline Report were produced using data for the EHS survey years of 2020-21 and 2021-22 and are referred to as findings for '2021'.

5.115 For the extrapolation of findings on carbon monoxide detectors, not all the control totals required for stock with solid fuel and those with non-solid fuel appliances were available (Although for 2020-21, frequencies of total

dwellings with solid fuel appliances/no solid fuel appliances can be obtained from the EHS database, these are for occupied dwellings only).

   a. Changes in frequencies derived from the '2019' combined data and the '2020' combined data were calculated and the frequency change was applied to predict what the number would be in the '2021' combined data.

   b. Using the same approach as in (a), interim percentages of homes with CO detectors by solid fuel/non-solid fuel appliance were calculated for the whole (occupied plus vacant) housing stock.

   c. Those interim percentages were then used to estimate the stock totals for dwellings with solid fuel and those with non-solid fuel appliances.

   d. The sum of the stock total of dwellings with solid fuel and those of dwellings with non-solid fuel appliances derived from (c) did not equate to the control total for whole occupied plus vacant stock derived from the dwelling stock estimates used to develop the 2021-22 EHS analysis weights. Therefore, the totals for solid/non solid fuel obtained in (c) were scaled up so that they summed to the occupied plus vacant stock total derived from the dwelling stock estimates used to develop the 2021-22 EHS analysis weights.

   e. Percentages with carbon monoxide detectors amongst dwellings with solid fuel and those with non-solid fuel appliances were then recalculated using the new scaled up stock totals in (d). A check was made to ensure that these were not too different to interim percentages in (b).

   f. A final check was made to ensure the distribution of dwellings with solid fuel and those of dwellings with non-solid fuel appliances amongst occupied stock was not dissimilar to the distribution among estimated occupied plus vacant stock totals.

Approach for producing the predictive modelled estimates on the presence of damp

5.116 The 2021-22 dwelling level predictive modelled estimates on damp indicate whether a dwelling had any form of damp or not. To maximise the accuracy of the predictive modelling, three different models were used to predict the presence of damp: one for houses; one for purpose-built flats; and one for converted flats. Serious forms of damp are rare attributes so modelling techniques were applied to take account of the low prevalence of those conditions.

5.117 The predictive model used was first developed for the EHS 2020-21 and the full details on the modelling methodology can be found in the English Housing Survey 2020 to 2021, technical report (English Housing Survey 2020 to 2021: technical report - GOV.UK (www.gov.uk)). The information below summarises

the process undertaken for the analysis of 2021-22 data and reports on the overall accuracy of the model.

Predictive model for damp in houses

5.118 The predictive model, which used the strongest indicators of damp in houses (disrepair to external building elements and length of residence in the accommodation) was run several times to determine:

    a. the probability threshold for classifying an individual house as having damp or not; and

    b. the success of the model in identifying correctly whether the house had damp or not when compared with the actual EHS surveyor assessments for 2018 and 2019 (used to 'train' the original model)

5.119 The chosen threshold which provided a good balance between the proportion of houses correctly identified as having damp or not, produced a prevalence of damp in the EHS 2021-22 data of 2.99% roughly the same as the proportion as that found in the 2018+2019 combined data (3.04%).

5.120 Additional quality assurance checks were made against the EHS predicted modelled finding for the 2020 EHS to ensure that the number and percentage of houses with damp were feasible.

Predictive model for damp in flats

5.121 The main predictors of damp in purpose-built flats identified through predicting modelling are disrepair to the external elements, length of residence in the accommodation and average room size. For converted flats the predictive model identified type of wall insulation and disrepair as the main predictors.

5.122 The same modelling process for predicting dampness in houses was applied separately to purpose built flats and converted flats.

5.123 The chosen threshold which provided a good balance between the proportion of purpose built flats correctly identified as having damp or not produced a prevalence of damp in the EHS 2021-22 data of 3.76% roughly the same as the proportion as that found in the 2018+2019 combined data (3.66%).

5.124 Additional quality assurance checks were also made against the EHS predicted modelled finding for the 2020 EHS to ensure that the number and percentage of purpose built flats with damp were feasible.

5.125 The chosen threshold which provided a good balance between the proportion of converted flats correctly identified as having damp or not produced a prevalence of damp in the EHS 2021-22 data of 8.56% roughly the same as the proportion as that found in the 2018+2019 combined data (9.45%).

5.126 Additional quality assurance checks were made against the EHS predicted modelled finding for the 2020 EHS to ensure that the number and percentage of converted flats with damp were feasible.

Predictive model for Category 1 hazards

5.127 The 2021-22 dwelling level estimates on Category 1 hazards also used predictive modelling. The predictive model used was first developed for the EHS 2020-21 and the full details on the modelling methodology can be found in the English Housing Survey 2020 to 2021, technical report; English Housing Survey 2020 to 2021: technical report - GOV.UK (www.gov.uk).

5.128 The main predictor of a dwelling having a Category 1 hazard is dwelling age. The second highest predictor was the presence of a Category 1 excessive cold hazard which could be created using the EHS 2021-22 data in the normal way.

5.129 The information below summarises the process undertaken for the analysis of 2021-22 data and reports on the overall accuracy of the model.

5.130 The predictive model, was run several times to determine

  a. the probability threshold for classifying an individual dwelling a Category 1 hazard or not; and

  b. the success of the model in identifying correctly whether the dwelling had a Category 1 hazard or not when compared with the actual EHS surveyor assessments for 2018 and 2019 (used to 'train' the original model)

5.131 The chosen threshold which provided a good balance between the proportion of dwellings correctly identified as having a Category 1 hazard or not produced a prevalence in the EHS 2021-22 data of 9.95% roughly the same as the proportion as that found in the 2018+2019 combined data (9.92%).

5.132 Additional quality assurance checks were made against the EHS predicted modelled finding for the 2020 EHS to ensure that the number and percentage of dwellings with a Category 1 hazard were feasible.

Housing Health and Safety Rating System

What is the HHSRS?

5.133 The HHSRS is the Government's evidence based risk assessment procedure for residential properties. It replaced the Housing Fitness Regime on the 6 April 2006 in England. The HHSRS also replaced the Fitness Standard as an

element of the Decent Homes Standard. The HHSRS is a means of identifying defects in dwellings and of evaluating the potential effect of any defects on the health and safety of occupants, visitors, neighbours, and passers-by. The system provides a means of rating the seriousness of any hazard so that it is possible to differentiate between minor hazards and those where there is an imminent threat of major harm or even death. The emphasis is placed on the potential effect of any defects on the health and safety of occupants, visitors, and particularly vulnerable people. Altogether 29 hazards are included, Annex Table 5.5.1.

Annex Table 5.5.1: The 29 hazards covered by HHSRS

Physiological requirements	Protection against infection
• dampness and mould growth • excess cold • excess heat • asbestos (and MMF) • biocides • carbon monoxide and fuel combustion products • lead • radiation • uncombusted fuel gas • volatile organic compounds **Psychological requirements** • crowding and space • entry by intruders • lighting • noise	• domestic hygiene, pests and refuse • food safety • personal hygiene, sanitation and drainage • water supply **Protection against accidents** • falls associated with baths etc. • falling on level surfaces • falling on stairs etc. • falling between levels • electrical safety • fire • flames, hot surfaces etc. • collision and entrapment • explosions • position and operability of amenities etc. • structural collapse and falling elements

5.134 The HHSRS scoring procedure uses a formula to generate a numerical hazard score for each of the hazards identified at the property – the higher the score, the greater the severity of that hazard. Potential hazards are assessed in relation to the most vulnerable class of person who might typically occupy or visit the dwelling. For example, for falls on stairs and falls on the level, the vulnerable group is defined as persons over 60 years, and for falls between levels it is children under 5 years old.

5.135 The hazard score formula requires the HHSRS inspector to make two judgements.

- The likelihood of an occurrence which could result in harm to a vulnerable person over the following 12 months. The likelihood is to be given as a ratio – e.g., 1 in 100, 1 in 500, etc.

- The likely health outcomes or harms which would result from the occurrence. From any occurrence there may be a most likely outcome, and other possible ones which may be more or less severe. For example, a fall from a second floor window could result in a 60% chance of a severe concussion, but there may also be a 30% chance of a more serious injury and a 10% chance of something less serious. The four classes of harms and the weightings given to them are listed in Annex Table 5.5.2.

Annex Table 5.5.2: Classes of harms and weightings used in the HHSRS

Class	Examples	Weightings
Class I	Death, permanent paralysis below the neck, malignant lung tumour, regular severe pneumonia, permanent loss of consciousness, and 80% burn injuries.	10,000
Class II	Chronic confusion, mild strokes, regular severe fever, loss of a hand or foot, serious fractures, very serious burns and loss of consciousness for days.	1,000
Class III	Chronic severe stress, mild heart attack, regular and persistent dermatitis, malignant but treatable skin cancer, loss of a finger, fractured skull, severe concussion, serious puncture wounds to head or body, severe burns to hands, serious strain or sprain injuries and regular and severe migraine.	300
Class IV	Occasional severe discomfort, chronic or regular skin irritation, benign tumours, occasional mild pneumonia, a broken finger, sprained hip, slight concussion, moderate cuts to face or body, severe bruising to body, 10% burns and regular serious coughs or colds.	10

5.136 From the judgements made by the HHSRS inspector, a hazard score can be generated for each hazard, Annex Table 5.5.3.

Annex Table 5.5.3: Calculation of HHSRS hazard score

Class of Harm Weighting			Likelihood 1 in		Spread of Harm (%)		
I	10,000	÷	100	X	0	=	0
II	1,000	÷	100	X	10	=	100
III	300	÷	100	X	30	=	90
IV	10	÷	100	X	60	=	6
					Hazard Score	=	196

5.137 To provide a simple means for handling and comparing the potentially wide range of scores and avoid placing too much emphasis on the exact numbers, a series of ten hazard score bands have been devised, Annex Table 5.5.4. Bands A, B, and C are the most serious and grouped together as presenting a Category 1 hazard; local authorities have a statutory duty to consider some form of action where these are present.

Annex Table 5.5.4: HHSRS hazard score bands

Band	Equivalent Hazard Scores
A	5,000 or more
B	2,000 – 4,999
C	1,000 –1,999
D	500 – 999
E	200 – 499
F	100 – 199
G	50 – 99
H	20 – 49
I	10 – 19
J	9 or less

5.138 A number of guidance documents for HHSRS practitioners and private landlords are available. For guidance published by gov.uk see:

https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/15810/142631.pdf

How does EHS measure and model Category 1 hazards?

5.139 For the EHS, surveyors are required to collect a wide range of information in what is a relatively short and non-intrusive property inspection. The survey cannot therefore replicate in full the HHSRS assessment that would be carried out by a local authority environmental health practitioner. The approach used has been developed by the Building Research Establishment working in close co-operation with experts from the University of Warwick who were involved in the development of the HHSRS methodology.

5.140 Of the 29 HHSRS hazards only three (which occur very rarely in the stock) are not assessed by the EHS in any year. These are asbestos (and manufactured mineral fibres), biocides and volatile organic compounds.

5.141 The EHS uses three different methods to assess whether any of the 26 Category 1 hazards exist in dwellings:

- Fully measured hazards as part of the physical survey for the most common types of hazards. The surveyor first assesses whether the risks presented for each of these hazards are significantly worse than average for the age and type of dwelling concerned. If this is the case, they then score both a likelihood of an incident occurring and the expected range of outcomes. An actual HHSRS score is not computed in the field but where risks are assessed as significantly worse than average surveyors obtain this score later during validation of their survey data prior to submission. Between April 2012 and March 2020, EHS surveyors, fully measured six hazards (falling on level surfaces; falling on stairs; falling between levels; flames, hot services; fire; and damp and mould growth). In 2020-21 and 2021-22, they assessed the
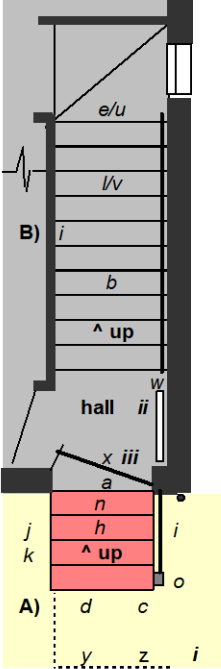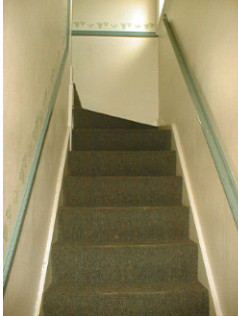
exterior of the dwellings assigned to them and produced an HHSRS score for those six hazards based only on the exterior features.

- Hazards flagged only when an 'extreme' risk is found as part of the physical survey. This approach is used for some of the rarer hazards where surveyors are instructed that 'extreme risk' equates to a Category 1 hazard. Between April 2012 and March 2020, EHS surveyors assessed 16 hazards by this approach (Annex Table 5.5.6). In 2020-21 and 2021-22, they assessed the exterior of the dwellings assigned to them and produced an HHSRS score for those 16 hazards based only on the exterior features.

- The remaining four hazards (excess cold, overcrowding, radon and lead) are usually modelled post fieldwork from other data collected on the physical survey form. This approach is used where the surveyor is less able to directly assess the risk from these hazards. The 2020-21 and 2021-22 EHS was unable to model lead post fieldwork because surveyors could not determine whether a dwelling had lead piping or not from an external inspection. Annex Table 5.5.7 summarises the assumptions and data used for the 4 hazards that are modelled from other data.

5.142 Annex Figure 5.5.5 shows a worked example of HHSRS assessment. In making their HHSRS assessments surveyors are instructed to ignore the current occupancy and assume a member of the group most vulnerable to the particular hazard occupies the property. Annex Table 5.5.6 shows how information on each hazard is collected. Please note that, to confirm with social distancing restrictions introduced during the COVID-19 pandemic, surveyors working in 2020-21 and 2021-22 would not be able to take account of the features associated with the main stairs and the small hall in the example.

Annex Figure 5.5.5: Worked example of HHSRS assessment

FALLS ON STAIRS ETC
HHSRS VERSION 2

Vulnerable group	Persons aged 60 years or over	**Multiple locations**	Yes	No
Related hazards	None	**Secondary hazards**	Yes	No

A) Front door steps **A/B)** Plan **B)** Main stairs

(Plan labels: e/u, l/v, B) i, b, ^ up, hall ii, w, x iii, a, n, h, i, j, k, ^ up, o, A) d, c, y z i)

C) Steps at gate

DESCRIPTION OF HAZARD/S

Dwelling: 1930s, Semi-detached house

A) **Front door steps:** These are of smooth painted concrete and have no top 'landing'. The bottom riser is high and uneven (300 mm max). There is a wobbly tubular steel handrail on one side but no guarding at all, despite the narrow width. There is no external porch light and little street lighting.

B) **Main stair:** The main internal stairs have two winders at the top and are moderately steep. There is a handrail only along the outside wall of the straight flight. There is a projecting radiator in the small hall and some glass in the front door close to the foot of the stairs.

C) **Steps at gate:** The steps close to the front gate are of rough spalling concrete. They have high uneven risers and a narrow tread. There is a crude rotten timber handrail but no guarding.

LIST OF RELEVANT MATTERS

LIKELIHOOD		A	B	C		OUTCOMES		A	B	C
a	Tread lengths	1	1	2		a	Length of flight	-	1	-
b	Riser heights	3	1	2		b	Pitch of stairs	-	2	-
c	Variation in T&Rs	3	1	2		c	Projections etc #	-	2	3
d	Nosing length	-	-	-		d	Hard surfaces #	2	1	2
e	Poor friction quality	3	-	1		e	Construction/repair	2	-	3
f	Openings - in stairs	-	-	-		f	Thermal efficiency	3	-	2
g	Alternating treads	-	-	-						
h-i	Lack/height handrails	3	2	2		**# Secondary hazards**		A	B	C
j-l	Lack/height guarding	3	-	1		i	Concrete kerb	2	-	-
m	Stair width	2	-	-		ii	Projecting radiator	-	2	-
n	Length of flight	-	1	-		iii	Glass in front door	-	1	-
o-q	Inadequate lighting etc	3	-	3		iv	Condition of paths	3	-	2
r	Door/s onto stairs	-	-	-						
s	Inadequate landing	3	-	-						
t	Construction/repair	2	-	3		**Key**	3 Seriously defective		1 Not satisfactory	
u	Thermal efficiency	2	-	1			2 Defective		- Satisfactory/NA	

COMPLETION OF SECTION 22 OF EHS FORM

LIKELIHOOD

Falling on stairs etc.

Significantly higher than average (Y) N

Likelihood of a person over 60 having a fall leading to harm

		1800	1000	560	320	180	100	56	32	(18)	6	2

Average / Pre 1919

Justification The main stairs are assessed as giving the same likelihood of a major fall as the average for inter-war houses, (i.e. around 1 in 320), the limited handrail provision cancelling out any benefits of the broad winders.   However, the added presence of the front access steps - particularly dangerous in icy weather and at night - substantially increases the overall annual probability of such a fall - to 1 in 18.

OUTCOMES

Likely outcome if a person over 60 should fall

Class 1 Extreme %	0.1	0.2	0.5	1	2.2	(4.6)	10	21.5	31.6	46.4	100
Class 2 Severe %	0.1	0.2	0.5	1	2.2	4.6	(10)	21.5	31.6	46.4	100
Class 3 Serious %	0.1	0.2	0.5	1	2.2	4.6	10	21.5	31.6	46.4	100

Action required

Justification The stairs are designed to be carpeted but the resulting lower harms are offset by the small hall, projecting radiator and single glazing in the door, albeit this is not at low level. However, the presence of the external front door steps and steps near the front gate, both flanked by rough tarmac and a concrete kerb, significantly increase the risk of a fatal or severe fall occurring, particularly in cold weather or at night.

Likelihood Class 1 Outcome	1 in 1800	1 in 1000	1 in 560	1 in 320	1 in 180	1 in 100	1 in 56	1 in 32	1 in 18	1 in 6	1 in 2
0.1%							E	D	C	B	A
0.2%						E-	E	D	C	B	A
0.5%						E	E	D	C	B	A
1.0%						E	E+	D	C	A-	A
2.2%				F	E-	E	D	C	B	A	A
4.6%				E-	E	D	C	B-	B	A	A
10.0%			E-	E	D	C	B-	B	A	A	A
21.5%		E	E	D	C	B	B	A	A	A	A
31.6%		E	D	C	C	B	A	A	A	A	A
46.4%	E	E	D	C	B	B	A	A	A	A	A
100%	D	C-	C	B	A	A	A	A	A	A	A

ACTION REQUIRED

Justification Replacing the steps to the front door and at the gate will be picked up under Section 18. This will bring the property's rating back to average for its age and type.

Action required?	Action	Coded elsewhere?		Quantity	
Y	Install handrail	Y	N	Metres:	
Y	Install balustrade	Y	N	Metres:	
Y	Cover dangerous balustrade/guarding	Y	N	Metres:	
Y	Repair/replace internal staircase (S5)	Y			
Y	Redesign internal, common or external staircase (design, not condition)		N	Number:	
Y	Repair/replace external/common staircase (S9)	Y			
Y	Repair/replace external steps (S11, S18)	(Y)	N	Number:	
Y	Cover slippery stairs	Y	N	Flights:	
Y	Repair/replace/provide additional lighting (S5, S9, S11)	Y	N	Number:	
Y	Remove obstacle		N	Number:	

Annex Table 5.5.6: Summary of how EHS collects and models information about HHSRS hazards

Hazard	How assessed	Average HHSRS score	Specified vulnerable age group
Excess cold*	Modelled	926	Age 65 or over
Falling on level surfaces*	Fully measured	181	Age 60 or over
Falling on stairs etc.*	Fully measured	134	Age 60 or over
Radiation*	Modelled	91	None
Collision and entrapment	Flagged if an extreme risk	57	Age under 5
Flames, hot surfaces etc.*	Fully measured	42	Age under 5
Crowding and space*	Modelled	19	None
Fire*	Fully measured	17	Age 60 or over
Dampness and mould growth*	Fully measured	11	Age under 14
Entry by intruders	Flagged if an extreme risk	11	None
Falls associated with baths	Flagged if an extreme risk	7	Age 60 or over
Noise*	Flagged if an extreme risk	6	None
Falling between levels*	Fully measured	4	Age under 5
Food safety	Flagged if an extreme risk	2	None
Electrical safety*	Flagged if an extreme risk	2	Age under 5
Carbon monoxide and fuel combustion products*	Flagged if an extreme risk	1	Age 65 or over
Personal hygiene, sanitation and drainage*.	Flagged if an extreme risk	1	Age under 5
Explosions	Flagged if an extreme risk	1	None
Position and operability of amenities etc.	Flagged if an extreme risk	1	Age 60 or over
Structural collapse and falling elements	Flagged if an extreme risk	1	None
Excess heat	Flagged if an extreme risk	0	Age 65 or over
Asbestos (and MMF)	Not assessed	0	None
Biocides	Not assessed	0	None
Lead*	Modelled	0	Age under 3
Uncombusted fuel gas	Flagged if an extreme risk	0	None
Volatile organic compounds	Not assessed	0	None
Lighting	Flagged if an extreme risk	0	None
Domestic hygiene pests and refuse.*	Flagged if an extreme risk	0	None
Water supply	Flagged if an extreme risk	0	None

Notes:
1) average scores are for all dwellings and taken from Version 2 of the HHSRS guidance. The averages have been calculated for the age range of the population most vulnerable to each type of hazard.
2) the 15 hazards which were scored or modelled for 2006 and 2007 are identified by an asterisk. This group is still used in the current reporting of the 'Decent Homes' HHSRS criterion.

5.143 In the 2006 and 2007 English House Condition Survey (EHCS), fewer hazards were fully scored and some of the hazards that are now measured or flagged were modelled using other data; see the EHCS technical report from 2007 for full details:
http://webarchive.nationalarchives.gov.uk/20120919132719/www.communities.gov.uk/publications/housing/ehcstechnicalreport2007).

5.144 From 2008, reporting of HHSRS covers all of the 26 hazards covered by EHS so figures are not strictly comparable with the 2006 and 2007 HHSRS data. Reporting on decent homes (see decent homes section later in this section), continues to use the 'old' (15 hazards) version of HHSRS for continuity over time. Annex Table 5.5.7 summarises the assumptions and data used for the 4 hazards that are modelled from other data.

Annex Table 5.5.7: Modelling HHSRS hazards using EHS data

Hazard	Category 1 hazard defined as:
Excess Cold	Estimates the number of households living in homes with a threat to health arising from sub-optimal indoor temperatures using the Standard Assessment Procedure (SAP).[18] This hazard is based on dwellings with an energy efficiency rating of less than 35 based using the original SAP 2001 methodology. The updated SAP 2009 methodology, used for the 2010-2012 EHS reports, recalculated the comparable threshold to be 35.79. From 2013 to 2017, the EHS report used the updated SAP 2012 methodology and the comparable excess cold threshold was recalculated to 33.52. In 2018, the SAP 2012 methodology was updated for the EHS and the comparable excess cold threshold was recalculated to 37.6. This approach ensured that the number and percentage of dwellings failing on excess cold would be the same under both the old and new SAP2012 methodology. Although the changes in SAP methodology and cut-off thresholds create difficulties in reporting on excess cold trends over time, the approach allows the findings to offer some degree of consistency for those who wish to look at HHSRS over time.
Radiation	The dwelling is located in one of the critical 16 post code sectors, based on a radon exposure map of England AND the dwelling was built before 1980.
Lead (not modelled in 2020-21)	The dwelling is located in one of 4 post codes with very soft water (based on the drinking water quality map of England) AND built before 1945 AND with lead piping present either before or after the mains stop cock.
Crowding and space	The occupants per habitable room ratio is calculated. If this exceeds 2 the dwelling has a category 1 hazard regardless of size. If it is equal to 2 and the number of habitable rooms is 2 or more the dwelling also has category 1 hazard.

Data quality and reliability

5.145 Surveyors working on the EHS have received extensive training and support to help ensure their HHSRS assessments are consistent and robust. This includes residential training involving classroom and field exercises together with e-learning exercises. Refresher programmes are provided annually, together with manuals providing benchmark examples for reference when making their judgements. New surveyors are accompanied in the field and there is on-going close supervision throughout fieldwork. Calibration exercises are also used to monitor variability in surveyors' HHSRS assessments over time.

---

[18] SAP is the Governments standard procedure for Energy ratings of dwellings.

5.146 While these measures ensure a good level of consistency in judgements, some surveyor variability is to be expected. The EHS approach to the HHSRS provides surveyors with a systematic approach with which to make these judgements.

Decent homes

5.147 Due to COVID-19 restrictions in 2020-21 and 2021-22, it was not possible to collect data on the internal condition of homes required to model whether a dwelling met the Decent Homes Standard or not. However, it was possible to develop a binary classification model to produce modelled estimates at dwelling level to indicate whether or not a dwelling met the Standard. This binary classification model included the predictive modelling of HHSRS hazards.

5.148 To provide context for the binary classification model this section gives a detailed definition of the four criteria that a dwelling is required to meet to be considered 'decent' under the Decent Homes Standard. A dwelling must meet all of the four criteria listed below to be classed as decent:

A) it meets the current statutory minimum standard for housing

B) it is in a reasonable state of repair

C) it provides reasonably modern facilities and services

D) it provides a reasonable degree of thermal comfort

Criterion A: the dwelling meets the current statutory minimum standard for housing

5.149 The current statutory minimum standard for housing is the HHSRS. To be decent, the dwelling must be free from Category 1 hazards (see previous section).

5.150 In 2020-21 and 2021-22, the predictive modelled estimate on whether a dwelling had a Category 1 hazard or not was used to indicate the presence of Category 1 hazards. Although the HHSRS criterion for the Standard is normally modelled for 15 hazards, the prevalence of 15 hazards or 26 hazards in the stock is very similar (less than 0.1% in the EHS 2018/19). Consequently, the predictive model for the 26 hazards was considered to be sufficiently robust.

Criterion B: the dwelling is in a reasonable state of repair

5.151 A dwelling satisfies this criterion unless:

- one or more key building components are old and, because of their condition, need replacing or major repair; or

- two or more other building components are old and, because of their condition, need replacement or major repair.

5.152 This criterion was modelled in 2020-21 and 2021-22 using only those building elements collected by the external plus approach (chimney, external doors, roof structure, roof covering, wall structure, wall covering and windows). Previous years' EHS data indicates that these external building elements are the main drivers of homes failing this criterion. Data items collected by the 'external plus' approach are highlighted in red boxes on the 2021-22 EHS physical survey form https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/1124155/EHS_2021-22_Physical_Survey_Form.pdf

5.153 Key building components are those which, if in poor condition, could have an immediate impact on the integrity of the building and cause further deterioration in other components. If any of these components are old, and need replacing or require immediate major repair, then the dwelling is not in a reasonable state of repair. They are the external components plus internal components that have potential safety implications and include:

- external walls;
- roof structure and covering;
- windows/doors;
- chimneys;
- central heating boilers (not included in the modelling in 2020-21 and 2021-22); and
- electrics (not included in the modelling in 2020-21 and 2021-22).

5.154 Other building components are those that have a less immediate impact on the integrity of the dwelling. Their combined effect is therefore considered, with a dwelling not in a reasonable state of repair if two or more are old and need replacing or require immediate major repair. Such components include kitchen and bathroom amenities, central heating distribution and storage heating. Due to the use to the external plus approach to the physical survey, these components were not included in the modelling in 2020-21 and 2021-22.

5.155 The terms 'old' and 'in poor condition' are also quite tightly defined as below:

- Old: the component is older than its expected or standard lifetime. The component lifetimes are listed in Annex Table 5.5.8.
- In poor condition: the component needs major work, either full replacement or major repair. The definitions used for different components are as listed in Annex Table 5.5.9.

5.156 Establishing whether dwellings surveyed in the EHS meet this criterion depends on the assessment of both the ages of key and other building components and of their condition.

5.157 The EHS surveyors record their assessment of the ages of the main external building elements together with key services and amenities. They are also given the shortcut option of recording whether elements are original i.e. the same as the building itself. Where the age of a component cannot be assessed, it is assumed to be original i.e. the same age as the dwelling. In the cases where components are the 'same age as dwelling', the probability that they have exceeded their lifetime is calculated based on dwelling age band. This is because in earlier years of decent homes modelling, dwelling age was only recorded in bands rather than as a single year. In recent years, although actual age of construction is recorded as a single year, the same method of calculating these probabilities is adopted for a consistent approach.

5.158 For example, for houses, windows are assumed to have exceeded their lifetime if they are more than 40 years old (for flats, the windows lifetime is 30 years old). Where houses were built between 1975 and 1980 (inclusive) and still had the original windows, some of these would have windows that were over 40 years old. A simple and robust approach is used, assuming that roughly equal numbers of dwellings were built in each year of this age band. In 2021, houses built in 1975, 1976, 1977, 1978, 1979 and 1980 (are over 40 years old and) represent all 6 years out of the 6 year age band, so all original windows in houses built in 1975-1980 are given a probability of 1.0 (6/6) of being over 40 years old.

5.159 For most dwellings, the assessment of whether or not they satisfy the disrepair criterion is clear cut. For the remainder, for each building component which is in poor condition, the probabilities of being beyond the normal lifetime are combined to give a total probability, taking into account the split into major and minor elements. If this total is greater than 0.5, the dwelling is classed as non-decent due to disrepair.

5.160 Annex Table 5.5.8 shows the lifetimes of building components used to assess whether the components are 'old' in the terms of the disrepair criterion. These lifetimes are used to construct the national estimates of the number of dwellings that are decent and those that fail.

Annex Table 5.5.8: Component lifetimes used in the disrepair criterion

Building components (key components marked *)	Houses and bungalows	All flats in blocks of below 6 storeys	All flats in blocks of 6 or more storeys
Wall structure*	80	80	80
Lintels*	60	60	60
Brickwork (spalling)*	30	30	30
Wall finish*	60	60	30
Roof structure *	50	30	30
Roof finish *	50	30	30
Chimney *	50	50	N/A
Windows *	40	30	30
External doors *	40	30	30
Kitchen	30	30	30
Bathrooms	40	40	40
Heating – central heating gas boiler *	15	15	15
Heating – central heating distribution system	40	40	40
Heating – other	30	30	30
Electrical system *	Modern	Modern	modern

5.161 As age of electrical system is not collected in the EHS, it is considered to be 'old' if it is not modern, i.e. it has lead or rubber covered wiring, there are separate fuse boxes for each circuit, or earthing wires are unsheathed/green covered.

5.162 Annex Table 5.5.9 sets out the definitions used within the disrepair criterion to identify whether building components are 'in poor condition'. For more detailed information on how surveyors are instructed to record disrepair, see the repair costs section.

Annex Table 5.5.9: definition of 'poor condition' used in disrepair criterion

Building components (key components marked *)	Definition of 'in poor condition' used on the EHS
Wall structure	Replace 10% or more, or repair 30% or more
Wall finish	Replace/repoint/repair 50% or more
Chimneys	One chimney needing partial rebuilding or more
Roof structure	Replace 10% or more, or strengthen 30% or more
Roof covering	Replace or isolated repair to 50% or more
Windows	Replace at least one window or repair/replace sash or member to at least two (excluding easing sashes, reglazing, painting)
External doors	Replace at least one
Kitchen	Major repair or replace 3 or more items out of 6 (cold water drinking supply, hot water, sink, cooking provision, cupboards, worktop)
Bathroom	Major repair or replace 2 or more items (bath, wash hand basin, WC)
Electrical system	Replace or major repair to system
Central heating boiler	Replace or major repair
Central heating distribution	Replace or major repair
Storage heaters	Replace or major repair

Criterion C: The dwelling has reasonably modern facilities and services

5.163   As surveyors were unable to inspect the interior of dwellings in 2020-21 and 2021-22, a dwelling was deemed to have failed the criterion on facilities and services if there were serious noise pollution issues assessed by the surveyor and a lack of any double glazing.

Criterion D: the dwelling provides a reasonable degree of thermal comfort

5.164   This criterion was modelled in 2021-22 in the normal way by the EHS energy model. The definition requires a dwelling to have both efficient heating and effective insulation. Both of these are defined very precisely in terms of what is present rather than by the overall energy performance of the dwelling.

5.165   Under this definition, efficient heating is defined as any gas or oil programmable central heating or electric storage heaters / programmable solid fuel, or communal heating or LPG central heating or similarly efficient heating systems. Heating sources which provide less energy efficient options do not meet this criterion.

5.166   Because of the differences in efficiency between gas/oil heating systems and the other heating systems listed, the level of insulation that is appropriate also differs:

- For dwellings with gas/oil programmable heating, cavity wall insulation (if there are cavity walls that can be insulated effectively) or at least

50mm loft insulation (if there is loft space) is an effective package of insulation.

- For dwellings heated by electric storage heaters / programmable solid fuel or LPG central heating a higher specification of insulation is required to meet the same standard: at least 200mm of loft insulation (if there is a loft) and cavity wall insulation (if there are cavity walls that can be insulated effectively).

5.167 Assessing whether the EHS sample dwellings pass or fail the decent homes thermal comfort criterion is complex because it involves an array of survey information related to insulation, heating and structural properties. The data collected on the form and the modelling assumptions have been changed and refined since the original 'baseline' figures were published in 2001. For more information on how these changed from 2001 to 2007 see the EHCS 2007 Technical Report: http://webarchive.nationalarchives.gov.uk/20120919132719/www.communities.gov.uk/publications/housing/ehcstechnicalreport2007

5.168 The 2008 data experienced modelling changes in the assumptions on cavity wall insulation to incorporate the use of the summary section on cavity wall insulation newly collected on the EHS 2008 physical survey form. There were no modelling changes in 2009.

5.169 The key modelling and form changes affecting thermal comfort since 2010 are:

- Where no loft insulation information is available for a room in the roof or a flat roof, appendix S of the SAP 2012 information booklet is used in conjunction with the actual date of construction or, if it is a loft conversion, the date of the loft conversion to determine an estimated amount of loft insulation (the banded construction date is used if the actual construction date is unknown).
- The type of loft insulation is now recorded and used to adjust insulation thickness based on the relative performance of the insulation material compared to mineral wool.
- Changes in assumptions on cavity wall insulation to incorporate extended use of the summary section on cavity wall insulation collected on the EHS physical survey form following a wording change to the overarching cavity wall insulation summary question.
- Homes built after 2002 with cavity walls are assumed to have full cavity wall insulation.
- Park homes are now identified through the EHS survey. A park home is a temporary dwelling which is located on a managed site, alongside other similar homes. Insulation parameters for these dwellings are calculated using specific park home assumptions from Appendix S.

Repair costs

5.170 Due to COVID-19 restrictions in 2021-22 it was not possible to collect data on the internal condition of homes required to undertake timeseries modelling of disrepair costs. The 2021 EHS Housing Quality and Condition report and the 2021 Private Rented Sector and Social Rented Sector Reports include analysis on the estimated cost to make all non-decent homes meet the Decent Homes Standard. These 'cost to make decent' normally use data from the repair cost modelling but for the 2021 reports, costs have been extrapolated using previous years data.

Cost to make decent

5.171 For the 2021-22 EHS the cost to make a non-decent home meet the Decent Homes Standard could not be modelled in the normal way due to the absence of repair cost data. The cost to make decent for each home that failed the Standard used the estimates created for the 2020-21 EHS and inflated the prices to 2021.

5.172 The 2020-21 EHS methodology to estimate the cost to make a non-decent home meet the Standard used the data from previous EHS surveys, specifically previous 5 years of single year survey data. From each of these 5 years, average costs for 10 dwelling type categories (using the variable 'alltypex'), split by private/social sector, were used for the calculation. These individual average costs were then 'sense checked' against the distribution of costs.

5.173 All average costs within each dwelling type and tenure sector from the EHS 2015-2018 single year data were then inflated to 2019 prices by applying an inflation factor used for repair cost models for the EHS 2019. An overall average cost to make decent was then calculated for each type and tenure category at 2019 prices. Where sample sizes were very small, an appropriate cost was imputed, for example, for social sector pre-1919 built and post-1919 built detached homes, an average cost was calculated for all social sector detached homes.

5.174 The estimated costs to make decent were based on the 26 hazard HHSRS definition to replicate the approach of the decent homes modelling.

5.175 The repair costs are for those items causing the dwelling the fail the Standard and do not take account of work that may arise in the future due to ageing components.

5.176 As the 2020-21 and 2021-22 costs to make decent were estimated using the average costs from previous years of the EHS, information on how it was modelled prior to 2020-21 is detailed below to provide additional context. For

the initial modelling, dwellings failing each of the four components of decent homes (modernisation, disrepair, thermal comfort and HHSRS[19]) are costed separately.

HHSRS component

5.177 Costs to mitigate Category 1 HHSRS hazards, known as costs to make safe, provide notional costs to remedy each type of Category 1 hazard.

I. For the fully measured hazards, this involves two processes;
- costing the required works identified by the surveyor in the HHSRS section of the physical survey form; a 'typical' specification of work has been devised by an experienced HHSRS practitioner for the remedial action
- using repair cost work described elsewhere on the form.

II. For costs to mitigate excess cold at a dwelling the EHS uses the latest EPC methodology to identify both the energy improvement work required and the associated cost.

III. For the modelled hazards, a typical action has been selected by an experienced HHSRS practitioner (as above for measured hazards). The exceptions are for: falls associated with baths, entry by intruders, noise and collision and entrapment. In 2012 those were assessed as extreme risks, having previously been fully measured by surveyors. For those hazards, the average costs of mitigating the hazard, based on previous years' data, is used to derive the costs.

5.178 The outputs from the three processes above are then combined to provide the notional overall costs for all 26 hazards collected by the EHS and the 15 hazards used in the cost to make decent modelling. Any double counting of work (e.g. heating improvements may be required to mitigate dampness and excess cold) is removed where feasible. Actions are grouped into broad work areas e.g. insulation, heating, kitchens, stairs. Remedial works that do not fit into these classifications, and are independent of any other jobs, are grouped as 'other' e.g. work to mitigate the risk of harm from lead or overcrowding.

Disrepair, modernisation and thermal comfort components

5.179 The costs for dealing with disrepair are based on repair costs derived from the repair cost model. Action to services and heating are costed slightly differently. Although those are based on repair cost model prices, the specifications of work are more specific than in the repair cost model.

5.180 For dwellings failing the modernisation component, the costs include work to remedy all items that currently fail the Standard (even if by fixing only one

---

[19] Based on the 15 hazard HHSRS definition

item makes the dwelling decent, costs here include remedying all items). Costs to modernise the kitchen and bathroom are based on 2001 Major Repairs Allowance (MRA) prices. Only kitchen costs are scaled by dwelling size. The costs to install double glazing are also based on 2001 MRA prices and scaled by total window area for dwelling. Costs for other works are specified and derived using information from price books, actual spending on work from other sources and professional experience.

5.181 Thermal comfort (heating) costs are derived from MRA prices. Thermal comfort (insulation) costs are derived from price books, actual spending on work from other sources and professional experience.

Total costs

5.182 The four component costs are then combined. Any double counting between all four components is removed where feasible. For example, bathroom disrepair (failing on age and action needed), bathroom modernisation and HHSRS personal hygiene/sanitation costs). In order to resolve that, actions are grouped into broad work areas e.g. windows, bathroom, kitchen, heating, common access areas. Types of work that do not fit into those classifications, and are independent of any other jobs, are grouped as 'other' (e.g. installing/repairing extractor fans to fix condensation).

5.183 Once the costs for all work areas are added together, additional costs are added to account for preliminaries and access equipment[20] to arrive at a total cost to make decent for each dwelling. This cost is then inflated to the current survey year, adjusted for region and whether the dwelling is in the private or social sector.

---

[20] See repair costs of this Technical Report for further information.

Annex 5.6: Energy efficiency

5.184 The English Housing Survey (EHS) collects a large amount of detailed information relating to building construction, heating, and insulation. This provides a detailed profile of the energy performance and carbon emissions of the existing housing stock and how far these could be improved using different types of measures. This annex sets out:

- Main components of energy efficiency – how the individual components contributing to overall efficiency are defined and modelled.
- Standard Assessment Procedure (SAP) of buildings energy performance – the methods and assumptions used to calculate SAP (energy efficiency) ratings and carbon dioxide emissions.
- Energy performance certificates (EPC) – how an assessment is made of what measures could be installed to improve energy efficiency and reduce carbon emissions, and what the impact of installing these possible improvements would be. It also briefly discusses households' awareness of EPCs and the improvements they suggest.
- Barriers to improving insulation – classifying the ease of installation or specific barriers to loft, cavity and solid wall insulation.

Main components of energy efficiency

5.185 Due to COVID-19 restrictions in place during the 2021 survey year, an alternative methodology was implemented to carry out the EHS physical survey. Surveyors collected the physical inspection data using an assessment of the exterior of the dwelling together with a short, socially distanced, interview carried out to ask the occupants for details of items usually captured internally by the survey. For more information on the data collected directly and indirectly by the surveyor see the EHS 2021 physical survey form https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/1039253/English_Housing_Survey_2020_to_2021_physical_survey.pdf . This included items such as the type of heating system present, and extent of low energy lighting. Whenever possible, surveyors also used information from energy performance certificates (EPCs) and web-based tools such Google Street View and Rightmove to inform their decisions.

5.186 Where data were missing (e.g. loft insulation present but no insulation thickness given) due to the data collection method changes, imputed default values were used consistent with the approach used in a full survey year.

Primary heating systems

5.187 The EHS records up to two forms of space heating system and all water heating systems present in each dwelling. Where two types of space heating system are present, the EHS designates the one that covers the largest proportion of the dwelling as the primary heating system and collects detailed information on its overall type, the fuel used, boiler details (where relevant) and heating controls. The primary space heating type is classified as central heating system, storage heaters or room heaters.

Central heating system

5.188 This is most commonly a system with a gas fired boiler and radiators, distributing heat throughout the dwelling. Also included in this definition are warm air systems, heat pumps, communal heating and electric ceiling/underfloor heating, (included in 'other systems' in the 2012 dataset). Central heating is generally considered to be a cost effective and relatively efficient method of heating a dwelling, although the cost effectiveness and level of carbon dioxide (CO_2) emissions will be closely linked to the type of fuel.

Storage heaters

5.189 These are predominately used in dwellings that have an off-peak electricity tariff. Storage heaters use off-peak electricity to store heat in clay bricks or a ceramic material; this heat is then released throughout the day. These are more cost effective than fixed or portable room heaters, but storage heating can prove expensive if too much on-peak electricity is used during the day. Their efficiency is calculated based on their age and the type of controls present.

Room heaters

5.190 This category includes all other types of heater such as fixed electric heaters . This type of heating is generally considered to be the least cost effective of the main systems and produces more CO_2 emissions per kWh than other heating systems.

Secondary heating systems

5.191 Where more than one space heating system or appliance has been recorded and the primary system identified as above, the additional appliance is coded as the secondary system and, along with the secondary fuel, used in the SAP calculation and other analysis. These systems may have been originally installed alongside the primary system, perhaps in a larger home. More often they would have been the only source of heating when the property was built, before being superseded by a new system, typically a form of central heating. The secondary space heating type is classified as follows:

- Fixed room heaters: the majority of secondary systems fall into this category, which includes various types of mains gas fires, solid fuel fires and stoves, and direct acting electric panels and radiators which are wired into the mains electricity.
- Storage radiators: individual storage heaters which are subsidiary to the main heating system are included here.
- Portable heaters: where the only secondary heating is through a portable electric heater. This includes cases where the SAP methodology has concluded that the main fixed heating is insufficient to heat the dwelling to a satisfactory level, so a portable secondary system is imputed to allow an energy efficiency rating to be calculated.

Boilers

5.192 Where the heating system has a boiler, the EHS collects basic information on its generic type. The EHS also collects information about the make and model of the boiler and its age so that an accurate estimate of its overall fuel efficiency can be derived. There are four main types of boilers:

- Standard boiler: these provide hot water or warm air for space heating, with the former also providing hot water via a separate storage cylinder.
- Back boiler: these older models are located behind room heaters and feed hot water to a separate storage cylinder. They are generally less efficient than other boiler types.
- Combination boiler: provides hot water or warm air for space heating and can provide hot water on demand, thus negating the need for a storage cylinder and therefore requiring less room.
- Condensing boiler: standard and combination boilers can also be condensing. A condensing boiler uses a larger, or dual, heat exchanger to obtain more heat from burning fuel than an ordinary boiler, and is generally the most efficient boiler type. Changes to Building Regulations have seen an increase in condensing boilers as they have become mandatory for all replacements.

Water heating

5.193 All existing water heating systems are recorded by the surveyor. Where more than one water heating system is present, the system used for analysis and modelling is selected in the order of the categories below. The categories of water heating systems used in the report are:

- With central heating: the water is primarily heated by the same system as the primary space heating, usually a standard boiler with a separate storage cylinder or a combination boiler heating water on demand.
- Dedicated water boiler: a separate boiler to the space heating system, possibly using a different fuel, provides the hot water. This category

includes dedicated back boilers and communal systems for water heating only.

- Immersion heater: hot water is provided by a single or dual electric immersion heater in the storage cylinder. These are less energy efficient than central or separate boilers but are often found as a 'top-up' system for other systems.
- Instantaneous water heater: these water heating appliances heat small amounts of water on demand in a similar way to a kettle and distribute the hot water to one or more points.
- Instantaneous showers: these are shower units in which the water is heated electrically on demand.
- Other: category can include solar water heating, or any other systems not captured by the categories above.

Wall types and wall insulation

5.194 The construction of the external walls and whether they contain any additional insulation is an important determinant of heat loss. The EHS collects detailed information on the overall construction type, age of the building, added wall insulation and what proportion of the external walls consists of different types.

Cavity walls

5.195 A cavity wall is one constructed of two brick or block walls separated by a cavity that is at least 50mm wide. They are generally found in houses dating from about 1930 onwards, although some older examples exist. Many dwellings (especially older private sector homes) have a mix of wall types because they have had one or more extensions added at different times.

5.196 Dwellings are only classed as having 'cavity walls' where at least 50% of the total external wall area is cavity brickwork. This means that a small house built with solid 9" brick walls in 1900 which had a cavity brickwork extension that was larger than the original building added in 1960 would be classed as having 'cavity walls'.

5.197 Dwellings with cavity walls can have none, part or all of the cavity wall area insulated. The insulation can be built into the original wall construction or installed later.

Post-1995 dwellings

5.198 In addition to cases that have been identified in the EHS physical survey as having evidence of cavity wall insulation and those without evidence of cavity wall insulation, a third category is established: post-1995 dwellings with predominantly cavity walls, without evidence of full cavity wall insulation. It is likely that these dwellings had cavity wall insulation installed at the time of construction (known as 'as-built' insulation), however the non-intrusive survey undertaken in the EHS would not always be able to identify this. In order to

provide a more realistic estimate of the number of insulated cavity walls, all post-1995 cavity wall dwellings are assumed to be fully insulated, even if this was not apparent in the physical survey.

5.199 The 1990 and 1994 Building Regulations both specify an external wall U-value, however compliance could be achieved through other mechanisms as an alternative to cavity fill. Therefore, an increasing proportion of new dwellings were built with cavity wall insulation in the early 1990s, and it is thought to have become the predominant practice after 1995. The age band in which as-built insulation is assumed to have been installed was revised for the EHS data from 2015, in order to bring the methodology in line with the SAP and RdSAP modelling methodologies. Details of the previous methodology can be found in the 2014 Technical Report https://www.gov.uk/government/publications/english-housing-survey-2014-to-2015-technical-report.

Solid walls

5.200 Where dwellings do not have cavity walls, external or internal wall insulation can be installed to improve energy efficiency, where the thermal properties of the external walls are poor. Where a surveyor has recorded that external wall insulation had been applied to at least 50% of a non-cavity walled dwelling, or that at least 50% of the rooms have had internal insulation applied, it was classed as having an insulated solid wall.

Park homes

5.201 From 2019 onwards, the EHS form allows for park homes to be identified as a distinct dwelling type (previously these were classed as detached dwellings). Where a park home is identified, the wall type for these dwellings is classified in the same way as for other dwelling types, using form data as recorded by the surveyor but specific park home U-values are used to calculate the thermal performance of the walls, as specified in appendix S of SAP 2012.

Drylining and internal wall insulation

5.202 Data on the presence of drylining and internal wall insulation were not collected in the 2021-22 physical survey. It was assumed there was no drylining present for all cases. The household questionnaire was used to determine whether there was any internal wall insulation present.

Loft insulation

5.203 Adequate loft insulation can make significant savings to both heating costs and CO_2 emissions, making this generally a cost-effective method of insulation. It involves fitting insulating foam or fibre between the joists or rafters in a loft, which reduces heat from within the thermal envelope below escaping through the roof. The thermal envelope of the dwelling is the

physical barrier between the warm interior and the cold air outside or in unheated spaces such as a loft for example. From 2015 onwards, the performance of loft insulation has been adjusted to reflect the performance of the material compared with mineral wool. Rigid foam board is assumed to perform twice as well as mineral wool and Vermiculite beads are modelled to perform two-thirds as well as mineral wool.

5.204 The EHS physical survey involves an inspection of the loft where the surveyor notes whether insulation is present and measures its thickness. The collection of loft insulation data was changed after the 2001 English House Condition Survey (EHCS), so analysis of data from 2003 onwards cannot be directly compared to previous data (see the EHCS 2003 technical report for details). In cases where surveyors are unable to access lofts or where the dwelling is a house or top-floor flat with a flat or shallow pitched roof, the amount of insulation in the dwelling was classed as unknown. However, for the purpose of calculating a SAP rating, an amount was imputed using the mean value for dwellings of that age, tenure and broad geographical area. These classifications were used because earlier regression analysis indicated that these factors were the main determinants of the amount of loft insulation present. In the case of park homes, insulation thickness is imputed by following RdSAP conventions. RdSAP consists of a number of conventions, default assumptions, and inference procedures which can be followed in the event of having fewer data inputs than is required to perform a full SAP calculation. These are listed in Appendix S of SAP 2012. The latest version (v9.93) can be found here: https://www.bre.co.uk/filelibrary/SAP/2012/RdSAP-9.93/RdSAP_2012_9.93.pdf.

Solid floor insulation

5.205 Energy calculations require the internal dimensions of the surveyed rooms (living room, kitchen, bedroom and bathroom) to calculate the extent of solid floors on the basement/ground level to inform the calculation of ground floor heat losses.

5.206 As surveyors were unable to collect internal room dimensions in the 2021-22 survey year, internal dimensions from the EHS 2015/16 and 2017/18 combined year datasets were used to calculate proxy room areas to calculate the extent of solid floors. Analysis of 2019 data indicated that this change decreased the annual mean SAP rating by around 0.05 SAP points.

Low energy lighting and conservatories

5.207 Analysis in the 2011 EHS Report on the housing stock (https://www.gov.uk/government/statistics/english-housing-survey-2011-homes-report) examined headline figures for homes which predominantly use low energy lighting and those with conservatories. This used the interior section to calculate the proportion of surveyed rooms with low energy lights at

the time of survey, whist data relating to the size, glazing type and heating of conservatories were also taken from the raw physical survey data. These data have not been re-analysed in subsequent EHS reports, but these areas may be re-visited in future.

Renewable energy measures

5.208 Since 2009, EHS surveyors have recorded the presence of solar photovoltaic panels and domestic wind turbines for electricity generation, whilst the presence of solar hot water panels has been collected since 2001. Since 2011, EHS reports have included analysis of any observed renewable energy technologies.

5.209 In the full SAP methodology, a calculation is used to determine the electricity production of PV panels, using peak power of the type of PV panel multiplied by factors based on dwelling orientation and over-shading. Prior to 2014, an assumed peak power of 2.5kWh was applied to each case. From the 2014-15 survey year onwards, the area of PV panels has been collected, which allows peak power to be calculated more accurately for each dwelling, using the equation: (Peak power (kWh) = 0.12 x PV area), in line with RdSAP conventions.

5.210 Prior to 2014, detailed information on biomass heating fuels was not available, so biomass heating systems were assumed to use wooden logs as their primary heating fuel. From the 2014-15 survey onwards, an update to the EHS form allows for more detailed information on the type of biomass heating fuel to be collected. From the 2017-18 survey onwards, information on the type of biomass fuel has also been collected for communal heating systems, where previously, the heating fuel was assumed to be gas. This allows heating systems with renewable heating fuels to be modelled more effectively.

Standard Assessment Procedure (SAP) of buildings energy performance

SAP ratings

5.211 The Standard Assessment Procedure (SAP) is the Government's recommended system for home energy ratings. SAP ratings allow comparisons of energy efficiency between different dwellings to be made. The SAP rating is expressed on a logarithmic scale, which normally runs from 1 (very inefficient) to 100, where 100 represents zero energy cost. The rating can be greater than 100 for dwellings that are net exporters of energy; however, these are extremely rare in the existing dwelling stock. In extremely inefficient cases the formula that defines the rating can result in negative values, but values less than one are coded to a SAP rating of one.

5.212 The Building Regulations require a SAP assessment to be carried out for all new dwellings and conversions. Local authorities, housing associations, and other landlords also use SAP ratings to estimate the energy efficiency of existing housing. The version of SAP used in the survey is currently SAP 2012. This version is used in the current EHS dataset (employed retrospectively to provide a consistent measure from 1996 to the most recent survey year). EHS reports dating from 2010 to 2012 used the previous (SAP 2009) version of SAP.

5.213 The 2012 changes in the SAP methodology are far less reaching than those which occurred following the move from SAP2005 to SAP2009 in 2010. Differences in SAP ratings calculated under SAP2009 and SAP2012 mainly occur for dwellings using solid fuel; for further details see The Government's Standard Assessment Procedure for Energy Rating of Dwellings 2012 edition (SAP worksheet Table 15 p.231 http://www.bre.co.uk/filelibrary/SAP/2012/SAP-2012_9-92.pdf).

5.214 In 2017, an update to Appendix S of SAP was released (RdSAP 2012 v9.93) which included revised U-values for solid and cavity brick walls and stone walls. The changes specified in this update were incorporated into the EHS modelling from 2019, with the full effect of the changes realised in the 2019-20 combined year survey. The RdSAP v9.93 updates have been applied in the modelling of Fuel Poverty by DESNZ since 2018. It is therefore worth noting that outputs produced for the 17-18 and 18-19 EHS surveys will not be directly comparable to the equivalent Fuel Poverty outputs, due to the staggered adoption of the methodological update.

5.215 The SAP ratings give a measure of the annual unit energy cost of space and water heating for the dwelling under a set heating regime which assumes specific heating patterns and room temperatures. The fuel prices used are averaged over the previous three years across the different areas of the UK. The SAP rating takes into account a range of factors that contribute to energy efficiency, which include:

- thermal insulation of the building fabric;
- the shape and exposed surfaces of the dwelling;
- materials used for construction of the dwelling;
- efficiency and control of the heating system;
- the fuel used for space and water heating, ventilation and lighting;
- ventilation and solar gain characteristics of the dwelling; and
- renewable energy technologies.

5.216 SAP is not affected by the individual characteristics of the household occupying the dwelling, nor by its geographical location. The calculation is based on a fixed heating pattern of 21°C in the main living area and 18°C

elsewhere. It is also based on standard occupancy assumptions with the household size correlating with the total floor area of the dwelling.

5.217 The EHS uses a computerised version of the SAP methodology to calculate the SAP rating for each dwelling included in the physical survey sample. Most of the data required to calculate SAP are available from the survey, either directly from the questions asked or as a result of further modelling. Those data items that are not collected have very little impact on the final calculated rating. Where data items are missing these are dealt with using default information based on information from dwellings of the same age, built form, tenure, number of floors and size.

5.218 The Energy Efficiency Rating (EER) is derived by translating the SAP ratings into an A to G banding system where band A represents low energy costs and band G represents high energy costs, Annex Table 5.6.1.

Annex Table 5.6.1: SAP rating and Energy Efficiency Rating (EER) bands

SAP rating	EER band
1 to 20	G
21 to 38	F
39 to 54	E
55 to 68	D
69 to 80	C
81 to 91	B
92 or more	A

Carbon dioxide emissions

5.219 The carbon dioxide (CO_2) emissions are calculated using the same SAP document and method as for the SAP rating except that it uses CO_2 emissions factors for each fuel in place of unit prices to derive the CO_2 emissions rate per m^2 of floor area. The same logarithmic scale as used for SAP converts the CO_2 emissions rate into the Environmental Impact Rating (EIR), which also runs on a 1–100 scale where 1 represents very high emissions per m^2 and 100 is achieved at zero net emissions. The EIR can rise above 100 if the dwelling is a net exporter of energy.

Comparison with actual energy data

5.220 The SAP methodology that is used to calculate both energy efficiency and CO_2 emissions tends to provide higher estimates of energy requirements and associated emissions for heating, lighting and ventilating dwellings than estimates derived from actual household energy consumption. This is primarily because the assumed heating regime (achieving a temperature of 21°C in the living area of the dwelling and 18°C in the rest of the dwelling for a standard number of hours), and the assumed hot water and lighting requirements (depending on a level of occupancy determined by the floor area of the home rather than actual occupancy) are more likely to result in an

overall over estimation than under estimation of actual energy consumption for most dwellings. However, such standardised assumptions are necessary in order to compare the energy performance of one part of the housing stock with another and over time.

Energy performance certificates (EPC)

Energy performance certificate (EPC) improvement measures

5.221 Following the implementation of the European Energy Performance of Buildings Directive in 2007, all homes are required to have an Energy Performance Certificate (EPC) when they are sold or let. The EPC provides an overall assessment of the current energy performance of the property and makes recommendations regarding a range of lower and higher cost heating, insulation and lighting upgrades and electricity generation technologies that would improve its energy performance. The EHS is able to provide a whole stock assessment of homes that could benefit from these measures.

5.222 SAP was updated to SAP 2012, version 9.92, in October 2013. In June 2014, Appendix S and Appendix T of the SAP booklet were updated to RdSAP as part of SAP 2012. The revisions significantly altered the way that RdSAP software implements improvements as part of the EPC production process. The methodology for assessing the EPC improvement measures using the EHS data was therefore substantially revised in light of these revisions, with the new methodology applied to the 2015 EHS data onwards. The change in methodology means that results for pre-2015 data should not be directly compared. Revisions to the RdSAP methodology as part of the version 9.93 update have not affected the improvement modelling process for the EHS data, other than to inform the U-values that are applied to dwellings where cavity or solid wall insulation is applicable.

5.223 Details of the upgrade measures recommended on an EPC are provided in Appendix T of the SAP 2012 specification, available at: https://www.bregroup.com/wp-content/uploads/2019/09/RdSAP_2012_9.94-20-09-2019.pdf. Compared to the previous methodology (last applied to the 2014 EHS data), measures are no longer categorised as low cost, higher cost and further measures and the order for considering recommendations has been updated. The most relevant being that Measure Q (solid wall insulation) is now third priority, and measures R, S and T (upgrading boilers with/without fuel switching) have moved up in priority. Seven additional improvement measures have been added to the specification and the criteria and/or improvement specification has changed for some existing measures.

5.224 From 2019 onwards, the EHS form allows for park homes to be identified as a distinct dwelling type and improvement measures were updated to ensure these dwellings were improved following SAP assumptions. When applying

insulation improvements to park homes the following calculation is used to determine the improved U-value, as outlined in Appendix S of the SAP 2012 specification, where R_{ins} is the thermal resistance and is provided in Appendix T for all relevant measures

$$U_{insulated} = \frac{1}{\frac{1}{U_{existing}} + R_{ins}}$$

Annex Table 5.6.2: List of improvements specified in SAP 2012 appendix T that are included in the updated methodology

Item	Measure	Calculated in previous methodology	Any change to measure since previous version?
A	Loft insulation	Yes	Improve to 270mm (previously 250mm) For a park home use R_{ins} = 1.5m²K/W
A2	Flat roof insulation	No	For a park home use R_{ins} = 1.5m²K/W
A3	Roof room insulation	No	
B	Cavity wall insulation	Yes	
Q	Solid wall insulation	No	For a park home use R_{ins} = 2.0 m²K/W
W1	Floor insulation (suspended floor)	No	For a park home use R_{ins} = 1.5 m²K/W
W2	Floor insulation (solid floor)	No	
C	HW cylinder insulation	Yes	
D	Draught proofing	No	
E	Low energy lighting	No	
F	Cylinder thermostat	Yes	
G	Heating controls for wet Central Heating system	Yes	TRVs without room thermostat can be also without programmer
H	Heating controls for warm air system	Yes	
J	Biomass boiler	Yes	
K	Biomass room heater with boiler	Yes	Water cylinder upgraded in addition to heating system

I	Upgrade boiler, same fuel	Yes	
R	Install condensing oil boiler	No	
S	Condensing gas boiler no fuel switch	No	
T	Condensing gas boiler fuel switch	No	
L2	Replacement/New storage heaters	Yes	Change from fan assisted with automatic charge control to high heat retention type. Electric secondary heating no longer included
M	Replacement warm air unit	Yes	Split by fuel type, applies to non-condensing, mains gas units, and LPG units installed before 1998
N	Solar water heating	No	
O	Double glazing	No	
O3	Glazing replacement	No	
X	Insulated doors	No	
U	Photovoltaics	No	
V2	Wind turbine	No	

5.225 In the method used in the EHS, measures are only recommended for implementation if that measure alone would result in the SAP rating increasing by at least 0.95 SAP points. The suggested measures do not necessarily imply that current measures in place in the home are defective or that the home is deficient in terms of any particular standard.

5.226 The calculation of Improvements N (solar hot water) and U (photovoltaics) deviate slightly from Appendix T in that additional data collected in the EHS on dwelling roof pitch and orientation are used.

5.227 The EHS does not include EPC measure T2 (flue gas heat recovery) or Y (waste water heat recovery) because the survey is unable to assess how effective they would be in improving the performance of individual dwellings. Measure B4 (party wall insulation) is also excluded, as the EHS does not provide enough information to assess whether the dwelling has satisfied the installation condition.

5.228 The SAP 2012 Appendix T also includes some 'Alternative measures' which are measures shown on an EPC if relevant but are never selected for inclusion unless an assessor specifically includes it. These alternative

measures (Q2, J2, Z1, Z2 and Z3) are not included in this update to the EHS EPC modelling. Measure P is considered in the same way as the "alternative" measures because it is only considered if the assessor de-selects measure O.

Notional costs of installing the recommended EPC measures

5.229   The EHS estimates the notional costs of installing the recommended measures. The methodology for estimating these costs has also been revised from 2015 onwards and are now calculated using indicative costs that have been produced and included within the Product Characteristics Database (PCDB). Indicative costs for EPC measures are provided in Table 181 of the PCDB (available at http://www.boilers.org.uk/download.php). For the majority of measures, a low and high range of costs is provided and the mean of these is used. The costs are also calculated using the low and high prices to give an indication of the impact that the range has on the overall cost of improving the stock. Measure E (low energy lighting), L2 (storage radiators) and X (insulated doors) have an associated cost per item dependant on a variable e.g. number of rooms for L2. The costs for these measures are therefore specific to each dwelling. The cost of measure U was revised down in May 2019 to better reflect the reduction in cost associated with installing photovoltaics.

Annex Table 5.6.3: Indicative costs of the EPC measures

Item	Measure	PCDB Indicative cost (£) (mean, (low, high))	
A	Loft insulation	£225	(£100, £350)
A2	Flat roof insulation	£1,175	(£850, £1,500)
A3	Roof room insulation	£2,100	(£1,500, £2,700)
B	Cavity wall insulation	£1,000	(£500, £1,500)
Q	Solid wall insulation	£9,000	(£4,000, £14,000)
W1	Floor insulation (suspended floor)	£1,000	(£800, £1200)
W2	Floor insulation (solid floor)	£5,000	(£4,000, £6,000)
C	HW cylinder insulation	£23	(£15, £30)
D	Draught proofing	£100	(£80, £120)
E	Low energy lighting	£5 per unit	(£5, £5) per unit
F	Cylinder thermostat	£300	(£200, £400)
G	Heating controls for wet Central Heating system	£400	(£350, £450)
H	Heating controls for warm air system	£400	(£350, £450)

J	Biomass boiler	£10,000	(£7,000, £13,000)
K	Biomass room heater with boiler	£10,000	(£7,000, £13,000)
I	Upgrade boiler, same fuel	£2,600	(£2,200, £3,000)
R	Install condensing oil boiler	£5,000	(£3,000, £7,000)
S	Condensing gas boiler no fuel switch	£5,000	(£3,000, £7,000)
T	Condensing gas boiler fuel switch	£5,000	(£3,000, £7,000)
L2	Replacement/New storage heaters	£500 per unit	(£400, £600) per unit
M	Replacement warm air unit	£1,875	(£1,250, £2,500)
N	Solar water heating	£5,000	(£4,000, £6,000)
O	Double glazing	£4,900	(£3,300, £6,500)
O3	Glazing replacement	£1,200	(£1,000, £1,400)
X	Insulated doors	£500 per unit	(£500, £500) per unit
U	Photovoltaics	£4,500	(£3,500, £5,500)
V2	Wind turbine	£20,000	(£15,000, £25,000)

Pre- and Post-improvement performance and costs

5.230   The EHS also estimates the total carbon dioxide emissions, primary energy use and fuel costs by end-use before and after any recommended improvements have been installed. The EPC methodology has been updated so that for running costs and savings, total emissions and primary energy, the calculations are done using regional weather and fuel prices taken from the SAP fuel price library for the July of the survey year (see file 'RdSAP fuel prices', available to download at https://www.bre.co.uk/sap2012/page.jsp?id=2759)

5.231   The SAP rating and Environmental Impact rating are still calculated using UK average climate data and the fuel prices or emissions factors published in Table 12 of the SAP 2012 document. The Environmental Impact (EI) rating is a representation of carbon dioxide emissions attributable to a dwelling. Similar to the SAP rating, the EI rating scale goes from 1 – 100, where 100 represents net-zero emissions. It can rise above 100 if the dwelling is a net exporter of energy.

5.232   It is also important to emphasise that these are *notional* estimates based on standard assumptions about occupancy and consumption patterns. What improvements would be realised in practice will depend critically on actual occupancy and consumption patterns.

Cost of improving to EER band C

5.233 Since 2019, analysis has been undertaken to determine the cost to improve dwellings with an EER band of D or lower to an EER band of C. For each of the dwellings identified as having an EER band of D or lower, improvement measures are simulated cumulatively following Appendix T of the SAP 2012 specification. After each improvement, the SAP rating is recalculated until the dwelling reaches the threshold for EER band C (a SAP rating of 68.5 or higher).

5.234 A dwelling's EER may also be improved beyond the target band in certain cases. This may occur where a dwelling's energy efficiency rating is already close to the band C threshold, and a measure with a high SAP improvement yield is installed (e.g. solid wall insulation), pushing the dwelling beyond band C and across the band B threshold.

5.235 A proportion of dwellings are not able to be improved to a EER band of C after applying the eligible improvement measures outlined in Appendix T. This occurs in two instances; where the dwelling qualifies for certain improvement measures but is not able to be improved enough to reach the band C threshold, and where the dwelling either does not qualify for any of the improvement measures outlined in Appendix T, or each measure applied does not make a greater than 0.95 SAP point improvement to the dwelling. In this case, the dwelling does not receive the measure or have the associated notional costs applied. That is not to say that it is impossible for these dwellings to reach a band C, just that measures included in the EPC modelling process are insufficient to reach the target EER band.

Barriers to improving insulation

5.236 EHS energy reports have often examined the potential to install loft, cavity wall and solid wall insulation, and explored the practical and other barriers to actual installation that can occur, in order to provide a more realistic indication of the potential for carrying out these improvements. Categories classifying the ease of installation or specific barriers for each insulation type were created from EHS physical data on dwelling fabric and shape.

Loft insulation

5.237 Categories for the ease of installing or topping up loft insulation were as follows. All lofts that were not designated as non-problematic were classed as hard to treat:

- Non-problematic: these were identified as potentially upgradeable under the EPC improvement measure analysis and in these cases installation would be straightforward with no barriers.

- More problematic: these were identified as potentially upgradeable under the EPC improvement measure analysis but where the loft was fully boarded across the joists, which would lead to extra work and expense.
- Room in roof: these cases may already have sufficient insulation installed when built or during the loft conversion, but if insulation is needed to be added between the rafters, very extensive work and considerable expense would be involved. For the EHS 2020 Energy Report onwards, the methodology was updated to use the EPC improvement modelling variable for Measure A3, so that rooms in the roof which were assessed as having sufficient insulation were no longer designated as hard to treat.
- Flat or shallow pitched roof: again, these cases may already have sufficient insulation installed when built but otherwise it is not feasible to install loft insulation as there is no access into the loft or no loft space. For the EHS 2020 Energy Report onwards, the methodology was updated to use the EPC improvement modelling variable for Measure A2, so that flat or shallow pitched roofs which were assessed as having sufficient insulation were no longer designated as hard to treat.

Cavity wall insulation

5.238 For the 2012 and 2016 EHS Energy Efficiency of English Housing Reports, the main classification and analysis for the 'fillability' of cavity walls aims, as far as possible, to provide a count of hard to treat cavity walls consistent with the Energy Companies Obligation (ECO) definition, although the EHS is unable to fully replicate this, https://www.ofgem.gov.uk/sites/default/files/docs/2014/05/eco_supplementary_guidance_on_hard-to-treat_cavity_wall_insulation_0.pdf

5.239 For the 2020 EHS Energy Report onwards, the same characteristics were continued to be used in the classification of hard to treat cavity walls, with a minor methodological update to use the wall thickness variable in the assessment of narrow cavities.

5.240 Categories for the 'fillability' of uninsulated cavity walls were created using information on the area of external wall finish as surveyed and other factors such as the presence of external features such as conservatories and the dwelling type. All cavity walls that were not designated as non-problematic were classed as hard to treat.

5.241 *Standard fillable:* With these cases, no compelling physical barrier to installation exists. These are typically houses with masonry cavity walls and masonry pointing or rendered finishes and no conservatory attached.

5.242 *Hard to treat cavity walls:* These are homes with cavity walls that could, in theory, be filled but which exhibit at least one of the following difficulties.

- They are in a building with 3 or more storeys, where each storey has cavity walls. The need for scaffolding to install insulation in these higher buildings would contribute to the complication and cost of improving these homes.
- The gap found in the cavity wall is found to be narrower than in standard walls, typically less than 50mm. Although an attempt could be made to insulate these homes by injecting foam, the limited cavity space may lead to an uneven spread of the insulating material, resulting in substandard thermal properties.
- The dwelling is of predominantly prefabricated concrete, metal or timber frame construction. Although more recent examples of these homes will have had insulation applied during construction, these are generally unsuitable for retrospective treatment. In the case of timber frame construction, the industry recommendation is not to inject insulation as this can hamper ventilation between the frame and the external wall that may lead to rot in the timber frame.
- The cavity wall includes one or both leaves formed of natural stone, or has an outer leaf finished predominantly with tiles or cladding. Natural stone can give an uneven cavity, causing difficulties when injecting insulation, whilst non-masonry finishes can also make the process more difficult.
- From 2015, the definition was widened to include cavity walls that were not suitable to insulate due to severe dwelling exposure to wind-driven rain, or where there was disrepair to the existing wall finish.

Solid wall insulation

5.243 Categories for the ease of installing external solid wall insulation were created using information on the area of external wall finish as surveyed and other factors such as the presence of external features such as conservatories, porches and bays and the dwelling type. All solid walls that were not designated as non-problematic were classed as hard to treat:

- Non-problematic: no serious barriers.
- Masonry-walled dwellings with attached conservatories or other features: these are otherwise non-problematic, but fixing the insulation round any projections like conservatories, porches or bays requires additional work and therefore additional expense.
- Dwellings with a predominant rendered finish: although dwellings with a rendered finish can be treated with external solid wall insulation, this may add to the costs of the work as the render may need to be removed, repaired or treated before the insulation can be installed.
- Dwellings with a predominant non-masonry wall finish: improving dwellings with wall finishes such as stone cladding, tile, timber or metal

panels would either add to the cost of the work or even preclude external solid wall insulation where the wall structure itself is stone or timber.

- Flats: if the dwelling is a flat, then this treatment can be problematic for two reasons. Firstly, there are likely to be issues related to dealing with multiple leaseholders (getting their agreement and financial contribution to the work). Secondly, the height of the module for high-rise flats would present significant complications in applying external solid wall insulation.

Annex 5.7 Dimensions

5.244 Once all EHS physical surveys have been submitted by the surveyors for the survey year, the data are converted into a series of raw physical survey SPSS files. The first complex model to run using the raw physical survey EHS data is the Quantities (or "Dimensions") model.

5.245 The purpose of this model is to take a series of observations and measurements recorded by the EHS surveyors carrying out the physical survey and convert those into all of the required dimensions e.g. floor area, external wall area, window area, roof area, etc. These dimensions are then used as inputs into a number of processes, including costing of repairs and energy modelling.

5.246 In the process of running the Dimensions model, BRE staff undertake a number of consistency and plausibility checks on the raw physical survey data. The purpose of these checks is to, firstly, detect and eliminate certain logical inconsistencies that could not be processed by the Dimensions model and, secondly, to identify highly implausible answers from the Dimensions model outputs which, if deemed necessary after detailed investigation, can lead to correction of the EHS raw physical survey data.

5.247 The raw EHS physical survey data are stored in an SPSS database format and, after edits due to HMO and data comparison validation are applied, the dataset is used as the input into the dimensions modelling process.

Raw physical file checks

Wall thickness

5.248 From 2013/14, the 'wall structure' section of the EHS form included a new variable 'Wall thickness (cm)' (Fexws1wt, Fexws2wt) for which surveyors were asked to measure the wall thickness for each wall type (eight types possible) recorded on the front and back views. This information is used within the Dimensions model in the calculation of the floor area, replacing the previous default wall thickness assumption of 0.2m.

5.249 A flag is created for cases with a warning for wall thickness, front view and back view; these are cases where the wall thickness is not typical of the wall type selected. Each case is investigated by looking at the physical survey form in detail in conjunction with the photos/surveyor comments and a note is made for each case of what is likely to have caused the warning on the form. Based upon the information gained, the action is decided upon for each case. This could be no action required or it could be that the physical survey data

look incorrect, either the wall thickness value or the way the surveyor has coded something as wall that should not be counted as wall. The spreadsheet is passed to a second analyst to quality assure the outcomes. The appropriate modifications are applied to the raw physical survey data. In 2021, 50 cases were flagged and investigated with 2 case subsequently being edited.

5.250 An average wall thickness is calculated, by multiplying the thickness of each wall type by its area (in tenths) and dividing by 10. The resulting values are rounded to the nearest whole number to maintain the level of accuracy present in the original data. The variable is used as an input into the Dimensions model for the conversion of dimensions from external measurements to internal measurements.

Levels

5.251 A series of checks, written in SPSS syntax, are performed in order to identify possible logical inconsistencies in the raw data associated with the number of storeys in the building and the floor occupied by the dwelling. If the inconsistencies were left unchanged it would result in problems in the running of the Dimensions model.

5.252 The EHS uses the British system of denoting floor levels, where the ground floor is designated G, then the next level 1, then 2, 3, etc. Therefore a dwelling with 3 storeys should only have rooms, flats, or measured levels on floors G, 1, and 2. The most common type of error relating to floors is one where the surveyor switches to the American designation and uses 1 to indicate the ground floor. Other types of error linked to floor levels include failure to identify an attic as habitable and failure to include the basement in the count of the number of storeys. The level checks for inconsistency include:

- a room on a level that does not exist (e.g. 3rd floor of a three storey module)
- a room on a level that is not part of the flat (e.g. room on the 3rd floor but flat on the 2nd floor)
- a measured level that is not part of the module (e.g. dimensions for 3rd floor when the dwelling only has three storeys)
- a flat on a level that does not exist (e.g. flat on the 3rd floor when the module only has three storeys)
- An additional part of the dwelling that starts on a higher storey than the main part of the dwelling

5.253 All cases that are identified as having inconsistencies are manually examined by inspecting the EHS physical survey form and surveyor photographs in order to determine why the error has occurred and how the data should be changed. The cause of the error and the actions required to resolve the issue are then documented. In 2021, 62 cases were flagged as having possible

inconsistencies. Of these, 40 were found to have an error and were therefore corrected.

5.254 When it is established that there is an error in the raw data, SPSS syntax is used to alter the required variable/s. The altered physical files are then saved in a new location. The levels checks are repeated in order to confirm that all issues have been resolved.

Dimensions calculations

5.255 The Dimensions calculation engine takes the altered physical files and uses the data to compute all the required dimensions e.g. floor area, external wall area, window area, roof area, etc. as outlined in flow chart, Figure 5.7.1.

5.256 The model begins by taking the surveyor measurements for the floors measured by the surveyor and then extrapolates from those the size of any unmeasured floors. Once a plan of each floor has been calculated, the ceiling heights can be used to calculate or extrapolate (depending on the location of the five measured rooms within the building) the overall height of each floor. This in turn can be used as the basis for an overall wall area.

5.257 Wall areas are complex as there is the potential for additions beyond the surveyor's initial measurements (through base walls, gables and the like) and reductions brought about by attachment to other properties. It is also necessary at this stage to consider the location of the additional part, which may affect the size of the dwelling envelope; the envelope in this context relates to the design and construction of the exterior of the dwelling and consists of its roof, sub floor, exterior doors, windows, and exterior walls.

5.258 With the areas of the walls calculated for each face of the dwelling the level of fenestration on each face is assessed by using the fenestration ratio (the proportion of the gross surface area of the external wall which is given over to windows, voids and wall) as recorded by the surveyor with reference to the exposed area of the face.

5.259 Measurements of rooms that are normally inspected as part of the full physical survey (kitchen, living room, bedroom, and bathroom) are used to calculate the area of partition walls, which in turn are used to calculate the total usable floor area of the dwelling (i.e. the area below the partition walls are subtracted from the total floor area to create the total usable floor area variable). In 2021-22, surveyors were unable to collect measurements for these internal rooms therefore default room floor areas for each dwelling type were calculated using average measurements from the EHS 2015/16 and 2017/18 combined year datasets. The methodology was tested using 2019 single year data and it was found that there was a minimal impact to the average usable floor area (an increase of just 0.06m2 compared with that

obtained by the full dimension modelling). Furthermore, from 2018 onwards measurements for bathrooms were recorded allowing for the calculation of bathroom dimensions. Prior to 2018 bathroom dimensions were assumed to be 4m². As surveyors were unable to inspect bathrooms in 2021, this assumption was reinstated.

5.260 The area of the roof is calculated by taking the area of each floor and comparing it to the floor above. Where the latter is smaller the difference is considered to be exposed roof. This is then increased to allow for an eaves overhang, and then the results are summed across all levels up to the top floor, creating a total roof area that will typically be larger than the footprint of the dwelling. The same piece of code is also used to look at the amount of eaves perimeter on each level and total these across all levels.

Figure 5.7.1: Dimension model process
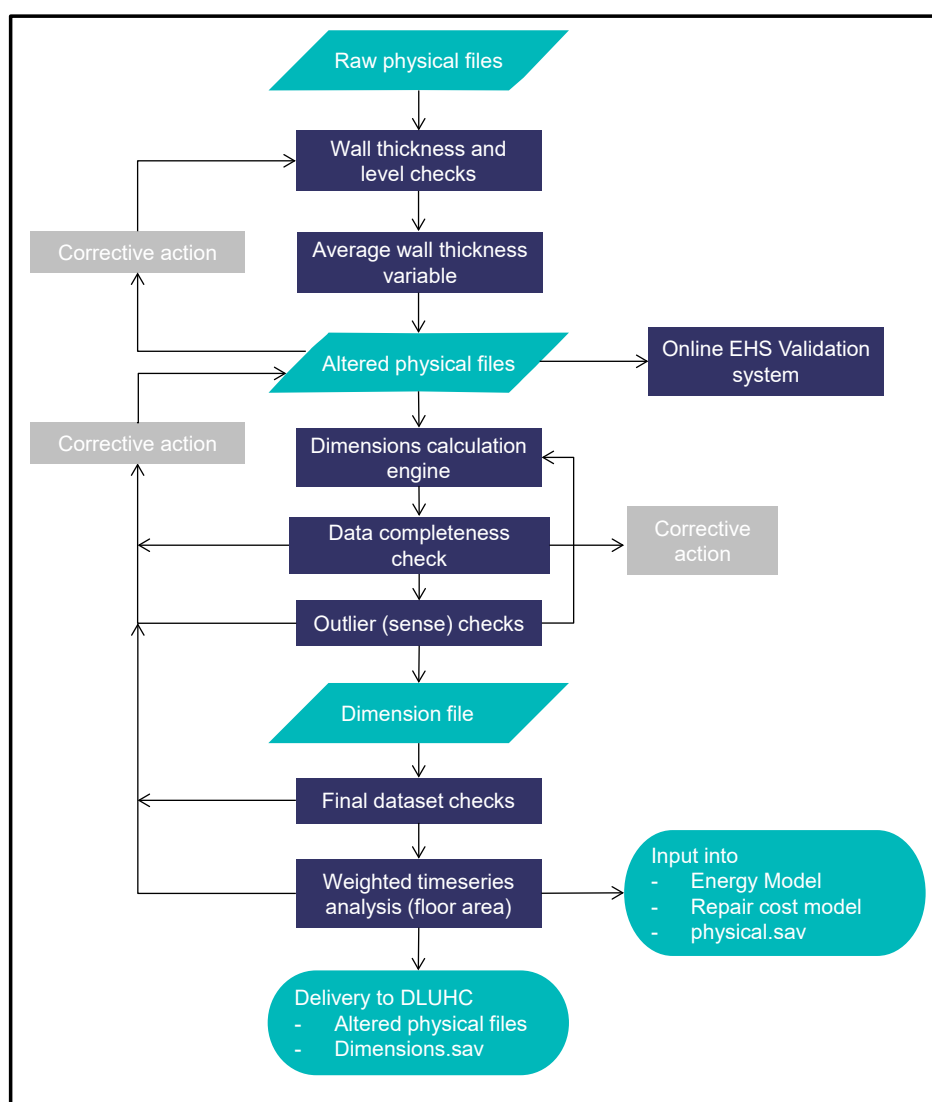
5.261 In the event that the dwelling is a flat, a number of additional algorithms are used to separate out the interior dimensions of the flat itself, and the exterior dimensions which relate to the "module". It is important for the energy modelling that dimensions relate to the flat itself, but for repair costs, the overall cost of repairs to the module must be shared among all dwellings within the envelope, and therefore external dimensions should apply to the module.

5.262 The preliminary results are stored in SPSS database format.

Data completeness checks

5.263 In previous years, some cases were not processed by the Dimensions model due to missing data for the flat floor level. In those instances, the missing information was added, the model rerun and a new validation check was added to the EHS online validation system. The validation check ensured that no cases were subject to this problem from 2016 onwards.

Outlier (sense) checks

5.264 A number of checks are then performed on the derived dimensions to validate the outputs. The checks are designed to identify:

    a) implausible wall areas

    b) implausible floor areas

    c) implausible window areas

    d) incorrect number of flats

    e) incorrect roof type

5.265 A statistical function written in SPSS syntax is used to identify outlying cases in terms of floor, wall and window area. Cases that have unusually large or small areas given the dwelling's size (in terms of the number of rooms) are flagged up. The survey form and photographs of each outlying case are manually examined to determine whether the outlier is genuine, or the result of an error. Errors can occur where key data have not been entered (such as an integral garage), dimensions mis-measured (such as entering 72 instead of 7.2) or features over- or underestimated. BRE defer to the surveyors judgement unless convinced the data are incorrect (i.e. the evidence from the photographs). If an error in the raw data is identified, the cause of the error and the actions required to resolve it are recorded. For 2021 around 420 cases were identified as outliers for floor, wall or window areas, and of these 105 cases were found to have an error of some kind.

5.266 For the check on the number of flats, the dimensions of the surveyed flat are checked against the total floor area of the module to determine whether the number of flats per module seems realistic. Where it does not, BRE staff will examine the floor plans and photographs in an attempt to understand the surveyor's error, which usually stems from the module definition, and to determine the correct number of flats.

5.267 Certain roof types (Chalet and Mansard) can only occur where the dwelling has an attic. On occasion surveyors may mistake steep pitched roofs for chalet roofs. In this situation, the data for pitched and chalet roofs is swapped over.

Alterations to physical data

5.268 In relation to the above checks, where there is missing information, BRE will attempt to fill in the missing data or otherwise resolve the error. For the applicable cases, the alterations syntax is updated to alter the required variable/s and the new altered physical files replace those created following the levels checks. If necessary, the EHS validation system may also be altered to help avoid similar issues occurring in future years. The Dimensions model is then rerun using the altered physical files.

Alterations to dimensions model code

5.269 Occasionally, the data completeness checks and the sense checks highlight amendments that are required to the Dimensions model code. Where necessary, the code is altered and tested to ensure it is working correctly.

5.270 The Dimensions model is rerun using the altered physical files and a series of first run 'final' dimensions files are created in SPSS database format.

5.271 The above procedure is repeated when the 'second run' of the EHS raw physical data is available (post-HMO edits, post-comparison edits, core EHS cases only) and once the dimensions file has been run through the first run of the energy model.

5.272 Basic checks are performed on the final derived dataset from the Dimensions model ('DimensionsXX.sav') using the previous year's dataset as a template e.g. a check that all variables are present and labelled correctly, that all variable values are present (via a frequency check) and that all missing values are set as missing following the EHS conventions.

5.273 On delivery of the single year paired sample grossing factor, timeseries analysis is carried out using the EHS floor area variables (floorx and floory) cross referenced by a number of dwelling characteristics to look at the trends

over time and to put the data in context with previous EHS data. Further investigation would take place should the weighted results highlight an unexpected change in the data.

5.274 The Dimensions modelling is complete. The 'DimensionsXX.sav' is used to derive the original EHS floor area variable (floorx) and the new Building Regulations floor area variable (floory) which are added to the derived physicalXX.sav dataset. The altered physical files are then delivered to DLUHC along with the 'DimensionsXX.sav' and these datasets are used internally as an input into the repair cost model and the energy model. Please see the data dictionary for 'DimensionsXX.sav' which explains more about the derivation of the individual variables.

Annex 5.8: Poor quality environments

5.275 Although analysis on local environments, based on data collected during the physical survey, is not included in the 2021 EHS annual reports this information is made available to assist users accessing the EHS datasets.

5.276 'Neighbourhood' or 'local environment' problems from the survey are based on the professional surveyors' assessments of problems in the immediate environment of the home on a scale of 1 ('no problems') to 5 ('major problems'). These assessments are based on observed problems (in some cases verified with the resident) rather than any specialised measurement instruments or recourse to other environmental data.

5.277 The survey assesses three types of problems contributing to a poor quality environment:

- Upkeep: the upkeep, management or misuse of the private and public space and buildings (specifically, the presence of: scruffy or neglected buildings; poor condition housing; graffiti; scruffy gardens or landscaping; litter, rubbish or dumping; vandalism; dog or other excrement; nuisance from street parking; condition of road/pavements and street furniture);
- Traffic and transport: road traffic and other forms of transport (specifically the presence of: intrusive motorways and main roads; railway or aircraft noise; heavy traffic; and ambient air quality);
- Utilisation: abandonment or non-residential use of property (specifically, vacant sites; vacant or boarded up buildings; intrusive industry; or non-conforming use of a residential area).

5.278 A home is regarded as having a significant problem of a given type if it is assessed to have codes 4 or 5 on the scale in respect of any of the specific environmental problems assessed and grouped under that type.

Chapter 6
Weighting

The English Housing Survey (EHS) is weighted to take account of the over-sampling of the less prevalent tenure groups and differential non-response, in order to provide unbiased national estimates. This chapter provides details of the weighting methodology and gives advice on which sets of weights to use when conducting analysis using EHS data.

Overview

6.1 The following weights have been calculated for the 2021-22 EHS data:

- Household weights for the full sample interview survey 2021-22 – aagfh21

- Average dwelling weights for cases in 2020-21 and 2021-22 that had both the interview and physical survey conducted (paired cases) [21] – aagpd2021

- Average household weights for cases in 2020-21 and 2021-22 that had both the interview and physical survey conducted (paired cases) – aagph2021

6.2 These weights adjust the sample to correct for the over-sampling of the less prevalent tenure groups and reduce the bias from differential non-response. The resulting weights sum to estimated population totals, enabling the survey to provide estimates of the total population of dwellings and households in England.

6.3 The weighting of the 2021-22 EHS data was undertaken by NatCen who managed the survey on behalf of the Department for Levelling Up, Housing & Communities (DLUHC).

---

[21] Cases which were eligible for PS (the paired sample) included addresses which the interviewers had determined to be vacant and where a physical survey had also been achieved. Vacant properties are usually identified by interviewers in field, based on the status of the property at their first visit. In 2020-21, it was not possible to identify vacant properties as interviewers did not visit the issued addresses in person. As a result, the dwelling weights for cases in 2020-21 and 2021-22 do not include vacant properties.

Application of weights during analysis

6.4   The EHS comes with its own weights covering the household and dwelling sample of cases. The weights for the household sample can be found in the file generalfs21.sav and is called agfh21. This should be used for any analysis for which the aim is to provide estimates of households, based on the data of the interview survey.

6.5   The weights covering the dwelling sample of cases can be found in the file general 20+21.sav. Aagpd2021 should be used for any analysis in which the aim is to provide estimates of dwellings and that includes physical survey data (e.g. percentage of non-decent dwellings) while aagph2021 should be used for any analysis in which the aim is to provide estimates of households and that includes interview survey data (e.g. percentage of households in non-decent dwellings). The weights can only be used on the full 2-year dataset. They cannot be used on the data split into separate years.

6.6   The recommended application of weights is summarised in Table 6.1.

Table 6.1: Application of weights during analysis

Weight	Description	Base
aagfh21	Household weights for the full sample interview survey 2021-22	9,752
Aagpd2021	Average dwelling weights for cases in 2020-21 and 2021-22 that had both the interview and physical survey conducted (paired cases)	10,572
Aagph2021	Average household weights for cases in 2020-21 and 2021-22 that had both the interview and physical survey conducted (paired cases)	10,572

Summary of weighting methodology

6.7   The weighting methodology[22] used a sequence of stages[23] described below. Some stages correct for the disproportionate sampling by tenure; others for the differential non-response in the interview survey (IS) and physical survey (PS) response process.

---

[22] The current weighting methodology was introduced in 2013-14 as a result of a weighting review which was carried out after the 2012-13 survey. The results of the review are published here: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/406771/EHS_2013-14_weighting_methodology_paper_FINAL.pdf.

[23] To avoid confusion with the numbering of stages under the previous methodology and for consistency with the weighting review report, the prefix "N" is used to describe the stages.

6.8   In 2021-22, face-to-face fieldwork was not possible because of restrictions introduced during the COVID pandemic. The EHS was therefore carried out over a reduced fieldwork period using a 'knock-to-nudge' approach and an 'external plus' physical survey.[24] Weighting was amended to support the new approach.

6.9   The following stages were applied to both the full household sample and the dwelling sample:

- Stage N1: Calculating the probability of selection for addresses

- Stage N2a: Adjusting for not-worked cases

- Stage N2: Filtering out office refusals

- Stage N3: Contact at IS

- Stage N4: Cooperation at IS

- Stage N5: Address-to-dwelling ratio

- Stage N6: Dwelling-to-household ratio

6.10   The following stage was applied to the full household sample:

- Stage N7: Calibration weighting for IS (full household weights)

6.11   The following stages were applied to the dwelling sample:

- Stage N8: Sub-sampling by tenure for PS

- Stage N9: Cooperation at PS

- Stage N10: Calibration weighting for PS

- Stage N11: Adjustment for new build (final paired dwelling weights)

- Stage N12: Creation of final paired household weights

6.12   Stages N1 to N6, N8 to N9 and N12 were implemented using a set of SPSS syntax scripts in combination with the specialist SPSS module AnswerTree.

---

[24] The 'knock-to-nudge' approach was a remote approach that allowed us to carry out the EHS while COVID restrictions were in place. EHS interviewers visited the sampled address to collect respondents' contact details and then conducted the interview over the telephone. Selected respondents were invited to take part in the 'external plus' physical survey. Surveyors did not enter the household but instead asked a small number of questions on the doorstep, carried out an external only survey of the property, and used supplementary information available from EPC and other data sources.

Stages N7, N10 and N11 used a combination of SPSS syntax and STATA's calibration command.

6.13   The weighting stages are described in more detail below.

Weighting the full household sample

6.14   Stage N1 generated the selection weights for the issued sample of addresses. This had two components: a grossing constant equal to the total number of delivery points on the Postcode Address File (PAF) divided by the number of delivery points sampled; and, to allow for the under-sampling of properties that were likely to be owner-occupiers, a correction factor to account for the unequal selection of addresses within each predicted tenure type (within each quarter). That correction factor was equal to the number predicted to be tenure X in the PAF divided by the number selected in predicted tenure X.

6.15   In previous years (when fieldwork was carried out face-to-face), a very small number of the addresses sampled were not issued to interviewers for operational reasons and these were filtered out at stage N2a by applying an adjustment to the Stage N1 weights so that the distribution of the weighted worked addresses by region matched that based on all the issued addresses. In 2021-22, a large number of the sampled addresses were not issued to interviewers because of challenges with fieldwork allocation during COVID-19. These addresses were filtered out at stage N2a by applying a CHAID model with the same predictors utilized for the Stage N2 Office Refusal model described below.

6.16   Non-response at the key fieldwork stages (refusal to co-operate prior to the interview, non-contact at the interview and refusal to co-operate at the interview) do not happen completely at random and the factors associated with each of these three processes may differ. Stages N2, N3, and N4 made a series of adjustments for non-participation. The adjustments were made using weighting classes (groups of cases) generated by the CHAID[25] algorithm of the SPSS AnswerTree software. The models generated from the algorithm were developed to identify the factors (or predictor variables) that were significantly associated with each of the three causes of non-response. These classes were derived at each stage using an unweighted CHAID model at address level which partitioned the sample of occupied dwellings based on predictor variables significantly associated with the propensity to respond at each of the key fieldwork stages above. Typical predictor variables for stage N2 included geographical area, predominant tenure, dwelling age and dwelling type in the area, and urban/rural classification; for subsequent stages, information collected by the interviewer was also used. Once the

[25] Chi-squared Automatic Interaction Detector

classes had been generated, the adjustment within each class was made based on the weighted estimates (using the combined weights up to that stage).

6.17　The EHS analyses are concerned with dwellings and households rather than addresses, and there is not always a one-to-one relationship between an address, a dwelling, and a household[26]. Usually there is only one dwelling at each address sampled from the PAF, but addresses are occasionally found to cover more than one dwelling (for example if a house has been converted into self-contained flats) or only part of a dwelling (for example a bedsit which shares facilities with a household at a separate postal address). As only one dwelling was selected at these addresses and one household was selected at dwellings that contain more than one household, weights were required.

6.18　Where an address refers to more/less than one dwelling, each dwelling at that address would have a lower/higher chance of selection. Similarly, when a dwelling contains more than one household, each household at that dwelling would have a lower chance of selection. Stages N5 and N6 corrected for the selection of one dwelling unit at the sampled address and one household at the selected dwelling by deriving address-to-dwelling and dwelling-to-household ratios designed to be applied with the weights developed further downstream to derive the final weight. The address-to-dwelling[27] and dwelling-to-household ratios are averaged (smoothed) within weighting (smoothing) classes defined by type of property (3 groups: house, low-rise flat, high-rise flat), tenure (4 groups) and region (9 groups) to remove large peaks and troughs within the classes.

6.19　The smoothing of the dwelling-to-household ratio is carried out over two years of the EHS to improve the stability of year-on-year estimates including household totals[28]. This is because the household totals are dependent on estimates of the "dwelling-to-household" ratio which is subject to sampling variation[29].

6.20　The previous stages account for the sampling and response probabilities. Applying the weights derived from the previous stages to the household-level

---

[26] For the purposes of the survey, a dwelling is defined as "a self-contained unit of accommodation where the occupants of that accommodation have sole use of all the rooms and facilities".

[27] Apart from occupied dwellings, the address-to-dwelling ratio was also computed for vacant dwellings as this was needed for the sub-sampling stage (stage N8).

[28] Because EHS 2020-21 was administered push-to-telephone there was no information on number of households at a dwelling (which would have been collected from the ARF/doorstep form), therefore it was not possible to use 2020-21 data when smoothing over 2 years; because of this EHS 2018-19 was used instead (which is the previous survey covering the same half of England as in 2020-21).

[29] Unlike the dwelling totals which are controlled to "external" DLUHC estimates of dwellings by tenure and region during the calibration stage N7. More details about the effect of the smoothing methodology in producing less variable dwelling-to-household ratios can be found in the 2015-16 Headline Report:
https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/595785/2015-16_EHS_Headline_Report.pdf

data would provide a preliminary survey estimate of the total number of households in England. However, this will differ from the true value because of sampling error, under-coverage of the frame and inability of the model-based process to allow for all the factors associated with non-response. Estimates for subgroups such as tenures will differ from their true values for the same reason. These differences in the survey estimates can be reduced by calibrating the weights.

6.21   Calibration takes an initial set of weights (e.g. the composite weights from stages N1 to N5) and then adjusts (or calibrates) it to given control totals. The control totals are usually a population count of a specific attribute or set of characteristic derived from a source external to the survey. The process generates weights which produces survey estimates that exactly match the population for the specific characteristics (the control totals) used in the adjustment.

6.22   The calibration of the interview sample was carried out on the part of the sample that contained occupied dwellings only. That section of the sample was first weighted by the composite weights from stages N1 to N5 and then calibrated to the control totals below.

- population proportions for age/sex based on ONS population projections (for dwellings with more than one household, the household counts were adjusted using the dwelling-to-household ratio from Stage N6);

- counts of occupied dwellings by tenure for each region (these control totals were estimated by adjusting the DLUHC dwelling counts, which include both occupied and vacant dwellings, using estimates of the proportion of dwellings that are occupied from the current and four previous years of the EHS).

The control totals were measured as of 1 October 2021.

6.23   The calibration weights from Stage N7 were combined with the dwelling-to-household ratio from stage N6 to give the weights for the full household sample.

Weighting the dwelling sample

6.24   A lot of analyses using the EHS data are carried out using the dwelling as the unit of analysis. These analyses usually make use of cases containing both the interview and physical survey data (or cases in the paired sample). Cases which were eligible for PS (the paired sample) included addresses which the interviewers had determined to be vacant and where a physical survey had

also been achieved. The inclusion of vacant addresses was thus necessary for the production of the final paired dwelling weights.

6.25   The weighting process started by using stages N1 to N5 as described above.

6.26   The PS sample included a disproportionately larger number of dwellings from renting tenure groups to enable detailed analysis of these. This was achieved by under-sampling of properties likely to be owner occupied (sub-sampling rates varied by quarter). Stage N8 calculated selection weights to take account of the under-sampling of owner-occupied properties.

6.27   Stage N9 adjusted for non-response to PS using weighting classes generated from a CHAID model designed to identify the factors significantly associated with non-response to the physical survey. Data collected during the interview survey were also used to help determine the weighting classes. Vacant cases were treated using a separate CHAID model because the process of participation for vacant dwellings is generally different from that for occupied dwellings and this needed to be reflected in the weights.

6.28   The occupied paired sample (i.e. cases with both IS and PS) was weighted by the composite weights from stages N1 to N5 and N8 to N9. The vacant sample (i.e. vacant addresses at PS) was weighted by the composite weights from stages N1, N2a, N5, N8, and N9. This gives the initial weights for dwellings.

6.29   At stage N10, both the occupied and vacant samples were then combined and calibrated to:

- population proportions for age/sex (for dwellings with more than one household, the household counts were adjusted using the dwelling-to-household ratio from Stage N6; for vacant dwellings the household counts were all zero);

- DLUHC counts of all dwellings (occupied + vacant) by tenure for each region;

- counts of vacant dwellings by tenure (RSL and LA collapsed) for each region (these control totals were estimated by adjusting the DLUHC dwelling counts using estimates of vacancy rates from the current and four previous years of the EHS).

6.30   The achieved sample of dwellings did not include any dwellings built since the sample was drawn; stage N11 adjusted the weights to allow for those new dwellings. Because of the small number of recently built dwellings in the survey, the weights of all cases with a construction date of 1990 onwards were weighted-up to cover for new addresses on the PAF by re-calibrating to the same overall control totals, but with re-estimated counts of post-1990

build. This was carried out separately for areas with a high/low rate of new build, and for private/social sector housing (excluding local authority housing because the rate of new building in this sector is negligible). This process generated the final paired sample dwelling weights.

6.31   At Stage N12, the final paired dwelling weights from Stage N11 (after removing the vacant dwellings) were adjusted using the dwelling-to-household ratio from stage N6 to generate the final paired sample household weights.

Calculating two year weights

6.32   Because of the smaller annual sample sizes involved, analysis of the dwelling sample is normally carried out using 2 years' weighted data. This section sets out how the combined weights were calculated.

6.33   The individual year (2020-21 and 2021-22) datasets with the dwelling weights after the PS cooperation stage N9 were merged together and two-year dwelling weights were created by dividing each year's weights by 2, so that each dataset has equal influence on the weighting process.[30]

6.34   The two-year dwelling weights were calibrated by repeating stages N10 and N11 of the annual weighting[31]. This process generated the final two-year paired sample dwelling weights.

6.35   Similarly to stage N12 of the annual weighting, the final two-year paired dwelling weights were adjusted using the dwelling-to-household ratio to generate the final two-year paired sample household weights.

---

[30] Vacant dwellings from 2021-22 were excluded from the two-year dwelling sample because vacant properties were not included in 2020-21.
[31] The reference date for the 2-year weights was 1st April 2021, therefore the estimated DLUHC dwelling counts used as control totals were adjusted accordingly.

Chapter 7
Standard errors

All survey estimates are subject to sampling error because they are derived from a sample of a population rather from the whole population. This chapter describes how the sampling errors of the key 2021-22 EHS variables were calculated. The sampling errors of the key 2021-22 EHS variables are available as a separate document on Gov.uk.[32]

Overview

7.1 The sampling errors associated with survey estimates are expressed in terms of standard errors which measure the uncertainty around survey estimates. Standard errors for the 2021-22 EHS were calculated on weighted data using the statistical package Stata. The variance of survey estimates is affected by features of the survey or sample design such as weighting, clustering and stratification. The effect of those features on the standard errors of the estimates from the 2021-22 EHS has been taken account of in the calculation of the standard errors.

Sources of error in surveys

7.2 The total error in a survey estimate is the difference between the estimate derived from the data collected and the true value for the population. Survey estimates are subject to various sources of error; the two main types of error are systematic and random error.

Systematic error

7.3 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 a certain section of the population is excluded from the sampling frame, because non-respondents to the survey have different characteristics to respondents, or if interviewers 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.

---

[32] The sampling errors of the key 2021-22 EHS variables are made available in the annex tables of this chapter on https://www.gov.uk/government/collections/english-housing-survey-technical-advice.

Random error

7.4 The most important component of random error is sampling error, which is the error that arises because the estimate is based on a random sample rather than a full census of the population. The results obtained for any single sample may, by chance, vary from the true values for the population but the variation would be expected to average to zero over a number of repeats of the survey. The extent of the variation from the population value depends on the size of the sample, the sample design and the weighting methodology.

7.5 Random errors may also result from other sources such as variations in respondents' interpretation of the questions, or variations in the way different interviewers ask the questions. The effect of these can be minimised through pilot work and interviewer training.

Standard errors for complex sample designs

7.6 Key features of the design of a survey such as weighting, clustering and stratification can have an impact on the standard errors. The effect of those complex design features has to be taken into account when assessing the reliability of estimates. Standard errors calculated on the basis of a simple random sample design will not reflect the true variation because the effect of the complex sample design has not been taken into account.

7.7 The sample design features of the EHS that has an impact on standard errors are the use of a two-stage stratified sample design to select a clustered sample and differential sampling probabilities. The methodology used to weight the data also affects the standard errors.

7.8 The two-stage sample selection process described in Chapter 1 can lead to an increase in standard error if the households or people within individual primary sampling units (PSUs) are relatively homogenous but the mix of households or people between PSUs differ from one another. As each year's EHS sample covers half of the PSUs in England, the loss in precision from clustering should be fairly small. In addition, half the PSUs in England are included in the survey in one year and the other half in the following year, so the sample combining two years of data is, in fact, unclustered. Stratification, on the other hand, tends to reduce standard error and is most advantageous where the stratification factor is related to the characteristics of interest on the survey.

7.9 In order to obtain a sufficiently large number of cases for detailed analysis, households in the rented sector are given a higher probability of selection compared with owner occupiers. The methodology used to weight the data to adjust for differential sampling probabilities will increase the standard errors.

On the other hand, the use of population totals to control for differential non-response tends to lead to a small reduction in sampling errors.

Design factors

7.10   The design factor, or deft, is the ratio of the standard error of an estimate from a complex sample to the standard error that would have resulted had the survey design been a simple random sample of the same size. The size of the design factor depends on the degree to which a characteristic is: clustered within PSUs, varies across the strata and is correlated with the weights. Design factors below 1.0 show that adopting the complex sample design has improved the reliability of the estimate compared with using a simple random sample design, probably due to the benefits of stratification. Design factors greater than 1.0 show that, due to the effects of clustering, weighting or other design features, the complex sample design has resulted in less reliable estimates than might be gained from a simple random sample.

7.11   The design factors for selected survey estimates are shown with the corresponding sampling errors in the tables released on Gov.uk. These can be used to estimate likely sampling errors for other variables on the basis of their similarity to one of the variables presented. As shown in the formulas below, the standard error (se) of a proportion (p) based on a simple random sample (srs) multiplied by the deft gives the standard error of a complex design.

$$se(p) = deft \times se(p)_{srs}$$

where:

$$se(p)_{srs} = \sqrt{\frac{p \times (100 - p)}{n}} \quad {}_{33}$$

7.12   The formula to calculate the standard error of the difference between two percentages for a complex sample design is:

$$se(p_1 - p_2) = \sqrt{\frac{deft_1^2 \times (p_1 \times (100 - p_1))}{n_1} + \frac{deft_2^2 \times (p_2 \times (100 - p_2))}{n_2}}$$

Where p_1 and p_2 are observed percentages for the two subsamples and n_1 and n_2 are the subsample sizes.

---

[33] The precise formula uses *n-1* as the denominator but this equates to *n* in large samples.

Confidence intervals

7.13 Although an estimate 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. The standard error can be estimated from the values obtained from the sample and allows the calculation of confidence intervals, which indicate the range of random variation in the survey estimates.

7.14 It is common, when quoting confidence intervals, to refer to the 95% confidence interval around a survey estimate. This is calculated at 1.96 times the standard error on either side of the estimated percentage or mean since, under a normal distribution, 95% of values lie within 1.96 standard errors of the mean value. If it were possible to repeat the survey under the same conditions many times, 95% of these confidence intervals would contain the population values.

7.15 The 95% confidence interval for the difference between two percentages is given by:

$$p_1 - p_2 \pm 1.96 \times se(p_1 - p_2)$$

7.16 If this confidence interval includes zero then the hypothesis that the two proportions are the same and the observed difference is due to chance alone is not rejected. If the interval does not include zero then it is unlikely (less than five per cent probability) that the observed difference could have occurred by chance and this constitutes a 'significant difference' at the 95% confidence level.

Sampling errors of the 2021-22 EHS estimates

7.17 The sampling errors of the key 2021-22 EHS variables are released on Gov.uk. Weighted data were used to calculate the standard errors so that the values of the percentages and means were the same as those in the published EHS reports. The method used to calculate the sampling errors correctly allows for the inflation in the sampling errors caused by the sample design and weighting methodology except for the positive effects of using population totals to control for differential nonresponse. As a result, the standard errors and design factors (defts) presented in the tables released on Gov.uk are likely to be slight over-estimates.

Chapter 8
Impact of COVID-19 on the English Housing Survey

8.1   The 2021-22 English Housing Survey data was collected during the COVID-19 pandemic. This chapter summarises the impact of COVID-19 on the English Housing Survey and sets out the changes that were made to the survey as a result. It also summarises the work undertaken to understand the impact of these changes on EHS data.

Overview

8.2   COVID-19 necessitated changes in the established English Housing Survey methodology. Face-to-face interviews were replaced with a 'knock-to-nudge' approach and the full internal and external inspection of dwellings was replaced by an 'external plus' approach where the inspection was restricted to an assessment of the exterior of the dwelling, supplemented by information about the interior of the dwelling the surveyor collected (socially distanced) at the doorstep. Compared with a typical face-to-face year, response rates were lower and there was a smaller sample size in 2021-22 with 9,752 interviews (down from around 13,300 in a usual year) and 5,478 physical surveys (down from 6,200 in a usual year).

Fieldwork

8.3   To conform with the social distancing restrictions, EHS interviews and physical surveys were carried out using approaches which minimise social contact with respondents. The interview survey used a 'knock-to-nudge' approach where an introductory letter was sent out followed by a socially distanced visit from an interviewer, and a telephone interview. The interviewer role was to collect interviewer observations, to make an appointment to carry out the EHS interview by telephone and to provide additional information about the survey if needed. Interviewers remained outside the home.. The physical survey adopted an 'external plus' approach where the visual inspection was restricted to an assessment of the exterior of the dwelling, supplemented by information about the interior of the dwelling the surveyor collected (socially distanced) at the doorstep.

8.4   Sampling was changed to support the new approach, as outlined in Chapter 1 of this report.

Response rates

8.5   Response rates were lower than for the standard face-to-face approach. For the interview survey, we achieved a response rate of 33% However, the knock-to-nudge approach led to a better response rate and sample profile than the push-to-telephone approach adopted in 2020-21: this was largely due to the return of interviewer visits to issued addresses (see EHS 2020-21 technical report for details).[34] Response to the physical survey remained high, at 74%, as the physical survey follows the interview survey.

Achieved sample

8.6   The change in survey approach makes it difficult to establish whether the changes in housing circumstances observed in the 2021-22 English Housing Survey compared to previous years were the consequence of real change, a function of the change in survey mode and response bias or both. For this reason, caution is advised when carrying out analysis and interpreting observed changes against findings from pre pandemic surveys.

Reporting

8.7   As the reliability of the results of sample surveys such as the English Housing Survey, is positively related to the unweighted sample size, the smaller sample limited reporting in 2021-22. Notably, there was increase in apparent changes that were not statistically significant; and there was a decrease in the amount of sub-group analysis that we could do (i.e., because the sample was too small to produce reliable estimates).

8.8   In addition, many data items on the condition of the interior of dwellings were not collected so detailed derived variables relating to the Housing Health and Safety Rating System, actual and standardised repairs costs are not available. For the EHS headline and annual reports, web tables and archived data, dwelling level modelled estimates were developed to fill in some key data gaps on the Decent Homes Standard; the cost of making a non-decent home meet the Standard; HHSRS Category 1 hazards; and damp. For the headline report, statistics on the presence of carbon monoxide detectors were extrapolated from previous EHS trends.

8.9   A full account of the modelling undertaken is provided in Chapter 5 of this report.

---

[34] English Housing Survey 2020 to 2021: technical report - GOV.UK (www.gov.uk)