Department of Energy & Climate Change

Smart Metering Early Learning Project:

Consumer survey and qualitative research

Technical Report

Report prepared by Ipsos MORI

March 2015

URN 15D/083

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

Contents

1. Introduction

1.1 Background to this document

This Technical Report sits alongside the main research report for the Smart Meter Early Assessment Study. It contains the background to the project phases and documents the following processes for both the qualitative and quantitative stages:

- Sampling
- Development of fieldwork materials
- Fieldwork stages
- Analysis

The final section in this report summarises the accompanying documents presented in the Technical Appendices (i.e. the full questionnaire, qualitative interview guides, recruitment specifications etc.).

1.2 Background to project (as stated in the main research report)

Smart meters[1] are the next generation of gas and electricity meters offering a range of intelligent functions. The Government's vision is for every home and smaller business in Great Britain to have their existing meters replaced by smart electricity and gas meters by 2020. The roll-out of smart meters will play an important role in Britain's transition to a low carbon economy and help meet some of the long term challenges faced in ensuring an affordable, secure and sustainable energy supply.

Smart meters are expected to deliver a range of benefits to consumers, which include bringing an end to estimated billing and providing consumers with near-real time information on their energy consumption through the use of an In-Home Display unit (IHD[2]) supplied with the smart meter. It is envisaged that this will provide consumers with better control over their energy use, help them to budget better and help make switching between suppliers smoother and faster. Some energy suppliers have already started installing smart meters in what this report refers to as the 'early roll-out stage'.

DECC's Early Learning Project (ELP) was set-up to explore the experiences of domestic consumers involved in this early roll-out stage and to investigate the outcomes they have experienced. DECC commissioned Ipsos MORI to undertake research which forms one part of the ELP. This research has involved a national quantitative survey with domestic smart meter customers (and a matched sample of legacy meter customers) as well as follow-up in-depth qualitative interviews with smart meter customers. The analysis phase of the research involved using a number of statistical techniques to explore the survey findings in greater depth, including matching techniques to enable comparisons between smart and legacy (traditional) meter customers, and Key Drivers Analysis (KDA) to examine the extent to which

[1] Smart meters are able to communicate directly with energy suppliers by sending and receiving information about the amount of energy being used.

[2] IHD's allows consumers to see what energy they are using and how much it is costing in near-real-time, including information about the amount of energy used in the past day, week, month and year.

different elements of the smart meter customer journey are related to subsequent behavioural and attitudinal outcomes.

In total, 4,016 customers aged 18 and over were interviewed face-to-face across Great Britain. These interviews took place between 4th October 2013 and 1st February 2014 with pre-selected named customers provided by two energy suppliers. This included a representative sample of 2,037 smart meter customers who had received a smart meter between 1st April 2011 and 28th February 2013. The other 1,979 interviews were conducted with a sample of legacy meter customers selected to match, as far as possible, the profile of the smart meter sample. In February 2014, qualitative in-depth interviews were conducted with 79 of the smart meter customers who had already been surveyed. These were commissioned to understand in more detail the experiences of these customers throughout the different stages of their smart meter customer journeys.

2. Quantitative sampling

This section sets out the approach to sampling households for DECC's Smart Meter Early Assessment Study. It presents the principles underlying the sampling process; the information known about the survey population; the inclusion and exclusion criteria; the agreed sampling approach; and the final make-up of the survey sample.

2.1 Overall objective

The aim was to create a representative sample of the two suppliers' smart meter customers, meaning customers who are adults aged 18+, living in Great Britain, and who are domestic premises customers.

The profile of this sample was matched in a second sample of the two suppliers' legacy meter customers. The latter is therefore not representative of any current population, but was selected to resemble the smart meter sample as far as possible.

The legacy meter sample was designed to provide a counterfactual to the smart meter sample. In an ideal world we would know how the smart meter households would have behaved had they not had smart-meters installed; the legacy meter sample was constructed in an effort to represent this notional counterfactual population.

The target sample size, to cover both smart meter and legacy meter customers, was 4,000. This number of interviews was selected to allow robust subgroup analysis of groups of customers within the smart and legacy meter sample and within the sample of customers from each energy supplier (e.g. based on demographics, attitudes and behaviours).The final target was slightly higher in order to ensure an even spread interviews across the postcode areas (see Sections 2.7 and 2.8 for further detail).

2.2 Survey population

The survey population consisted of the two suppliers customers, drawn from their databases. The raw files contained records for over 13 million properties, and included over 250,000 smart meter customers. The first step in producing the sample for the main stage was to receive anonymised datasets from the two suppliers that provided key details of their customer base, but with no identifying variables (e.g. customer name or address). This provided a sampling frame for the survey.

The main steps taken by Ipsos MORI on receipt of these files are presented below.

2.3 Re-formatting raw sampling frame files (Step 2 above)

The raw files provided by the two energy suppliers were reformatted for the following reasons:

- To merge variables which were provided by both Company A and Company B but with unique names
- To remove variables which were either not necessary for sampling or which did not have good coverage for both Company B and Company A samples (a complete list of these is provided below)
- To add further variables which could help improve sample (a complete list of these is provided below)
- To aid in sample frame cleaning (e.g. removing duplicates, identifying invalid data)

Variables added to sampling frame

The following variables were added to the sampling frame data to help ensure the final sample comprised a distribution of households by a range of *geographical and socio-economic measures as well as energy consumption:*

- **Administrative Geography**
  - Postal sector, Government Office Region, Local Authority, Census Output Area[3]
  - These variables were used to analyse the geographic distribution of smart meter installations. They were then used to cluster the sample into small local area interviewer points and to ensure these cover a representative spread of customers across the population as a whole

- **ACORN classification**
  - The ACORN classification segments postcodes and neighbourhoods into 6 Categories, 18 Groups and 62 types. The classification is based on demographic and social factors such as tenure, property type, property value, ethnicity, family structure, proportion of students and company directors, etc.
  - It is a very useful tool for segmentation, stratification and in this case in profiling in order to check for representativeness of the final sample

- **Population statistics**: residential property count
  - The statistics used are the Royal Mail residential household count at postcode level, summed up to other levels of geography.
  - These are used to determine the penetration rate of smart meters at various geographic levels

[3] Lowest geographical level at which census estimates are provided – created from clusters of adjacent postcode units and designed to have similar population sizes and to be as socially homogenous as possible based on tenure and property type

Energy consumption data: annualised energy consumption deciles were added and later converted into quintiles for the final groupings

- The suppliers gave consent for meter point electricity and gas consumption data collected for DECC's sub-national consumption statistics[4] to be used by DECC to provide a single annualised figure for each meter. DECC are already provided with annual estimates of consumption in order to produce and publish sub-national statistics on average consumption and it would have taken a long time to complete the complex calculations required to re-create annual consumption from actual and estimated reads provided from the energy suppliers for the wider Smart Metering Early Learning Project . The existing estimates were therefore used to minimise the amount of time between the period in which smart meter installations (covered by the energy supplier datasets) took place and the period in which respondents were interviewed about their experiences.
- Although 2011 data were available, the 2010 data were used because they covered up to the point at which smart meters began to be installed.
- This does present a limitations due to potential changes since then that could have affected consumption (e.g. property renovations or extensions, household composition changes, new occupants, etc.), but 2011 data included consumption after smart meters were installed. However, it was likely that both the smart and legacy samples would be affected in similar ways by this time lag. The consumption data were also linked to the meter ID at each property rather than the occupant; we knew that the property structure/size is a key driver of consumption, and very generally speaking we could assume that similar types of household composition might live in a certain type/size property even if there had been changes in the occupants of the property.

2.4 Eligibility criteria for inclusion in master sample file (Step 3 above)

Based on discussions with DECC, Ipsos MORI set the following conditions for inclusion of a household within the smart meter sample:

Inclusion criteria	Rationale
1. Smart meter for electricity installed between, and including, 1st April 2011 and 28th February 2013	This study was commissioned to examine installations that occurred during the Foundation Stage of the Programme (which started in April 2011). Cut-off dates were chosen to ensure installations from a range of different time periods were covered. This allowed for an assessment of the experience of impacts over time whilst still capturing data from more recent installations where installation experiences would be less affected by recall issues. The end cut-off

[4] https://www.gov.uk/government/statistics/regional-energy-data-guidance-note

	date reflects the latest installation dates for which data was available at the time of the information request to energy suppliers. During this period it was not possible to have only a gas meter installed for technical reasons, so sampling was based on the presence of an electricity smart meter.
2. Properties with both smart electricity and gas meters allowed, but properties with only a smart gas meter and not a smart electricity meter excluded	As it was not possible to have just a gas smart meter installed, any records which indicated a household had a gas smart meter only were considered to be erroneous entries; only a few such records were identified.
3. Pre-existing smart meters (installed during the cut off period) allowed. Pre-existing defined as smart meter installation date prior to customer start date at property.	New occupants in properties that had already had a smart meter installed (during the cut off period) were included to enable data to be collected on their level of awareness of the smart meter, their understanding and usage of it (this is discussed further in section 7 below).

The following conditions were therefore the basis for exclusion from the smart meter sample:

1. Smart meter installation date outside period stated for inclusion criteria 1. above

2. Information in sample files suggesting household has a gas-only smart meter (this is indicated where the property is gas only but flagged as having a smart meter installed, or where property is listed as having both gas and electricity but is flagged as only having a gas smart meter. We know that gas only installations had not taken place during this period due to technical reasons so this indicates an error in the sample file). As a result, given that there were only electricity only or electricity and gas smart meter installations, all sampling was carried out on based on the electricity meter information.

The following conditions were the bases for exclusion from both the smart meter and legacy meter samples[5]:

1. No annualised meter consumption from 2010

2. If annualised energy consumption in 2010 is greater than 4.5 standard deviations greater than the mean consumption

3. If annualised energy consumption in 2010 is less than 1.5 or greater than 2.5 standard deviations than the mean consumption then cap at those figures[6]

[5] The method used here differs from that deployed during Energy Consumption Analysis (ECA). This is because the current method was based on annual estimates rather than actual meter read data that was used in the ECA.

[6] Where the consumption figure was greater than the mean plus 4.5 standard deviations the record was excluded. Where the consumption figure was less than the mean plus 4.5 standard deviations, but greater than mean plus 2.5 standard deviations, or less than mean minus 1.5 standard deviations

4. Pre-payment meter customers as smart meters had not yet been installed in sufficient numbers for these customers. (DECC commissioned a separate piece of work to explore the experiences of customers with smart-type prepayment meters).

2.6 Key sampling decisions for smart meter sample, suggested approach and rationale

A number of key decisions were made in relation to the final sample splits between different customer groups and rules around which records should be included and which excluded. The rationale for these decisions is outlined below:

50:50 split between suppliers
A 50:50 split was chosen in order to maximise the range of different approaches/ experiences being captured in the data. The data were re-weighted to the real split between the two suppliers in the final event in order to ensure that the final data were representative of the early roll-out population.

50:50 split between smart and legacy meter customers
A 50:50 split was chosen because it would help to increase the robustness of matches and the number of appropriate legacy customers available for matching. Moreover, and as with the split between the suppliers, a 50:50 split between smart and legacy customers in order to maximise the range of different approaches/ experiences being captured in the data.

Inclusion of legacy IHD customers
In order to robustly assess the impact of a smart meter and IHD on energy-related attitudes and behaviours it was important to include legacy customers who already owned a clip-on IHD as excluding this group may have over-inflated the impact of the smart meter (with IHD) roll-out. Beyond this theoretical rationale, there would also have been a practical challenge of identifying (and then excluding) legacy meter customers with an IHD before, or even at the start of, the survey. The survey avoided asking about IHDs until an initial set of attitudinal and behavioural statements had been asked. This was to prevent questions about IHDs affecting the collection of data about potential impacts.

Inclusion of customers who had moved into a property which already had a smart meter installed
These customers were included because it was important to find out whether householders in this position were aware of the smart meter installed in their property, whether or not they were using an IHD, and whether or not they were consciously benefiting from the installation or not. Moreover, excluding this group would over-inflate the difference in energy use between smart and legacy customers.

---

the record was kept in. However, in these cases the consumption figure was capped at mean plus 2.5 standard deviations (or mean minus 1.5 standard deviations) for decile calculations. The upper cap of plus 2.5 standard deviations was selected as it allowed sufficient variation (for over 99% of cases) whilst reining in the more extreme values

No variations of sampling fractions by time of smart meter installation
The respondents selected for the final sample were selected randomly across the smart meter installation period. The result was a spread of households from early to later installation dates. This allowed for the possibility of later analysis by installation date. Furthermore, the variable was not present for the Legacy sample and therefore could not be used as a control variable in the final sample selection.

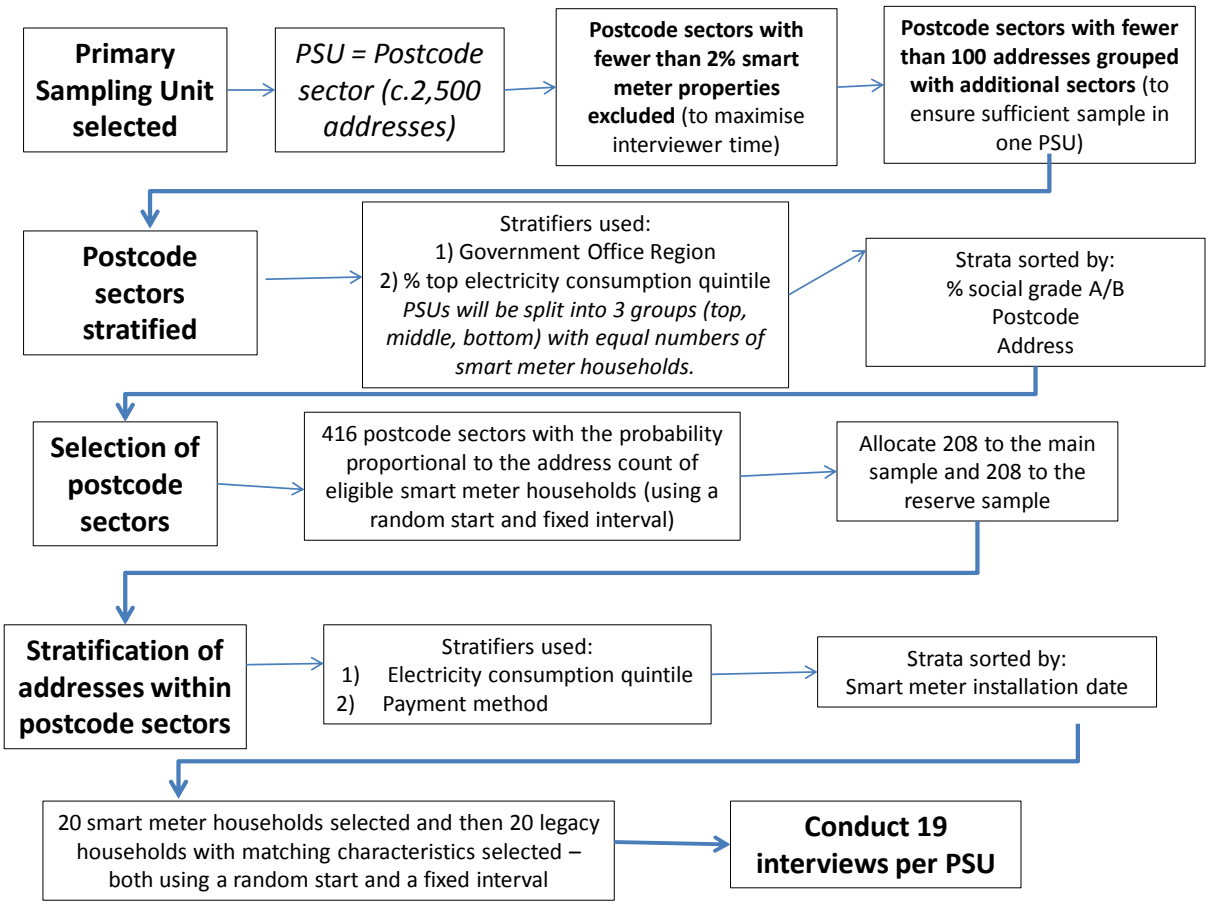
Legacy meter customers sampled from same areas as smart meter customers
No reasons were found for excluding legacy meter customers from the same areas as smart meter customers, either when considering possible sources of difference before receipt of sample, or at the point when the sample was received and analysed.

Excluding the rural/ urban split from the stratifying variables
The PSU selection process ensured that the postcodes selected were representative of the split of urban and rural households and areas. A check was run on the final selected sample to ensure sufficient numbers of rural and urban areas were included in order to be nationally representative

Exclusion of pre-payment customers
Pre-payment customers were intentionally excluded. The two energy suppliers which provided the sampling frame for this research were not generally installing smart meters for prepayment customers at the time of the research. The few instances that were identified in the sampling frame are most likely where a household had smart prepayment as part of a small-scale trial and therefore there would not be enough of these households to sample for this research and they are unlikely to be sufficiently representative of wider installations.

2.7 Summary of sampling approach

The flow diagram below sets out the main stages in the final selection of respondents for interview.



Deadwood

When conducting a pre-selected survey a proportion of the respondent records selected to be part of the sample are likely to have to be excluded either at the formatting stage, or later on in field. This may be for reasons such as the status of the respondent changing from legacy to smart meter customer in the interim period, incorrect addresses recorded in the sample files or finding the property to be vacant.

It is necessary therefore at the outset to estimate how many respondents will be excluded and to increase the number of records selected in the sample to cover the losses. This is referred to as survey deadwood. The weighting applied to the survey results corrected for any bias in the profile of records that are excluded for any of these reasons. Dealing with the exclusions in this way means the achieved number of interviews is unaffected as interviewers were provided with more addresses to achieve their targets. The exclusions listed below were/ are all treated as deadwood, and the number of sample records selected were increased to account for this.

Deadwood calculations

The deadwood estimate took into account the following issues:

Movers or switchers between May – August 2013 – between 10% and 16% (depending on supplier): This is any customer (smart or legacy) that has moved property, or switched supplier, between the time the original sample was issued showing their postcode sector and the final matching stage with their current address. This was estimated with help from the suppliers.

Known smart meter installs between 1st March – August 2013 – 3%: Legacy meter customers who suppliers know have had a smart meter installed since issuing the original sample. These customers will be dropped as their smart meter installation date will be after the cut-off date for the period of interest. The 3% figure was estimated by DECC from roll-out figures.

Claimed smart meter installs after 1st March 2013 – 2%: Legacy meter customers who claim during the survey that they have had a smart meter, or new meter + IHD, installed since on or since 1st March. These customers are dropped because their smart meter installation date will be after the cut-off date. The 2% figure is a rough estimation. Although it was known that around 7% in a previous smart meter tracker survey claimed to have a smart meter when they are likely to do so, this was based on any time period and did not ask about whether an IHD was provided at the same time.

New occupants who have switched supplier – 4%: This includes any smart or legacy customers who have moved between the address matching stage and fieldwork (could also include a few who may have moved before but not told their supplier). It is also only switchers within this group that will be dropped as new occupants were eligible to be surveyed if they were still customers of the targeted energy suppliers. The estimate of 2% is taken from figures on how many people move each year and figures on how many people switch supplier.

Standard in-field deadwood – 5%: This covers all other, fairly standard, reasons records may be ineligible and applies to both smart and legacy records. This could be an address that no longer exists because a block of flats has been knocked down etc. The 5% estimate was recommended by the Ipsos MORI Field team based on their experience working with pre-selected samples.

Final estimates

Table 1 sets out the target numbers of addresses that were issued overall.

Table 1: Target sample design

Sample design	**40 minutes average**
Issued addresses	**10,608**
PSUs	208
Issued addresses per PSU	**51**
Estimated interviews	**4175**
Estimated interviews per PSU	**20**
Estimated adjusted response rate	52.4%
Ratio (issued addresses per estimated interview)	**2.5**

2.8 Selection of PSUs and addresses

The sample was clustered in order to minimise interviewer travel costs. Essentially clustering involved drawing the sample in two stages:

1. Drawing a sample of defined geographical units ('clusters' – e.g. postcode sectors);

2. Drawing a sample of households within these units.

Selection of PSUs

Postcode sectors as a basis for PSUs

We proposed that the primary sampling units (PSUs) used for the survey were postcode sectors (standard on most government surveys).[7]

[7] These are areas comprising on average 2,500 addresses which are of a manageable size for interviewers to get around and sufficiently large to avoid overly large damage to survey design effects from clustering due to intra-cluster homogeneity (smaller areas typically being associated with higher levels).

Distribution of Smart Meters in the sample

The distribution of smart meters within the sample allowed us to keep to our original specification that assumed a clustering level of 1 in 50, noting that we would issue smart meter and legacy meter addresses in the same PSUs, meaning that a clustering level of 1 in 100 smart meter households was sufficient to deliver this.

Distribution of Smart Meters across PSUs

Owing to the nature of the Smart Meter roll-out to date, the distribution of Smart Meters across PSUs did vary, with a higher proportion of the Smart Meter sample coming from certain areas of the country – notably the East Midlands and the North West of England. As a result the sample distribution was skewed when compared to general population, but was representative of the smart meter population.

Exclusions

Postcode sectors that contained less than 2% properties with smart meter customers were excluded: 7,754 postcodes, covering 181,404 properties, were therefore excluded at the outset, leaving 1,474 eligible postcodes and 65,404 properties from which to sample. The approach was designed to maintain a high level of coverage of smart meter households whilst removing large numbers of the postcode sectors prior to sampling, simply because large numbers contain none or few meters (based on where they have been installed). In so doing, the approach greatly increased the survey's efficiency as the frame was restricted to more concentrated areas. The final sample was therefore geographically biased in that it was non-representative at a regional level, but was representative of the Smart Meter population at that time.

Any remaining postcode sectors with below the minimum number of 100 households were then combined with neighbouring sectors prior to sampling. This was done to ensure that the areas were of a sufficient size to limit the potential for sample efficiency damage from intra-cluster homogeneity.[8]

The postcode sectors/sector combinations formed the sample frame of PSUs for the survey.

Use of PPS to select PSUs

Probability Proportional to Size (PPS) methods were used in the selection process based on a weighted size measure. This measure was calculated based on the agreed stratification profile and number of each stratum group in each PSU. PPS allows the selection of equal numbers of sample units (e.g. households) from a set of selected PSUs (e.g. postal sectors) regardless of the size of each PSU. This method allows each fieldwork assignment to have a constant size and ensures that each sample unit has an equal probability of selection by increasing the probability of selection for larger PSU whilst decreasing the probability of selection of sample units within those larger PSU, and vice versa for smaller PSU. In short, PPS is an efficient method able to deliver equal sampling fractions within strata and hence optimise the efficiency (precision) of the sample.

[8] Intra-cluster homogeneity is a measure of how similar each sample unit (e.g. household) is to each other within a PSU (e.g. postal sector).

Reserve sample

Following these methods a sample of 208 PSUs was selected alongside a reserve sample to provide contingency. However, the required sample size was achieved from the main sample and so in practice further sampling from the reserve sample was not required.

Selection of households

The PSU selection probabilities and household selection probabilities ensured that within each stratum all households would have an equal chance of selection. A list of the customer IDs of those selected at random was then sent to the suppliers to append the contact details of the customers. These formed the sample issued to interviewers.

Selection of respondent for interview

In total, 8,035 addresses were issued across all interviewers in the 208 postcode sectors. In each postcode sector there are 39 issued addresses, and interviewers aimed to achieve 19 interviews from these addresses. This was based on a response rate of 52.4%, estimated using response rates from similar previously conducted surveys. In the field, interviewers approached the named person in the first instance; however the survey started by asking questions to ascertain the most appropriate person within the household to complete the interview with.

In the first instance the interviewer ascertained if they were speaking to the person named on the sample as a customer of Company A or Company B for their electricity:

- If they were speaking to the customer named on the sample, the interviewer checked whether this person was best placed to answer questions on their household's electricity and(if applicable) gas usage
    - If this person was the right person to talk to, then the interview could proceed
    - If it was someone else in the household then the interviewer would speak to the alternative person, check whether they were aged 18 or over, and were at least jointly responsible for decisions about household energy bills/use. As long as the person met these criteria then the interview could proceed

- If they were unable to speak to the customer named on the sample, the interviewer checked whether the named person still lived at the property
    - If the named person still lived at the property, then the respondent either arranged an alternative date for an interview, or checked whether the alternative person was aged 18 or over, and were at least jointly responsible for decisions about household energy bills/use. As long as the person met these criteria then the interview could proceed
    - If the named person no longer lived at the property, and there was a new occupant, the interviewer would check if the new person was a customer of Company A or Company B for their electricity, aged 18 or over, and at least jointly responsible for decisions about household

energy bills/use. As long as the person met these criteria then the interview could also proceed

At no point were referrals allowed. Interviews could only occur with occupants of the properties specified in the sample.

2.9 Approach to stratification

Stratification was required to ensure that the sample of smart meter customers was as representative as possible of the smart meter populations of the two suppliers across Great Britain.

Ultimately stratification was required to match the profiles of the two samples in order to help discount third variable explanations of differences between the smart and legacy samples on key survey variables. For this to be effective stratification variables needed to correlate with both key survey variables and smart-meter use.

From a review of the variables included in the raw data files, we recommended the following variables for use in the stratification due to their likely correlation with the outcomes being measured (that is reduced energy consumption, improved energy literacy and greater engagement in the energy market, all of which are hypothesised to be achievable through smart meter usage). The following variables were chosen on the basis of our knowledge of previous research conducted by Ipsos MORI, as well as others:

- **Annualised consumption data***: this was one of the main outcomes being measured and so any known data should be included in the stratification approach*
- **Payment type***: we know this is often correlated with engagement with energy bills and usage levels as well as household income, tenure and work status*

The other variables included in the final master file (listed in full in Appendix 3) were in the main factual measures such as customer, property and meter IDs, and customer start dates, and which were less likely to have an impact on energy use, literacy and engagement. The remaining variables also included a range of profiling variables (such as ACORN which was appended to the file, as stated in section 2.3). While this information was used to check the profile of the final drawn sample it was not felt to be vital to the stratification process which was limited to the most valuable variables. For example, the ACORN variable provided a further check on the likely average household income level of the area (it can be used as a proxy for this). However, it did not need to be used during the stratification process as census data on social grade was used for this purpose.

There were two main stages to the stratification process:

1. **Stratification of the PSUs** to ensure the selected postcode sectors are spread around Great Britain and are broadly representative of different population profiles, and;

2. **Stratification of addresses within the selected postcode sectors** to ensure the smart meter households selected are representative of the range of customers included in the original population. We present the recommended approach to these two stages below.

Stratification of PSUs (postcode sectors)

Stratifier 1: Government Office Region: to ensure the sample was representative of the spread of the smart meter population across the different regions of Great Britain.

Stratifier 2: Percentage of households in top electricity consumption quintile: (based on capped annualised consumption data) to ensure the postcode sectors chosen cover a range of household consumption profiles. We recommended using the top quintile (rather than decile) to reduce the number of cells in the sample as it would prove difficult to select a household in every cell if this is divided too finely. For the same reason, it would have been inadvisable to use both gas and electricity consumption values. We selected electricity consumption given that the smart meter customers would have had at least an electricity smart meter (but may not have had a gas smart meter).

Ordering the strata by the percentage of households in social grade A or B (based-upon census data) helped to ensure that the PSUs were broadly representative of the region. Following this, 416 postcode sectors were selected with the probability proportional to the address count of eligible smart meter households (using a random start and fixed interval).

Stratification of addresses within the selected postcode sectors

Stratifier 1: Capped electricity consumption quintiles: as explained above, quintiles are preferable to deciles so that fewer cells are created. As such the deciles were combined into quintiles for the final sampling.

Stratifier 2: Payment type: we recommended setting this as a binary variable based on the customer either making direct debit payments or not. Again, this was to prevent too many cells being produced.

Stratifier 3: Smart Meter Installation Date: the clusters were then also **sorted by smart meter installation date**.

Producing a matched legacy sample

The profile of smart meter customers based on these variables formed the target profile for both sample groups: i.e. the stratification scheme for the legacy group was proportional to the smart meter sample.

2.10 Quality Assurance

Once the records had been selected, a number of checks were run as a guarantee that the selection process had been correctly followed and that no records had been incorrectly included or excluded.

2.11 Final sample selection figures

Over 13 million records were eligible for selection; running a randomised selection process (based upon the splits and selection criteria discussed above). Just over 21,000 records were requested from the suppliers. Records were returned unless one of the following problems was identified: Customer drawn for the legacy meter sample, but had since had a smart meter installed

- Customers who were no longer with either supplier (i.e. those who had left in the intervening time between sample selection and delivery)[9]
- Deceased customers
- Duplicates
- Records that were identified as problematic due to incomplete or erroneous information

At the end of this process, the final clean sample consisted of 18,142 records, with over 4000 smart meter records in both the main and reserve batches. Suppliers were not able to identify whether customers had been selected for the main and reserve sample, so were unaware which customers were approached or subsequently participated in the survey.

[9] Ex-customers accounted for 73% of the records that were dropped at this stage

3. Development of quantitative fieldwork materials

3.1 Introduction

There were several stages to the development of materials for the various stages of the quantitative side of the project. All materials were tested prior to their use as part of the research, in order to test comprehension and ease of use in field.

3.2 Questionnaire development

The questionnaire was formulated over several stages and involved inputs from DECC, Ipsos MORI and academic researchers based in the Environmental Change Institute (ECI) in Oxford.

Draft questionnaire

Questionnaire development began with a list of topics and key areas to be covered; discussions were had concerning the final aims of the wider Early Learning Project and how the survey data would feed in to the project. From these initial discussions broad themes were formulated and these formed the basis for the sections of the survey itself.

Ipsos MORI developed questions for these distinct sections, using previous related surveys and existing information on smart meters in order to guide the use of terminology and the development of the key concepts to be covered as part of the survey. A draft questionnaire was then formulated, along with timings, and shared with the teams at DECC, and the ECI to gather initial feedback. Further drafts were then developed taking on board feedback from these teams and the energy suppliers involved until a version was ready for testing with consumers.

Cognitive testing

Once the questionnaire had received initial sign-off it was tested with members of the public in order to test comprehension of the terminology and concepts covered by the survey, as well as to gather general feedback on the other materials that would form part of the interview (see Advance letter and Data linking leaflet below).

The cognitive interviews produced feedback on a variety of issues which helped inform the draft questionnaire to take to piloting. For example, the cognitive interviews showed some areas where the question wording needed minor edits to improve understanding for respondents or helped to add possible response options to incomplete pre-code lists.

A wide range of individuals of different ages and backgrounds were recruited for the cognitive interviews. In order to test all sections of the questionnaire, quotas were set on legacy and smart meter customers. In total, 24 cognitive interviews were conducted. Face to face interviews took place in Nottingham, London and Derby,

and telephone interviews with respondents from across the country. The customers selected were a mixture of smart and legacy meter customers with either Company A or Company B. Some smart meter customers were recruited directly by the energy suppliers, whilst other smart meter customers and the legacy customers were free-found (i.e. randomly recruited).

Interviews were conducted through a mixture of face to face and telephone contact, with a selection of key questions covered. All interviews began by covering the contents of the Advance letter and concluded with a discussion of the content of the Data Linking Leaflet.

Testing the Advance letter and Data linking leaflet

The **Advance letter** was designed as a form of “warm-up”, alerting selected households to the survey and advising them of what it would cover and why it was being conducted in the first place. During the cognitive interviews respondents were asked about whether there was anything within the letter that was unclear, whether it prompted any questions, and if there were any improvements that they felt could be made.

The **Data linking leaflet** was designed to inform all those who completed the survey about how the answers that they had given would be used by DECC, and in particular about how, if they had given permission, their answers would be linked to other data already held about their properties. The leaflet also contained contact details if they changed their mind about either having or not having their data linked.

As with the Advance letter respondents were asked about anything in the leaflet that was unclear, whether there was anything that they felt should be added or whether it prompted further questions or concerns.

Once the cognitive interviews had been concluded, feedback was collated into a report for DECC. Amendments to the materials were then discussed and implemented before the survey was piloted.

Pilot

A pilot survey of 81 interviews with a mix of smart and legacy meter customers was run prior to sign-off. The interviews were conducted in areas identified by suppliers as having concentrations of smart meter customers because a list of all smart meter properties was not available at this stage. The pilot was a quota survey meaning that respondents were free-found by Ipsos MORI interviewers, and were surveyed as long as they fulfilled a number of criteria (i.e. to control the number of interviews conducted with smart and legacy meter households, and with customers of the different energy suppliers). The main purpose of the pilot was to test the survey in a live environment to refine the question wording and any aspects of the interviewer approach.

Prior to the pilot interviewers were briefed by researchers at Ipsos MORI about the purposes of the survey and the main aims of the pilot. The briefing also covered the

content of the survey materials and interviewers were given the opportunity to ask about any questions.

The pilot represented the first point at which the questionnaire was to be tested using Computer Assisted Personal Interviewing (CAPI), and it therefore also represented an opportunity to ensure that the survey had been correctly programmed to replicate the final paper version.

Pilot feedback

As with the cognitive interviews feedback was collated into a report for DECC. Some small amendments relating to routing and terminology were made, and the timings for different sections were reviewed. At this point the questionnaire and accompanying CAPI script, along with the other fieldwork materials, were signed-off ready for fieldwork.

4. Quantitative fieldwork

4.1 Key fieldwork facts

Quantitative fieldwork launched on 4th October 2013 and concluded on 1st Feb 2014.

4,016 interviews were completed in total: 2,037 with smart meter customers and 1,979 with legacy meter customers. Details of response rates can be found in Section 4.6. Each survey lasted approximately 40 minutes.

Interviews were conducted by the Ipsos MORI face-to-face interviewing team. Interviewers were based within the region in which they were assigned interviews, with fieldwork coordinated by a central London-based team

4.2 Interviewer pack

Before the launch of fieldwork all interviewers were sent a pack that included:

- Interviewer notes
- Advance letters
- Spare warm-up letters
- Contact sheets
- Showcards
- Data linking leaflets

Interviewer notes

The interviewer notes contained a summary of the survey as a whole, listed all of the points covered during the briefing, and highlighted the key questions in the survey. They acted as a first point of reference and answered FAQs that interviewers may have about the survey.

Advance letters

Interviewers were sent the advance letters to mail themselves. This approach was preferred to a central mailing-out the letters because a) the fieldwork start was slightly staggered with launches taking place at slightly different times from region to region and b) to achieve the optimal response rates interviewers were aiming to follow-up within a week of letters being sent-out, and in being able to mail them out themselves they had a little more flexibility in when this would take place.

Showcards

All interviewers received a paper pack of showcards. Some questions had a long list of possible pre-coded responses and being able to see this list would assist customers in answering questions. All showcards were numbered and linked to the question in the main survey.

Data linking leaflets

All interviewers received sufficient data linking leaflets to leave with customers who participated in the survey.

Interviewer briefings

Once they had received their pack, interviewers also participated in a briefing conducted by a member of the research team.

4.3 Efforts to boost response rates

In order to ensure that the response rate was as high as possible, in addition to sending out advance letters at least one week before attempting to conduct an interview, all interviewers had to make a minimum number of attempts to contact a household before listing it as a dead lead:

- Record at least 5 attempts to contact a household before abandoning an address
- Make at least three calls to an address in the evening (6pm or later) and/ or weekends
- Leave a minimum of 3 weeks between first and final calls for all non-contactable addresses

In addition, in the final month of the survey non-completed addresses were re-allocated to new interviewers to try to maximise chance of getting a response.

If a household refused to participate then no further attempts were made at re-contact, either during the main stage of fieldwork, or when working through the non-completed addresses at the end of fieldwork.

4.4 Tracking fieldwork

The Ipsos fieldwork team made use of a special portal that allowed all interviewers to upload progress in contacting the addresses that they had been issued. In recording interim outcomes the fieldwork and central research teams were able to closely monitor overall progress and identify any problematic regions.

4.5 Responding to customer questions and complaints during fieldwork

A number of steps were taken to ensure that customers were given the opportunity to contact the survey team should they have a question, comment or complaint that they wished to lodge.

- A telephone helpline was set up to receive opt-out calls, answer any queries about the research, and take appointment bookings
- An e-mail address was also supplied for the same purpose

Both were managed by the central research team and where necessary a follow-up phone call was made or e-mail written to ensure that the matter was properly dealt with.

4.6 Final fieldwork outcomes

Table 2 below shows the outcomes for the fieldwork stage. In total, the target of 4,000 interviews was achieved; however, the final overall adjusted response rate of 45.01% was lower than the target of 52.3%. The adjusted response rate is calculated from the total number of interviews achieved divided by the eligible addresses issued. There were a number of factors contributing to the lower response rate, including a higher than anticipated refusal rate and fewer ineligible addresses than estimated via the deadwood calculations (set out in section 2.7). The table below also shows, however, that 100% of the addresses issued were worked by interviewers. This means that despite the over-issuing of addresses (due to high deadwood calculations which did not come true in the field), the full amount of sample issued was targeted for interview meaning we do not need to worry about bias being introduced by missing some addresses.

Steps were taken from the outset, and also repeated during fieldwork, to prevent refusals as far as possible. This included thorough interviewer briefings, advance letters sent to respondents, scripted re-assurances about where the sample had come from and that the survey was being conducted on behalf of the UK Government who could use the results to review policies around energy. While the majority of refusals were recorded to be due to standard reasons (such as being too busy to be interviewed) it may be that the timing of the research, which was over the same period that energy companies were prominent in the media due to price rises, may have had an impact.

Table 2: Fieldwork outcomes

	Final
Target Number of Interviews	4,000
Target Adjusted Response Rate	52.3%
Final number of Interviews Achieved	4,016
Final overall Adjusted Response Rate	45.01%
Final Smart Meter Customer Adjusted Response Rate	42.41%
Final Smart Meter Customer Adjusted Response Rate	48.05%
1st Issue Adjusted Response Rate on Completed Points	42.9%
Number of Final Outcomes - 1st Issue	9,287
% of Addresses Covered	100.00%
% of Target Achieved	100.50%
% Ineligible Sample	4.01%
% Refusals - 1st Issue	41.41%
% Non Contacts - 1st Issue	12.37%
Final number of Reissue Interviews Achieved	192
Final Shortfall / Overage	-16

Interviews by customer type:

Overall: 4,016

Smart meter customer: 2,037 (51% of total)

Legacy meter customer: 1,979 (49% of total)

A breakdown of completed interviews by region is provided overleaf (Table 3) as well as details of the final re-contact and data linking consent rates (Table 4).

Table 3: Interviews by region

	OVERALL	South East	East Anglia	Midlands	North West	North	South West	London	Scotland	Wales
Target	4000	***58***	***423***	***1615***	***904***	***385***	***115***	***250***	***154***	***96***
Achieved	4016	66	433	1579	846	398	127	252	189	126
Percentage of individual target	*100*	*114*	*102*	*98*	*94*	*103*	*110*	*101*	*123*	*131*

Table 4: Data linking consent: overall and by customer type

	Overall		Smart meter		Legacy meter	
	Data linking	**Recontact**	**Data linking**	**Recontact**	**Data linking**	**Recontact**
Achieved	3508	3260	1786	1646	1722	1614
Percentage of individual target	*87*	*81*	*88*	*81*	*87*	*82*

5. Analysis

The analytical stage of the project comprised several phases and forms of analysis.

The final report represents a synthesis of the qualitative and quantitative stages, bringing together survey results, statistical analysis and qualitative data. The information below covers the rationale for the decision taken relating to the methods used in analysing the quantitative data.

5.1 Data Processing and Coding

All data processing and coding was undertaken internally at Ipsos MORI. The main data processing aspects of the project were the drafting and production of topline reports, data tables and an SPSS file. Responses to “Other (please specify)” codes at either prompted or unprompted questions were assigned codes. These responses were either back-coded (i.e. assigned a previous code in the questionnaire) or given a new code. The code frame was developed and reviewed continuously throughout fieldwork, to ensure the accuracy and validity of the codes.

5.2 Weighting

Weighting the smart meter sample

The survey responses from the smart meter customer sample were weighted to correct for non-response. These weights were applied to ensure the achieved sample of completed interviews matched the profile of the eligible sample. It tends to be necessary to weight survey data in this way as it is likely that the characteristics of respondents not responding to a survey are different from those that participate. For example, achieved samples tend to be lower on full-time workers and males.

Variables can be used to develop a non-response model if they are available both on responders, and also on non-responders. A range of household-level variables (such as payment type, energy supplier, customer start date and energy consumption band) and area-level variables (such as GOR, proportion of social grade A or B households and ACORN group) were considered as candidate variables for the weighting. A logistic regression model was created to understand which of these variables were either predictive of non-response or were related to the outcome variables being measured. The final model included the following variables: GOR, high electricity consumption band, energy supplier, gas grid status, energy bill payment type and the proportion of social grade A or B households. The weight derived at the end of the non-response modelling step was then calibrated to known population totals on four variables: the high electricity consumption band, GOR, the social grade proportions and the supplier name. This ensured the weighted sample matched the eligible population on these variables.

The final weight was created by multiplying the design weight (created at the sampling stage to equalise the probability of selection) with the non-response weight and rescaling it to an average of one. Scaling to have a mean of one is common in

social research because it means the weighted sample size equals the actual sample size (rather than the population size, which would also be possible but is not often the approach taken).

The weighting was designed to reflect the relative size of the current population totals of the smart meter customer bases of each energy supplier. This means the data was not weighted 50:50 to give an even split between Company A and Company B but instead was weighted in proportion with the size of their respective customer base. This approach was taken to ensure the data was representative of the early roll-out.

Weighting the legacy meter sample

The survey responses from the legacy meter customer sample were not weighted in the same way as the smart meter sample as these customers were not drawn to represent any known population. Instead part of this data (that relating to key questions recording attitudinal and behavioural outcomes) was weighted using propensity score weights as explained in section 5.3.

5.3 Propensity score matching

Rationale for Propensity Score Matching

One key aim of this study was to ascertain whether having a smart meter installed is likely to lead householders to adopt certain attitudes and behaviours. For example, whether having a smart meter installed leads to a feeling of greater control over energy bills or a higher level of engagement with different energy tariffs. A straight forward comparison of the smart and legacy meter results would not identify whether any differences observed where due to systematic differences in the profile of the two customer groups (for example, by age, gender or social grade) or whether this was due to the presence of a smart meter and/or IHD. A more sophisticated analytical technique was therefore required to feel more confident that any differences between the two groups could be attributed to the smart meter.

At the outset of the study, the sample of legacy meter customers was selected to match as closely as possible to the profile of the sample of smart meter customers. This matching was limited by the available data (energy supplier, fuel type, payment type, energy consumption data, region, % social grade AB in output area). For instance, it was not possible to match customers on other variables likely to affect energy use such as property type, household composition, work status and so on. It was likely therefore that noticeable differences would remain between the two samples. The survey offered the opportunity to collect more detailed attitudinal and demographic information. A technique called Propensity Score Matching (PSM) used this additional information to improve the quality of the matching between the smart meter and legacy meter respondents. PSM is essentially a weight which is applied to the legacy comparison group to match it as closely as possible to the smart meter group. While PSM improves the matching of the two groups it can only do so on the basis of differences that have been observed, either through the survey or known from the sample. It is likely that some differences may remain between the smart and

legacy meter groups which are unobserved, but PSM offers the best opportunity to minimise these.

Overview of Propensity Score Matching method

The idea of balancing or conditioning on matching variables in observational (non-experimental) studies comes from work done by Rosenbaum and Rubin (Biometrika, 1983)). They show that matching can be sufficient to remove bias due to all observed characteristics. In other words any bias in survey outcomes due to observed differences on matching variable characteristics between smart and legacy meter cases will be removed through matching. An efficient way to do this is through Propensity Score Matching (PSM).

The PSM process involved three key steps:

1. Choosing the covariates to be used in the matching process;
2. Applying the matching algorithm to form matched sets;
3. Diagnosing the matches obtained; and

The matched cases obtained were then used to estimate the effect of smart meters on the outcome variables.

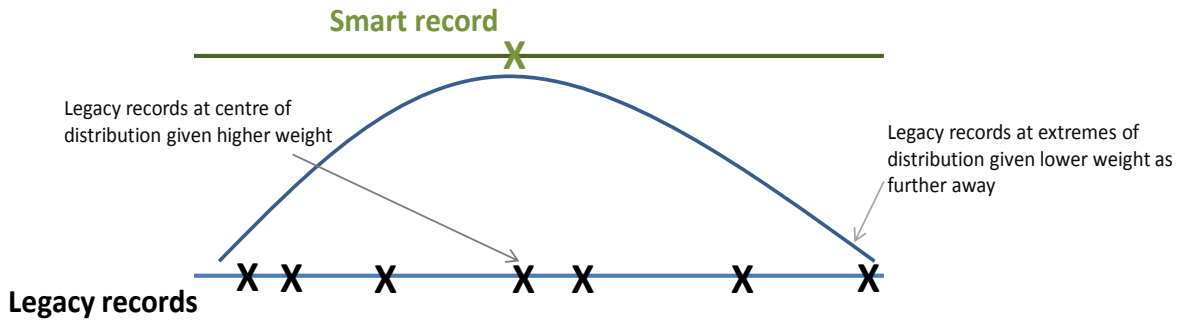
Step 1: Choosing the covariates to be used in the matching process

The survey variables that were used for matching needed to be those which have a relationship with the key desired outcomes but which were not the desired outcomes themselves, or could theoretically be expected to have been influenced by a smart meter installation. For example, matching could not be conducted on feeling in control of energy use or the frequency of practising energy efficient behaviours as these are the outcomes the analysis is attempting to measure. In order to choose the covariates to use in the matching process any variables such as these were firstly excluded. Analysis was then conducted on the remaining survey variables to identify those which were predictive of a respondent being in the smart meter group[10]. This found that variables such as household size and tenure should be used in the matching process. A full list of the variables used in the final matching are presented in Table 2 below.

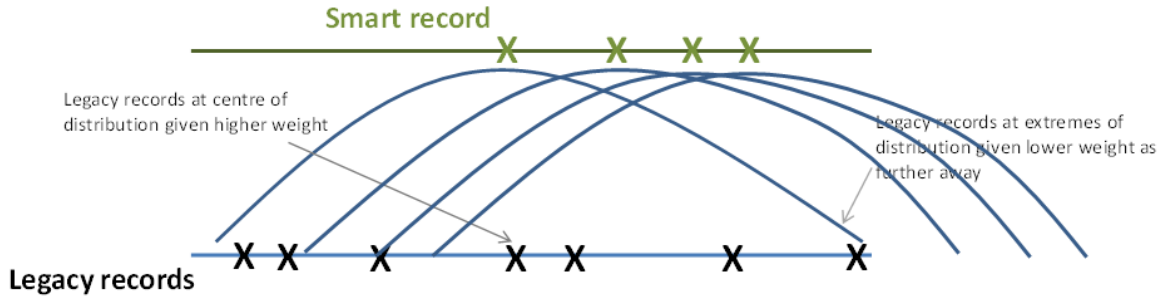
Step 2: Applying the matching algorithm to form matched sets

As the SMEA survey involves a large, but equal number, of treatment and control records the PSM method used was a kernel approach, as it was deemed the most suitable. This looks at the distribution of the control records in relation to a treatment record and assigns progressively greater weights to the control records closest to the position of the smart record. This is illustrated in the diagram below.

[10] This analysis was conducted using a stepwise regression approach to test the relationship between each survey variable and being in the smart meter group.



This weighting process was repeated for every single record in order that the records match as closely as possible.



Matching was performed in Stata 13 using the programme psmatch[2.11] A binary logistic regression model was fitted. The binary variable used as the dependent variable in this model was a 0/1 flag identifying whether the case was a smart meter (1) or legacy meter (0) customer. The data was matched on the log-odds of the predicted probability of being in the smart meter group. The smart meter flag was based on the information held in the sample obtained from the energy suppliers i.e. any respondent flagged as having a smart meter according to their supplier was assigned a smart meter flag regardless of whether they recognised themselves to have a smart meter when responding to the survey.

The original list of candidate variables was then reduced to a shorter list of variables related to treatment group. A variable should be omitted at this stage if it is not related to treatment group (or if it is related to treatment, but not to outcomes). Although it would not be an error to include such a variable, and would not introduce bias, it would possibly introduce noise. The variable GOR was included, as it was used in the stratification, and variables for the supplier, and (for gas users) the presence of a central thermostat were included (this being because a good match was needed on this variable to ensure no bias due to routing was introduced on the questions asked only of those with a central thermostat). After these three variables were included, the modelling strategy used a forward selection method. Variables that improved the model were added one at a time, until no further significant improvement could be obtained.

[11] Leuven, E and B Sianesi, B (2003). "PSMATCH2: Stata module to perform full Mahalanobis and propensity score matching, common support graphing, and covariate imbalance testing"http://ideas.repec.org/c/boc/bocode/s432001.html This version 4.0.5

Two models were developed, depending on whether the household had mains gas. As the sample size for the mains gas matching was quite large (*n* = 3784) it was possible to include a larger number of variables than in the smaller sample of non-gas households (*n* = 232)[12]. The final list of variables used in the matching was as shown in Table 5 below.

Table 5: Matching variables used in Propensity Score Matching

Matching variables	HHs with mains gas	HHs with no mains gas
Geographic		
GOR	X	
Area classification i.e. urban, rural etc.	X	X
Property		
Energy supplier	X	X
Dual / single fuel	X	
No. bedrooms	X	
Property type		X
Age of property	X	
Presence of central thermostat	X	
Presence of wall insulation	X	
Presence of loft insulation	X	
Household		
Age of CIE	X	X
Work status of CIE	X	
Household income	X	
Tenure	X	X
Length of residence	X	X
Household size (no. people)	X	X
Age of people in household	X	
Individual		
Attitude towards new gadgets	X	

Attitude towards improving the home		X
Attitude towards saving money		X
Education status		X

For each of these models the probability of being in the treatment group was estimated, and the groups were matched on the estimated probability. Kernel matching (using the Epanechnikov kernel) was used, and a small number of large weights were trimmed. Fifteen smart meter households were discarded as being outside the range of common support and so the analysis was restricted to the rest of the sample.

Step 3: Diagnosing the matches obtained

The most important step in using matching methods is to diagnose the quality of the matched samples. This was assessed by examining the weighted frequency tables (or weighted means for numerical variables) of the treatment and comparison groups. The differences between the two samples was reduced on a number of variables (most notably the number of bedrooms, the work status of the head of household, frequency of internet usage, the family type, tenure and the length of residence). The variables which were not improved by the matching tended to be those where the two samples were very similar to start with. For 98% of the variables the difference between the matched samples was within 2%. For a very small number of variables the difference between the profile of the smart meter group and the matched legacy meter group was between 2% and 3%. These differences were felt to be within acceptable limits.

Using the matched cases to estimate the effect of smart meters on the outcome variables

The impact on an outcome for any single respondent in the smart meter group was estimated by the difference between that individual's score for that outcome and the weighted score of the legacy meter respondents they were matched with. The impact of the smart meter on the smart meter group is simply the mean of the individual impacts weighted by the non-response weights. The analysis consisted of comparing mean scores in the smart meter and comparison groups, and estimating impact by looking at the difference.

Variance estimation on these results is not straightforward, and researchers have suggested different methods of calculating variance estimates. The approach taken here has been to treat the matched data as being similar to a randomized experiment and estimate standard errors and confidence intervals using a cluster-robust variance estimator. The difference between the means of the two groups is then deemed significant if the 95% confidence interval for the mean difference does not include zero.

The impacts of the PSM were considered in relation to a subset of the survey questions which focused on assessing any outcome of having a smart meter installed.

5.4 Factor analysis

Overview and objective

Factor analysis is a statistical technique that examines whether a large number of different variables can be simplified into a smaller number of core 'factors' without losing a significant amount of the information reflected in the full list of variables.

Factor analysis was necessary in this study as it enabled the extensive survey to be condensed down into a small number of variables. These were then used in the Key Drivers Analysis (discussed in section 5.5.).

Exploratory vs. Confirmatory Factor Analysis

There are two main forms of factor analysis: Exploratory and Confirmatory.

Exploratory Factor Analysis (EFA) is as it suggests a more experimental approach that involves looking at a wide range of factors and being guided by the results. This approach is used to reveal patterns among the inter-relationships of survey items.

Confirmatory Factor Analysis (CFA) takes a pre-determined structure (each statement down to just one factor); i.e. the factors to be used are chosen in advance, based on existing understandings about potential relationships between different variables. CFA is usually used if the aim is to create an interactive model that predicts what would happen to X if individual statements were increased.

The approach taken on the factor analysis for this study was somewhere between an exploratory and confirmatory approach. The steps taken to create a final list of factors are described below.

Key stages of factor analysis

Step 1: Suggested factor list created from survey

The first step was to review the survey questionnaire to identify groups of questions which may, in theory, group well together. This was done by the core Ipsos MORI research team using their knowledge of the rationale behind the design of different questions to group together those which covered similar types of attitude or behaviour.

Step 2: Testing of the initial factor list

Correlations were run on the statements selected to be in each of the groups identified at Step 1 were. The methodology used to test the correlations depended on the type of response scale used in each group of questions. For example, pairwise polychoric correlation was used for questions with a five-point response scale while tetrachoric correlation was used for questions with a binary scale. Eigenvalues were also run to show how many factors the statements could be spread across i.e. whether they would support being spread across only one or two

factors, or whether a greater number of factors existed with the data for those questions.

Step 3: Reviewing the initial factor analysis outputs

Analysis of the correlation and Eigenvalue results showed which of the statements chosen during Step 1 held together when tested (rather than just in theory). This led to a new set, of more robustly created, factors. The factors produced during Step 2 were compared against the original factors identified directly from the survey questionnaire during Step 1. This comparison highlighted where changes needed to be made to the groupings created by the Ipsos MORI team – for example, perhaps one statement did not fit within the rest of the grouping, or perhaps one of the suggested groupings actually worked better split into two sets of statements rather than one.

It was also important at this stage to consider the substantive meaning of the factors i.e. did the groupings make sense? Each variable is usually assigned to the factor with which it is most strongly correlated but if it appears more naturally associated with another factor (with which it is also significantly correlated) then it may be moved.

The revised factors were discussed and signed off by DECC before being run as the final groupings of statements. The final list of factors is presented below in Table 6.

Table 6: Final factor list (see overleaf)

Factor name	Constituent statements
Environment not a priority *(agree/disagree statements)*	The environment is a low priority compared with other things in my life It's not worth me doing things to help the environment if others don't do the same
Feel in control *(agree/disagree statements)*	I feel in control of how much electricity I personally use I know what uses the most electricity in my home I feel in control of what I spend on my energy bills
Feel in control (attributed) *(attitudes encouraged by having smart meter)*	I feel in control of how much electricity I personally use I know what uses the most electricity in my home I feel in control of what I spend on my energy bills
Thought given to energy use *(amount of thought given to energy usage)*	Amount of gas used in home Amount of electricity used in home
Engagement with energy market *(frequency of action in last couple of years)*	Found out more information about electricity or gas tariffs Changed the tariff you are on while staying with the same electricity or gas supplier and staying at the same address
Engagement with energy market (attributed)	Found out more information about electricity or gas tariffs Changed the tariff you are on while staying with the same electricity or gas supplier and staying

(behaviour encouraged by having smart meter)	at the same address
Household discussion of energy *(frequency of action in last couple of years)*	Discussed 'How you can save energy' with household Discussed 'Who uses the most energy' with household
Change in payment amount *(increase/decrease statements)*	Overall amount paid for gas Overall amount paid for electricity
Change in usage *(increase/decrease statements)*	Amount of gas used Amount of electricity used
Installing insulation measures *(already installed in home)*	Loft insulation or top-up loft insulation Wall insulation Under floor insulation
Installing heating controls/systems *(already installed in home)*	Heating programmer / timer Central thermostat to control temperature of whole home Replacing an older gas boiler with a more efficient condensing gas boiler
Satisfaction with installer explanations *(level of satisfaction rating)*	The ability of the installer to answer the questions you had The explanation provided by the installer of how the smart meter(s) / meter works

	The explanation provided by the installer of how the in-home display(s) works
Satisfaction with information received pre-, during and post-installation *(level of satisfaction rating)*	Any information you received in advance advising you what would happen during the installation The booklet explaining how the smart meter(s) / meter works The booklet explaining how the in-home display works Any follow-up contact (e.g. telephone call) you received from the energy supplier after the installation
Satisfaction with energy-related advice at installation *(level of satisfaction rating)*	Energy efficiency measures such as loft insulation, cavity wall insulation, etc. Steps you could take to use less energy in your home The range of energy tariffs available Where to receive independent advice on energy issues How to get more detailed information about the energy your household uses The availability of advanced in home display units with extra features or a different design How to use the in home display to identify how much energy different appliances are using
Satisfaction with practical aspects of installation *(level of satisfaction rating)*	Ease of arranging the appointment for the installer to fit your smart meter / new meter Installer arriving on time for the scheduled appointment How long it took to install The installation taking place without any major difficulties or problems

	How your property was left after the installation visit
Recent general engagement with IHD	Is your in-home display generally still plugged in and in use? Thinking about the last few weeks, how often have you personally tended to look at your in-home display?
Ease of using IHD *(extent of knowledge carrying out these actions)*	Switching between the display showing how much you are spending and the display showing how much energy you are using Accessing information on your household's current energy use at that moment in time Accessing information on how much energy your household has used in the past (e.g. past day, week or month) My in-home display is easy to use
In-depth personal use of IHD *(frequency of using feature)*	Information on your past electricity usage (the kilo-watts measure) i.e. how much electricity you have used over the last week or month Information on your current electricity usage (the kilo-watts measure) i.e. how much electricity you are using at that point in time Information on your past gas usage (in meters cubed) i.e. how much gas you have used over the last week or month Information on your current gas usage (in meters cubed) i.e. how much gas you are using at that point in time Information on how much you have spent in the past on electricity (the money display) i.e. how much you have spent on electricity over the last week or month Information on how much you are currently spending on electricity (the money display) i.e. how

	much you are spending on electricity at that point in time Information on how much you have spent in the past on gas (the money display) i.e. how much you have spent on gas over the last week or month Information on how much you are currently spending on gas (the money display) i.e. how much you are spending on gas at that point in time
Initially used IHD to check usage *(agree/disagree statements)*	I looked at it when I first received it to see how much electricity different appliances use I looked at it when I first received it to see how much gas I was using
Used IHD information to understand and manage energy usage *(agree/disagree statements)*	I've used it to work out what a normal level of energy use is for my household i.e. what we use on a typical day, week or month I've used it to check that nothing is left on in the house when I go out or when I go to sleep I use it to encourage others in my household to reduce their energy use Have you used the in-home display to try to estimate what your energy bills might be?
Experienced delays with IHD *(yes/no statements)*	A delay to how quickly your in-home display responds when you turn electrical appliances on and off A delay to how quickly your in-home display responds when you turn gas appliances on and off
Experienced issue with IHD *(yes/no statements)*	A loss of signal between your meter and the in-home display A loss of information on your past energy use when you unplug your in-home display The battery which powers your in-home display running dead Having to keep your in-home display in a part of the house that you do not want to keep it in due

	to problems receiving a signal from the meter
Receipt of information during installation *(yes/no statements)*	I asked questions of the installer during the installation visit The installer explained / showed me how the smart meter(s) /meter works The installer explained / showed me how the in-home display(s) works I received a document explaining how the smart meter(s)/meter works I received a document explaining how the in-home display works
Receipt of information after installation *(yes/no statements)*	I received a booklet in post / by email explaining how the smart meter(s)/meter works I received a booklet in the post / by email explaining how the in-home displays works
Receipt of energy-related advice at installation *(yes/no statements)*	Energy efficiency measures such as loft insulation, cavity wall insulation, etc. Steps you could take to use less energy in your home The range of energy tariffs available Where to receive independent advice on energy issues How to get more detailed information about the energy your household uses The availability of advanced in home display units with extra features or a different design How to use the in home display to identify how much energy different appliances are using

5.5 Key drivers analysis

Overview and objective

Key Drivers Analysis (KDA) is a statistical technique which aims to understand which factors, or “drivers”, most influence a given outcome. It was used in this study to investigate whether particular aspects of the customer experience during the early smart meter roll-out were more or less likely to contribute towards certain behavioural and attitudinal outcomes.

Key stages in KDA

Step 1: Selecting outcome (dependent) variables to explore

Through close consultation with DECC seven outcome variables were chosen for KDA analysis. These covered a range of outcomes related to customer satisfaction, use of the IHD and energy-related attitudinal and behavioural changes. These are explained fully in Chapter 6 of the main research report.

Step 2: Selecting the predictor (independent) variables

There were a number of stages to selecting the final factors to use as drivers in the KDA. As a result of the steps set out below to assess the explanatory factors in the context of each outcome variable, the number and range of factors chosen was different for each of the seven models.

1. An initial master list of survey variables and factors was selected by Ipsos MORI and DECC including all variables thought likely to have a theoretical relationship with the entire set of outcome variables listed above.

- Amount known about smart meters before contact with supplier
- Active responder
- Passive agreement
- Proactive requesters
- Prior installation
- Meter replacement
- Time since installation of SM and IHD
- B/D. I received a letter/phone call advising me what would happen during the installation
- F/G I received a booklet explaining how the smart meter(s)/meter/IHD would work
- Receipt of energy-related advice at installation
- Receipt of information during installation
- Receipt of information after installation
- Received follow-up telephone call after installation
- Satisfaction with practical aspects of installation

- Satisfaction with installers' explanations
- Satisfaction with information received pre-, during and post-installation
- Satisfaction with energy-related advice at installation
- Recent general engagement with IHD
- Initially used IHD to check usage
- In-depth personal use of IHD
- Used IHD information to understand and manage energy usage
- Ease of using IHD
- I'd like further information about how to use it to manage how much energy I use
- Discussed info shown on IHD with household
- IHD has shown: Using energy at different times of day (response code 7)
- IHD has shown: Using more than expected (response code 1, 3, 5)
- IHD has shown: Using less than expected (response code 2, 4, 6)

2. The strength and direction of the relationship between each driver variable and the outcome variable was explored using Pearson correlations. This was conducted as bivariate analysis – that is only two variables were analysed at one time; the outcome variable against each individual explanatory variable. This preliminary testing was conducted, prior to running the final KDA models, so that the list of explanatory factors included in the final models could be reduced down to only include those known to have a significant relationship with the outcome being explored. Any explanatory factors shown through this analysis to not have a significant relationship[13] with the outcome variable were discarded at this stage.

3. A number of other tests were conducted on the remaining driver variables to assess their suitability for inclusion in the final KDA models:

Checking correlations between explanatory factors

It is usually important to check whether explanatory factors are correlated with one another as if a correlated pair are used in the model it is likely that the results will be distorted i.e. the relative importance of one of the correlated pair may be inflated relative to other factors whilst the other may be ranked lower than is accurate. A correlation value of -1 or +1 denotes perfect correlation. The threshold to consider something as highly correlated is generally taken to be +/- 0.8 although a more lenient threshold would be +/- 0.6. The results of the correlation tests on the explanatory factors found that most did not reach either of these thresholds. There were three variables however with a correlation test statistic of between +/-0.6 and +/-0.7. Where more than one of these factors was included in the list of significant factors for any of the outcome variables action was taken to remove one of them

[13] A test statistic from the Pearson's correlation of 0.05 or below was taken as statistically significant. This means that 95 times in 100 the variable will be found to have a significant relationship with the outcome.

from the final model. The decision on which to remove involved discussions with DECC and consideration of the theoretical importance of each to understanding the outcome in question, as well as an assessment of the proportion of missing values (as discussed below).

While this check on correlations was conducted during this analysis it was less critical than may usually be the case as a non-standard regression approach was taken to minimise the effect of even moderately correlated variables. The regression approach is described under Step 3 below.

Checking the proportion of missing values for each explanatory factor

Missing values are an issue for KDA, and affected the models run for this study given the filtered nature of many of the survey variables (i.e. those asked of a subset of customers, for example only those with an IHD, or who had personally used their IHD). Diagnostics were therefore run on the explanatory variables being considered for inclusion in the final models and it was advised that those with a high proportion of missing values (around the 30% level or higher) should be dropped. In a couple of cases however, these factors were felt to be of significant theoretical importance for understanding the outcome and so the factor was retained and the missing data was imputed.

For each model where missing values on one or more factors was an issue this involved: filtering the sample for the model so it was only based on respondents who had provided a response to at least half of the explanatory factors (this is a general rule of thumb and in practice it may sometimes have been more than half depending on what the data allowed – the objective was to keep in as many people as possible within the sample for the model). The next step was to impute the missing values on each statement. This was run in SAS software using the MI procedure (multiple imputation) based on maximum likelihood estimates and MCMC chains. Following imputation, the distribution of the responses was compared to the pre-imputation distribution to ensure this had not changed. During this process, the distributions remained constant which provides reassurances that the imputation did not change the results, and thus the models can still be considered robust. Another check run on some models was to compare the Pearson correlation results from the bivariate stage for the imputed data to the bivariate stage results for the non-imputed data. For the most part, this again confirmed that the imputation had not had a distorting effect on the data. However, in the case of a couple of factors, the bivariate results were shown to be distorted. These were subsequently removed and excluded from the final models.

The variables were imputed at the raw data level. The imputation happened once on the overall sample included in the overall model. This sample was then filtered to create the subgroup models (in the same way as if the data had not been imputed).

As a result of this variable assessment and selection stage, the number and range of variables chosen was different for each of the seven models.

Step 3: Run final models

The aim of the final models was then to show the relative importance of the different significant explanatory factors relative to one another. This output provided a ranked list of significant factors by their relative importance.

The method used to run the final KDA models was dominance analysis (Budescu 1993)[14]. This approach computes regression models with all the possible combinations of drivers i.e. it adds a new explanatory factor in combination with one other variable, and then two other variables and then three other variables and so on until all possible subset regression models have been run. Each explanatory factor is then given its general dominance weight, and which is the average contribution of that factor to all the models. This directly addresses a factor's contribution by itself and in combination with the other factors, while overcoming any correlation between the explanatory factors. If the explanatory factors are inter-related this approach ensures they are both still shown to be significant, and important, by providing the unique amount of variance contributed by each without overinflating the importance of one factor in relation to another. This is the main advantage of this approach over a standard linear regression method which includes all the factors together in a single model.

A final model was run for each of the seven outcome variables on the basis of the overall smart meter customer sample, as well as for a number of specific subgroups of smart meter customers (e.g. those aged 65+, those with higher or lower total annual household incomes, those from a household where someone lives with a long-term condition, those with a greater or lesser interest in technology etc.). The subgroups were chosen by Ipsos MORI and DECC to include some customers who may be considered more vulnerable (due to low incomes or poor health) as well as to cover issues of key interest (for example, how a pre-existing interest in technology may affect the impacts of smart metering). The subgroup models were run using the same set of explanatory factors selected for the final model of each outcome variable based on all smart meter customers. The analysis of the subgroup models focused on the likely range in ranking position of each explanatory factor (i.e. the highest and lowest possible position shown through its confidence interval), rather than on the absolute portion of variance explained by each factor (i.e. the individual R-squared value for the factor). This was due to the relatively small sample sizes for these models and the low overall R-squared values. Further explanation of R-squared values and the relative ranking positions is provided under Step 4 below.

Step 4: Interpreting the KDA results:

Three main statistics are produced in the final KDA models:

- ***Overall R-squared value (R-sq.)****:* this provides a measure of how much of the variation observed in the outcome variable being explored is explained by

[14] Further information about this method is provided in the paper: Tonidandel, S., and LeBreton, J, M. (2011). Relative Importance Analysis: A Useful Supplement to Regression Analysis, Journal of Business and Psychology, Volume 26, Issue 1 , pp 1-9, (http://link.springer.com/article/10.1007%2Fs10869-010-9204-3)

the explanatory factors included in the model. This means the extent to which a change in the value of the outcome variable (for example, the level of satisfaction with a service) can be explained by changes in the explanatory factors (for example, the friendliness of the person delivering the service). The R-squared value can range from 0% (i.e. the explanatory factors included in the model do not explain any of the variation in the outcome so it must all be being driven by something that has been unmeasured) to 100% (i.e. the explanatory factors included in the model explain the full amount of variation in the outcome)[15].

In the academic and research literature there is no agreed threshold on what constitutes a good or bad R-squared value. Whether or not an R-squared value is taken as acceptable or not is highly dependent on a number of factors, such as the outcome being explained and the coverage of predictors collected by the survey. In general, it is likely that a model designed to explain levels of satisfaction with a service or product will achieve higher R-squared values than models run to explain the drivers of attitudes and behaviours. This is because these outcomes can be very complex and it is unlikely that a survey will be able to collect information on all the possible factors influencing the specific attitude or behaviour.

- ***Proportion of R-squared (PR2):*** each individual explanatory factor is assigned a PR2 value which across all the factors sums to the R-squared percentage. This statistic shows the relative importance of one driver variable compared to another. This is the result presented in Chapter 6 of the main report.

- ***Importance value (%):*** The importance % for any one factor is calculated by dividing the % it represents of the total R2 over the R2 value. E.g. if a model containing three factors has an R2 of 40% and the three variables each contribute 20%, 15% and 5% to the R2 (adding to 40% total) then the importance % of the first factor will be 50%. The importance % therefore shows how much a single factor contributes to the explanatory power of the model.

When interpreting the results of the KDA the main comparison made is in the relative ranking of the explanatory factors included in a model. This has been done within each model (i.e. within the overall model for ‘Ease of IHD use’) as well as between subgroup models (i.e. between the model based on those aged 65+ for ‘Ease of IHD use’ and the model based on those aged 64 and under). While it is not possible to apply statistical significance testing to these results as has been done for the survey results, it is possible to compare the confidence intervals around the relative rankings. For example, a factor ranked 2nd in a model may have a confidence interval showing it could be ranked anywhere between 1st and 5th in the model. Where comparisons between variables has been made in the report this has been done on the bases of the confidence ranges differing (or having only a small amount

[15] The R-squared value can also be presented as a value between 0 and 1. It is formally known as the coefficient of multiple determination.

of overlap between the extremes of the ranges) across the factors being discussed (i.e. one has a range of 1-4 and another a range of 3-7 which shows that these factors can be assumed to occupy different ranking positions relative to one another).

6. Qualitative sampling, fieldwork material and fieldwork details

6.1 Sampling of smart meter customers for qualitative research

The sampling approach for the qualitative research was purposive, seeking to achieve symbolic representation of a series of specific characteristics of interest, combined with diversity across the population of smart meter recipients.

Given the diversity of the population and the broad range of characteristics of interest, alongside the considerations of data saturation and of budget, it was decided that 80 interviews would allow for sufficient exploration of the relevant research questions. This included a target of 20 interviews to be held with multiple members of the same household present, and who contributed to the research.

Study population

The sample frame comprised recipients of smart meters who had participated in the quantitative research, and who had consented to re-contact for the purpose of further research. This represented 81% of the smart meter population, or 1,644 people. As interviews were to be conducted face-to-face, four geographical clusters of smart meter customers were identified: East Midlands, West Midlands, London, and the North West. All those outside of these areas were excluded from the sample frame, leaving a final frame of 1,237 people.

Selection criteria

The quantitative data set comprised demographic, attitudinal, behavioural and installation information for all potential interviewees. On some points, however, it was felt that more current information might be needed, or that further definition or exploration of points raised in the interviews would be needed at the point of recruitment. A screener questionnaire was therefore devised for use alongside the quantitative data already held on potential respondents. Recruitment then took place over the phone.

Quota targets

Table 7 below shows the target number of interviews for each quota set, as well as the criteria monitored during the recruitment process.

Table 7: Comparison of qualitative target and completed quotas (see overleaf)

	Target number of interviews	Final Completes	
		Based on screening/ interview data	Based on original survey data
Priority quotas (outwith matrix)			
IHD Never plugged in	**5**	4	5
Low income	**15**	Not gathered during qualitative interviewing process	20
Elderly (respondent or other in household 75+)	**8**		9
Health condition	**20**		24
Not present at installation / new occupant	**8**		8
Presence of children in household	**16**		20
Monitoring criteria			
No-one looks at IHD	**8**	Not gathered during qualitative interviewing process	6
Low education	**25**		33
Low tech	**10**		6
Single person household (excluding elderly)	**5**		6
Criteria of interest			
Busy lifestyle	**Monitored**	25	Not collected
Resistant to external advice	**Monitored**	19	Not collected
Environmental interest – high	**Monitored**	27	26

Environmental interest – low	**Monitored**	5	5
Gadget-lovers	**Monitored**	18	21

6.2 Developing the qualitative research materials

Discussion guide

Discussion guide development began with the hypotheses and points of interest arising from interim analysis of the quantitative research data. Initial discussions within the project team led to the decision to centre the guide on the customer journey, and the sections of the discussion guide were therefore structured around this journey.

Because of the variation in customer journeys apparent from the quantitative research, the guide contained detailed routing, with different questions, for example, for those who had and hadn't ever had their IHD plugged in for any length of time, and for those who were and weren't present for the installation of the smart meter and IHD.

A draft topic guide was formulated and refined within the Ipsos MORI team. This draft was then shared with the team at DECC to gather feedback. Revisions were then made according to this feedback, and a final version of the guide was agreed. The first few interviews formed a form of mini-pilot to test the discussion guide. No changes were made as a result of the feedback on the use of the guide during the first few interviews.

Variations on the discussion guide

The research design for the qualitative interviews allowed for the inclusion of further householders alongside the original respondent in the quantitative research. Variations on the topic guide were therefore included for:

- Adults living in the home with the primary research subject;
- Children living in the home with the primary research subject;
    - Aged 8-11 years;
    - Aged 12-15 years.

Adult variations on the discussion guide were included in the main guide, as a series of additional prompts for household interviews, which explored, amongst other things, different perspectives on the smart meter and IHD, differences in ownership and use, and perceived impacts on household dynamics of the IHD, if any.

The guides for children and young people explored understanding of energy use in general, and sources of information on the subject; use of the IHD specifically, and discussions about the IHD, both inside and outside of the home.

These variations were drafted by the Ipsos MORI team, and shared with DECC to gather feedback. Revisions were then made according to DECCs feedback, and final topic guide variations agreed.

Customer journey map

To aid exploration of the customer journey in the course of the interviews, a simple customer journey map was devised, on which interviewers could record significant points in the customer journey, as experienced by the respondent. This journey could then be reviewed by interviewer and respondent at the end of the interview, to act as both a sense-check and a prompt for further exploration.

This customer journey map was produced in PowerPoint by the Ipsos MORI team, reviewed by DECC, and revised accordingly. To aid the interviewers, an example of a completed customer journey was also provided.

Example materials

In-depth interview respondents were shown examples of tailored home energy information in order to gauge their interest in, and receptiveness to, the sort of information that smart meter data may enable suppliers to provide to customers. The visual materials used included:

- A graphic displaying the gas and electricity use (in kWh) of a single household against a regional benchmark and accompanied by a statement on how much more or less this usage was costing than a local average user and energy efficiency advice
- A graph showing a home's energy use (in kWH) against 'similar homes' and against 'efficient similar homes' with a rating on how good or not the comparison was using smiley face icons.
- A graphic showing a detailed breakdown of a household's energy use – for example showing the percentage of the energy bill spent on heating.

The examples were sourced from various online tools and energy supplier customer portals. The visuals selected were agreed with DECC before being used in the qualitative interviews.

IHD use diary

Where it was known that respondents had their IHD plugged in, it was decided that use of a diary pre-task might allow us to capture, relatively unprompted, household behaviour towards the IHD. A diary task was therefore devised, to simply capture the following information:

- Who looked at the display and when;
- What prompted them to look at the display;
- What information was sought;
- What information was found; and
- Reactions to what was found.

The diary was designed by the Ipsos MORI team, then reviewed by the team at DECC, and revised accordingly.

Introductory letter

Those who received a diary also received an introductory letter explaining how and when the diary should be completed, and how it would be used, both during the interview, and afterwards as part of the reporting process. The division of incentives between the diary and interview tasks was also reiterated in the letter, having first been explained at the point of recruitment.

The introductory letter was written by the Ipsos MORI team, reviewed by the team at DECC, and revised accordingly.

Consent forms

Where children and / or young people were to be present at the interview, consent was requested in three ways:

- At the point of recruitment, verbally, from the parent or guardian;
- At the point of interview, through completion of a written consent form by the parent / guardian;
- At the point of interview, through completion of a written consent form by the child / young person.

Parents or guardians who agreed in principle to children or young people taking part in the research were informed at the point of recruitment as to what the research would entail, so that informed verbal consent could be given in principle. Consent forms were written, in accordance with MRS best practice, for parents / guardians and children to be completed at the start of the interview. These consent forms contained details of the research and how findings would be used. These were reviewed by DECC and revised accordingly.

6.3. Qualitative fieldwork

Qualitative fieldwork followed the conclusion of the quantitative stage. All interviews were conducted by members of the Ipsos MORI research team.

7. Summary of materials in Technical Appendix

All of the materials referenced within this document are saved in the Technical Appendix. The materials in the Appendix are as follows:

Appendix A: Quantitative Materials

A1: Advance Contact Letter

A2: Data Linking Leaflet

A3: Questionnaire

A4: Codeframes

Appendix B: Qualitative Materials

B1: Qualitative interviews – Recruitment screener

B2: Qualitative interviews - Letter and IHD usage diary

B3: Qualitative interviews topic guide – Smart meter customers

B4: Qualitative interviews topic guide – legacy meter customers

B5: Qualitative interviews topic guide – children’s guide

Department of Energy & Climate Change

3 Whitehall Place

London SW1A 2AW

www.gov.uk/decc

URN 15D/083