Department of Energy & Climate Change

Development work for a longitudinal survey of energy use

Report, prepared by Ipsos MORI and University College London for DECC

03 August 2015

The views expressed in this report are those of the authors, not necessarily those of the Department of Energy and Climate Change (nor do they reflect Government policy).

Credits
This report was written by Sam Clemens, Kevin Pickering, Antonia Dickman and Amy Wheeler of Ipsos MORI and Professor Gary Raw
The wider research team comprised Dr. David Shipworth and Dr. Adam Cooper of UCL and Dr. Ben Anderson of University of Southampton.

Citation
Clemens, S., Pickering, K., Dickman, A., Wheeler, A., & Raw, G. (2015) “Development work for a longitudinal survey of energy use - Report”, A report for the Department for Energy and Climate Change. Ipsos MORI and University College London. DECC, London.

Acknowledgments
The authors of this report wish to thank the many contributors across government and academia for their engagement with this study and offering their views and advice on the most appropriate scope and design for a potential future longitudinal survey of energy use

URN: 15D/409

Any enquiries regarding this publication should be sent to us at:Fraser.Macleod@decc.gsi.gov.uk

Contents

Executive summary

Study background, aims and approach

Ipsos MORI and University College London (UCL) were commissioned by the Department of Energy and Climate Change (DECC) to continue scoping the potential design for a future longitudinal survey of domestic energy use. This aimed to build on, and further progress, the previous Feasibility Study carried out by RCUK Centre for Energy Epidemiology at the UCL Energy Institute, UCL Department of Science, Technology, Engineering and Public Policy UCL and NatCen Social Research in early 2014[1]. No commitment has been made by RCUK or DECC to commission the survey but this project will help inform decisions on whether it is a cost-effective investment.

The aim of this further development work was to:

- begin questionnaire development work to understand and prioritise topics on which it will be feasible to collect data;
- prioritise coverage of sub-groups and provide a more detailed sampling strategy; and to
- provide a more detailed assessment of costs.

Final outputs from each of these three tasks accompany this final report. The report itself presents a discussion of the overall recommendations for a future survey of domestic energy use. The recommendations are based on the team's own survey design expertise as well as engagement with a broad range of central government, devolved administration, academic, research council and regulatory stakeholders. Stakeholders were invited to engage with this development work through a variety of channels including in-depth interviews and written responses. Their views helped shape an understanding of how the survey would need to be designed to be of maximum value for future policy making, modelling and analysis around domestic energy use.

Potential value of a future energy use survey

Overall there exists significant appetite across Government and other stakeholders for a longitudinal survey of domestic energy use. The primary benefit recognised by many is the opportunity to have a new source of socio-technical data to complement the existing strong evidence base on the technical potential of buildings and energy efficiency measures. It is widely hoped that a new survey would provide information on how different households use energy in their homes and how they interact with energy efficiency measures and technologies in order to answer questions around *what works, for whom, in what circumstances*.

Collecting data on household-level behaviours around heating and keeping warm was seen as a priority for the survey, and to a greater extent than information on hot water use or keeping cool, although both of these topics should be covered to a degree. Exploring other energy uses (e.g. related to appliance use or lighting) was also important to most stakeholders, and

[1] The Executive Summary of the Feasibility Study report is available at http://www.ucl.ac.uk/steapp/docs/lukes-reports/exec

particularly for academic groups. Some major energy uses such as electric vehicles and home working were identified specifically as being of growing importance to start exploring and tracking. The survey could also offer a good opportunity, particularly in the minds of Government stakeholders, to explore the impact of energy efficiency installations (for example, on actual bill savings, perceptions of warmth and comfort and any rebound effects) and the future policy levers for encouraging further retrofit activity.

While levels of support for a new domestic energy use survey were fairly high across the stakeholders engaged, a few concerns were raised which would need to considered further, and communicated about in order to address and allay concerns (especially if being discussed with potential funders). Concerns related to identifying a clear focus and "added value" for a new survey which avoided overlap with existing surveys or analytical possibilities through existing datasets. Uncertainty also exists about how the sample could be designed to meet the needs of many disparate potential users of the data. This report seeks to provide DECC with recommendations to inform their response and discussion of these concerns with stakeholders.

Key recommendations for survey design

A number of different approaches to the survey design are considered in this report. Following an examination of different options, and in light of stakeholder preferences and needs, the final recommended design comprises the following key elements:

- The survey should be longitudinal in design, but should also remain representative across each year for cross-sectional analyses. There should be a limit on the number of waves that any single household stays within the panel for: in the opinion of the research team, the maximum length of involvement should be four waves for any participating household. (The survey as a whole can continue for longer than four years, but each household should not take part for more than four years before being replaced by a new household).
- Interviews with the panel to take place face-to-face and on an annual basis in the first instance. It is recommended that the fieldwork period is spread across the year for greatest efficiency and to allow for seasonal variation.
- The sample should be selected to be nationally representative at the level of Great Britain, with each country sampled in proportion to its number of homes (dwelling and households but sampled by address).
- No oversampling of specific subgroups of the population, or of property types, should be attempted.
- Use a clustered design with Primary Sampling Units (PSUs) defined by postcode sectors; limit the expected number of achieved interviews in wave 1 to 15 or fewer in each PSU; do not reselect PCSs for four years (assuming a four year panel design).
- For the panel component, it is the dwelling rather than the household (which might change) that should be followed up.
- Maintain the assumption set in the Feasibility Study that there is a limit of 2,750 fresh Level 2 cases each year, owing to possible limits on the availability of surveyors and/or monitoring equipment. Ideally the targeted sample size for Level 1 and Level 2[2] would be the same so that as many cases as possible have both sets of data collected from the same time period. However, DECC are keen for the survey to start with a large sample size, rather than build over time, so a boost of Level 1 interviews is included in the

[2] Level 1 covers the household survey only while Level 2 involves the household survey plus the technical building survey, as well as low level monitoring (around 4 devices installed in participating dwellings monitoring weekly gas use, weekly electricity use and hourly temperature data for three rooms).

recommended design to ensure an overall sample size of 10,000 household interviews each year, with a proportion of these having a Level 2 follow up.

The "Survey Cost Calculator"[3] developed as part of this work allows indicative costs for the study to be generated (see Appendices). The calculator focuses on a face-to-face survey of different designs (for example, the sample size, questionnaire length or the proportion of follow-up cases can be changed) but also includes costs for the surveyor (Level 2) visits and allows monitoring costs (generated through the separate "Monitoring Equipment – Cost Calculator (See Appendices)" to be added. The overall cost incorporating both the survey costs and the monitoring equipment costs suggests that the likely resource requirement for the recommended design set out above is as follows.

Year 1 £3,706,983 (includes £491,304 for Level 2 monitoring cost[4])

Year 2 £3,639,152 (includes £491,304 for Level 2 monitoring cost)

Year 3 £3,251,409 (includes £366,602 for Level 2 monitoring cost)

Year 4 £3,600,865 (includes £366,602 for Level 2 monitoring cost)

Year 5 £3,487,015 (includes £366,602 for Level 2 monitoring cost)

TOTAL £17,685,424 (includes £2,082.414 for Level 2 monitoring costs)

In addition, resource would also be required for any more intensive Level 3 monitoring. It is recommended that the separate "Monitoring Equipment – Cost Calculator" developed specifically for designing an appropriate monitoring approach is used to generate these additional costs[5].

More information about how these costs have been calculated and what they include is presented in Section 5.

Table 1 below shows how various elements of the recommended design align with stakeholder needs from this potential future survey.

Table 1: Aligning the research design with stakeholder needs

Key stakeholder need/preference	Summary of survey design response
A design enabling robust measurement of how and why energy use changes over time (including understanding how households move in and out of fuel poverty, the ongoing impact of energy efficiency installations, etc.).	Longitudinal panel element ensuring sufficient sample size among groups of key interest (e.g. an achieved annual sample size of 10,000 would include around 1,100 households in fuel poverty, 2,900 solid wall homes, 5,000 with cavity wall insulation).
Maintaining the ability to generate nationally representative cross-sectional estimates, and not compromising this through panel	Representativeness of panel maintained through up to 4 waves involvement for each household and refreshing the sample

[3] Only available to DECC staff internally due to the data supplied being commercial in confidence - (See Appendices for illustrative representation)

[4] The monitoring costs included here are based on recommended Package 1 from the "Costing Monitoring Equipment for a Longitudinal Energy Survey" final report. The recommended sample design includes 2,750 Level 2 dwellings (i.e. those receiving a surveyor visit) in Year 1 and Year 2 and then 2,052 in subsequent study years. The monitoring costs provides here have been calculated on the basis of 80% of this Level 2 sample agreeing to the installation of the Package 1 monitoring equipment (at a cost of £227,880 per 1,000 dwellings).

[5] It is likely that Level 3 monitoring will apply to a far smaller proportion of dwellings (in the region of 100 dwellings for example) although the extent of extensive monitoring required could be adjusted to meet available resources.

conditioning or Hawthorne effects.	each year.
Increase understanding of how households interact with the physical form, fabric and contents of their property.	Close linking of Level 1 and Level 2 data collection to provide maximum valuable data for each case. Sociotechnical focus on what households do and why. In the case of movers, attempt to recruit the new household moving into the original surveyed dwelling. Substantial information will be known about the physical form and fabric, enabling exploration of how energy use in that space changes with a new household.
Need for large sample size from the start in order to immediately generate data of policy-interest (i.e. rather than allowing the sample to build up more slowly over time).	Level 1 boost sample included in initial waves to ensure household survey sample size always reaches 10,000.

Table 2 below presents a summary of the advantages and disadvantages offered by different approaches to the study design, providing a comparison against the recommended approach. These points are discussed in greater detail in the main body of the report.

Table 2: Summary comparison of recommended design against alternative designs:

Design option	Detail of design option	Pros	Cons
RECOMMENDED OPTION: Balancing cross-sectional and longitudinal design **(Design 1** in "Sampling design options spreadsheet"**)**	As described above. Assume Level 2 limit & include Level 1 boost.	Provides robust and representative longitudinal and cross-sectional data. Achieves 10,000 Level 1 interviews per year, and maximises timely collection of Level 2 data where possible.	Some Level 1 cases will not be issued for Level 2.
Maximising value of all cases (**Design 3**)	Level 1 interview only conducted if case being issued for Level 2 (assuming Level 2 limit).	All cases have Level 1 and Level 2 data, maximising the value of the data for each dwelling.	Level 1 sample size builds over time so 10,000 household interviews not reached until Year 4.

Pure longitudinal approach	Cases retained indefinitely in panel, with sample refreshed over time only to replace drop-outs.	Provides measure of change over time, identifying triggers for change in energy use and picking up range of trigger points across sample.	Increasingly low composite response rate means this design is less suitable for cross-sectional estimates. With no limit to the number of waves of involvement the sample becomes increasingly unrepresentative.
Biennial data collection	4 waves of data collection per case spread over 7 years (data collection could take place every other year or each wave could be spread across two year period).	May pick up a greater number of changes occurring in a household/to a property over 7 year period.	Poorer recall of information received or triggers experienced resulting in less reliable data (potential to overcome this through interim check-up telephone interviews).
Boosting key subgroups	Disproportionate sampling of key groups of interest to inflate their size above that expected in a nationally representative sample.	Increased sample size for analysis of subgroup results.	Reduced precision of national estimates and oversampling of other groups outside those targeted (e.g. oversampling young people through attempts to boost private renters).
Multi-mode approach	Later waves (e.g. from Wave 3) collected by web and/or telephone rather than face-to-face. Could also apply to interim follow-up waves between main biennial waves of face-to-face data collection.	Potential cost efficiencies.	Telephone or web will have lower response rates, will require shorter questionnaires, and may lead to mode effects affecting the analysis of data collected by different modes.

Future work recommended to progress potential longitudinal survey

The following next steps would be required to pursue further development work in relation to a future domestic energy use survey. The indicative timescales given for delivering these stages are informed by the study team's experience of developing and conducting other studies of this type and assume that funding negotiations have already been agreed at the starting point.

Design and testing phase: a suggested questionnaire structure has been delivered as part of the current development work reported here. This questionnaire structure puts forward recommended topics for inclusion in the household survey and also provides existing examples of tried and tested questions that could be used. Further work would be needed to refine this questionnaire design and to finalise the exact question wordings used. It is recommended that the survey is cognitively tested, particularly to test any newly drafted questions but also more generally to test the flow of the complete survey instrument. Once a final cognitively tested questionnaire has been developed, this should be tested through a live survey pilot conducted face-to-face in-home among a small sample of households (estimated 6 – 9 months required).

Fieldwork: it is recommended that this is conducted on a rolling basis throughout the year.

Data production: this could take the form of a quarterly roll-out of data reports (i.e. in the first year the quarterly data is interim and then becomes accumulating each quarter).

1st complete Year 1 data set: this is likely to be available 3 to 4 months after the close of Year 1 fieldwork.

1. Introduction

1.1 Study background and aims

This development work for a future longitudinal survey of domestic energy use follows the Feasibility Study carried out by the RCUK Centre for Energy Epidemiology at the UCL Energy Institute, the UCL Department of Science, Technology, Engineering and Public Policy and NatCen Social Research in early 2014[6]. This Feasibility Study found that there is limited data about energy demand and what drives it in homes and that a new longitudinal survey of households could provide better evidence on this to enable better targeting of policies to meet long term goals to decarbonise and reduce energy used in homes by 2040. The Feasibility Study recommended further consultation with stakeholders on the most appropriate design for such a survey and, drawing on this, further development of a recommended approach.

Ipsos MORI and UCL were subsequently commissioned by DECC to continue scoping the potential design for a future longitudinal survey of energy use, building on the previous Feasibility Study. No commitment has been made by RCUK or DECC to commission the survey but this project will help inform decisions on whether it is a cost-effective investment. The aim of this further development work was to:

- Begin questionnaire development work to understand and prioritise topics that will be feasible to collect data on;
- Prioritise coverage of sub-groups and provide a more detailed sampling strategy; and to
- Provide a more detailed assessment of costs.

This report, prepared by Ipsos MORI and UCL, provides a discussion of the findings from Contract 1 of DECC's development work around a potential future domestic energy use survey. This report references these other key outputs from this study:

- Sampling design options spreadsheet [See Appendices]].
- Survey Cost Calculator tool [See Appendices].

A parallel study (Contract 2), also commissioned by DECC and led by Cambridge Architectural Research (CAR), developed the technical element of the longitudinal energy use survey. The Contract 2 project reviewed the availability and cost of appropriate monitoring equipment which would take measurements to complement the information collected via a householder survey. The Contract 2 report (Costing Monitoring Equipment for a Longitudinal Energy Survey) is available as a separate appendix

1.2 Overview of study approach

Ipsos MORI and UCL delivered five key tasks in order to answer the aims of the study set out above. The main elements of each of the five tasks are set out below.

Task 1: Inception phase (January 2015)

- Inception meeting with DECC to agree remit of study.
- Review of Feasibility Study design and meetings with Feasibility Study authors at UCL to confirm understanding of, and rationale for, the proposed approach.

[6] The Executive Summary of the Feasibility Study report is available at http://www.ucl.ac.uk/steapp/docs/lukes-reports/exec.

- Identification and collection of existing questionnaires for review.
- Mapping of stakeholders and prioritisation of initial contacts.

Task 2: Stakeholder engagement (February and March 2015)

- In-depth telephone interviews with stakeholders across Government departments (DECC and other central Government departments with a significant interest in the possibility of a potential future survey of this kind) and Devolved Administrations.
- Wider engagement with academics, research councils and additional Government stakeholders and regulators through written pro formas and topic lists for prioritisation.

All stakeholders were initially sent a briefing pack which provided them with context about the study. It included a summary of findings from the previous Feasibility Study, an overview of a potential survey framework, and key questions to help direct their feedback on the potential value of such a survey being commissioned, and their needs and preferences around its design. The stakeholder briefing pack can be found in Appendix 1. The telephone interview discussion guide and written pro forma template are provided in Appendix 2 and 3.

Stakeholders (either following the in-depth interview, or alongside the written pro forma) were also sent a questionnaire topic list for them to assign priority to a range of topics which could be included in the householder survey. This topic list is shown in Appendix 4. The total numbers of stakeholders engaged are detailed in Table 3 below.

Table 3: Stakeholder engagement in development work

Engaged via interview		**17**
Government stakeholders		14
	DECC	9
	Other departments	3
	Devolved administrations	2
Academic stakeholders		3
Invited to interview		23
Declined interview but engaged in writing*		5
Engaged in writing**		**39**
	Government responses (inc. devolved administrations)	19
	Academic responses	20

*As they wished to engage, and seek comment from, wider members of their team or organisation.

**Across 14 pro formas and 11 questionnaire topic lists.

Task 3: Questionnaire Development

- Review of literature around domestic energy use and behaviour to identify key question topic areas.

- Review of 20 existing questionnaires to develop question topic areas and identify potential question wording and harmonisation possibilities.
- Drafting of an example questionnaire structure, setting out the topics that could be covered in a 60-minute core survey as well as additional modules of interest for future waves, specific subgroups, surveyor observation or self-completion by the householder.

Task 4: Developing potential sampling strategy

- Review of the Feasibility Study design and refinement of its suggested approach.
- Development of a range of alternative sampling strategy options based on different trade-offs around cost, ease of administration and reflecting stakeholder needs.

Task 5: Cost Calculator

- Development of a calculator covering the range of costs involved in designing and administering the survey. The Survey Cost Calculator enables a range of costs to be calculated for the interviewer visit and surveyor visit based on a variety of survey design and sampling approaches. Costs for the monitoring equipment are provided by a separate calculator. These costs can then be added to the Survey Cost Calculator to provide an overall cost.

2. Understanding the potential value of a future domestic energy use survey

2.1 Introduction

Drawing on the Feasibility Study conclusions and discussions with DECC, the aim of a potential future survey of domestic energy use being commissioned can be summarised as follows.

- To fill gaps in knowledge about how and why people use energy in homes as they currently do, what is changing (and why), and what drives changes in energy use in the home and variation between homes (dwellings and households).
- Understand how the building physical form and fabric, heating technologies and their controls, and building occupants and their behaviours, needs, wants, practices and social contexts interact to shape energy demand spatially and temporally.
- To improve strategic oversight and strategic capacity to understand how best to manage energy demand in buildings and identify opportunities to reduce it and/or shift it away from periods of peak demand (e.g. via innovation or more effective deployment of current technology or behaviour change interventions).

Ultimately, it is therefore hoped that such a survey would help enhance the evidence base for effective policy-making in DECC to meet energy security and climate change challenges within the legislated time-frames and across Government. A future survey may be able to achieve this by having a core capacity to support the policy areas that DECC is likely to be responsible for (e.g. fuel poverty), but not be tied closely to any specific current policy, thus retaining relevance over an extended period. Supplementary elements to the core survey may address more specific, “issues of the day”.

The Feasibility Study concluded that essential to meeting these aims will be the collection of high quality longitudinal data to map energy use in homes accurately at the national level and observe how and why it changes over time. One of the aims of this development work was to test this assertion and understand in more detail what the value of such a survey would be to potential users of the data and how it could most appropriately be designed to meet key evidence needs.

2.2 Stakeholder views on the value of a future domestic energy use survey

The stakeholder consultation conducted during this development work was seeking to understand the ambition of a range of stakeholders around what a longitudinal survey could seek to achieve and to prioritise the potential data needs. A consultation on energy and environmental monitoring requirements and equipment was covered by a separate Contract.

Stakeholders were asked about their general views on whether the survey was needed and what value it may be able to offer them. Across the stakeholders consulted there was a high level of support for a future survey, at least in principle. It was felt to be of particular value if it was able to cover the following key areas of desired further evidence.

- **Socio-technical information**, seeking to pull together an understanding of how, for example, energy efficiency measures perform in theory with evidence on how this is

affected by the circumstances in which they are installed both in terms of the property and household type. This was felt to be valuable to develop a better understanding of "what works, for whom, and under what circumstances".

- **Behavioural data** to complement the strong evidence base on the technical potential of buildings and energy efficiency measures with information on how different households use energy in their homes and how they interact with energy efficiency measures and technologies. It was reported that only around 40-50% of domestic energy use can be explained via the technical potential of measures and so behavioural data was felt essential to understand the remaining variation between homes. In particular, there was a lot of interest across Government stakeholders for a more detailed understanding of heating behaviours, including heating patterns over the course of a year and by room.
- Data on **the long-term impact of energy efficiency measures**, to provide evidence on the extent to which energy bill savings and improvements in thermal comfort occur as a result, and the timescales over which these changes are realised. This was considered to be of particular value in understanding the impact of current DECC policies encouraging uptake of energy efficiency measures.
- **Rebound data** to provide firm evidence for the assumptions underpinning modelling around the balance between energy savings and raised indoor temperatures following retrofit.
- **Insight into potential policy levers** that might be used to reach future targets and identifying ways to measure progress towards these.

While levels of support for a new domestic energy use survey were fairly high across stakeholders consulted during this study, a few concerns were raised as detailed below. Looking to the potential future of this survey, and if it was being discussed with funders, these may be the types of issues that need to be communicated about in order to address and allay concerns.

- **Avoiding too much overlap with existing surveys** and identifying a USP for this survey and the evidence base it would produce. Related to this is a significant challenge of working out the focus for a potential new survey **without firm understanding of context** in which it could be commissioned, i.e. whether in place of, or additional too, existing surveys, as this can considerably change the priorities stakeholders have for its design and coverage.
- **Potential to maximise value of existing datasets instead** was a view offered by a couple of stakeholders who felt funding for such a survey could be better spent making best use of existing datasets.
- **Smaller niche target and beneficiary audiences** for many energy-related policies and programmes would not be captured in sufficient numbers through a national survey (Table 8 provides sample size figures for key groups of interest);
- **Likely security of future funding**: a few stakeholders were concerned about the extent to which Government would be able to commit to long-term funding for such a survey and therefore whether establishing a longitudinal design was worthwhile.
- **'Contamination' of longitudinal sample** over time through involvement in multiple waves of research and monitoring reducing the degree to which the sample is representative of the wider population. Stakeholders' perceptions of the pros and cons of a longitudinal approach are discussed further below. How the recommended sample design will address these concerns is presented in section 3.3.2.

2.3 Assessing the cost-benefit of a potential future longitudinal survey of domestic energy use

In discussing the potential value of a new domestic energy use survey, stakeholders considered what evidence gaps it could fill and what efficiencies it could create, for example by covering areas currently covered by other surveys and thus saving resources. While stakeholders found it difficult to quantify the benefits, or savings, a new survey could offer, some were able to indicate where existing, or future planned surveys, could be rationalised by a new large domestic energy use survey being commissioned. For example, DECC's smart metering team predicted efficiencies in the primary research they would need to commission if this new longitudinal survey were to proceed (with the caveat that this may still depend on the final sample design, questionnaire content and timing of the new survey). Stakeholders also spoke in general terms about the added value of improved accuracy in energy-related modelling activities – for example, from more accurate estimations of rebound effects, or more accurate estimations of the bill savings realised following energy efficiency installations.

Any future cost-benefit analysis around a new survey of domestic energy use would need to seek information on:

- Other strands of planned primary research activity and the anticipated financial spend on these, alongside an assessment of the extent to which the new domestic energy use survey might replace the need for these;
- Time, and any direct financial spend, allocated to linking and analysing existing datasets, alongside an assessment of the extent to which the new domestic energy use survey might replace the need for this activity (though it is acknowledged that it is unlikely that a new survey would ever negate the case for better data linking); and
- Estimations of the financial value of improved accuracy in modelling estimations, for example in informing policy and technology development.

3. Exploring potential study designs

3.1 Introduction

The Feasibility Study put forward a potential design for a five-wave longitudinal study on domestic energy use. Over the course of this contract for further development work, this design has been revisited and discussed with experts from Ipsos MORI, UCL, stakeholders from DECC, other Government departments, Devolved Administrations and academics. This section sets out the recommendations that have been formulated through these discussions. Whilst this recommended design is based on survey expertise and stakeholder engagement, the final survey design will be subject to its wider context, such as other surveys and data being collected, and the funding streams available. Recommendations here are around geographic coverage of the survey, sample designs, boosts of key groups of interest and frequency of the waves in a longitudinal study.

It is likely that the study would involve a number of different types of data collection and that different portions of the overall sample would be subject to each of these. The various elements include a face-to-face interview conducted in the household by a survey interviewer; a technical survey of the building; and the collection of monitoring information from installed devices. These elements were divided into three levels of intensity in the Feasibility Study and the same definitions of these levels are referred to this in this section of the report:

Level 1: Household survey only - Interviewer collects data from household via one visit per wave. No technical survey by a surveyor or monitoring is undertaken. Smart metering is accessed when available.

Level 2: Household survey plus technical building survey and low level monitoring - Surveyor also attends to undertake survey of building (in Wave 1). Some monitoring devices are installed (around 4 devices monitoring weekly electricity and gas use and hourly temperature data in around 3 rooms).

Level 3: Household survey plus technical building survey and extensive monitoring - Monitoring technician also attends to fit around 20-30 devices around the home.

The ways in which these varying intensities of data collection, first set out by the Feasibility Study, should be organised have been revisited and refined through this further development work. For example, while the concept of the three levels has been retained, the timing and sample sizes attributed to each level have been revised.

3.2 Moving on from the initial Feasibility Study

Table 4 summarises the main ways in which the Feasibility Study design has been progressed further through Ipsos MORI and UCL's further development work. A detailed exploration of each element of the proposed design is then presented in the rest of this chapter.

Table 4: Summary of changes to Feasibility Study design

Feasibility study recommendation	Revised suggestion following further development work	Rationale for revised recommendation
Cases retained indefinitely in the sample.	Set limit of 4 waves of involvement in the study per case.	Indefinite involvement will skew the sample, affecting the characteristics of the panel (and possibly lead to panel conditioning, which could affect the energy use behaviour of the panel members). A limited length of participation helps maintain representativeness in the panel as it becomes refreshed on a rolling basis. This is important to provide the balance between the value of a longitudinal element whilst offering policymakers with robust cross-sectional estimates.
Fresh sample of 2,000 cases added each year to the overall panel.	Issue more than 2,000 fresh sample in each year of the survey (this is possible as balanced against 4-wave participation limit).	Improves robustness of cross-sectional estimates
Level 1, 2 and 3 data collection planned to occur sequentially with (at least) a year's gap between Level 1 and Level 2.	Carry out Level 2 as soon as possible after Level 1 (similarly to EHS approach).	Reduces likely attrition between Level 1 and 2, thereby maximising information likely to be collected about each case. Also allows combined interview and physical data to be available in first year and to relate to the same time period for the household.
Assume limit of ~2,750 for Level 2 sample owing to availability of surveyors and/or monitoring kit, but still conduct large Level 1 sample in Year 1 of 10,000.	Match targeted sample size for Levels 1 and 2, but owing to potential Level 2 limit include a boost of Level 1 interviews.	Aims to achieve Level 1 and Level 2 data collection for as many cases as possible. The Level 1 boost is included to prevent a small starting sample (due to Level 2 limit determining lower number of Level 1 surveys).

3.3 Longitudinal versus cross-sectional design

Key recommendation: the survey should be longitudinal in design, but should also be representative in each year for cross-sectional analyses. A limit on the number of waves in the panel should be made: the research team's recommendation is that the maximum should be four waves for each respondent.

3.3.1 Stakeholder preferences for survey design

A longitudinal approach (that is, tracking the same households over time) was considered by most stakeholders to be an essential feature of the design – there is a need for such a study and this type of evidence is not covered by the studies and data currently available. For DECC teams a longitudinal approach would be especially useful: it could help to track the impact of various measures and policies such as smart meters, Green Deal and ECO, and identify their impacts against a range of other contextual factors and household-level changes. Longitudinal data could also help DECC stakeholders, and those in Devolved Administrations, to understand how households move in and out of fuel poverty over time and therefore help to identify 'triggers' that affects people's ability to pay for their energy needs.

The longitudinal element of this survey would therefore be a unique benefit in comparison to other existing, or past, surveys. However, the addition of a cross-sectional element would maximise its value to a wide group of stakeholders with different needs for the data, and also ensure the survey maintained its value in a number of different future contexts related to other cross-sectional surveys. Concerns were raised that longitudinal survey data could become 'contaminated' by respondents whose energy behaviour is influenced by taking part in the survey year after year. Some stakeholders also raised concerns that a longitudinal study would require Government to confirm long-term funding, which could be difficult to achieve. Also, a number of stakeholders saw this potentially as another source of cross-sectional data, for example to measure fuel poverty.

Summary of research questions addressed by longitudinal design

- In what ways, and for what reasons, is energy use in homes changing?
- How do occupants interact with building physical form and fabric, heating technologies and controls over time to shape their energy demand?
- How can changing energy demand best be managed and what impact are various behaviour change interventions, new technologies or other factors having on this? What is the impact of these various measures and policies in the context of other factors and other household-level changes?
- What is long-term impact of energy efficiency measures and to what extent do energy bill savings and improvements in thermal comfort occur as a result, and over what timescales are these changes realised?
- What can be learned about rebound effects?
- How do households move in and out of fuel poverty over time and what triggers can be identified that affects people's ability to pay for their energy needs?

3.3.2 Recommendations for survey design

The study should be designed as a longitudinal survey, but with the aim of keeping the cross-sectional sample as representative as possible for each year.

One of the concerns of stakeholders was to be able to use the data for each year to obtain national estimates. This means that the core design needs to have a limit on the number of waves for which participants are retained. As there will be drop out at every wave, the composite response rate of the panel component decreases. This increases the risk of that component of the sample not being representative of the population. Non-response weighting might reduce some of the bias, but weighting is not a panacea and some residual bias will always remain.

The composite response rates shown in Table 5 below are based on response rates from the EHS and Understanding Society study. This demonstrates that the composite response rate for the household survey (Level 1) at Wave 4 is likely to be about 40% and is likely to fall to about 35% in Wave 5 (and lower than this for the Level 2 data collection). Although there is no universal rule of thumb on what is an acceptable response rate, a composite response rate below 40% would reduce confidence in the accuracy of the estimates.

Table 5 shows that if a pure longitudinal design was adopted (with no limit on the number of waves of participation) then by Wave 8 the composite response rate for Level 1 would be 26% and 17% for Level 2. This would represent a very compliant group keen to be in the study and to respond (shown by the 90% response rate estimated to be achieved per wave) but which are, by this stage, likely to be highly unrepresentative of the wider national population. The longitudinal sample would also increasingly over-represent households that move home less often, which may include more mature households in owner-occupied and social housing.

Table 5: Composite response rate estimates

	Response rate per wave	Composite: Level 1	Composite: Level 2
Wave 1	60%	60%	39%
Wave 2	80%	48%	31%
Wave 3	90%	43%	28%
Wave 4	90%	39%	25%
Wave 5	90%	35%	23%
Wave 6	90%	31%	20%
Wave 7	90%	28%	18%
Wave 8	90%	26%	17%

There is another benefit to restricting the number of waves for which participants are retained and that is that it will reduce the likely impact of panel condition effects. Participants' attitudes and behaviours may be affected through the regular surveying asking them to consider their energy use in detail, and due to the installation of monitoring equipment (this is the Hawthorne Effect). While there may be short-term effects pre and post interview the evidence is less clear on the long term effect of discreet background monitoring. This survey will offer an ideal opportunity to test the strengths of this effect given the natural split into "treatment" and "control" groups, given that only a portion of households will have monitoring equipment installed, and with varying levels of intensity of monitoring taking place. Overall, it is felt that the low composite response rate and the implications this has for representativeness is likely to be the most problematic issue with indefinite retention in the sample, and more so than the possible Hawthorne Effect.

There is no single correct answer for the length of time that participants should be retained in a panel and instead this largely depends on the level of composite response rate the users of the

survey data are comfortable to accept. In the opinion of the authors, based on the composite response rates set out in Table 5, the recommendation is that participants are retained for a maximum of four waves. This is the assumption used in the rest of this report. This does not mean that the study can run for only four years, but that any individual household is only retained for four years before being replaced by a new household.

Sampling designs have also been produced for a three-wave panel design; this has the advantage that the risk of bias from non-response and the impact of conditioning effects is reduced compared to a four-wave design, but with a reduction in the size of the panel component of the sample and potentially less opportunity to observe key changes in households or properties and the impacts of those.

Section 3.4 of this report presents debate and recommendations around the frequency with which panel data should be collected and whether the 4 waves of panel involvement should represent 4 annual waves or 4 biennial waves spread over 7 years.

There is also an issue to consider around whether to re-issue cases that took part in the Level 1 in their first wave but for which there was no participation in Level 2. This could be resolved with a simple rule. For example, if there was no participation in Level 2 at wave 1, then re-issue the case for one more wave. If the case still does not participate in Level 2 then do not issue it for any future waves.

3.3.3 Understanding the value of longitudinal design for identifying and measuring the impact of trigger points

A key aim of a longitudinal study is to identify, and understand the impact of, trigger or change points that, in the context of this survey, might lead to a change in the energy dynamics of a household. It is therefore important to consider how the survey can be designed to fulfil this need. Key trigger points of interest will be the installation of home energy efficiency improvement measures as well as changes in factors such as household income, health, size and structure.

Focusing here on the likely pick-up in this survey of home energy efficiency improvements, DECC's official statistics[7] (up to end of December 2014), have been used in Table 6 to estimate the likely coverage that can be expected in the overall 10,000 sample. In reality some of these factors will be geographically skewed due to the clustering of similar property types, but Table 6 presents national average figures as an indication.

Table 6: Estimated coverage of home energy efficiency improvements

Residential dwellings with,,,	National incidence	Approx. sample size
Lofts	87%	8,700
Cavity walls	71%	7,100
Solid walls	29%	2,900
Loft insulation	70% (of homes with lofts)	6,090
Cavity wall insulation	73% (of CW homes)	5,183
Solid wall insulation	4% (of SW homes)	116

[7] https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/414614/Quarterly_Statistical_Release_-_GD_ECO_and_insulation_levels_in_Great_Britain_-_19_Mar_Final.pdf

Although DECC also publishes national statistics on the number of installations in any one given year, these can vary a great deal depending on weather patterns, promotional activity by suppliers (linked to Government policies), and the remaining potential. For example, in the annual period between December 2013 and December 2014, 3% of applicable residential dwellings had cavity wall insulation installed (i.e. of those with cavity walls), 2% of dwellings with lofts had loft insulation installed that met the 125 mm recommended depth, and 26% of 'eligible' homes had solid wall insulation installed. These are unlikely to be constant rates year on year and it also should be noted that DECC's statistics include new build properties as well as retrofit activity.

When considering how likely the survey is to pick up installations, and in the context of a 4-wave panel design, it is also important to remember that not all dwellings in the sample will be at "Time zero" at the start of their involvement in the panel but instead will be at various different stages in the lifecycle of these types of home improvement. As it is a rolling panel there will always be properties in the sample at different stages so that changes are likely to be captured.

Having households in the survey for longer periods of time (i.e. for more than the recommended 4 waves) does not increase the likelihood of picking up change, as different households will enter the survey at different points in their "change cycle" and so it is the number of longitudinal households in the sample rather than how long each one has been in the sample that is more important. The period of retention in the survey instead affects the period of data available before and after the installation.

The following tables show the number of longitudinal cases (the "panel" sample size in any given year) and how these build up under the recommended design. These figures demonstrate that it is not just one 4-year period and then another 4-year period with new dwellings, but instead these are overlapping periods and so the survey will be picking up greater levels of change across the whole sample.

For example, Table 7 shows that in Year 2 there will be a longitudinal panel sample of 8,000 participants i.e. 8,000 households that have taken part in two waves, rising to 11,385 that have taken part in two waves by Year 3 and in the same year a sample of 7.200 that have taken part in three waves, and so on. This is the sample size for which year-on-year data will be collected, including before and after measurements where their involvement in the panel coincides with a potential "trigger point".

Table 7: Build-up of longitudinal panel size over time

	Year 1	Year 2	Year 3	Year 4	Year 5
LEVEL 1					
X-sectional	10,000	12,231	14,816	13,404	13,404
2 waves		8,000	11,385	14,770	18,154
3 waves			7,200	10,246	13,293
4 waves				2,742	5,483
LEVEL 2					
X-sectional	2,750	4,950	6,930	8,712	8,712
2 waves		2,200	4,400	6,600	8,800
3 waves			1,980	3,960	5,940
4 waves				1,782	3,564

3.4 Options for survey frequency and mode of data collection

Key recommendation: Interviews with the panel to take place face-to-face, and on an annual basis (at least initially, with a review after a few years of survey data have been collected).

3.4.1 Key considerations around data collection frequency

The Feasibility Study suggested annual data collection (spread across the year to allow for seasonal variation), but this is not a fixed requirement. It is important that the data collection is spread across a whole year, given the importance of the season in energy use, and it is also important to ensure that each participant is interviewed at a similar time of year in each wave. However, how often participants are interviewed is an issue to consider.

Data collection (other than remote monitoring) more frequently than annually would be burdensome for participants, expensive and would probably not lead to much gain in terms of useful data, and so anything more often than annual is not recommended. This does not preclude focused low-cost interim data collection, by post, online or by telephone (for example to check meter readings or whether anything major has changed in the home or household, or to follow up atypical monitoring data).

However, there are precedents for less frequent data collection in panels and cohorts. For example, the English Longitudinal Study of Ageing (ELSA) interviews the panel members every two years. The main birth cohort studies often leave gaps of 5-7 years between interviews.

The optimum frequency of the data collection will largely depend on what the survey is trying to measure, and in what detail. Cohort studies, for example, often have more frequent data collection in the early years of life, when there is a lot of developmental change and then move to less frequent interviews when the rate of change slows down.

National policy-making takes place over periods of years, and the impact of policy initiatives are unlikely (at a macro level) to be seen in a short timescale. Individual changes that have a significant impact on energy use are also likely to be relatively rare, but may have a large impact when they do occur (e.g. changes in household composition, changes in income or working status, changes in heating technology/insulation). When these changes occur, the survey will want to be able to gather details so that the impact of the change can be assessed, as well as how and why it had this impact. Asking about change and behaviour retrospectively becomes more difficult the longer ago the change happened, and is more difficult for changes that have less salience for the participant.

Therefore, for a study that wants to focus on energy use behaviour and how this changes, there is a good argument to interview the participating households annually, so that it picks up relevant changes more frequently and gathers more reliable information about the change. Across a two year period, more changes will be picked up, but it would be more difficult to get reliable information about some of the changes that took place longer ago.

However, the cost of annual data collection is high, especially if that data collection is carried out face to face. Conducting fieldwork every two years will save money and so is a scenario that is worth considering. However, there are logistical issues around having fieldwork that is not continuous. If the fieldwork takes place every second year, with no fieldwork in between, this is difficult to manage, as the skilled staff taken on to conduct the interviews, surveys and installation can be difficult to retain in the “off” year if there is no alternative work to offer them. It may be possible for a large research organisation to move the interviewers to other work, but it will be less easy to retain surveyors and/or installers. This means that every fieldwork year there will need to be recruitment, training and re-engagement of staff which is both costly and risky in terms of delivery.

Given this, a better option for having two-yearly interviewing would be to spread the fieldwork across a two year period. Cost wise, this would be slightly more expensive than collecting all the data in one year (due to increased project management costs), but not by a great deal. A drawback of this approach is that any data and reporting for the whole sample is delayed by a year compared to the annual approach.

This approach would therefore mean that the sample would build up more slowly, so would not be recommended for the first few years of the study. However, once the survey has bedded in, this component of the design could be reviewed. If it was found that running the survey every two years was a more efficient design and captured change more effectively, then the survey could switch to collecting data every two years (with other modes used to maintain contact in between). Worked examples of a one-year and two-year rotational design are presented in the accompanying sampling design spreadsheet (See Appendix).).

3.4.2 Key considerations around data collection mode

Consideration has been given to using different modes of data collection for panel members in the later waves (say Waves 3 and 4, when the panel members are relatively compliant and full contact details are available). If the questionnaire for these waves is very focused on change, and so can be kept fairly short, it would be possible to consider making use of web or telephone interviewing. This could negatively impact on the response rate, and possibly the data quality (depending on the nature of the questions), and so would need careful consideration. There is some evidence from Understanding Society that use of sequential mixed mode approaches on longitudinal studies (starting with web, then telephone and finally face-to-face) can lead to a lower overall response rate than a full face-to-face approach[8], and that large cost savings are not achieved. Use of just web and telephone interviews could lead to larger cost savings, but would lead to a lower response rate, and possibly greater attrition from the panel. There is also the risk, if you use both web and telephone interviews, that data quality will be affected due to the mode effect (some people are answering questions by reading them, with no interviewer present, while others have an interviewer present and are hearing questions. This can lead to differential responses to the same questions and can lead of overestimation of individual change[9]). Less data would be able to be collected using web or telephone interviews due to the need to have shorter length surveys compared with a face-to-face in-home interview.

3.5 National coverage

Key recommendation: the sample should be selected to be nationally representative at the level of Great Britain, with each country sampled in proportion to its number of dwellings.

3.5.1 Stakeholder preferences for survey coverage

Three broad options for survey coverage were considered by stakeholders: UK-wide, Great Britain, or England only. Stakeholders highlighted a number of key considerations when thinking about suitable survey coverage, including:

- Differences in housing stock;
- Differences in weather;
- Differences in governance; and
- Differences in energy supply.

Overall, most agreed that Great Britain coverage was the most appropriate, even those DECC teams that have an England-only remit. Including Scotland in the survey was felt to be important. Stakeholders recognised that Scotland's relatively old housing stock, its harsher weather, and remote, dispersed population in some areas influenced its energy needs, use and costs and meant these varied from the rest of the UK. The different governance arrangements

[8] Role of mode in respondents' decisions to participate in IP5: findings from a qualitative follow-up study Understanding Society Working Paper Series 2014-03.

[9] Impact of mixed modes on measurement errors and estimates of change in panel data Understanding Society Working Paper Series 2014-05.

in devolved administrations were also felt to provide valuable scope for exploring the impacts of different types of policies and sharing learnings.

Policymakers and analysts in the Northern Ireland Executive were consulted and expressed interest and need from the data that could be collected by such a survey (particularly seeing value in longitudinal data). However, the inclusion of Northern Ireland in the sample for this study is likely to depend on funding decisions. Fewer stakeholders located outside Northern Ireland made a case for its inclusion, in comparison to Scotland for instance, citing significant differences in its energy use and needs and policy administration from Great Britain.

3.5.2 Recommendations for survey coverage

The conclusion from the stakeholder interviews was that the study should cover Great Britain – DECC is still responsible for some elements of energy policy in Wales and Scotland as powers have not been fully devolved in either country. From a statistical perspective it is not recommended to boost the sample for either country; in other words, both should have sample sizes in proportion to their number of dwellings. Boosting the sample size for either country (or both) would reduce the statistical precision for all other subgroup estimates and reduce the power for analyses of the data collected. As it is not currently an aim of the study to analyse the three countries separately (or compare them against each other), boosting the sample size for Scotland or Wales should only be done if either of the Devolved Administrations required separate analyse for their country (and funding arrangements could be put in place to enable this).

If additional funding for the study was available from one or both of the Devolved Administrations, then the sample size for Scotland or Wales should simply be increased to a required level. Under a proportionate design, the 10,000 sample in each year would be expected to split between 8,570 in England, 930 in Scotland and 500 in Wales. So, using Scotland for example[10], if additional funding was made available, then one could select a larger sample in Scotland to boost the sample size to 2,000. This would require an additional issued sample of about 1,940 addresses to achieve the additional 1,070 interviews required. This would increase the sample size for Scotland to 2,000 and for Great Britain to 11,070, and have no impact on the samples for England and Wales. Therefore the estimates for Scotland and Great Britain would be more precise although, because of the unequal selection probabilities within Great Britain[11] (addresses in Scotland having a higher probability of being selected than those in England and Wales), the increase in precision for Great Britain would be equivalent to having a proportionate sample of about 10,500. This is an acceptable loss in precision for estimates in Great Britain, as the motivation for oversampling in Scotland would be to increase the precision for estimates in Scotland separately.

The sample for Great Britain would still be a random probability sample because every stage of the sampling is done at random – it is just the case that addresses in Scotland would have a higher probability of selection. Design selection weights would be calculated that corrected for the over-sampling of Scotland – with these applied there would be no bias in the estimates for Great Britain resulting from oversampling Scotland.

It should be noted that an increase in the sample in Scotland from 930 to 2,000 would allow some additional subgroup analyses to be carried out. As an example, a subgroup with prevalence of 15% in Scotland would have a sample of about 300 with a boosted design compared with 140 with the proportionate design. There is no limit to the level to which the

[10] Similar arguments would apply for Wales.

[11] Selecting a sample with unequal selection probabilities results in a loss in precision for the total sample. This loss in precision can be estimated in advance using a measure called a design effect (deff) – this is explained in more detail in the section on clustering.

sample in Scotland could be boosted – and there would be no impact on the estimates for England, Wales and Great Britain as long as the sample sizes for England and Wales were unaffected.

3.6 Boosting subgroups

3.6.1 Stakeholder preferences for subgroup coverage

There was a view among some stakeholders (particularly those based in Government) that a future domestic energy use survey should allow disaggregation of data for certain subgroups. Key issues that stakeholders ideally would like to see addressed in the sample design include:

- Reaching niche audiences: policy research and evaluations often need to focus on specific groups of beneficiaries which may not be possible within the remit of a national survey;
- Providing sufficient coverage of fuel-poor households;
- Capturing specific groups of interest: such groups could include solid-walled homes, properties off the gas grid, those on heat networks or in dense urban areas; and
- Ensuring representation of the private rented sector; although this group should have sufficiently high prevalence in a simple random sample, response rates may be low and attrition high (owing to frequent changes of address).

However, stakeholders did also raise concerns around achieving a balance between covering some of these different groups of interest and achieving representative samples across nations.

The Feasibility Study also noted that the study could be used to assess the effectiveness of the Building Regulations - to do this in detail would probably entail a need to boost homes built to recent editions of the Approved Documents. Even with a boost sample however, evaluation would be difficult, given the complex relationship between household, choice of home and physical parameters of the home.

There was no interest from the stakeholders engaged during this study for the potential future survey to include communal residential premises such as military housing, student residences or care homes.

3.6.2 Recommendations for subgroup coverage and considerations for boosting

Key recommendation: no oversampling should be attempted.

The sampling frame for this study will be the Postcode Address File (PAF) which contains no useable information for sampling other than the address itself. This does mean that any geographic boost, such as regional and country boosts, can be delivered by using appropriate sampling fractions in the relevant geographies – this would be the approach used to oversample addresses in Scotland (or Wales) described above. These are relatively straightforward to carry out as they only require adjustment to the selection probabilities.

The sample sizes for a particular subgroup can be boosted by giving a higher chance of being selected to the detriment of other cases within the constraint of a fixed total sample. This approach however has statistical consequences. First, by oversampling a particular subgroup within a fixed sample size, it means that the sample sizes are smaller in the rest of the sample. So, for example, oversampling Scotland would reduce the sample sizes for England and Wales. There is a second more subtle impact on precision which is that oversampling a particular subgroup results in a loss of precision in the survey estimates at the national level and for all other subgroups that cut across it. So, for example, oversampling Scotland would reduce the precision for Great Britain for the full sample, and for analyses for all the tenure groups. Therefore the benefit from oversampling a particular subgroup needs to outweigh the loss in

precision and power for the analyses of all other groups. This is why it is recommend to boost the samples in Scotland or Wales only by selecting additional samples of addresses from the PAF in those countries.

One approach to oversampling is to merge external data into the PAF and use that to oversample the relevant subgroup. There is, however, an additional problem with that approach which is that it only works well if the external data are accurate, which is usually not the case. Oversampling dwellings in rural areas and in areas of high deprivation would be relatively straightforward as the definition is based on the area itself. Oversampling dwellings in rural areas would involve merging in the area-level indicator of rurality and then calculating selection probabilities based on whether the address is in a rural or urban area – in fact, in much the same way as addresses in Scotland may be oversampled. So, if it was required to oversample rural dwellings, this would be achievable; there would be a loss in precision for estimates of other subgroups, but that would be weighed against having more precise estimates for rural dwellings. The oversampling would be precise because the sampling frame information perfectly predicts the subgroup of the dwelling.

For most subgroups the external data are not sufficiently accurate to obtain such neat boosted subgroups. Sufficiently accurate external information is not available to identify households that are fuel poor, or dwellings that are off the gas grid, for example. Even external information that is fairly accurate, such as the Residata data on tenure, results in a sample with reduced efficiency. The EHS uses Experian data on likely tenure to sample disproportionately based on tenure. This dataset allocates each address to its most likely tenure group based on its postcode. This works well for the design of the EHS for which obtaining a boosted sample of social housing is a key requirement. However, because the external data are not perfect, this results in samples with reduced efficiency as a large proportion of addresses identified by the external data as being (likely to be) social housing, and therefore selected with higher probability, are actually owner occupied or private rented. In effect, this approach generates a boosted sample of owner occupied and private rented dwellings in areas of high level of social housings. As weights have to be applied to correct for this, there is a loss of statistical efficiency for the sample of owner occupied and private rented dwellings. This is a hit that the EHS is willing to take, given the necessity for a boosted sample of social dwellings. However, for this study, we would not recommend that any boosts are selected using disproportionate selection. The only viable option to obtain a boost sample for a particular subgroup would be to select it as an addition sample. This is because boosts of, for example, households that are fuel poor or dwellings that are off gas grid cannot be directly sampled from the PAF. Instead, these groups can only be boosted by means of in-field screening of large samples for the presence of eligible cases. An alternative approach could be to follow up eligible respondents from a different survey such as the English Housing Survey that collected the required information but this would be more complicated, subject to permission from the survey owner and result in lower effective participation rates.

The viability of doing these boosts very much depends exactly which group is to be boosted. If the group to be boosted is relatively rare, such screening can prove to be extremely expensive. By way of example, if it was planned to boost households that were off the gas grid or were fuel poor then over 15 PAF addresses would have to be issued in order to achieve a single productive case. Some rare groups are spread unevenly geographically and it is possible to use this to devise a more cost-effective screening methodology. Essentially this involves restricting fieldwork to areas of higher concentration which has the effect of increasing the screening efficiency, but at the cost of reducing the coverage of the target population. This would be a feasible option for off-gas dwellings as lists of postcodes are available that identify postcodes

that are likely to be off grid[12]. Not all the addresses in the list of postcodes would be off grid, so screening would be still be required, but fewer addresses would need to be issued for the screening exercise than if a full PAF sample was used so it would be more efficient. Likewise, not all dwellings that are off grid would be included in this list of postcodes – this is referred to as a coverage error. The extent of this coverage error would be unknown in advance, but could be estimated because off grid dwellings would also be identified in the main survey. Indeed, re-weighting the sample of dwellings that were off gas but not in the list of postcodes would correct for this coverage error. So, obtaining a boost sample of off gas dwellings would be possible through this approach, but would have to be selected as an additional sample.

Obtaining a boost sample of households living in fuel poverty would be more difficult. Oversampling addresses in areas that were identified as having high levels of deprivation based on the Index of Multiple Deprivation would seem like a sensible option to reduce the screening rate, but would not be that productive in practice as the vast majority of households living in those areas would not satisfy the criteria to be defined as fuel poor. In addition, a large percentage of households that are fuel poor are likely to live in areas that are not defined as being deprived, so there would be a coverage issue with this approach. The main problem though would be a logistical one. This is because the criteria for being defined as being fuel poor are very complicated[13] and so it would not be feasible to identify households that are fuel poor using a set of simple questions asked on the doorstep. If boosting is required, it is recommended that a separate piece of work is conducted to explore the most appropriate design for a suitable screening instrument, and the costs involved in the required level of screening.

Boosting either off gas grid or fuel poor households would, as noted earlier, have a negative effect on statistical efficiency.

Analyses of the likely sample sizes for the subgroups identified in the Feasibility Report and from the stakeholder interviews suggest that, for most, sufficient numbers would be obtained from a sample without any boosts. Table 8 below sets out the excepted achieved sample sizes in key groups of interest based on an achieved annual sample size of 10,000.

Table 8: Estimated coverage of key subgroups

Key subgroup of interest	***Estimated subgroup sample size***
Households living in fuel poverty	1,100
Off-gas grid household	1,100
Private rented dwellings	1,800
Off-gas grid households living in fuel poverty	700
Rented properties (private or social) with households living in fuel poverty	600
Owner-occupied properties with households living in fuel poverty	500

[12] http://www.xoserve.com/wp-content/uploads/Off-Gas-Postcodes.xlsx

[13] See https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/319280/Fuel_Poverty_Report_Final.pdf

Solid wall homes	2,900

Given the aims of this study and the adequate sample size for most subgroup analyses, it is therefore recommended that no oversampling is attempted. If funding were made available to generate a specific boost of off gas grid dwellings, then this could be achieved as noted above. Looking at subgroups over time shows the sample size build-up for subgroup analyses. Looking at the numbers of cases that are fuel poor (or off gas gird, as the two groups have similar overall incidence of level) the cross-sectional sample will have between 1,100 and 1,474 cases in any year (the design settles down at year 6 to 1,100 cases). For analyses of change over two waves the samples stack over the years of the study, so that by year 5 of the study there would be an analysis pool of nearly 2,000. For private rented dwellings, the sample size for analysis of change at year 5 would be over 3,200.

Table 9: Sample size for off-gas or fuel poor subgroups in the recommended four-wave panel design[14]:

	Year 1	Year 2	Year 3	Year 4	Year 5
LEVEL 1					
X-sectional	1,100	1,345	1,630	1,474	1,474
2 waves		880	1,252	1,625	1,997
3 waves			792	1,127	1,462
4 waves				302	603
LEVEL 2					
X-sectional	303	545	762	958	958
2 waves		242	484	726	968
3 waves			218	436	653
4 waves				196	392

Table 10: Sample size for the private rented subgroups in the recommended four-wave panel design:

	Year 1	Year 2	Year 3	Year 4	Year 5
LEVEL 1					
X-sectional	1,800	2,202	2,667	2,413	2,413
2 waves		1,440	2,049	2,659	3,268
3 waves			1,296	1,844	2,393
4 waves				494	987
LEVEL 2					
X-sectional	495	891	1,247	1,568	1,568
2 waves		396	792	1,188	1,584
3 waves			356	713	1,069
4 waves				321	642

Even the counts of the number of cases of fuel poor households in off gas dwellings, which has a relatively low prevalence (7%), would be sufficient to perform fairly detailed analyses of change after a few years of the study. For example, after year 4, there would be over 1,000 cases that would have responded to a wave 1 and wave 2 interview and so could be used for an analysis of change.

[14] These figures are based on no differential attrition for the particular subgroup being explored.

Table 11: Sample size for fuel poor off-gas grid household subgroup in the recommended four-wave panel design:

	Year 1	Year 2	Year 3	Year 4	Year 5
LEVEL 1					
X-sectional	700	856	1,037	938	938
2 waves		560	797	1,034	1,271
3 waves			504	717	930
4 waves				192	384
LEVEL 2					
X-sectional	193	347	485	610	610
2 waves		154	308	462	616
3 waves			139	277	416
4 waves				125	249

It is possible in practice that the sample sizes for some of the important subgroups will fall short of the numbers expected. For example, it would seem plausible that it will be more difficult to recruit private rented dwellings (compared with owner occupied dwellings) for the fitting of monitoring devices for Level 2, given the need to contact and then get permission from the owners to fit devices. One approach to boost the numbers would be to retain the PRS dwellings for more than four waves. This would help because they would add to the sample for longitudinal analysis. To reduce the risk of bias in estimates, those PRS dwellings that were in their fifth or later wave could then be included in analyses of the panel component (i.e. the longitudinal analysis), and maybe for analyses of the PRS subgroup, but be removed from the national estimates (i.e. the cross-sectional analysis). The impact of varying the number of waves on analyses would be controlled by the weighting. This approach could be used to boost the sample sizes for other important subgroups if required.

There was some discussion with stakeholders around using this study to carry out evaluation work, including selecting subgroups on which to carry out the evaluation. It would be challenging to meet this objective alongside the requirement to generate representative cross-sectional estimates. It is possible that panel participants could be transferred into more specific evaluative investigations and testing as they roll off the panel. Obtaining the evaluation sample in this way would allow a representative sample to be defined, with much data being already available on the current and recent household and dwelling characteristics.

It would also be appropriate to use this survey to monitor interventions that are being set up on such a large scale that a sufficient number of the panel would receive the intervention anyway in the natural course of events. This could apply to a range of domestic energy efficiency policies promoting home energy efficiency improvements or to the national smart meter roll-out.

The smart meter roll-out period, being implemented in different regions and through different customer approaches by individual suppliers, offers a natural experiment to compare treatment and control groups within the sample. The impact of smart meter installation should be observable in Wave 1 as long as sufficient cases do not have a smart meter installed at this point (this will depend on the timing of this survey and the progress of the mass roll-out period). It is difficult to predict the level of smart metering among dwellings that may participate in this study as the mass roll-out period is likely to be under way, making current smart meter installation rates less useful for gauging this. It is also difficult to predict how the smart meter roll-out will affect the sample over time given the flexible approach being adopted and the autonomy given to individual suppliers to manage this process for their customers.

The monitoring benefits from accelerating the installation of smart meters in the dwellings making up the sample for this survey are discussed in the separate "Costing Monitoring Equipment for a Longitudinal Energy Survey".

3.7 Dealing with moving households

Key recommendation: For the panel component, the dwelling and not the household should be followed up.

Based on EHS 2012/13 estimates[15], we would expect about 10% of households to move dwellings over a one-year period. By tenure, this is 4% of owner occupiers, 10% of social renters and 34% of private renters, For this study, if a household (or part of a household) leaves a participating home, then the recommendation is to follow the dwelling for the panel sample rather than the household (or part household). This seems a natural approach given that:

- treating this as a sample of dwellings (and the households that live within them) would simplify the survey procedures;
- the dwelling will have already been surveyed with the original occupants; and
- the monitoring devices will have been fitted to the dwelling and so it will be more efficient to leave them in place where this can be arranged with the incoming tenants and landlord (although it is likely that they will need to be removed and re-installed in many cases).

Following up both the household and the dwelling would lead to a very complicated panel design with the eligible sample for any particular analysis being different depending on whether it was at the household or dwelling level. This would require two sets of weights: one for household-level and one for dwelling-level analyses. In fact, in practice it would be impossible to calculate the correct weights for the household sample if it included both the moved and replacement households. There would also be the issue that the moved household would have no data collected on their dwelling, so it would be debatable how much use following them up would be anyway.

There would be cost implications for tracking households to their new dwelling as it would be unlikely to be within the areas included in the clustered sample – hence an interviewer and surveyor would have to make a specific visit to the new area, which would not be efficient. Given that 10% of households move each year, and the proposal is to use four waves of study, a sample of households would become more unclustered over the course of the study. If there was specific interest in following up households after they moved, then this should be done as a separate survey and the data collected should not be included in this study.

The implications of new households entering the study for questionnaire design is discussed in Section 4.

3.8 Sample design – extent of clustering required

Key recommendation: use a clustered design with PSUs defined by postcode sectors; limit the expected number of achieved interviews in wave 1 to 15 or fewer in each PSU; do not reselect PCSs for four years (assuming a four year panel design).

The standard approach for large scale face-to-face surveys is to cluster the samples within geographic areas, usually defined by postcode sectors[16]. An approach called sampling with

[15] https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/335751/EHS_Households_Report_2012-13.p

[16] A postcode sector covers approximately 2,500 addresses and is defined by the 1st half or a postcode plus the 1st number of the second half of the full postcode.

probability proportional to size (PPS) is used to sample areas (the primary sampling units or PSUs); within each selected PSU, a fixed number of addresses are sampled with equal probability. This generates an equal probability sample of addresses, so is statistically efficient – everything else being equal, sampling with unequal selection probabilities reduces the efficiency of the sample and hence the precision of survey estimates. In addition, clustering is efficient for fieldwork as the addresses within each PSU are relatively close to each other, so the travel times between the addresses (and hence travel costs) are dramatically reduced compared to an unclustered sample.

There is however a price to be paid for clustering the sample, which is a loss in precision for survey estimates. This is because people, households and dwellings in the same area tend to be more similar than those spread across the country. This similarity within the clusters results in a small loss of information, which is reflected in a reduction in the precision of estimates and the power for analyses. The more clustered the sample, and the more similar the participants, the greater the loss in precision and power. This is the reason why samples are usually designed so that the number of units within each PSU is kept fairly small (less than 20 participating cases), and are not clustered in very small geographical areas such as output areas - in general, the smaller the areas, the greater the level of similarity. In practice, samples are designed to be clustered because that maximises the level of precision and power per unit cost; in other words, clustered samples are more cost-efficient.

It should be noted that a clustered sample is still a random probability sample as both stages of the sampling are carried out at random. Therefore, there is no sampling bias resulting from the approach, it is just the case that there is usually a reduction in the precision for point estimates (i.e. wider confidence intervals) and a loss in power for analyses.

The design of the English Housing Survey (EHS) deviates from this standard design. For the EHS 2012/13 onwards, England was divided into 1,808 relatively large areas which were defined by merging coterminous LSOAs so that each contained about 12,500 addresses. These 1,808 areas were randomly allocated to two sets of 904 areas and the sample was selected within each of these sets on alternate years. This design means that the effect of clustering in each year of the survey is fairly small, as the areas within which the sample is clustered are relatively large. In addition, it means that the sample will be geographically unclustered over a two year period – in other words by combining two years of data, there is no loss in precision for clustering. This was the main reason for adopting this design, as the EHS had always previously had an unclustered design in the years prior to EHS 2012/13.

In a pilot study to examine the likely clustering effects that would result from running the EHS as a clustered sample (Pickering et al., 2003[17]), it was found that the level of the clustering effect was acceptable for most of the measures collected in the EHS. However, large clustering effects were found for a few of the measures: tenure, age of property, construction date, dwelling type, and building materials. As an important role of the EHS is to generate national estimates for these measures, a clustered design was not adopted. However generating national estimates for these measures is not an aim of this study: they will be used as covariates in analyses, but there is no requirement to maximise the precision of their national prevalence estimates as is the case for the EHS. Employing an unclustered design is expensive and so should only be employed if there is a strong argument for doing so; that is not the case for this study. Therefore it is not recommended to use an unclustered design for this study.

The pilot study for the EHS tested a wide range of measures, including: Health and Safety ratings, fitness assessments, dampness and disrepair, refurbishment date of

[17] Pickering, K., Lynn, P. & Thomas, R. (2003) Clustering Pilot for the English House Condition Survey. London: National Centre for Social Research.

kitchen/WC/bathroom, age of roof components and doors, age of boiler, types and thickness of insulation, structural defects and presence of faults. For all these measures, the level of clustering was at a level that was fairly typical for a large-scale quantitative survey. This therefore gives confidence that the level of clustering for most of the key outcome measures for this survey would also be acceptable. As described above, the features that tended to be more highly clustered would only be included as covariates in analyses of this energy study, not as outcome measures. Therefore, it is recommended that the sample for this study should be selected as a clustered sample, with postcode sectors[18] (PCSs) used as the PSUs.

To balance fieldwork and statistical efficiency, the sampling should be designed to generate no more than 15 achieved Level 1 interviews (on average) in each PCS – this would mean that there be about 9 or 10 interviews in each PCS in the Wave 4, which would still be a feasible workload for interviewers. Other candidates for the PSUs would be wards and Lower Layer Super Output Areas (LSOAs). They are all about the same size and would be associated with similar levels of clustering effect – although there is some evidence that the effect is higher in LSOAs because they were purposively designed to be homogenous. The advantage of using PCSs is that, because they are part of the postcode hierarchy, they are much more practical to use in practice: SW12 0 (PCS) is more informative as an area identifier than E05000430 (ward) or E01033207 (LSOA).

A fresh sample of PCSs would need to be selected for each year of the study. It would be possible to rotate PCSs over the years of the study, but this would increase the clustering effect for cross-sectional analyses as the average number of cases in each PCS would be increased. As an extreme example, any PCSs that were used for four years of the study would contain (on average) about 50 cases in the fourth year. In addition, a design that rotated PCSs would be harder to manage: it would cause needless confusion to have a mix of waves in each PCS in a particular year. Therefore, once a PCS has been selected for the study, it should not be re-selected for four years (assuming a four year panel design).

The use of postcode sectors could mean that a sample point is dominated by a single factor, for example if it covers one large social housing estate which had experienced the same roll-out of energy efficiency measures. While this could be perceived as a limitation to the design and something to control for, it is the view of the authors that it could also work to the advantage of the longitudinal design as the similarity in dwellings could be used to model differences in energy use arising from household factors. As fresh sample would be collected from new postcode sectors, this issue would also not generate long-term or widespread bias in the overall sample.

3.9 Sample design spreadsheets

Based on the recommendations in this section, and the work done for the Feasibility report, six potential sample designs were produced – these are available as a set of worksheets in Excel [See Appendices]]. The designs are generated in the spreadsheets by the user setting a number of parameters and, from these, the issues sample sizes and achieved sample sizes for Levels 1 and 2 are presented in separate tables. Note that these designs assume that the Level 2 work would be carried out as soon after the Level 1 interview as possible, which is the approach used for the EHS and is recommended for this study to ensure as high a participation rate in Level 2 as possible, and to ensure that interview and survey data relate to the same period.

[18] Postcode sectors with a small number of addresses are merged with neighbouring PCSs. For brevity, when referring to PCSs used as PSUs, we mean postcode sectors and merger postcode sectors.

The default wave 1 deadwood rate (8%), response rate (60%) and agreement to participation in Level 2 (65%) were based on the corresponding estimates from the EHS 2012/13; the follow up rates for wave 2 (80%) and all subsequent waves (90%) were based on Understanding Society. These estimates can all be changed in the spreadsheet if desired.

The first spreadsheet gives instructions for using the sampling design spreadsheets. The second spreadsheet (Design 1) shows our recommended design. This design assumes that there is a limit on the number of fresh Level 2 cases each year[19], but allows the initial waves to be boosted for Level 1 only. This design allows analysis of Level 1 interview data after the first year of the study. The spreadsheet has been set to achieve 10,000 Level 1 cross-sectional interviews (minimum) in each wave - that was the sample size assumed in the Feasibility Report - but that can be adjusted to any figure that is consistent with the Level 2 constraint. Note that the design in the Feasibility Report was developed assuming that there would be a limit to the number of Level 2 work that could be carried it in a year. It was felt that this would be limited either because of the availability of adequately trained surveyors or by restrictions on the number of measuring devices available each year. For the Feasibility Report the limit was set at about 2,750 new Level 2 cases each year – the default estimate has therefore been set at 2,750 in line with the Feasibility Report, but a different estimate can be entered in the spreadsheet.

Using these defaults with Design 1 gives the following number of Level 1 interviews at each wave.

Table 12: Number of Level 1 household surveys in each year and cohort under the recommended sample design (Design 1)

	Year 1	Year 2	Year 3	Year 4	Year 5
Cohort 1	10,000 →	5,769 →	3,459 →	2,742	
Cohort 2		4,231 →	3,385 →	3,046 →	2,742
Cohort 3			3,157 →	2,525 →	2,273
Cohort 4				3,157 →	2,525
Cohort 5					3,157
Fresh	**10,000**	**4,231**	**3,157**	**3,157**	**3,157**
Panel	**0**	**5,769**	**6,843**	**8,313**	**7,540**
Total	**10,000**	**10,000**	**10,000**	**11,470**	**10,696**

Using these defaults with Design 1 gives the following number of Level 2 achieved cases at each wave.

Table 12: Number of Level 2 technical surveys in each year and cohort under the recommended sample design (Design 1)

	Year 1	Year 2	Year 3	Year 4	Year 5
Cohort 1	2,750 →	2,200 →	1,980 →	1,782	
Cohort 2		2,750 →	2,200 →	1,980 →	1,782
Cohort 3			2,052 →	1,641 →	1,477
Cohort 4				2,052 →	1,641
Cohort 5					2,052
Fresh	**2,750**	**2,750**	**2,052**	**2,052**	**2,052**
Panel	**0**	**2,200**	**4,180**	**5,403**	**4,901**
Total	**2,080**	**4,950**	**6,232**	**7,455**	**6,952**

[19] If this is found not to be the case, then Design 3 would be the recommended approach.

If it turned out in practice that there was no limit to the number of Level 2 cases in each year, then Design 2 (presented in the third spreadsheet) would be recommended. This removes the limits on the number of Level 2 cases, so the only design parameter is the number of interviews that are required each year (this has been set to the default estimate of 10,000). The design maximises the number of Level 2 cases in wave 1, assuming that Level 2 data will be collected for all households that participate with Level 1.

The fourth spreadsheet (Design 3) assumes that a case would only be issued for the Level 1 interview if it was also going to be included in Level 2. A feature of the recommended design (Design 1) is that some Level 1 respondents will not be issued for Level 2 – in effect it is a Level 1 only boost. This was done to allow analysis of interview data from a large sample at Wave 1. Design 3 is included for completeness and should only be used if the decision was made not to issue any addresses without householder agreement to Level 2 data being collected. Note that for this design it takes a few years for the sample to accumulate and reach the required Level 1 sample size – in the default example, the Level 1 sample would reach 10,000 in Year 4.

Three further designs are included in the sample design spreadsheets which follow the same principles as the three designs described above but based on a three-wave panel design.

3.10 Encouraging participation

It is likely that financial incentives will be required to maximise participation in any future domestic energy use survey, particularly if it is longitudinal in design and if a household agrees to multiple forms of data collection and observation (i.e. taking part in a household interview as well as permitting a technical survey and installation of monitoring equipment).

It is well documented that financial incentives increase response rates. The EHS has, for example, been run both with and without incentives and has found that even a small incentive enables target response rates to be met.

It is therefore recommended that a small unconditional incentive is used to maximise the response to the initial approach on the doorstep. A conditional incentive may also be needed to maximise uptake of the technical survey visit, and monitoring equipment and also to encourage ongoing participation in the study in future years. The value of this incentive will depend on the level of monitoring chosen, level of intrusion it represents for a household, the number of visits required for installation and maintenance and the level of household involvement in this maintenance e.g. any requirement to document and report readings, replacing batteries. The Survey Cost Calculator does not automatically include an incentive, but it does allow an incentive amount to be added for each achieved interview and surveyor visit. Generally, the level of a conditional incentive for a standard interview would tend to be in range of £5 to £20. The amount for a surveyor visit would tend to be higher than this, and would depend on what the surveyor visit involved.

4. Developing a future longitudinal domestic household survey questionnaire

4.1 Introduction

This development work conducted by Ipsos MORI and UCL for DECC has involved four main exercises to guide the possible questionnaire content.

1. The Feasibility Study was reviewed for topics that might be included in an interview survey of householders.
2. A major recent review of the literature on domestic energy facilities and behaviour was examined to identify relevant topics for the study as a whole and the questionnaire in particular. Supplemented by the team's knowledge of the field, this exercise also provided a structure for reviewing other sources of topics and questions.
3. Twenty questionnaires used in other recent studies of domestic energy were identified, obtained and reviewed to identify specific questions that could be adopted or adapted for the new survey. The purpose of this was to identify tried and tested questions and also opportunities for harmonisation between different surveys. The topics covered are summarised in a 'survey question bank' provided to DECC.
4. Engagement with a wide range of stakeholders to explore their preferences, and priorities, for coverage of energy-related topics in the questionnaire.

Stakeholders' preferences for questionnaire coverage are detailed first in this section, before an explanation of how all four tasks above have been brought together into a recommended questionnaire structure to take forward at the end of this development work contract.

4.2 Stakeholder preferences for questionnaire coverage

A key aim for this survey, identified in the Feasibility Study, and supported across the range of stakeholders engaged in this development work, is to integrate social and technical elements of domestic energy use so that data on both aspects is collected within the confines of a single study. All energy use involves some combination of behaviour and technology, in a social context. Failure to integrate understanding of technology, behaviour and social context risks seriously misunderstanding what householders are doing, why, and why they might change. As discussed in section 2.2, stakeholders recognised that while there are existing surveys which collect technical data, and surveys that collect behavioural and attitudinal data, there is not a survey that combines these two things, and which enables this to be explored at the level of individual households and household members.

The questionnaire development work reported here therefore focuses on bringing together socio-technical information. For example, the questionnaire approaches the subject of heating, by asking what people do to keep warm, rather than simply describing the heating system and the times of day when it is in use.

The Feasibility Study provided a list of possible research questions for a future longitudinal survey of domestic energy use (a long list was provided in Annex A and a shorter list in the Synthesis report). It is unlikely that a future survey would be able to answer all these questions. Supplementary work (such as the suggested omnibus or Innovation Lab facilities) or independent projects that make use of the survey data or findings might offer alternative

channels through which to assess some of the research questions listed. Therefore, part of the development work was to facilitate the definition of a priority set of research questions for the survey, through the stakeholder consultation.

A question topic list was completed by eleven stakeholders, split across Government and academics. This question topic list aimed to identify areas of high priority, low priority but of interest and those that are of little importance. There are patterns both overall and by Government versus academic stakeholders. These are discussed here and summarised in the diagram below. Appendix 6 provides a full list of the potential topics that could be covered by this survey. Appendix 7 and 8 then detail how these topics could be covered in a 60-minute Wave 1 household survey as well as through subsequent wave questionnaires, self-completion modules and the technical surveyor observations.

Overview of stakeholder questionnaire priorities

	Priority	Interesting	Not important
Heating and keeping warm *Overall priority*	Heating systems present and fuel Heating controls and when/how/why they are used for temperature and time Heating behaviours Meeting needs for warmth	Instructions and guidance on how to control heating Circumstances leading to change	N/A
Hot Water *Overall priority*	Hot water systems, tank and meters present Hot water controls and when/how/why they are used for temperature and time Hot water appliances	Circumstances leading to change Meeting needs for hot water	Guidance received
Keeping cool *Academic priority*	Cooling controls present Cooling behaviours	Cooling systems present Circumstances leading to change	N/A
Other energy uses *Overall priority*	Appliances present Lighting present	Other major energy uses (e.g. charging electric vehicles)	N/A
Power generation *Overall priority*	Micro-generation systems present	N/A	N/A
Retrofit *Government priority*	Energy efficiency improvements made Future home improvements	Non-energy improvements made	N/A
Attitudes *Government priority*	Attitudes towards the home Attitudes towards climate change and environment Attitudes towards gadgets and technology	N/A	Attitudes towards energy suppliers Attitudes towards district heating
Household and dwelling characteristics *Overall priority*	Participant's role in managing energy and bills Building characteristics Problems with dwelling Householder characteristics Household routines	Energy suppliers, billing and tariffs Decision dynamics Presence of smart meter / IHD and use Participation in schemes	N/A

Heating and keeping warm is a topic area that is seen as a priority by stakeholders across Government and academia. Top priorities in this section are the heating system(s) present, including the fuels used; the controls present and how, when and why these are used to control timing and temperature; typical heating behaviours; and views of the respondent about keeping warm enough. This information would not all have to be collected though the interview and the surveyor and/or installer visit would be a more appropriate means in many cases. Information about instruction received on how to control heating is less of a priority for Government stakeholders and not important for academics but such data could be important in relation to understanding the roles of manufacturers and installers, the importance of householders' understanding the heating system rather than the heating system per se.

Energy use for hot water is less of a priority than heating. The system present, the controls present and how, when and why they are used and appliances using hot water were seen as a priority area, slightly more so for academics than Government stakeholders. Three subtopics within hot water were rated only as of interest but not essential. These were the instructions received on how to use the system, circumstances that lead to changes in behaviour and respondents' views on their ability to heat water as desired. These were of less importance across Government stakeholders, although they saw more value in questions around instructions.

Keeping cool was generally seen as less important by those responding, especially by Government representatives with academics split across priority and of interest but not essential priority. Prioritisation across subtopics was even, there were no key elements of keeping cool were identified as higher priority than any other. This area may become more important before 2050.

Other energy uses were fairly important for most stakeholders. Only one respondent from the Government contact list felt that asking questions about appliances and lighting in the home and there usage was of no importance for their policy area. Other major energy uses were seen as valuable areas of questioning, with a Government representative and an academic making comment that this area would grow in importance in relation to electric vehicles and working from home.

It will be important for the study to understand the diurnal timing of energy demand. In some dwellings this data will be available from monitoring equipment. It is something that can also be investigated in the household survey although this would involve a fairly lengthy and complex set of questions, given that time of use shifting tends to vary with end use (for example, it will be different for heating, cooking, TV use and laundry), as well as varying depending on the energy basis of shifting (e.g. tariff versus maximising the benefit of PV) and the day of the week. Variation among household members and from week to week would also need to be considered. It is therefore recommend that this is added to a self-completion questionnaire to be completed by each household member who is involved in one of the major end uses (e.g. members who do the laundry or cooking or control the heating hours). There may also be a case for asking more at Wave 2 at monitored homes, based on the monitoring data or a one-week diary.

Improvements made to the home in relation to energy efficiency were the area of highest prioritisation within the **retrofit** topic area. Other improvements made were more mixed, with planned improvements being rated as of slightly higher priority. Comments made by respondents make reference to the relationship between home energy improvements and refurbishment of the home, the former often being dependent on the latter.

There is a widely varied view on the value of asking about householders **attitudes**. Of most value is asking about the underlying lifestyle choices and what people feel they need from their home. Also important is asking about attitudes and beliefs around more general energy and

environmental issues, although this is less important for academics. It should be borne in mind that stated attitudes are not the best predictor of behaviour and objective information on the dwelling and household is likely to be more useful. Therefore it is recommended that attitude questions should not take up a lot of space in the questionnaire, given the overall 60-minute length of the recommended draft survey. The example questionnaire (see Appendix 8) therefore does not currently have time saved for attitudes, and questions are not currently detailed but instead left to DECC and stakeholders to define, depending on "issues of the day". The most useful application of such questions may be to use selected items from the national attitudes tracker survey, to check whether the panel sample remains representative.

Household and dwelling characteristics were important for stakeholders in order to understand variations in domestic energy usage. All respondents rated the household characteristics and demographics as a priority for the survey. Also of a high priority was the building's characteristics. The importance of household decision-making dynamics was even across priority and by the type of stakeholder. The importance of participation in energy programmes was clearly split across priority by the stakeholder type. Those from the government felt this was a high priority whilst academics felt it was not important, this reflects the desire for this survey to evaluate the success of government policy.

This section of the survey will also cover topics relating to paying for energy, and where possible will seek information on the respondents' energy tariffs and demand-side response products, This information may be easier to collect if the householder is able to access, and show, a recent energy bill or other communication from their supplier confirming this information. An alternative approach would be to obtain this data through data matching or asking householders to provide periodic meter readings.

It is worth noting that, in general, Government stakeholders found it difficult to consider what future priorities the survey might need to cover and start baselining. As an example, academic stakeholders were relatively more likely to believe the survey should cover behaviours around cooling than their Government counterparts who instead placed a major focus on heating. The boundaries of these topics need to be considered carefully because, for example, cooling in winter can entail opening windows with the heating on, and some insulation measures to reduce heating energy use can increase the likelihood of overheating. In the eventuality that this survey development is taken forward, it may be advisable therefore to further interrogate the potential future issues that could be covered in order that baseline measures are taken. This will help maximise the power of a longitudinal study and future-proof its findings.

Appendix 7 describes a recommended questionnaire structure for a new, longitudinal panel survey of domestic energy use. The linked Appendix 8 shows example outline questionnaires and explains the logic of the items included and possible order of questions. Together, these two appendices represent the culmination of stakeholder engagement work on key evidence needs from such a survey and a review of existing literature and research. These appendices include estimated timings for each section of the questionnaire. The weight given to different sections of the questionnaire reflects the relative priorities for different areas of evidence, and the extent of questioning likely to be needed to gather this. The timings relate only to the Wave 1 household survey, and not any content suggested for subsequent waves of the household survey, the self-completion questionnaire or the surveyor visit. A summary of the proposed timing for the key sections of the Wave 1 household survey is as follows.

- **Heating and keeping warm** ~ 35 minutes.
- **Cooling and keeping cool** ~ 5 minutes.
- **Heating water and using hot water** ~ 10 minutes.

- **Other uses of energy** ~ 3 minutes (further coverage of other energy uses included in self-completion).
- **Retrofit –** 0 minutes (all collected via self-completion in Wave 1).
- **Attitudes** (section to be defined as described above).
- **Household characteristics** ~ 5 minutes.
- **Dwelling characteristics** ~ 2 minutes.

4.3 Developing a recommended questionnaire structure

The development work undertaken over the course of this study has resulted in proposed questionnaire content which can be understood under four general headings.

- **Context**: what are households and their homes like? A factual account of a household's resources, demographics, dwelling, appliances and building services.
- **Purpose**: what are people's principles and needs? Why do they use their home as they do? A focus on the household attitudes, needs and values that might underpin their daily activities and purchase decisions.
- **Actions**: what do we do in our homes that uses energy? How and why are we doing these things? Recording details of household behaviours that impact household energy use.
- **Desires**: what do people want to change in their homes? What unmet needs do they have? Exploring barriers to change and understanding potential mechanisms to enable change.

A potential questionnaire structure is included in Appendix 7. This is structured according to topic content, which is not necessarily the logical sequence of questions. Under each topic listed in the questionnaire structure, suggested existing questions from previous surveys are noted. This is to provide DECC and others with an indication of the type of question that could be asked to collect data on each specific topic. Appendix 8 provides the full example questions which are referenced in the questionnaire structure for Wave 1.

The topic coverage, and example questions, suggested for the Wave 1 household interview has been timed at approximately 60 minutes on average. This was the survey length recommended in the Feasibility Study. A 60 minute face to face questionnaire is manageable, but it is not recommended that it extends to be any longer than this due to the burden it places on the participant. The cost calculator allows costs for questionnaires ranging from 40 to 70 minutes to be calculated. To achieve a 60 minute interview, questions within the interview have been prioritised according to what should be:

- core questions for the household interview, asked (or confirmed) each wave of the survey;
- asked only in a second or subsequent wave of the household interview;
- asked only to particular subgroups;
- deferred to a technical surveyor; or
- appropriate for inclusion in a self-completion questionnaire (which could, for instance, be given to the householder by the survey interviewer and collected by the technical surveyor at their visit, or posted back).

Altogether, the data collection is intended to provide DECC and other stakeholders with understanding of:

- the profile of how households use energy, and how this varies in different contexts;
- the drivers of the way in which energy is used, and how this relates to people's capability, motivation and opportunity to act;
- opportunities for assisting households and homes to use energy efficiently in a way that meets the range of needs they have;
- a focus on energy used in domestic settings, with non-domestic buildings excluded from the survey. This could include behaviours that transfer energy consumption to locations outside the home (e.g. to launderettes or workplace showers), and there will need to be a plan to address charging of electric vehicles at home, and mixed-use premises (e.g. where the household has a home office makes domestic use of the non-domestic part of the property).

The data collection instruments would also need to be designed to deal with new occupants joining the study as a result of moving into a dwelling that is already part of the sample. The recommended design suggests continuing the study with the new household in the previously surveyed dwelling wherever possible. The new tenant household survey would differ from the mainstream sample questionnaire as it will need to collect new household data and may need to also collect updated property data (for example, to record the installation of any new appliances or home improvement measures taken). The new tenant survey would also be adjusted to include questions asking the new householder to compare their use of energy in their new home with their behaviours and practices in their previous home. This will be interesting to explore how and why energy use changes in difference circumstances and settings.

4.4 Household reference person

A household reference person needs to be selected as the main respondent to answer the household survey. While a randomised selection approach is often taken, for example on the basis of the last birthday in the household or similar, this is not the recommended approach in relation to this survey. Instead, the aim should be to identify the person best placed to answer questions about their household behaviours and practices and who will, therefore, help generate the most reliable data. Often studies around household energy use target the person responsible for paying household energy bills and the person who deals with the energy supplier. However, as this study is less about customer-supplier relationships or perceptions of billing and more about day-to-day behaviours around the home, such as laundry, cooking, cleaning and window opening, it may be more appropriate to target the person who describes themselves as "doing the most around the home" or who tends to be in the home most often in an average week. Some checks and balances would be needed with this approach, however, to ensure there is not a gender bias and to prevent unemployed adult children being disproportionately selected. Alternatively, the index person may be the sole occupant, sole adult occupant, sole head of household or randomly selected joint head.

Qualitative research is likely to be valuable in testing this hypothesis around the most appropriate household reference person and refining the precise descriptions used at recruitment. It is also recommended that the selected approach is tested in a pilot survey prior to the main fieldwork launch to check the ability of the introduction text to select the most appropriate person and to check the survey questions are answerable by this type of household member.

In addition to the household survey it has been recommended that a self-completion questionnaire is left behind for the household to complete and return by post (alternatively it

may be possible for the surveyor to collect this should the timing of Level 1 and Level 2 data collection allow). The self-completion element will include questions where information may need to be collected from or checked with other household members.

4.5 Potential value in data linking

The Wave 1 interview questionnaire, together with the self-completion questionnaire, allows for substantial data collection but by no means everything that could be of interest. One way of supplementing the questionnaires (or at least checking selected responses) would be to link individuals, households or dwellings to other databases. In principle, this could improve the efficiency of the study and quality of data, provide historic data (including data on the situation prior to the current household taking up residence), and facilitate additional policy-relevant analysis. Aside from the possible benefit to the study, it could also serve to validate administrative data sources against the new survey's observed/surveyed data.

Appendix 5 discusses some options for such data-linking, to obtain data relating to the sampled dwelling's local area, the dwelling itself and the individuals present.

Data on the local area should certainly be sought since it will provide valuable background data. However, it will not affect the questionnaires as currently proposed because they do not include questions about the local area. Neither should area-level data be substituted for case-specific data because substantial errors could be made (e.g. an affluent family could be living in an otherwise deprived neighbourhood, or a new home may be built among much older homes). Area-level data would need to have close to complete national coverage in order to be most useful.

Data on the dwelling and/or household and/or individuals could also be useful but a number of issues would need to be addressed.

- Coverage of other databases is not always national.
- Even where there is close to complete national coverage, respondents may not all give consent to obtaining the data (especially household and individual data) and there could be bias in the giving of consent.
- In theory, some data could be obtained without respondent consent but not if consent had already been denied (and there would be reputational risks to the survey in seeking to obtain the data without consent).
- Even with consent, obtaining data can take time and is not guaranteed; hence it is risky to omit a question from the questionnaires in the hope of obtaining the data in other ways.
- Some data may not be current or accurate or recorded using the same breakdowns as the domestic energy use survey (and data are likely to have been obtained at a range of points in time, hence making linking uncertain).
- The basis of linking is not always reliable (e.g. varying ways of representing addresses or energy meters).

At present, it is therefore safest to assume that questionnaire items cannot be omitted because of the potential for data-linking but:

- Consent should be sought anyway, since data-linking can provide useful supplementary data and allow additional analysis for the subsample that give consent;
- Where consent is given, and good data reliably obtained, this could be taken into account in the questionnaire design for later years of the survey.

Further work is needed to confirm what data, AND from which databases, should be requested.

5. Resource requirements

As part of the development work, a cost calculator has been provided to help estimate the resource required to fund a longitudinal study for five years (See Appendices for details). This cost calculator is similar to the one provided by the Feasibility Study; it allows the costs for a face-to-face survey to be estimated for up to 5 years, based on:

- the sample size each year (1000+);
- the length of the questionnaire (40, 45, 50, 55, 60, 65 or 70 minutes on average);
- from Wave 2 the percentage of the sample that is longitudinal (0 – 100);
- whether a leave behind self-completion is required;
- the amount of the incentive paid to participants for the interview and for the surveyor visit.

The calculator also provides costs for surveyor visits, and allows costs to be input to include the cost of the monitoring equipment.

The monitoring equipment cost can be specified using the separate monitoring calculator to generate a cost for the preferred monitoring approach, or one of the recommended monitoring packages can be chosen.

In the recommended design, the monitoring costs based on the recommended Package 1 have been included from the “Costing Monitoring Equipment for a Longitudinal Energy Survey” final report. This covers the installation of around 4 monitoring devices in Level 2 participating homes monitoring weekly gas use, weekly electricity use and hourly temperature data for three rooms. The Costing Monitoring Equipment Cost Calculator provides details of the cost for Package 1a in the Report Packages tab. This is a cost of £207 per home, including installation costs and maintenance costs over two years. If there is no internet, the cost is estimated to be around £294. Based on 76% of homes having internet, this leads to a cost per 1000 homes of £227,880. This cost has been used to calculate the recommended options presented in this report. The recommended sample design includes 2,750 Level 2 dwellings (i.e. those receiving a surveyor visit) in Year 1 and Year 2 and then 2,052 in subsequent study years. The monitoring costs provided here have been calculated on the basis of 80% of this Level 2 sample agreeing to the installation of the Package 1 monitoring equipment.

The cost calculator provides indicative costs (not precise prices) based on a detailed survey specification. A list of the assumptions that have been made to produce the costs are included as part of the cost calculator. The cost calculator only provides costs for face-to-face interviews, it does not provide costs for other modes such as telephone or online interviews.

Department of Energy & Climate Change
3 Whitehall Place
London SW1A 2AW
www.gov.uk/decc
URN: 15D/409