Energy Follow Up Survey 2017

Methodology report

Department for
Business, Energy
& Industrial Strategy

bre

The EFUS 2017 has been undertaken by BRE on behalf of the Department for Business, Energy & Industrial Strategy (BEIS).

Project directed by: Matt Custard and John Riley

Authors and analysts: Justin Allen, Adele Beaumont, Matt Custard, Helen Foster, Helen Garrett, Andrew Gemmell, Sami Kamal, Susie Margoles, Emma Munkley, and Felicity Taylor.

Additional thanks to the wider team of reviewers and contributors at BRE, BEIS and elsewhere, including GfK, Ipsos MORI, NatCen, Gemini Data Loggers and Technolog Ltd.

Table of Contents

1 Introduction

There is an ongoing requirement to keep our knowledge and understanding of domestic energy use up to date. This is essential to ensure that policies, and policy interventions, are directed in the most efficient and effective manner; that legislation and standards are based on principles and assumptions that reflect how people are actually using energy in their homes; and that models and statistics which provide the underpinning evidence base in this area are as accurate as possible. Of particular relevance at the moment are policies relating to fuel poverty, decarbonisation of heat, smart metering and minimising household energy bills.

The Energy Follow-Up Survey 2017 (EFUS 2017) was a follow-up survey of a sample of respondents from the 2014/15 to 2016/17 English Housing Surveys. In comparison to the previous EFUS 2011, the EFUS 2017 provided more detailed data on energy consumption, and the use of heating, hot water and appliances. In addition, there was a greater focus on the energy consumption and behaviours of fuel poor households.

The Department of Business, Energy and Industrial Strategy (BEIS) had several overarching aims which needed to be addressed by the new EFUS 2017. These were:

1. To determine current domestic energy consumption and heating patterns in England and to investigate how they change over time through time-series comparisons

2. To understand how and why there are variations in energy consumption between similar dwellings, and similarities in energy consumption between different dwellings

3. To understand how households in fuel poverty use energy and how their energy consumption patterns and behaviours compare with non-fuel poor households

The EFUS 2017 comprised three interview surveys and a short text message survey during a particularly hot spell. In addition, a sub-sample of households was selected to have temperature loggers, gas and electricity monitors, and humidity loggers installed.

This report outlines the survey methodology used in the EFUS 2017 including sampling, data collection, data quality, weighting, and data analysis. Figure 1.1 provides a timeline of the EFUS data collection.

Figure 1.1: EFUS 2017 data collection timeline

STAGE	Dates (d/m/y)
Interview 1	
Pilot	05/05/17 - 05/06/17
Face-to-Face	03/08/17 - 31/10/17
Online	25/11/17 - 02/01/18
Interview 2	
Pilot	21/11/17 - 07/12/17
Telephone	17/01/18 - 05/03/18
Online	09/03/18 - 07/04/18
Interview 3	
Pilot	23/11/18 - 29/12/18
Telephone	18/02/19 - 31/03/19
Online	31/01/19 - 25/03/19
Summer Text Survey	
Text Survey	02/07/18 - 08/07/18
Temperature & Humidity	
Monitoring	01/10/17 - 30/04/19
Hottest Week	23/07/18 - 29/07/18
Coldest Week	28/01/19 - 03/02/19
Meter Point Data	
Electricity	31/01/17 - 30/01/18
Gas	06/17 - 06/18 [1]
Detailed Energy Consumption	
Installations - pilot	17/07/17 - 14/08/17
Installations - main stage	22/01/18 - 26/10/18
Monitoring	01/05/18 - 30/04/19

Bars are inclusive of month

[1] **Exact dates are not available for the gas meter point data**

1.1 The English Housing Survey and the Energy Follow Up Survey

The English Housing Survey (EHS) is a national survey commissioned by the Ministry of Housing, Communities and Local Government (MHCLG) that collects information about people's housing circumstances and the condition and energy efficiency of housing in England. It has a complex multi-stage methodology consisting of two main elements: an initial interview survey of around 13,300 households with a follow up physical inspection of a sub-sample of approximately 6,200 of these dwellings, including vacant dwellings.

For more information regarding the EHS see:

https://www.gov.uk/government/collections/english-housing-survey

The EFUS 2017 project was a collaboration between BEIS and: BRE; GfK; NatCen; Loughborough University and Ipsos MORI. BRE was the lead partner and project manager, overseeing all aspects of the data collection, analysis and reporting from householder surveys, as well as temperature and energy monitoring. Research partners at NatCen, GFK and Ipsos MORI conducted the interviews with householders and Loughborough University analysed the summer temperature and thermal comfort data, as well as information on the use of smart appliances. A summary of the roles undertaken by each partner is given in Appendix A.

Households who participated in the EFUS 2017 were sampled from the 2014/15 to 2016/17 English Housing Surveys. The EFUS 2017 used a subsample of households who had received both the physical and interview survey. During the EHS interview, households were asked if they would be willing to participate in a further MHCLG or BEIS study. Those who agreed were selected as potential households for the EFUS.

For each household participating in the EFUS 2017, the EHS interview and physical survey data was then supplemented with the data collected specifically for EFUS 2017 via:

- Three interview surveys
- Summertime text survey
- Internal temperature data loggers
- Monitored electricity and gas consumption data
- Metered electricity and gas consumption data
- Humidity data loggers
- External temperature data from the Met Office MIDAS dataset

1.2 EFUS 2017 reports

For the EFUS 2017, there are five main reports:

- Heating Patterns and Occupancy report
- Thermal Comfort, Ventilation, Damp and Mould report
- Household Energy Consumption and Affordability report
- Lights, Appliances and Smart Technologies report
- Fuel Poverty report

Each report includes details of the key research questions examined within it, a summary of the methodology undertaken, an executive summary of the findings, and conclusions of the analysis.

2 Household interview survey methodology

2.1 Questionnaire design

Three interview questionnaires were undertaken for the EFUS 2017, each covering a number of topics. Some of these topics were covered in two interviews to allow for longitudinal analysis. A summary of the topics covered is provided in Table 2.1.

One of the key objectives of the project was to better understand how householders heat their homes and use energy during the heating season. It was considered essential that occupants were asked about their heating behaviours at a point in time that their heating was being regularly used. The timings of the three interview surveys were designed to ensure that householders were able to recall the relevant information and were reporting on their behaviours and experiences at the time of the interview itself or the preceding weeks.

Table 2.1: Summary of topics in each interview survey

Interview 1	Interview 2	Interview 3
Summer thermal comfort	Use of main, alternative and supplementary heating systems	Use of main heating systems including the heating season
Cooling behaviours	Winter thermal comfort	Proportion of the house heated
Hot water use	Winter ventilation behaviours	Occupancy patterns
Appliance ownership	Damp and mould	Smart technologies
Lighting	Winter appliance and hot water use	Method of payment and tariffs
Method of payment and tariffs	Lighting	Dwelling improvements and changes to the household
Other fuels	Trade-offs made by the households unable to afford to heat their homes	
Dwelling improvements and changes to the household	Occupancy patterns	

2.2 Interviewer briefings

All interviewers attended a face-to-face briefing prior to starting field work for each interview survey. Separate briefings were conducted for the pilot and main stage interviews for each survey. Additionally, debriefing sessions were carried out after each pilot study. The briefings were led by a GfK NOP researcher and BRE's field trial managers, with active participation throughout from BRE staff providing expert input into interview topics as required.

The briefing began with an outline of the aims of the survey, and a brief explanation of the sample. Instructions were given on handling the sample, including strategies for dealing with such a dispersed sample and techniques to maximise response. Interviewers were given a broad overview of the coverage of the questionnaire and an explanation of any technical terms that were helpful for interviewers to understand and which they may have needed to explain to respondents, as well as instructions as to how to access the additional support information. Interviewers carried out a dummy interview using the CAPI (Computer Assisted Personal Interviewing) system, with each interviewer taking it in turn to ask a few questions; all had the chance to raise queries at any time. At the end of the dummy interview BRE staff explained the purpose of the temperature loggers and the process of installing them.

2.3 Pilot studies

For each of the three interview surveys, an EFUS pilot questionnaire was undertaken with households. The purpose of each pilot study was as follows:

- To provide an accurate assessment of the length of the survey
- To identify any problematic questions or sections
- To identify any problematic words or phrases
- To assess the pre-coded response list and identify anything that was missing
- To identify questions or sections that worked well
- To gather any other useful feedback e.g. potential difficulties around the installations of the temperature loggers (Interview 1)

Households tested each of the interviews pilot surveys two to three months ahead of the main data collection phase. For Interview 1, a pilot survey of 94 households was carried out between May and June 2017. In addition to the face-to-face interviews, 41 of the pilot households had temperature monitors installed in up to five rooms in their homes to test the installation process. For Interview 2, a pilot of the telephone interview was carried out with 35 households between November and December 2017. For Interview 3, a pilot survey of 63 households was carried out between November and December 2018. Of these, 20 surveys were completed online and 43 were completed by telephone. As well as testing both the online and telephone versions of the survey, the pilot also tested the order of presentation (i.e. online first with a telephone follow up).

Feedback on each pilot survey was collated and assessed by the survey development team (BRE, GfK, NatCen and BEIS) through a series of project development meetings and discussions. Feedback was obtained through four main mechanisms:

- **Recording notes**. Interviewers actively recorded notes for enhancing the survey in the 'notes' field which was part of the CAPI software.
- **Specific analysis of the pilot survey data**. On receipt of the pilot survey data, BRE examined frequency distributions of the variables to determine the quantity and quality of the data being collected and where possible, compared the outputs of key household descriptor variables with other established surveys to establish that the responses were within the bounds of expected values.
- **Interviewer debriefing**. Immediately following the fieldwork period, a full debrief was held with all pilot survey interviewers. This involved assessing the pilot questionnaire question-by-question and gathering comments from the interviewers directly. Comments were made

on individual questions, with suggestions from both respondents and interviewers for ways to improve the questionnaire for the main stage of the survey. The session was recorded to ensure that the interviewer's observations were fully captured.

- **Questionnaire review**. The pilot survey was then reviewed by the project steering group. This included BRE's social research team and the EFUS 2017 project team at BEIS.

2.4 Sampling strategy and response rates

The objective of the EFUS 2017 sample design was to deliver a representative sample of at least 2,300 households at the first wave of the fieldwork. At least 800 fuel poor households were required, which necessitated an over-sampling of fuel poor households above their national prevalence of approximately 11% of all households. In order to achieve this objective, two slightly different sampling frames were used: the first being a nationally representative sampling frame and the second a 'fuel poverty boost' sample (referred to going forward as "FP boost"). The nationally representative sample was taken from EHS 2016/17. The additional fuel poverty cases were sampled from the fuel poor households who took part between EHS 2014/15 and 2015/16.

2.4.1 Interview survey 1

The first of the householder surveys was undertaken in the autumn of 2017 and is referred to as 'Interview 1'. This survey was conducted via a face-to-face interview conducted in the householders' homes between August and October 2017. In order to boost the sample, an online version of the same survey was completed by a further 671 households between November 2017 and January 2018, giving a total sample of 2,632.

The EFUS 2017 core sample consisted of the 4,950 households from the EHS 2016/17 that had agreed to be re-contacted. The FP boost sample consisted of 1,265 households from the EHS 2014/15 and EHS 2015/16 datasets that were flagged (at the time of sampling) as being in fuel poverty and who had agreed to be re-contacted; all households in the FP boost sample were contacted face-to-face.

In total 6,215 households were issued to the interviewing team. Productive interviews were achieved with 2,632 of them (2,058 in the core and 574 in the FP boost respectively). Just under half of the productive households (1,020) consented to the installation of temperature monitors. Only households who took part in the face-to-face survey were asked if temperature monitors could be installed.

2.4.2 Interview survey 2

The second of the householder surveys, a follow-up survey to Interview 1, was conducted between January and April 2018 and is referred to as Interview 2. All households from the Interview 1 sample, who had given their permission to be re-contacted for subsequent surveys, formed the sample population for Interview 2. In total 2,145 households agreed to be re-contacted (87% of the face-to-face sample and 77% of the online sample). To minimise disruption to the householders the survey was conducted via a telephone interview and 1,060 households completed the telephone survey. As with Interview 1, in order to boost the sample an online version of the Interview 2 survey was completed by a further 280 households giving a total sample of 1,340. Therefore almost 51% of the Interview 1 households also completed the Interview 2 survey.

2.4.3 Interview survey 3

The third of the householder surveys was conducted between January and March 2019 and is referred to as Interview 3. The same sample population that was contacted for Interview 2 was re-contacted for Interview 3 (excluding cases who had opted out of the study after Interviews 1 and/or Interview 2). The survey was conducted via a telephone interview and online survey; 447 households completed the telephone survey and a further 739 households responded online, giving a total sample of 1,186. Some 80% of the Interview 3 surveys had an Interview 1 and Interview 2 survey (944 households), while the remaining 242 households had an Interview 1 survey only.

2.5 Data quality

As in any interview survey, some data can be incorrect or missing. Most of the raw outputs from the interview surveys were generally complete and considered good quality. Where data was clearly incorrect, the approach has been to 'correct' the data where it was obvious what the response should have been. For example, in the householder reported heating on and off times, occasional problems resulted from confusion with the 24-hour clock or random input errors. There was also an inherent problem in validating the quality of this specific data as it was based on householders' responses to questions rather than to actual timeclock or programmer settings recorded by inspection of these controls by the interviewers.

A summary of the data cleaning and validation process for each of the interview surveys was as follows:

- Data from the pilot surveys was recoded and cleaned for consistency with variables in the main dataset, to incorporate changes between the pilot and main stage surveys

- Variables were recoded into a usable format for analysis, for example heating hours were combined and rounded, and 'other' responses were recoded into original variables. Where similar answers were provided by multiple households, new categories were created, and these are highlighted in the interview questionnaires. For example, at Interview 2, in q51 there were 459 responses in 'other' and in q53 there were 336 responses in 'other', both were recoded, and extra categories added to allow for more insightful analysis

- Checks were completed on survey routing, feed-forward variables, and multiple-response variables. Corrections were applied in line with how the survey or questions were intended

- Final inspection of extreme values and 'sense' checking was completed, before the combination of datasets ready for analysis

3 Other data collection

3.1 *Summer text survey*

In addition to the household interview surveys, a short text message survey was conducted at the beginning of July 2018 during a particularly hot spell. The purpose of the survey was to assess the effects of the high temperatures on householders' thermal comfort and behaviour.

The text survey began on 2nd July at 6pm and ran for seven days. Householders were sent four questions via text message. The response rate was highest in the first 24 hours. Of the 1,308 people to whom the text message was sent, between 353 and 376 responses were obtained for each question (27-29% of those sampled). There is the prospect of response bias in the results reported from the text survey, in particular, that those experiencing overheating might be more likely to respond to questions about the subject.

3.2 *Temperature monitoring*

During the Interview 1 survey, householders were invited to take part in a temperature monitoring study. Households were asked *'As part of this project we will be monitoring the temperature and humidity levels in some homes to enable us to better understand how the temperatures differ in different parts of the home and areas that are particularly cold or hot at different times of the year. We would like to install up to 5 small temperature sensors, and for these to remain in place for up to two years. Households who take part will receive a payment of £20 as a thankyou, £5 when the monitors are installed and £15 when they are returned at the end.'*

The temperature loggers used were modified TinyTag Transit 2 data loggers, produced by Gemini Data loggers[1], with an accuracy of +/- 0.4°C and a range of -40°C to 70°C. Each logger was marked for use in specific rooms to reduce the likelihood of loggers getting swapped around. To maximise the number of readings possible in the EFUS monitoring period, each logger was programmed to record data every 30 minutes until reaching capacity. An example of a living room logger is shown in Figure 3.1.

Figure 3.1: Modified Tinytag Transit 2 temperature logger

[1] http://gemini2.assets.d3r.com/pdfs/original/2058-tg-4081.pdf

The installation process replicated the approach undertaken for the EFUS 2011. Interviewers were given instructions on proper placement of the loggers during the interview briefings, emphasising the need for the loggers to be placed on an internal wall, away from heat sources and out of direct sunlight and at a height that could be reached by the occupant for removal of the logger (but out of the reach of small children). Because of practical issues placing loggers into the diverse circumstances found in homes, guidance on logger placement could not be overly prescriptive.

Up to five loggers per household were installed; in the living room, main bedroom and the hallway and up to two additional bedrooms where these existed. Loggers in living rooms and hallways were installed by interviewers. Where permission was not granted by the household to install a logger in the bedroom(s), the householder could install it. The loggers were given to the interviewers in sets of three (living room, main bedroom, and hallway). These were stored in sealed bags, one for each house and each bag had a bag number that was recorded in the CAPI system by the interviewer when prompted. For the fourth and fifth loggers used in properties with a second and third bedroom, four digit serial numbers for each logger were entered into the CAPI system when they were installed.

The first temperature loggers were installed during the first pilot study in May/June 2017 followed by households in the Interview 1 face-to-face survey in August to October 2017. In total 3,853 loggers were reported to be installed.

- 1,016 loggers were reported to be installed in living rooms (all by interviewers)
- 1,003 loggers were reported to be installed in hallways (6 by occupants)
- 1,009 loggers were reported to be installed in main bedrooms (444 by occupants), 529 in a second bedroom (266 by occupants), and 296 in a third bedroom (154 by occupants)

Of these, 2,580 loggers from 750 households were suitable for analysis after data cleaning. There were some issues with the loggers that necessitated corrective action or, if necessary, the removal of a household temperature profile from the analysis. These included:

- A logger ID recorded by the interviewer not matching the returned ID of a logger. This was sometimes caused by loggers being switched during their placement in the household, and these were corrected where possible
- Loggers recorded as being in a room that either did not exist in the dwelling or was not one of the five rooms intended for surveying were removed
- The pilot loggers ran out of memory and stopped recording, in March/April 2019; loggers were kept but data from March were removed from analysis
- A few loggers were not used because the householder returned them well after the expected return date, after data processing and cleaning was completed
- Extreme values were checked. Loggers failing outright during the monitoring period, recording -45°C onwards from the point of failure, or loggers recording extreme temperatures of more than 50°C for long periods of time, assumed to be a malfunction, were removed

- Cases with solar spikes, short spikes in temperature to over 40°C, that lasted usually for only half an hour, were retained, as this would have a minor effect on the averages calculated[2].

After data cleaning and validation, the internal temperatures collected were matched with the external temperature data (section 3.3) and used to produce a combined dataset, and separate temperature profiles for each household.

3.3 External Temperatures

The external temperatures came from the Met Office MIDAS-Open archive, which has data taken from weather stations across England on an hourly basis[3]. Households were paired with the station closest to them. Around 150 weather stations were used, with each station paired, on average, with five households, for which internal temperature monitoring data was available. The details of the data and details on its quality can be found on the CEDA archive[4].

As recommended by the Met Office, cleaning of the raw data was necessary, and some values were removed from this data. The steps taken to clean the data were:

1. Cases with a 'VERSION_NUM' of 1 were kept, all others were discarded
2. Where there were duplicate readings (more than one reading from the same station at the same time), those with a more recent 'METO_STMP_TIME' were retained
3. For any remaining duplicates; if all duplicates had the same 'AIR_TEMPERATURE' value then one of the cases were kept; if any two duplicates had different 'AIR_TEMPERATURE' all values were discarded[5]

Checks for missing values by time and by station identified stations with many missing values; in such cases data from more reliable stations was used. The final validation steps included checks on minimum and maximum values and extreme changes in temperature.

3.4 Gas and electricity metered consumption data

Two approaches were taken to obtain the annual gas and electricity consumption of households who took part in the EFUS study: firstly, up to three meter readings were gathered in the EFUS study; and secondly data from the external BEIS meter point data was obtained.

3.4.1 Meter readings

During Interview 1, electricity meter readings were provided by 1,647 households, and gas meter readings were provided by 1,361 households. Households were asked if they would consent to a further reading taken by a professional meter reading company[6], and in total 65% of the sample agreed to this. Follow-up meter readings were attempted in two batches by professional meter reading company Morrison Data Services: the first between 3rd August 2018 and 4th September

---

[2] An exception to this was for the analysis undertaken by Loughborough University on summer over-heating analysis, where solar spikes were removed.

[3] In contrast to the temperature monitoring data which was collected half-hourly

[4] https://www.metoffice.gov.uk/research/climate/maps-and-data/data/index

[5] An exception to this was for the analysis undertaken by Loughborough University on summer over-heating analysis, where the un-rounded value was retained

[6] At Interview 1 and Interview 2

2018 where 811 electricity and 780 gas meter readings were provided; and the second between 29th April 2019 and 31st May 2019 where 899 electricity and 827 gas meter readings were provided. In total, 397 households had three meter readings and 810 households had two meter readings.

Initial checks on the data indicated a high number of cases to be removed from the meter readings dataset, either due to missing meter readings, discrepancies with gas meter units, or based on the differences between the available meter readings. The following cases were removed from the dataset: cases with negative consumption figures; cases with implausibly high or low consumption; and duplicate cases. Due to the small number of cases after validation, the meter readings were annualised so they could be compared and used as a validation dataset for the meter point data (section 3.4.2) rather than for analysis and reporting.

3.4.2 Meter point data

At Interview 1, households were asked '*Would it be acceptable for the Government Department for Business, Energy & Industrial Strategy to obtain the reference numbers for your electricity and gas meters, known as the MPRN and MPAN numbers, from the agency that stores these and to match these numbers with data on energy consumption that they hold?'.* In total, 2,252 households provided permission for data matching. The data matching of EFUS cases to the meter point data was conducted by the BEIS Fuel Poverty and National Energy Efficiency Data-Framework (NEED) teams. In total 2,217 households were successfully matched to a Unique Property Reference Number (UPRN) required for matching to the meter point data. Finally, households were matched to the gas and electricity meter point datasets for 2017, and duplicates or cases with missing data were removed.

The 2017 electricity meter point data covered the period 31st January 2017 to 30th January 2018, and the 2017 gas meter point data covered the period mid-June 2017 to mid-June 2018. Annualised estimates of consumption were provided by Xoserve. Meter reads were taken as close to a year apart as possible, and if insufficient meter readings were available, the previous year's Annual Quantity (AQ) was brought forward. For the gas data, the AQ calculation was adjusted to account for bias of reads towards winter or summer, and for differences in seasonal conditions. Adjustments are referred to as 'weather-desensitisation' and are made based on weather sensitivity and actual daily weather. Further guidance on how this data was collected can be found here: https://www.gov.uk/government/publications/regional-energy-data-guidance-note; while information on the weather correction of gas data can be found here: https://www.gov.uk/government/publications/overview-of-weather-correction-of-gas-industry-consumption-data

The 2017 meter point data was validated against the EFUS meter readings and the previous year's meter point data, and a dataset was produced after completing final corrections based on main fuel type and gas connection status. Where households had no electricity or gas consumption data, these cases were set to 'unknown' if they were connected to the electricity or gas network and set to 'not applicable' if they were not connected to the network. In addition, the consumption data was set to missing in the following scenarios: households with PV or negative electricity consumption; households with 0 kWh/year for electricity, or less than 10kWh/year for gas; households with implausibly high or low consumption values based on main fuel type.

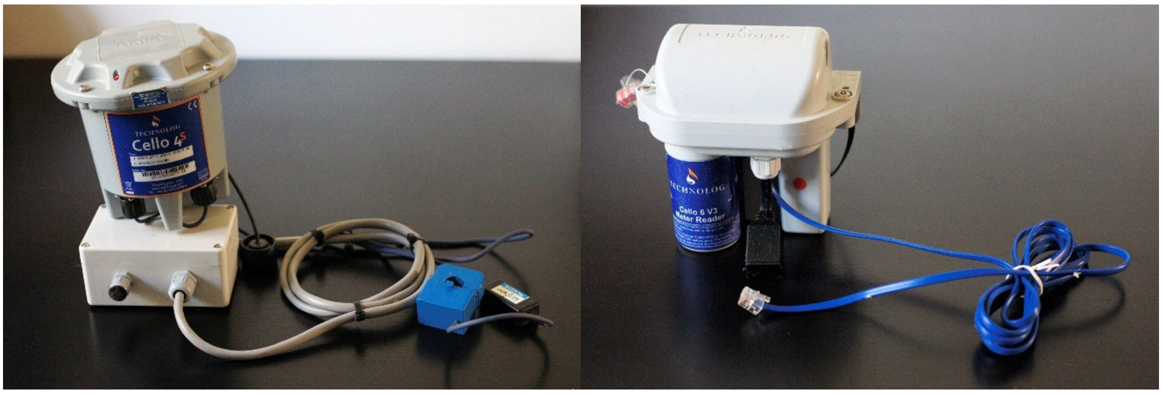
It was assumed that all households should have some associated electricity use, and if households had evidence of gas use while being recorded as not being connected to the gas network, the gas connection status was corrected, else these cases were set to missing. The main fuel was corrected where there was a discrepancy between the interview and consumption data, if there was enough evidence to support this change. For households with no recorded main heating system, the fuel with the greatest obvious use has been coded as the main fuel, otherwise this was left as provided from the EHS. Where there were multiple meter readings associated with a household, these were assessed on a case-by-case basis depending on available data.

After cleaning and validation, the resultant datasets contained annual electricity consumption for 1,994 households, and annual gas consumption for 1,770 households. The total gas and electricity consumption figures have been produced for 1,919 households with available gas and electricity data (or just electricity data where the household is not on the gas network).

3.5 Detailed electricity and gas monitoring

Detailed energy consumption data was collected by remote monitoring of mains gas and electricity consumption. Gas consumption was monitored at 30-minute time intervals and electricity at two second time intervals. The objective of monitoring the gas and electricity consumption was to examine average gas and electricity use across the day, for different times of the year. Such monitoring allows for the identification of peak usage times during the day and how these change across different seasons, particularly heating seasons, and months of the year.

Figure 3.2: Technolog loggers for electricity (left) and gas (right)



The supplier market for remote loggers was reviewed at the start of the project and a solution was provided by Technolog[7,8]; an experienced manufacturer and installer of remote monitoring equipment. The device chosen for the monitoring of gas consumption was a GPRS enabled automated meter reading (AMR) device called “CelloC6V3”. These monitors work by recording an electrical pulse output from a gas meter and logging the number of pulses detected over each 30 minute period, with the data sent to a secure data storage site via GPRS. As with the gas monitoring, a GPRS enabled electricity AMR device called “Cello4S” was used, in addition to a current clamp, to collect electrical consumption every two seconds. The data was sent to a secure data storage site via GPRS.

[7] Gas logger https://www.technolog.com/wp-content/uploads/2019/01/Cello_6S_Gas2137DS9100C.pdf
[8] Electricity logger https://www.technolog.com/wp-content/uploads/2016/07/Cello_4S_2099DS9000I.pdf

At Interview 1 households were asked how likely they were to move within the next two years. Households who responded, '*very unlikely*', '*unlikely*' or '*neither likely nor unlikely*' were then asked the following: '*As part of this project we aim to better understand how households across the country use energy. Where possible we would like to install equipment that monitors gas and electricity use. This would involve some equipment being attached close to the electricity and gas meters which records the amount of energy used each day for around two years. The equipment is silent and unobtrusive and will be installed and removed by professional engineers*'. There was a financial incentive for taking part in the gas and electricity monitoring and a total of 1,234 households consented (47%).

As part of the Interview 1 pilot survey the installation of electricity and gas monitoring was tested with those households that consented and had the appropriate meter. Technolog engineers installed 13 gas and 25 electricity monitors in households in July and August 2017[9] and these remained in place for the whole study. For households from the Interview 1 main stage survey, gas and electricity monitors were installed between January and October 2018. A total of 151 gas monitors and a total of 487 electricity monitors were installed (12% and 39% respectively of those that consented to having the gas and electricity monitored as part of Interview 1). Installation of monitors was unsuccessful in a number of households, and the most common reasons for unsuccessful installation are summarised in Table 3.1.

Table 3.1: Summary of EFUS gas and electricity monitor installations

	Gas consumption monitors installed		**Electricity consumption monitor installed**	
	Sample size	Percent (%) of available sample	*Sample size*	Percent (%) of available sample
Site not visited by an engineer [1]	*508*	41	*553*	45
No LF Output [2]	*126*	10	*1*	0
No access	*12*	1	*35*	3
No space	*0*	0	*30*	2
Refused access	*12*	1	*14*	1
No signal	*3*	0	*11*	1
No live meter	*5*	0	*0*	0
Other	*9*	1	*21*	2
Successful insulation	*151*	12	*487*	39
Total	***826***	**67**	***1,152***	**93**

Base: all households at Interview 1 that agreed to install monitors

Notes: [1] Sites were not visited by an engineer for several reasons including: not being able to get hold of the occupant to book an appointment, changes in circumstances since Interview 1, or they no longer wanted to be part of the project. [2] The LF output is a pulse output produced by meters when a given amount of energy or (in the case of a gas meter) fuel has passed through the meter. The output can be monitored in order to log consumption and is required by the Cello loggers used to monitor gas consumption for this study.

[9] Two electricity monitors were installed in March 2017

As monitors were installed in households between January 2018 and October 2018, there was a large time frame over which installations occurred, meaning there were different amounts of gas and electricity consumption data recorded for each household. In order to minimise large variations in the amount of data being analysed, a detailed consumption monitoring period was defined, chosen to run from 1st May 2018 to 30th April 2019. This period maximised the data available for analysis over the EFUS heating season, which was particularly relevant for the gas consumption data, while also covering a full year of data.

All consumption data for each household over the detailed consumption monitoring period was screened for errors and invalid readings, such as extended periods of zero consumption or very low consumption, and invalid readings were set to missing. Where this screening process removed the majority of the data for a household, over the detailed consumption monitoring period, that household was removed from the dataset.

After the data cleaning process for gas consumption, 143 households remained with valid gas consumption data over the chosen monitoring period; 93% of households had consumption data for the full year, 4% had consumption data for between 6 and 12 months, and 3% had data for 6 months or less. For electricity consumption, after data cleaning 436 households remained with valid electricity consumption data over the monitoring period; 96% of households had consumption data for the full year, 3% had consumption data for between 6 and 12 months, and 1% had data for 6 months or less.

3.6 Humidity data

A small number of humidity monitors were placed in 40 homes. These were homes where the EHS indicated there may be existing problems with condensation or dampness. Replicating the approach of the temperature loggers, humidity loggers were placed in up to five rooms per household and recorded both humidity and temperature values. Only a small number of loggers were returned and after data cleaning only 13 households had data suitable for analysis. Owing to this small sample size it was not possible to include any findings in the 'Thermal comfort, Ventilation, Damp and Mould' report.

4 Derived variables

A set of key EHS derived variables for dwelling and socio-economic characteristics were modelled for use in analysis across all EFUS 2017 reports. Where required, additional EHS/EFUS derived variables were created and reported on for specific reports. As far as possible, all EFUS 2017 derived variables were modelled to be consistent with the equivalent EFUS 2011 derived variables to facilitate timeseries comparisons.

4.1 Key derived variables for dwelling and household characteristics

Key EHS derived variables were predominantly modelled using EHS data obtained from the End User Licence (EUL) datasets 'physical.sav' and 'interview.sav', matched to the EFUS datasets (Table 4.1). The derived variables were updated using EFUS data from Interview 1 (to account for any reported dwelling and/or household changes since the original EHS survey and the first EFUS survey) and from Interview 3 (to account for any reported dwelling and/or household changes since the first EFUS survey). Derived variables were similar to those used in analysis in EFUS 2011, however new derived variables include: wall type, EPC rating band, household composition, long-term sickness or disability and fuel poverty gap. Furthermore, the derivation of income and employment status derived variables were not analogous to EFUS 2011. At Interview 3, the EFUS derived daytime occupancy variables was used as a key variable in analysis, replacing the employment status of households as this could not be updated at Interview 3. The EFUS 2017 dataset documentation provides further details on how these derived variables should be used in analysis[10].

4.1.1 Fuel poverty status and fuel poverty gap

Fuel poverty calculations were made at Interview 1 and at Interview 3. Changes that would have affected a households' fuel poverty status since the EHS included dwelling improvements, changes to household composition, changes in tenure, and changes to income. Householders were asked a variety of questions on these items and their responses were used, in conjunction with the original base EHS data, to recalculate the fuel poverty status of each case at the time of the original EHS survey. For more detailed information on fuel poverty derived variables refer to the EFUS 2017 'Fuel Poverty' report.

Note, at Interview 3, due to the smaller effective sample size, analysis of fuel poverty split by region and household composition should be interpreted as indicative only[11].

[10] Readers are also advised to refer to the EHS documentation on the UK Data Archive for more detailed information on the original EHS derived variables

[11] At Interview 3, when weighted there was a decrease in fuel poor households in Yorkshire and Humber and an increase in fuel poor households in the South West compared with the EHS; also there was a decrease in fuel poor households composed of lone parents with dependent children, and an increase in fuel poor multi-person households, when compared with the EHS.

Table 4.1: List of key dwelling and household derived variables

	Interview 1 & 2		Interview 3		
	Modelling matches	Data used	Modelling matches	Data used	Notes
Dwelling characteristic:					
Dwelling type	EFUS 2011 and EHS	EHS	EFUS 2011 and EHS	EHS	No changes between surveys
House or flat	EFUS 2011 and EHS	EHS	EFUS 2011 and EHS	EHS	No changes between surveys
Dwelling age	EFUS 2011 and EHS	EHS	EFUS 2011 and EHS	EHS	No changes between surveys
Useable floor area	EFUS 2011 and EHS	EHS	EFUS 2011 and EHS	EHS	No changes between surveys
Urban or rural location	EFUS 2011 and EHS	EHS	EFUS 2011 and EHS	EHS	No changes between surveys
Government office region (EHS version)	EFUS 2011 and EHS	EHS	EFUS 2011 and EHS	EHS	No changes between surveys
Main heating system	EFUS 2011 and EHS but also includes district, communal, heat pumps and micro-chp systems	EHS and Interview 2	EFUS 2011 and EHS but also includes district, communal, heat pumps and micro-chp systems	EHS, Interview 2 and 3	Updated at Interview 2 and then again at Interview 3
Fuel type of main heating system	EFUS 2011 and EHS	EHS and Interview 2	EFUS 2011 and EHS	EHS, Interview 2 and 3	Updated at Interview 2 and then again at Interview 3
Predominant wall type	EHS. Not used in EFUS 2011	EHS	EHS. Not used in EFUS 2011	EHS	No changes between surveys
Are the walls of the dwelling insulated?	EFUS 2011 and EHS	EHS and Interview 1	EFUS 2011 and EHS	EHS, Interview 1 and 3	Updated at Interview 1 and then again at Interview 3
Loft insulation thickness	EFUS 2011 and EHS	EHS and Interview 1	EFUS 2011 and EHS	EHS, Interview 1 and 3	Updated at Interview 1 and then again at Interview 3
Is dwelling fully double glazed?	EFUS 2011 and EHS. Not updated from EHS due to change in question on dwelling improvements since EFUS 2011	EHS	EFUS 2011 and EHS	EHS and Interview 3	Updated at Interview 3 only
Number of insulation measures	EFUS 2011 and EHS. Does not include updates to double glazing extent (see above)	EHS and Interview 1	EFUS 2011 and EHS	EHS, Interview 1 and 3	Updated at Interview 1 and then again at Interview 3
Energy performance certificate (EPC) rating band	EHS. Replaces energy efficiency rating used in EFUS 2011	EHS and Interview 1	EHS. Replaces energy efficiency rating used in EFUS 2011	EHS, Interview 1 and 3	Updated at Interview 1 and then again at Interview 3
Household characteristic:					
Tenure	EFUS 2011 and EHS	EHS and Interview 1	EFUS 2011 and EHS	EHS, Interview 1 and 3	Updated at Interview 1 and then again at Interview 3
Household composition	Not comparable to EHS as based on household responses rather than modelled household relationship data. Not used in EFUS 2011	Interview 1, missing data input from EHS	Not comparable to EHS as based on household responses rather than modelled household relationship data. Not used in EFUS 2011	Interview 1 and 3, missing data input from EHS	Interview 1 and then updated at Interview 3
Household size	EFUS 2011 and EHS	EHS and Interview 1	EFUS 2011 and EHS	EHS, Interview 1 and 3	Updated at Interview 1 and then again at Interview 3
Age of HRP (household reference person)	EFUS 2011 and EHS	EHS and Interview 1	EFUS 2011 and EHS	EHS, Interview 1 and 3	Updated at Interview 1 and then again at Interview 3
Pensioner present	EFUS 2011 and EHS	EHS and Interview 1	EFUS 2011 and EHS	EHS, Interview 1 and 3	Updated at Interview 1 and then again at Interview 3
Children present	EFUS 2011	EHS and Interview 1	EFUS 2011	EHS, Interview 1 and 3	Updated at Interview 1 and then again at Interview 3
Employment status of household	Not comparable to EHS or EFUS 2011 as includes all adults in the household as reported at Interview 1	Interview 1	-	-	Interview 1 only
Daytime occupancy	-	-	Not comparable to earlier surveys due to question wording	Interview 3	Interview 3 only
After housing costs equivalised income	EHS. Replaces HRP and partner gross income used in EFUS 2011	EHS and Interview 1	EHS. Replaces HRP and partner gross income used in EFUS 2011	EHS, Interview 1 and 3	Updated at Interview 1 and then again at Interview 3
Long-term sickness or disability	EHS. Not used in EFUS 2011	EHS	EHS. Not used in EFUS 2011	EHS	No changes between surveys
Under-occupancy status	EFUS 2011 and EHS	EHS and Interview 1	EFUS 2011 and EHS	EHS, Interview 1 and 3	Updated at Interview 1 and then again at Interview 3
Fuel poverty status	EFUS 2011 and EHS	EHS and Interview 1	EHS. Not used in EFUS 2011	EHS, Interview 1 and 3	Updated at Interview 1 and then again at Interview 3
Fuel poverty gap	EHS. Not used in EFUS 2011	EHS and Interview 1	EHS. Not used in EFUS 2011	EHS, Interview 1 and 3	Updated at Interview 1 and then again at Interview 3

4.2 Other derived variables

In addition to the variables tabulated in Table 4.1, further derived variables were modelled from the following datasets.

4.2.1 EFUS 2017 interview derived variables

Much of the data collected in the interview surveys required some derivation to provide variables to facilitate the in-depth analysis of precise aspects of the survey questions and present the findings in the individual reports. For example, derived variables were produced for the analysis of main heating from the interview survey. The reported on and off times in each heating period for the main heating system were used to calculate the numbers of hours the system was on for each weekday and the weekend (i.e. subtracting the on-time from the off-time for each period provides the number of hours in each period). Summary information on all the derived variables produced is provided in the EFUS 2017 dataset documentation that accompanies the datasets.

4.2.2 Temperature data derived variables

- Internal temperature data recorded from the temperature loggers was used to calculate mean daily averages for each room with data, zone 2[12] and the whole house from October 2017 to April 2019. Adjustments were made to account for the change from British Summer Time prior to the averaging of the data. Seasonal, monthly and heating season averages were calculated. Additional processing of this data was carried out by Loughborough University as part of their work to assess summer overheating (see the EFUS 2017 'Thermal Comfort, Ventilation, Damp and Mould' report for further details)

- External temperature data (from the Met Office MIDAS data) was used to calculate mean daily averages for each household (based on the closest weather station) from October 2017 to April 2019. Seasonal, monthly and heating season averages were calculated.

4.2.3 Detailed energy consumption data derived variables

Detailed energy consumption data was collected for both gas and electricity via monitors installed in the household. For both gas and electricity consumption data:

- Daily averages for each household were calculated by summing the consumption each day to a daily total, then averaging these daily totals across a time period of interest

- Hourly averages were produced for each hour in the day for each household, calculated by averaging the values into an hourly value, then averaging across all days in which data is recorded for that hour over a time period of interest

- Time periods of interest included: the total monitoring period, weekdays and weekends, the heating season, seasons and months of the year, and the coldest week of the year.

[12] The BREDEM calculation assumes two zones: a living area (zone 1) and a non-living area (zone 2).

4.2.4 Meter point data derived variables

Meter point data was provided by the BEIS National Energy Efficiency Data-Framework (NEED) team as annualised electricity consumption data (kWh/year), and annualised and weather corrected gas consumption data (kWh/year). In addition to the gas and electricity consumption, the following were derived:

- Total (combined) gas and electricity consumption (kWh/year) where either: cases have both gas and electricity data, or they are not on the gas network and have electricity data
- Corrections to the derived variables for 'main fuel used' and 'mains gas present' variables, for use with the 2017 meter point data
- The 2017 gas costs, electricity costs, and gas and electricity costs (£/year) calculated from available meter point data and tariff comparison rates from Interview 1
- Other fuel costs used for heating the home, heating water or for cooking, from Interview 1, and total energy costs (£/year) from combining the other fuel costs with total gas and electricity costs
- Additional processing of the consumption data was carried out for the 'Household Energy Consumption and Affordability' report to determine over-consumption (when comparing metered energy consumption with modelled energy consumption). Further details are provided in the report.

5 Weighting and errors

To enable the EFUS 2017 data to be scaled up to represent the national population, and to correct for non-response bias and sample design (stratification), weighting factors were calculated to align EFUS 2017 totals for key dwelling and socio-economic variables with the national totals reported in the latest available national statistics. For practical reasons, weighting can only control for a few factors at a time. As a result, biases will exist among groups which have not been controlled for as part of the weighting process.

Several sets of weights were created for the EFUS 2017. These weights provided national totals for the following households:

- Interview survey 1 (2,632 cases)
- Interview survey 2 (1,340 cases)
- Interview survey 3 (1,186 cases)
- Temperature monitoring data (750 cases)
- Electricity metered data (1,994 cases)
- Gas metered data (1,770 cases)
- Gas and electricity metered data (1,919 cases)

Due to the small sample sizes, no weighting factors were produced for the detailed gas and electricity consumption data, nor for the text survey data.

This section reports on how the various weightings for the EFUS were derived before examining the use of the weighting factors for reporting and analysis.

5.1 Interview data weighting

Weighting factors for the three interview surveys were derived using a RIM (Random Iterative Method) weighting method and logistic regression models. Weights were trimmed, re-based and grossed to population targets.

5.1.1 Interview 1

The weighting factor for the Interview 1 survey was derived using a RIM weighting method[13], based on the following variables from the EFUS Interview 1 dataset (i.e. not those from the original EHS fieldwork) and with targets taken from the latest available national statistics: age of HRP, sex of HRP, dwelling type, government office region, tenure and fuel poverty flag. The weights were trimmed to the 0.5 and 99.5 percentiles[14] and re-based to a mean of 1. Finally, the trimmed weights were grossed up to match the population of households in England in 2017 (23.95 million[15]).

[13] RIM weights were generated for all households using the SAS Raking macro
[14] Weights were trimmed to reduce the variance of the non-response weights and avoid any extreme values, therefore reducing the possibility that large weights would lead to a high impact on results
[15] Dwelling stock estimates 2017, MHCLG

A logistic regression modelling process was used to establish whether the set of RIM weights could be improved upon. Using the original EHS dataset, a stepwise approach was used to identify which EHS variables were most effective in predicting response at Interview 1, using the EFUS participation flag as the dependent variable. The only variable which provided a significant improvement to the logistic model, over and above the six variables used in the RIM weighting method, was household composition.

The RIM weighting method was repeated using the new set of seven variables from the EFUS dataset: age of HRP, sex of HRP, dwelling type, government office region, tenure, fuel poverty flag, and household composition. The revised weights were trimmed to the 0.5 and 99.5 percentiles and re-based to a mean of 1. Finally, the trimmed weights were grossed up to match the population of households in England in 2017 (23.95 million). The weighting efficiency was 74%. The reduction in efficiency was due to fuel poverty and tenure which were significantly out of line in the EFUS sample, due to over-sampling of fuel poor cases (see section 2.4.1).

5.1.2 Interview 2 (follow-up to Interview 1)

Using the EFUS dataset as the basis for the analysis, a stepwise logistic regression approach was used to optimize the prediction of whether an Interview 1 respondent had taken part in the Interview 2 follow-up survey. The Interview 2 participation flag was the dependent variable, and a wide range of relevant variables were included in the model. The stepwise modelling process produced a model with six predictor variables: age of the HRP, fuel poverty status, employment status, government office region, Interview 1 mode of completion (face-to-face or telephone) and tenure.

The weights for Interview 2 were calculated as the inverse of the probability resulting from the logistic regression model. In a similar way to Interview 1, the weights were trimmed at the 0.5 and 99.5 percentiles and re-based to a mean of 1. The weighting efficiency was 94%.

The final weighting factor for Interview 2 households was then created to match the population profile, by multiplying the final RIM weight at Interview 1 with the weights created from the logistic regression model at Interview 2. The weights were then grossed up to match the population of households in England in 2017 (23.95 million). The final weighting efficiency was 74%.

5.1.3 Interview 3 (follow-up to Interview 1)

A similar method to Interview 2 was used to create the weighting factor for Interview 3 households. Using the EFUS dataset as the basis for the analysis, a stepwise logistic regression approach was used to optimize the prediction of whether an Interview 1 respondent had taken part in the Interview 3 follow-up survey. The Interview 3 participation flag was the dependent variable, and a wide range of relevant variables were included in the model. The stepwise modelling process produced a model with five predictor variables: age of the HRP, government office region, fuel poverty status, tenure and Interview 1 mode of completion (face-to-face or telephone).

The weights for Interview 3 were calculated as the inverse of the probability resulting from the logistic regression model. The weights were trimmed at the 0.5 and 99.5 percentiles and re-based to a mean of 1. The weighting efficiency was 92%.

The final weighting factor for Interview 3 households was then created to match the population profile, by multiplying the final RIM weight at Interview 1 with the weights created from the logistic regression model at Interview 3. The weights were then grossed up to match the population of households in England in 2018 (24.17 million[16]). The final weighting efficiency was 72%.

5.2 Temperature and metered data weighting

The weighting strategy was the same for the creation of the weighting factors for the monitored data. Four sets of weights were derived for the following: all dwellings with temperature monitoring data; all dwellings with electricity meter point data; all dwellings with gas meter point data; and all dwellings with total energy consumption from meter point data (both gas and electricity readings for on-grid dwellings, and electricity readings only for off-grid dwellings).

Using the EFUS dataset as the basis for the analysis, a stepwise logistic regression approach was used to optimize the prediction of whether an Interview 1 respondent had a temperature monitor or metered consumption data. The outcome measure was the dependent variable, and a wide range of relevant variables were included in the model. The stepwise modelling process produced the following models:

- **Temperature monitor**: age of HRP, dwelling age, fuel poverty status
- **Electricity consumption**: dwelling age, fuel poverty status, pensioner present, partner present, main fuel type
- **Gas consumption**: house or flat, dwelling age, floor area, pensioner present, partner present
- **Total energy consumption**: dwelling age, dwelling type, main heating system, partner present, fuel poverty status, floor area, pensioner present

The weights for each of the monitored datasets were calculated as the inverse of the probability resulting from the logistic regression model. The only non-response weights that were trimmed were the temperature monitor non-response rates, which were trimmed at the 2.5 and 97.5 percentiles and re-based to a mean of 1[17]. The weighting efficiency for each of the weights ranged between 93 and 99% (see Table 5.1).

The final weighting factors were created by multiplying the final RIM weight at Interview 1 with the weights created from the logistic regression models. The weights were then grossed up to match the population of households in England in 2017 (23.95 million). The efficiency of the final weights varied from 63% for the temperature monitor weights to 73% for the gas consumption weights (see Table 5.1).

5.3 Use of weighting factors in analysis

For each of the sets of weighting factors produced, two weighting factors were created for use in analysis. The first was the household weighting factor (labelled as 'grosswt' in each dataset) where each household in the EFUS dataset represented the number of households in England. Therefore, when summed, this represented the total number of households in England (24.17 million for Interview 3, 23.95 million for Interview 1 and all other datasets).

---

[16] Dwelling stock estimates 2018, MHCLG

[17] Trimming of the other weights was not necessary as they were not that varied

For statistical analysis, a second weighting factor was derived (labelled as 'statwt' in each dataset). This was done firstly to allow for statistical analysis to be conducted at the sample-level, and secondly for the incorporation of the design effect, and follows a method detailed by Walker and Young (2003)[18] as outlined in section 5.3.1 below.

5.3.1 Design effect

Standard errors are greatly understated when using the household weighting factor for statistical analysis (as this represents the number of households in England, rather than the sample size). The weights were therefore re-scaled to a mean of 1, by dividing the weight for each household by the mean of the household weighting factors. Therefore, when summed, this represented the total number of households in the dataset (e.g. for Interview 1 this was 2,632 households). The weighting factor was thus still representative of the population proportions but has been re-scaled to allow for statistical testing.

The scaled weighting factor however does not account fully for the complex survey design, a design effect (sometimes called a 'weighting effect') needs to be included, else standard errors will still be underestimated and there exists a greater likelihood of finding an erroneously 'significant' result. Design effects provide a measure of the impact of weighting on the variance of the sample and can be used to adjust for the complex survey design within the weighting factor.

The design effect ($DEFF$) was calculated based on the following formulae[19]:

$$DEFF = \frac{n \sum_{i=1}^{n} w_i^2}{(\sum_{i=1}^{n} w_i)^2}$$

Where, n is the unweighted sample size, w_i is the scaled weight of the i-th household.

The design effect was divided into the weighting factor[20], reducing the effective sample size for analysis and thereby increasing the standard error and size of the confidence intervals when running statistical analysis. Table 5.1 shows the efficiency of the weighting factors produced, along with the effective sample size for analysis. The table should be interpreted as follows; where at Interview 1 if the survey was derived by simple random sampling, the necessary sample to observe the same result would have been 74% of households (or 1,955 households).

[18] Walker and Young, 2003. Example of the impact of weights and design effects on contingency tables and chi-square analysis. Journal of Modern Applied Statistical Methods, 2(2), 425-432.
[19] Kalton et al., 2005. Estimating components of design effects for use in sample design. Household sample surveys in developing and transition countries, United Nations, New York, 95-121
[20] The exception to this was when using the 'custom tables' package within SPSS. The data was weighted using the household weighting factor as the SPSS custom tables algorithms scales the weight to include the design effect for significance testing and producing the table outputs. For more information please see the SPSS statistics algorithms manual available from the IBM documentation website: https://www.ibm.com/support/pages/node/874712

Table 5.1: The efficiency of the weighting factors and effective analysis sample

	Non-response weight			Final weight[1]		
	Sample size	Effective sample size	Weighting efficiency percent (%)	*Sample size*	Effective sample size	Weighting efficiency percent (%)
Interview 1	-	-	-	*2,632*	1,955	74.3
Interview 2	*1,340*	1,256	93.7	*1,340*	989	73.8
Interview 3	*1,186*	1,096	92.4	*1,186*	849	71.6
Temperature monitor	*750*	694	92.6	*750*	475	63.3
Electricity consumption	*1,994*	1,973	99.0	*1,994*	1,420	71.2
Gas consumption	*1,770*	1,739	98.3	*1,770*	1,289	72.8
Total energy consumption	*1,919*	1,877	97.8	*1,919*	1,342	69.9

Note: [1] The Interview 1 RIM weights were combined with the non-response weight to create the final weight.

5.4 Measurement error

Like all estimates based on samples, the results of the EFUS are subject to various possible sources of error. The main source of error is sampling error, which can be broken down into systematic error and random error[21].

Sampling error is where the estimates obtained from a sample vary from the true values for the population (which is unknown) due to how the sample was arrived at. The two main sources of sampling error in EFUS are non-response bias and sample stratification (i.e. over-sampling of the fuel poor). The error introduced by this was minimised by the weighting strategy (described in section 5.3). Systematic error (or bias) covers those sources of error which will not average to zero over repeats of the survey. One example of systematic error is interviewers influencing how respondents answer questions, highlighting the importance of survey design, surveyor briefing and pilot studies. Systematic errors from interviewer differences have been minimised by applying an upper limit for the number of interviews undertaken by each of the interviewers. Each of the interviewers were assigned a fixed number of addresses to contact. Random error is where responses vary from one observation to another. An example of random error is the accuracy of temperatures measured by TinyTag loggers, as the temperatures recorded will vary randomly around the true value of the temperature (described in section 3.2).

5.5 Confidence intervals

Standard errors measure the uncertainty around the survey estimates. Confidence intervals are calculated from standard errors and provide a method of assessing the magnitude of sampling errors by indicating the range of random variation in survey estimates.

Appendix B provides sample comparisons for the percentage of households within each dwelling type, for each Interview survey and monitoring dataset. This provides an indication of the accuracy of the statistics when sampled from the Interview 1 dataset or subsequent interview or monitoring datasets.

[21] Further information on standard errors can be found in Chapter 7 of the EHS technical report: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/898113/2018-19_EHS_Technical_Report.pdf

6 Analysis

Statistical analysis was used to measure the significance of the findings presented in the EFUS 2017 reports. Bivariate and multivariate analysis was conducted to measure differences between key groups and to establish the key drivers that best explain those differences.

The key dependent variables used in reporting have been analysed by the defined set of EFUS social demographic and dwelling characteristic variables (see section 4.1). As a rule, only statistically significant results at the 99% level (where $p < 0.01$) have been included in the text, although there are some instances when results that are significant at the 95% level ($p < 0.05$) were reported[22].

Post-hoc testing and effect sizes were included in the statistical analysis, to determine where differences occur, and the relative magnitude of the differences, respectively. Throughout the reports, the effect size was used as a guide to ordering significant findings, for example when analysis was conducted on the core dwelling and household characteristics.

Missing values were excluded from analysis and have been defined as follows:

- Not applicable (-99) = Question or data not relevant to the household e.g. household not routed through to interview question
- Unknown (-88) = No answer to question or monitoring data e.g. household had missing temperature data
- Don’t know (-77) = Responded ‘don’t know’ to interview questions[23]
- Refused (-66) = Responded ‘refused’ to interview questions, where provided as an option

6.1 Bivariate analysis

The statistical tests used in significance testing were dependent on the research question, the data type, the sample size and number of variables used in the analysis. For all statistical analysis, it was assumed that observations are independent, and the data after weighting was collected from a random sample of the population. In addition to these, the following conditions were required for parametric analysis: populations are normally distributed; and homogenous variances. Where these conditions failed, non-parametric analysis was conducted. Brief descriptions of the parametric and non-parametric tests are included below, separated based on the data type.

6.1.1 Tests for categorical data

- The Chi-Squared (X^2) test was used when comparing two categorical variables to determine if they are independent. Alongside this the Z-test for proportions was used to determine where the differences occur, with a Bonferroni correction. Cramer’s V test was used to analyse the effect size. The Chi-Squared and Z-tests were conducted in SPSS using the ‘custom tables’ package. Cramer’s V[24] was calculated and added to tables using python

[22] In addition, all post-hoc tests have been conducted at the 95% level.
[23] The exception to this is where the proportion of ‘don’t know’ responses was greater than 5% of the unweighted sample, these were then included in analysis.
[24] Cramer’s V = SQRT (X^2 / (n * df)), where X^2 is the Chi-Squared test value, n is the sample size, and df the degrees of freedom for the variable with the least categories.

- McNemar's test was used when comparing two categorical variables, for a repeated measures design. This test was produced in R based on the formula in the SPSS statistics algorithms manual[25] to allow for the test to be run on weighted data (using 'statwt')

6.1.2 Tests for continuous (scale) data

6.1.2.1 Parametric

- Pearson correlations (R) have been reported for the correlation between two continuous variables. This was conducted in SPSS using the 'statwt' weighting factor

- Analysis of Variance (ANOVA) was used with continuous data to determine the impact of categorical variables, and the Tukey post-hoc test was used to determine where the differences occur. In addition, the effect size Eta-squared (η^2) has been calculated. Where assumptions for homogenous variances were violated, the result of the Welch test has been reported, and post-hoc testing has been conducted by independent t-tests. This was conducted in SPSS using the 'statwt' weighting factor

- Paired T-tests were used when comparing two continuous variables, for a repeated measure. This was conducted in SPSS using the 'statwt' weighting factor

6.1.2.2 Non-parametric

- The Kruskal-Wallis test was used for non-parametric analysis of continuous or discrete data, to determine the impact of categorical variables. The Mann-Whitney U test was used to determine where differences occur, and the effect size was approximated based on the r statistic. These tests were produced in R using the 'survey' package[26] to allow for the tests to be run on weighted data (using 'grosswt')

- The Wilcoxon-signed rank test was used for non-parametric paired analysis. This test was produced in R using the 'survey' package to allow for the test to be run on weighted data (using 'grosswt')

6.2 Multivariate analysis

Multivariate analysis, using regression models, were used in the analysis of gas and electricity consumption in the 'Household Energy Consumption and Affordability' report, and the analysis of thermal comfort and damp or mould in the 'Thermal Comfort, Ventilation, Damp and Mould' report.

Linear regression models were used to determine the extent to which different groups of factors explained variation in **annual gas consumption** and **annual electricity consumption**. Logistic regression models were used to determine whether any subset of dwelling or household variables were good predictors of the binary variables for: self-reported **summer overheating** in the living room and main bedroom, self-reported **winter thermal discomfort** in the living room and main bedroom and reported **damp and/or mould** growth. All regression models were created in R, using the base package.

---

[25] For more information please see the SPSS statistics algorithms manual available from the IBM documentation website: https://www.ibm.com/support/pages/node/874712

[26] https://cran.r-project.org/web/packages/survey/survey.pdf

Variables were selected to be included in models based on previous published research; and whether relevant to the research question. Principal Component Analysis (PCA) was used to look at how the data could be reduced, based on correlations between sets of variables. The results were used as a guide for selecting variables for the regression models, however final decisions were made based on where variables were important for the research question. Finally, variables were excluded if there were many missing values and if variables were highly correlated, based on the variance inflation factor (VIF)[27]. Where variables had groups with small samples, these were recoded where possible to reduce model complexity.

The normality of continuous variables was assessed, and variables were transformed (log transformations or square root transformations) where appropriate. Baseline categories (within the intercept) were selected based on the group with the largest number of cases, or for ordinal variables the 'lowest' category.

Across all models, due to the large number of variables of interest to the research question, variables were broken down into groups to reduce model complexity. The groups of similar variables used in individual models were: dwelling variables; household variables; and EFUS derived variables[28]. Variables selected in the individual models were then used in final combined models.

Stepwise regression was performed for each of the groups, using AIC[29] (forward and backward selection), and using the weighting factor appropriate for the dependent variable. Models were checked by looking at residual plots, variance inflation factors, and 'sense' checking. How much variance each model explains was reported using the R^2 value. To aid understanding of the contribution of each variable to the final model, a measure of relative importance was calculated for each variable using the LMG[30] method, using the 'relaimpo' package in R[31].

6.3 Timeseries comparisons

Timeseries analysis looks at changes over time for the same variable of interest, using different samples, for example EFUS 1998, EFUS 2011 and EFUS 2017. Longitudinal analysis similarly looks at changes over time, however the key difference is this is only applicable for the same households and so change is at the individual level. Within EFUS 2017 this is applicable for looking at changes between households surveyed in the EHS (2014/15 to 2016/17) with the same households in the EFUS 2017 survey (Interview 1 and 2: 2017; Interview 3: 2018).

6.3.1 EFUS 2011 and EFUS 1998

Two previous national surveys to inform the detailed use of heating systems and other sources of energy use in homes were undertaken through the EFUS 1998 and EFUS 2011. EFUS 2011 was a larger scale survey compared with EFUS 1998 and consisted of a follow-up interview survey of a sub-set of households first visited as part of the 2010/11 English Housing Survey (EHS). Additionally, a sub-sample of these households was selected to have temperature loggers and electricity monitors

[27] Variables with a VIF of three or greater (as a guideline) were excluded from the models and the VIF of variables included in the models indicated only negligible collinearity between them.
[28] Only dwelling and household characteristics were used in the logistic regression models.
[29] Akaike Information Criterion
[30] Grömping, Ulrike. (2006). Relative Importance for Linear Regression in R: The Package relaimpo. Journal of Statistical Software. 17. 1-27. 10.18637/jss.v017.i01.
[31] https://cran.r-project.org/web/packages/relaimpo/relaimpo.pdf

installed. A further stage of the EFUS 2011 involved the collection of gas and electricity consumption data from meter readings.

Data collected during the EFUS 2017 interview survey was compared back to data collected in both 2011[32] and 1998[33], where appropriate, to measure change over time. To allow for statistical analysis, the 1998 and 2011 weighting factors were re-derived using the original survey weights, and then adjusted for the design effect (section 5.3.1). In order to carry out comparisons between EFUS 2017 and previous EFUS surveys, the survey year was used as the dependent variable in significance testing to measure changes over time.

Timeseries analysis was particularly relevant to the 'Lights, Appliances and Smart Technology' report, which focused on appliance ownership and use habits. Details of the timeseries analysis conducted for each of the reports is highlighted below, with the exception of the 'Fuel Poverty' report, as this brought together key findings from the other four reports.

Heating Patterns and Occupancy report

Variables on the main heating system and controls, thermostat temperature, heating season and heating patterns were derived using the same method as in EFUS 2011, due to similarities in the design of the survey and the variables collected. This allowed for comparison of main heating data. In addition, information on unheated habitable rooms in EFUS 2011 was re-analysed to be comparable with how this data was collected in EFUS 2017, and comparisons were made between 2011 and 2017 on the use of supplementary heating.

Information on occupancy patterns was not collected in earlier EFUS surveys and there were too many differences in how information was collected on hot water usage. In addition, accurate comparisons could not be made with the EFUS 1998 heating patterns data.

Thermal Comfort, Ventilation, Damp and Mould report

The EFUS 2011 interview only collected limited data on both summer and winter thermal comfort. There were substantial differences between the EFUS 2011 and EFUS 2017 interview surveys both in terms of the questions asked[34] and when the surveys were conducted[35] meaning it was not possible to conduct meaningful comparisons between the 2011 and 2017 summer or winter thermal comfort surveys. In addition, questions were not asked relating to ventilation, damp and mould in EFUS 2011.

The method for monitoring internal temperatures, obtaining the external temperature data, and derivation of variables for analysis of internal and external temperatures in EFUS 2017 was similar to the method used in EFUS 2011, allowing for comparisons to be made between surveys. Questions on thermal comfort were not covered in the 1998 survey and only limited data was available on

[32] https://www.gov.uk/government/statistics/energy-follow-up-survey-efus-2011
[33] Unpublished data. Accessed from BRE archives.
[34] In EFUS 2011 "During the typical winter (December to February), can you normally keep comfortably warm in your living room?"; in EFUS 2017 "This winter, how often has the living room felt uncomfortably cold?".
[35] The EFUS 2017 Interview 1 survey was conducted immediately after the 15th hottest summer since 1910, whilst the EFUS 2011 was conducted during the winter following a much cooler summer meaning householders' recollection of summer thermal comfort was likely to be less accurate.

ventilation and damp and mould. EFUS 1998 temperature data was not collected in a way that allowed for comparisons with EFUS 2017.

Household Energy Consumption and Affordability report

Questions on the affordability of energy bills were not covered in EFUS 2011 or EFUS 1998 nor was detailed information collected on tariff costs or energy switching habits. However, data was available on energy suppliers in both previous surveys. In addition, information on other fuels used for heating, hot water and cooking were available in both earlier EFUS surveys.

Metered consumption data was available in both the EFUS 2011 and EFUS 1998 datasets, however in EFUS 2017 the meter point data (section 3.4.2) was used in analysis of annual energy consumption data and therefore could not be compared with previous surveys. Instead, the EFUS 2017 meter point data was compared with the 2010 meter point data (the dataset available to the project team), as this allowed for a more accurate comparison in terms of the methodology for data collection and processing, which included the weather correction of gas data.

Lights, Appliances and Smart Technologies report

When looking at appliance ownership, the questions asked across the three surveys were generally consistent, meaning that comparison was possible for common household appliances such as laundry appliances and use, dishwashers, cold appliances and cooking appliances and fuel. Questions around lighting types were also generally consistent, although the range of bulb types collected in 1998 was smaller[36]. Comparisons were possible for cooling appliances with EFUS 2011. Detailed electricity consumption data was collected in EFUS 2011 (albeit from a smaller sub-set of cases), and therefore indicative comparisons with EFUS 2017 data was possible.

Areas where timeseries comparison was limited were around changing technologies; electrical entertainment appliances and energy intensive electrical appliances. Questions around smart technologies; smart lighting, smart appliances and smart meters, were not asked in any capacity in 1998 or 2011.

6.3.2 Longitudinal analysis

Additional to timeseries comparisons, indicative longitudinal analysis was conducted in the 'Fuel Poverty' report, to analyse the number of households moving in and out of fuel poverty (at the sample level), based on reported household, dwelling and income changes since the EHS. In addition, changes in smart lighting devices between households at Interview 1 and Interview 3 was analysed in the 'Lights, Appliances and Smart Technologies' report[37].

[36] EFUS 2017 'Lights, Appliances and Smart Technologies' subject report

[37] Other survey questions that were repeated for longitudinal analyses, but were not possible to analyse due to small samples/differences in survey structure, routing and timing were: smart technologies (Interview 1 and 3); energy switching (Interview 1 and 3); heating patterns (Interview 2 and 3) and affordability of energy bills (Interview 2 and 3).

7 Glossary

Age of dwelling: This is the date of construction of the oldest part of the dwelling. Recorded by surveyors in the EHS physical survey.

Age of HRP: 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. The age of the HRP is derived from:

- variables obtained from the EHS Interview survey for households that had not changed since the earlier EHS interview.
- householder responses to questions 45-50 in EFUS Interview 1 and questions 41-45 in EFUS Interview 3 for new households.

Alternative heating: Heating system present in a room (or rooms) used as an alternative to the main heating system.

After housing costs equivalised income – weighted quintiles: This is calculated based on the fuel poverty income (from 2015 & 2016 fuel poverty datasets) and updated to account for any changes to income at Interview 1 and Interview 3 EFUS questionnaires. Validation of income based on reasons why household income had changed for the Interview 3 questionnaire provided increased confidence and reliability of the income.

Boiler type: Derived from the EHS data.

Children Present: Anyone in the household who is 16 years old or younger at the time of the EFUS interview. This is derived from;

- variables obtained from the EHS Interview survey for households that had not changed since the earlier EHS interview.
- householder responses to questions 45-50 in Interview 1 and questions 41-45 in Interview 3 in the EFUS questionnaires for new households

Daytime Occupancy Derived from the EFUS survey. A household has been classified as being 'in during a weekday' if they indicated being generally in the house on weekdays during the winter, for both the morning and afternoon periods. A household is classified as 'not in during the day' if they responded as not being in for both the morning and the afternoon periods. Households who were in for either the morning or afternoon period were coded as 'Variable' occupancy.

Dwelling insulation: The number of insulation measures (0 to 3) where positive responses for 'fully double glazed', 'insulated walls' and having loft insulation greater than 200mm count as insulation measures. EFUS Interview 1 and interview 3 questionnaires asked respondents about new insulation measures installed since the EHS survey. New windows installed since the EHS survey are excluded from the analysis as it cannot be assumed that this resulted in the dwelling being fully double glazed.

Dwelling type: Classification of dwelling on the basis of the surveyors' inspections during the EHS physical survey.

Employment status of the household: Derived from W1_q56 of EFUS Interview 1, and the modelling assumes responses are for all adults in the household (HRP, partner and any other additional adults in employment). 'Don't know' responses were coded as having no employment. Households either have at least one person employed, or all adults are unemployed.

Energy Performance Certificate (EPC) band: Energy Performance Certificate band, also sometimes known as the Energy efficiency rating (EER) band (SAP 2012) of the dwelling. Bands from A to G that are used in the Energy Performance Certificate. 'A' is the most efficient and 'G' is the least efficient. Derived from the SAP 2012 methodology used for the 2016 EHS. SAP2012 was re-modelled for dwellings which have had improvements between the EHS and EFUS Interviews 1 and 3.

Fuel poverty (LIHC) status: Based on the 'Low Income High Cost' (LIHC) definition, a household is considered to be fuel poor if: they have required fuel costs that are above average (the national median level); were they to spend that amount, they would be left with a residual income below the official poverty line. Each household's fuel poverty status has been updated using EFUS data on household changes, incomes and modelled fuel costs due to dwelling improvements.

Fuel poverty gap: The difference in pounds between the required energy costs for each fuel poor household and the nearest fuel poverty threshold.

Fuel type of main heating system: As recorded by surveyors in the EHS physical survey. Grouped into 'mains gas', 'electricity' and 'other', which includes bottled gas, bulk gas, solid fuels, oil and community schemes. The data was updated at Interview 2 and Interview 3 if a household reported using a different main heating system.

Assumptions for households reporting having central heating but did not answer about fuel type:

- Set to mains gas if a mains gas connection was recorded in the EHS

- If not on mains gas set to EHS recorded main fuel

- If reported not on gas in EFUS Interview 1, then categorised as 'other' gas (e.g. bottled).

Fully double glazed: Derived from the 'dblglaz4' EHS variable as recorded by surveyors in the physical survey. Fully double glazed is defined as 'entire house double glazed'. Not fully double glazed is anything less than fully double glazed.

New windows installed since the EHS survey were excluded from the analysis as it could not be assumed that this resulted in the dwelling being fully doubled glazed.

Heating season: The months when there is a requirement for the main heating system to provide heat. For the EFUS 2017 survey this is calculated based on householder responses to a question in Interview 2 (what month heating began every day) and a question in Interview 3 (what month heating stopped every day), both asked in relation to Winter 2017/18.

Household size: Number of persons in the household, banded into 5 groups, derived from the 'hhsizex' variable from the EHS Interview survey. The data was updated following any changes to household composition recorded in EFUS Interview 1 and Interview 3 questionnaires.

Insulated walls: Derived from the 'wallinsx' variable as measured by surveyors in the EHS physical survey and refers to any insulation for the predominant wall type. The 'solid uninsulated' category includes non-cavity other wall types such as timber, steel or concrete framed. EFUS Interview 1 and Interview 3 questionnaires asked the household about the installation of wall insulation since the EHS survey and the 'wallinsx' variable was updated.

Loft insulation: Banded variable of 'loftinsx', the level of loft insulation recorded by surveyors in the EHS physical survey. EFUS Interview 1 and Interview 3 questionnaires asked the household about the installation of loft insulation since the EHS survey and the 'loftinsx' variable was updated.

Long-term sickness or disability: Whether anyone in household has long-term illness or disability that limits their activities. And/or whether anyone in the household is registered disabled. This is self-reported by EHS interview respondents.

Pensioner Present: Anyone in the household who of state pension using data from the EHS Interview survey. Updates using responses to questions 45-50 in Interview 1 and questions 41-47 of Interview 3 EFUS questionnaires.

Region: Government Office Region that the dwelling is located in. Obtained from the EHS.

Rurality: Is the dwelling in a rural (village or isolated hamlet) or urban (urban or town or fringe) location. Derived from the 'rumorph' variable in the EHS.

SAP rating: The energy cost rating as determined by Government's Standard Assessment Procedure (SAP) and is used to monitor the energy efficiency of dwellings. It is an index based on calculated annual space and water heating costs for a standard heating regime and is expressed on a scale of 1 (highly inefficient) to 100 (highly efficient with 100 representing zero energy cost). An updated SAP rating was modelled for dwellings which had improvements between EHS and EFUS Interviews 1 and 3.

Supplementary heating:	Heating systems used in addition to the main heating system to boost internal temperatures.
Tenure:	Derived from the EHS but updated from householder responses in EFUS to q52 in Interview 1 and Q51 of the Interview 3. Cases responding 'don't know' left as the original EHS category. The modelling assumes a response of 'renting' to be a household living in the private rented sector.
Type of (main) heating system:	Derived from the EHS but adjusted for EFUS Interview 2 and Interview 3 responses (question 02). Grouped into central heating or non-central heating categories. Non-central heating includes storage radiators, gas fires, electric heaters, coal/wood/ smokeless fuel fires or stoves and other less common systems.
Under-occupying:	A household is considered to be under-occupying if the dwelling is more than large enough for the number (and type) of occupants living there. For the full definition of under occupancy, see the fuel poverty methodology handbook, which is available at: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/829010/Fuel_Poverty_Methodology_Handbook_2019.pdf
	Derived from EHS data and updated based on age and household changes at EFUS Interview 1 and 3.
Useable floor area:	The total usable internal floor area of the dwelling as modelled for the EHS 'floorx', rounded to the nearest square metre. It excludes integral garages, balconies, stores accessed from the outside only and the area under partition walls. Grouped into 6 categories.
Water heating system	Derived from EHS data. Categories are: 'with central heating', 'dedicated boiler', 'electric immersion heater', 'instantaneous', 'other'.

Appendix A: EFUS 2017 roles of each organisation

BRE (lead partner and project manager)

- Design and development of all the interview surveys (including the pilot studies)
- Data cleaning and validation of all the interview survey data and creation of data files
- Preparation of temperature and humidity loggers, and data processing, cleaning and analysis of internal and external temperature data
- Processing, cleaning and analysis of detailed gas and electricity consumption data and metered consumption data
- Creation of the EHS and EFUS derived variables for all the interview surveys and development of analysis and final weighting methodology
- Analysis and reporting for all EFUS reports and preparation of final datasets

NatCen

- Interview 1 sampling methodology
- Conducted half the interviews in the Interview 1 pilot study and some face-to-face surveys for Interview 1

GFK NOP

- Input into the design and development of Interview 1 and 2 (including the pilot studies)
- Conducted half the interviews in the Interview 1 pilot study, conducted face-to-face surveys for Interview 1 and all the Interview 1 online surveys
- Conducted the Interview 2 online and telephone surveys
- Conducted the summer text message survey
- Developed the weighting factors for all interview surveys

Loughborough University

- Input into the development of all the surveys and the summer text message survey
- Analysed, and reported on, the interview and temperature data related to summer thermal comfort. Also contributed to the analysis of data related to smart appliances

Ipsos MORI

- Input into the design and development of Interview 3
- Conducted the Interview 3 online and telephone surveys
- Developed the weighting factors for all monitoring data

Appendix B: Sample comparisons

		Unweighted Count	Adjusted Count	Column N %	Standard Error of Column N %	95.0% Lower CL for Column N %	95.0% Upper CL for Column N %
Dwelling type	End terrace	261	184	9.3%	0.7%	8.1%	10.6%
	Mid terrace	516	369	19.1%	0.9%	17.4%	20.9%
	Semi detached	714	530	25.3%	1.0%	23.5%	27.3%
	Detached	380	335	16.7%	0.8%	15.1%	18.4%
	Bungalow	233	180	8.9%	0.6%	7.7%	10.3%
	Flat	528	364	20.6%	0.9%	18.8%	22.4%
	Total	2632	1955	100.0%			

Base = all households, Interview 1

		Unweighted Count	Adjusted Count	Column N %	Standard Error of Column N %	95.0% Lower CL for Column N %	95.0% Upper CL for Column N %
Dwelling type	End terrace	123	87	9.5%	0.9%	7.8%	11.4%
	Mid terrace	235	174	17.6%	1.2%	15.4%	20.1%
	Semi detached	359	266	26.1%	1.4%	23.4%	28.9%
	Detached	236	210	18.1%	1.2%	15.8%	20.5%
	Bungalow	119	95	7.3%	0.8%	5.8%	9.0%
	Flat	268	175	21.4%	1.3%	18.9%	24.0%
	Total	1340	989	100.0%			

Base = all households, Interview 2

		Unweighted Count	Adjusted Count	Column N %	Standard Error of Column N %	95.0% Lower CL for Column N %	95.0% Upper CL for Column N %
Dwelling type	End terrace	102	71	8.8%	1.0%	7.0%	10.8%
	Mid terrace	219	153	20.0%	1.4%	17.4%	22.8%
	Semi detached	324	235	25.6%	1.5%	22.8%	28.6%
	Detached	213	185	17.7%	1.3%	15.2%	20.3%
	Bungalow	110	87	7.9%	0.9%	6.2%	9.8%
	Flat	218	139	20.1%	1.4%	17.5%	22.9%
	Total	1186	849	100.0%			

Base = all households, Interview 3

		Unweighted Count	Adjusted Count	Column N %	Standard Error of Column N %	95.0% Lower CL for Column N %	95.0% Upper CL for Column N %
Dwelling type	End terrace	78	47	10.6%	1.4%	8.1%	13.6%
	Mid terrace	130	81	18.7%	1.8%	15.4%	22.4%
	Semi detached	208	123	25.5%	2.0%	21.7%	29.6%
	Detached	116	94	18.5%	1.8%	15.2%	22.1%
	Bungalow	79	54	7.9%	1.2%	5.8%	10.6%
	Flat	139	84	18.7%	1.8%	15.4%	22.4%
	Total	750	475	100.0%			

Base = all households, temperature data

		Unweighted Count	Adjusted Count	Column N %	Standard Error of Column N %	95.0% Lower CL for Column N %	95.0% Upper CL for Column N %
Dwelling type	End terrace	196	134	9.3%	0.8%	7.9%	10.9%
	Mid terrace	407	274	19.8%	1.1%	17.7%	21.9%
	Semi detached	573	409	25.9%	1.2%	23.7%	28.2%
	Detached	266	230	16.6%	1.0%	14.7%	18.6%
	Bungalow	160	120	8.2%	0.7%	6.9%	9.7%
	Flat	392	260	20.2%	1.1%	18.2%	22.3%
	Total	1994	1420	100.0%			

Base = all households, electricity meter point data

		Unweighted Count	Adjusted Count	Column N %	Standard Error of Column N %	95.0% Lower CL for Column N %	95.0% Upper CL for Column N %
Dwelling type	End terrace	189	131	10.0%	0.8%	8.4%	11.7%
	Mid terrace	389	275	21.2%	1.1%	19.0%	23.5%
	Semi detached	549	403	27.5%	1.2%	25.1%	30.0%
	Detached	253	224	16.6%	1.0%	14.6%	18.7%
	Bungalow	150	114	8.4%	0.8%	7.0%	10.0%
	Flat	240	163	16.3%	1.0%	14.4%	18.4%
	Total	1770	1289	100.0%			

Base = all households, gas meter point data

		Unweighted Count	Adjusted Count	Column N %	Standard Error of Column N %	95.0% Lower CL for Column N %	95.0% Upper CL for Column N %
Dwelling type	End terrace	192	130	9.3%	0.8%	7.9%	11.0%
	Mid terrace	398	271	19.1%	1.1%	17.1%	21.3%
	Semi detached	561	400	25.3%	1.2%	23.1%	27.7%
	Detached	264	229	16.7%	1.0%	14.8%	18.8%
	Bungalow	157	117	9.1%	0.8%	7.6%	10.7%
	Flat	347	219	20.4%	1.1%	18.3%	22.6%
	Total	1919	1342	100.0%			

Base = all households, gas and electricity meter point data