

Energy Demand Research Project: Final Analysis

Appendix B: Quality Assessment



Prepared by: Gary Raw and David Ross
Regional Directors

Checked by: David Ross
Regional Director

Approved by: Andrew Cripps
Regional Director

Energy Demand Research Project: Final Analysis

Rev No	Comments	Checked by	Approved by	Date
2.2	Final version	David Ross	Andrew Cripps	21/06/11

AECOM House, 63-77 Victoria Street, St Albans, Hertfordshire, AL1 3ER
Telephone: 01727 535000 Website: <http://www.aecom.com>

Job No 60163857

Date Created June 2011

This document has been prepared by AECOM Limited for the sole use of our client (the "Client") and in accordance with generally accepted consultancy principles, the budget for fees and the terms of reference agreed between AECOM Limited and the Client. Any information provided by third parties and referred to herein has not been checked or verified by AECOM Limited, unless otherwise expressly stated in the document. No third party may rely upon this document without the prior and express written agreement of AECOM Limited.

Capabilities on project:
Building Engineering

Appendix B: Quality Assessment

In any research, particularly field research on human behaviour, there will be risks of bias or noise in the data – while it may not be feasible to remove them all, documenting them can aid understanding of the findings and their limitations. This Appendix provides a summary of the quality issues that remain with the trial design and methods at the end of AECOM's analysis, for each of the suppliers' trials; it is intended as an aid to interpretation and application of the findings. This assessment is based on an extensive review of all aspects of the trials, from the overall design to the details of equipment and letters to customers. It draws on material in the suppliers' own reports, additional material obtained as part of the final analysis, and discussions between AECOM and the suppliers.

Capabilities on project:
Building Engineering

B1 EDF quality assessment

B1.1 Overview

There is a complex mixture of issues with the EDF findings for overall reduction in consumption, as described in detail below. The most important concern is the lack of pretrial baseline data for gas and, for some groups, electricity. The findings from other trials show that, even with strenuous efforts to match trial and control groups during the recruitment process, large differences in pretrial consumption can remain.

The Web and TV consumption findings should not be seen as representing the overall potential of such interventions. Any further consideration of using web- or TV-based feedback would need to develop more effective ways of targeting those households most likely to be interested and benefit; and delivering a more effective system, including real-time feedback. The study tested only one approach in each case, with no targeting of those customers most likely to use the technology provided. For the other trial groups, while there are some concerns over aspects of the trial design, there is a good case for further analysis if the remaining baseline data can be obtained. In relation to wider application of similar interventions, design improvements identified through the EDF trial could achieve greater impact.

B1.2 Sample bias

All groups, including the control group, had to opt in to the trial because all were offered a changed meter or the RRR plan, and only 9% of the customers contacted agreed in principle to take part. This has the potential to create a bias for all groups to be composed of customers more interested in energy saving (for financial or other reasons), this bias increasing with the demands of participation but decreasing with the value of the intervention as perceived by the customer. While such bias might generally be expected to make the interventions more likely to succeed, it is also possible that those most interested in energy savings had already taken action and therefore had less potential for further action. These influences may at least partly balance each other.

Perhaps the most disruptive element of the interventions, and the one with least obvious benefit to the householder, was having the smart meter, hence bias would have been reflected in the control group too (since they also had smart meters). Furthermore, the fact that all comparisons were with a smart meter control would make it more difficult to see significant effects (compared to using non-smart meter controls, as in the other suppliers' trials). This reduces concerns over selection bias but these concerns remain for those groups that did not have sufficient baseline data to make pretrial/in-trial comparisons.

In addition to the simple fact of opting in, several aspects of the trial have the potential to increase bias towards customers who are more interested in, and/or more aware of, energy issues.

- Customers were excluded if they had fewer than two meter readings in 12 months or less than a year in the home or expecting to stay less than 2 years. These exclusions could create a bias to less mobile households and those with a more positive view of their energy supplier, hence more likely to take action on energy efficiency issues. They might also introduce bias to households with someone at home during the day, hence more interest in saving energy.
- Customers were excluded if they had no telephone number on record or were below a lower survival band of estimated annual energy consumption (could create a bias away from households with lower income and/or less potential to reduce consumption).
- Diverse selection criteria for different trials could create bias or noise in the data. In effect, it can change the nature of the intervention. Customers excluded on trial-specific criteria were not transferred to other trials, which mitigates potential bias.

Capabilities on project:
Building Engineering

- Recruitment was by telephone (including questionnaire administration to gather basic data on the household and dwelling). Calls were made by an agency – EDF often used this agency but staff may not have had the in-depth knowledge to address all customers' questions.
- The recruitment script included a range of motivational points to attract customers, including offer of new technology; more information about energy use; no need for estimated bills; help with energy, carbon and cost saving; backed by Government and Ofgem; no additional cost; one of the first UK households to benefit. This might have mitigated bias towards energy-aware households but it could equally be a lot to take in and mention of Government might have been a disincentive to some customers.
- The style of initial letters and the customer information pack might have been offputting because of jargon, "small print" terms and conditions and the prominence given to hazard warnings (which seem excessive). Also there is excessive emphasis on the acceptance form and needing customers to read the information pack first.
- Households were allocated to groups before they were invited to join the trial.

There were also aspects of the trial that would tend to create bias in the opposite direction.

- Customers were excluded if they were on the EDF Mailing Preference Service or had taken "Green products", i.e. special tariffs or consumption reduction products (could create a bias away from households with a more positive view of energy saving in general or of their energy supplier, hence less likely to take action on energy efficiency issues, but this should be the same for trial and control groups, and small numbers were excluded on this basis).
- The control group was paid to take part (£50 supermarket voucher). Payment itself could make people (temporarily at least) more interested in energy or more positively inclined to the supplier, although it was not sufficient to attract recruits to the trial: after a short initial period, recruits to the control group were given the additional incentive of accurate bills (only around the first 10 recruits were not offered accurate bills).
- Recruitment success rate was higher (12%) in the earlier phases of the study, creating less bias in the Control group and in intervention groups that were first to commence (Paper and Wall Panel), and it is those interventions that showed significant savings rather than later interventions.
- Only medium consumption households were recruited in the TOUT groups, which therefore had the most scope to change, either up or down.

Paper and Wall Panel, being the first trial groups to be recruited were also the ones that were most comparable to the Control group in other ways.

- There was initially a stricter requirement for number of historic meter readings.
- There were higher survey/installation failures than later in the trial when there was improved technology, increasing survey experience and reduced delay between recruitment survey and installation.
- First contact with the customer emphasised environmental benefits more than in later recruitment.
- The recruitment scripts for the Control, Paper and Wall Panel groups were similar in the sense of being neither highly motivating nor demanding. Mention of the alarm in recruitment script for the URA group was unpopular, which would tend to increase bias towards those most interested in energy. Similarly, the Web recruitment script made the intervention seem more demanding than the Wall Panel and hence might have attracted participants with more interest in energy.

For several of the interventions, the recruitment script included motivational points that were not specifically about saving energy:

- RRR – rewards just for customers reading their own meter and communicating the reading to EDF (rewards that were easier to achieve and of higher value than that for year-on-year reduction in energy use);

Capabilities on project:
Building Engineering

- TV – a "free, top of the range TV set-top box" with HD capability;
- TOUT – opportunity to save money;
- HEC¹ – heating and hot water programmer (hence an interest in the technology itself and/or saving gas rather than electricity).

Whereas this might have biased the sample towards people with less interest in saving energy, there would have been some counterbalance because of the relative complexity and/or demand associated with accepting these interventions.

B1.3 Intervention quality issues

Reviewing each of the interventions, it is now possible to see ways in which each could have been improved. The implication of this is that some interventions were successful even with their limitations, and greater energy savings might be achieved with improved interventions. In other cases, interventions were not successful but might have been if designed and executed differently.

RRR

- Differed from control in not having a smart meter and in the exclusion criteria (no exclusion based on meter distance, pacemaker, hearing aid or Economy 7) and customers were not excluded if already signed up to RRR. So both the control group and the baseline are suspect.
- Consumption data quality was generally poor.

Advice (used in all groups except RRR and TOUT)

- Tip for summer months was sent with February energy usage report. This included advice to use "solar lighting" in the garden; customers who understand energy demand might recognise that this would entail embodied energy if they would otherwise use no garden lighting at all. If this is typical, it would reduce the effectiveness of the advice.

Additional bill data (used in Paper group)

- Additional motivation of receiving a monthly summary mentioned in brief statement of the intervention, emphasising saving energy and money. Increases the likelihood of participants at least being aware of the intervention.
- Some logistical issues affected customer experience (difficulties sending initial paper statements; graphs with "an enlarged spike" in one of the monthly paper statements in late 2009; summaries ceased in 2010).
- Information sent was very busy and has several problems: meaning of colour codes is not immediately clear; statement says "this summary is for information only and does not ... require any further action from you"; tells people to use kWh (when they might find it easier to use cost, even if this does vary over time).
- Guides produced in small font and do not explain how to use comparisons over time.

Smart meter (used in all groups except RRR)

- Installers' training did not focus on customer interaction; they gave out spare user guides and answered questions, otherwise customers were referred to a remote customer services team. Interviews suggest that

¹ This analysis is supported by the *customer surveys* to the extent that the intervention was positively received but mainly as a heating controller rather than an aid to saving energy; nevertheless, 75% agreed that it had helped them plan or budget for energy use – the highest figure of any EDF intervention.

Capabilities on project:
Building Engineering

customers expected, and could have benefited from, more engagement and instruction at installation stage. Only the most engaged and motivated customers read, absorbed and referenced the booklet. In the absence of training, installers might have invented their own advice, which could increase variance or even give customers wrong messages.

Wall Panel RTD (used in Wall Panel, URA and HEC groups)

- Generally a good design but some confusion created by meter display in m³ and RTD in kW.
- Earlier project (Warmplan) had shown that, given the choice, customers tend to want the RTD in a cupboard. The installer tried to get it fitted in a visible position in the living space but the final choice was the customer's.
- User guide starts with opening page then 2 pages of health and safety material, including some details that are of doubtful necessity (e.g. avoid placing loose papers underneath your device).
- Explanation of energy usage information starts on page 14 out of 24. No explanation of what to do with the information. Operating the new equipment could have exhausted householders' interest before getting to the energy saving element. Probably too complex for many users – does not take them direct to how to display the most relevant information, or how to use this information.

Alarm (used in URA group)

- The alarm was set as a quiet beep and fixed so that customers could not change it. Hearing the alarm therefore depends on proximity, background noise and hearing acuity; it is not known how often was it actually heard.
- The alarm threshold was set by EDF, not the customer, at 6% below the previous year's usage in the first year (until a profile for individual customer was established). The choice of 6% is not explained but it may be too challenging to offer realistic motivation. It is also more difficult for customers to do the sums than for (say) 1, 2 or 10%.

Heating and hot water controller (used in HEC)

- Additional exclusions applied in this group – the intervention was offered only to owner occupiers with gas central heating who did not have a combi boiler, integrated programmer or more than one room thermostat. This could exclude the people who are most interested in (or able to use) advanced controls.
- There were more technical problems during the trial period than with the other interventions, particularly in the early trial period or at the start of the heating season: with installations earlier in the trial, a number of controllers had intermittent internal communications with other items in the home area network (this limited where the display could be placed in later installations); a faulty batch of temperature sensors caused delays (months) and cancellations of installations; some customers got into problems with the sequence of displays which required telephone calls (and sometimes site visits) to fix.
- The equipment user guide had four extra pages for the heating controller; the system and/or the instructions might be too complex for some users when added to the Wall Panel guide. On a specific point, the guide says how to use frost protection and "Economical" settings, but not why.

Basic RTD (used in TOUT group)

- Single line display with five buttons to toggle between display functions – it is less likely that customers would access all functions. But this may be offset by the simpler user guide, with no health and safety section and energy efficiency tips on page 3, before explanation of functionality and also alongside the later explanation.
- Same issues as Wall Panel RTD with positioning in the home and display in kW.

Capabilities on project:
Building Engineering

Time of use tariff (used in TOUT group)

- The major factor undermining the effectiveness of the intervention is probably the decision that the tariff deal should ensure that customers would not be able to lose money.
- It was difficult to recruit for this group because customers did not understand load shifting or believe that EDF would want to help them save money. The recruitment script did not explain load shifting and explanations may therefore have depended on recruiters' understanding.
- The customer information pack refers to three tariffs in four periods, which could be confusing because a separate tariff sheet shows four rates, one of which is the daily standing charge. The ratios of charges are not round numbers and they differ slightly between electricity and gas (N.B. this group purchased only electricity from EDF), further increasing the complication.
- The explanation of load variation was rather technical and confusingly referred to "times when usage is abnormally high" to refer times when usage is usually high.

TV intervention

- Feedback did not include real-time data.
- Technical limitations meant that a secondary TV was often used – i.e. not the one that would most often be viewed.
- Access to the energy screens was easy, using dedicated buttons on-screen and on the remote, with simple user instructions. However the impact may still depend on the person with control of the TV being interested in energy.
- Freeview boxes were delivered before final user acceptance testing and installation refinement. Owing to lack of standard communications control with TVs, quality of display would have been variable.
- The energy screens were rather cramped in layout with many words and much information; it might be better to decide what the main message is and concentrate on that – no logo, picture or multiple figures.
- Some logistical issues affected customer experience: energy data were not sent to customers during the earlier part of the trial (it is unclear how many customers were affected or for how long).
- A key issue might be that more interesting information is always available through this medium.
- Any further consideration of using TV-based feedback would need to develop more effective ways of targeting the households most likely to benefit.
- The set-top box is estimated to use 200 kWh per year; if sufficient data could be gathered to analyse the effect of the TV intervention, this energy consumption should be taken into account but this would be difficult because it is not known what equipment the customers had prior to the intervention.

Web intervention

- Feedback did not include real-time data.
- Uses customer's own interface and is therefore dependent on broadband availability. In theory allows remote access but there is no mention of this in the guidance to customers.
- There is no record of internet payment basis (unlimited or per minute) or whether the internet user is also the energy controller; both of these could determine the effectiveness of the intervention.
- Access to the web site requires a username of 14 characters and password of 12 characters – this might have been offputting and/or introduced errors in logging on.

Capabilities on project:
Building Engineering

- Some temporary technical problems occurred during the trial, regarding automated daily data collection (therefore availability to customer) and forecast usage being out of line with actual demand.
- The user guide is not always clear: crosses and ticks have the same meaning for consumption relative to forecast, hints on responding to the daily feedback might be little used because of delayed delivery of consumption information (although there are also ideas on how to interpret variations retrospectively); second year forecasts are based on first year usage but not adjusted for weekday vs weekend or variable public holidays.
- A key issue may be that more interesting information is always available through this medium.
- Any further consideration of using web-based feedback would need to develop more effective ways of targeting the households most likely to benefit and delivering a more effective system, including real-time feedback. The study tested only one approach, with no preselection of customers.

B1.4 Noise in the data

Several aspects of the trials potentially increase or decrease noise in the data, which would increase or decrease confidence in the significant differences or non-significant differences respectively.

- *Fuel type.* There is no record of whether the electricity-only customers had gas supplied by another company, hence this cannot be taken into account in the analysis.
- *Geographic control.* All the homes are in south-east England, including London, but the locations of trial and control homes could be different and/or within-group variance could be increased or decreased depending on how widely dispersed the homes are in each group. The variance would arise from a combination of weather/climatic differences (although the homes are probably sufficiently well clustered for this not to be a major concern) and local extraneous changes.²
- *Data cleaning.* EDF reports no systematic procedures for data cleaning (e.g. no outlier exclusions) because all data were from smart meters. The statistical technique of log transformation of data was used instead but this corrects for skewed distribution, not data processing or measurement errors.
- *External factors.* Consumption will be affected by factors extraneous to the project such as the economy, energy price changes, changing outdoor temperatures, credit crunch, press coverage and national campaigns. It is assumed that the effects of such factors are felt equally across all groups. This assumption probably has to be made but it is potentially flawed, e.g. because these factors create a pressure to reduce energy use in the control group, thus making the baseline change more difficult to exceed in the trial groups, or because trial participants who would have made investments in energy saving became concerned about spending capital.
- *Recruitment scripts.* There was some variation in wording, e.g. statement that smart meters will "reduce" (TV), "avoid" (Web) or "remove any need for" (other smart meter groups) estimated bills.
- *General logistics.* There were many early delays with the installation programme due to changes in the customer database. Even if a smart meter was installed, some customers had a meter reader visit because the visit was already in the system. Owing to the date of a price change and the percentage change being kept secret in advance, there was a delay in updating customer smart metering with new tariff rates. Data processing issues initially resulted in less than 70% of bills showing actual consumption and the remainder were given estimated bills.
- *Survey participation.* Another possible source of noise in the data is reduced by the customer surveys being universal.

² E.g. a local authority or newspaper running a campaign or a major employer making redundancies. Sociodemographic variance needs to be considered separately because it arises within small areas.

Capabilities on project:
Building Engineering

B1.5 Wider applicability of EDF's findings

Several aspects of the trials reduce confidence in applying the findings more widely to the GB population, without entailing a bias in any obvious direction. It is also possible that these same aspects would either increase or decrease noise in the data, which would increase confidence in the significant or non-significant differences respectively.

- *Stratification*. The sample was designed to be demographically representative of EDF's customer base but groups were stratified in a complex way that would make it difficult to gross up findings to national level.³ Although EDF was theoretically able to gross up to its own customer base, this does not appear to have been done and the consumption strata are not defined or coded in the database, hence grossing is not possible. "Green" customers are first excluded then used in stratification – this must also be complicated for grossing and interpretation.
- *Customer exclusions*. Customers were excluded from the trial if they: were supplied with gas only; had certain medical equipment in their homes (hearing aid or pacemaker); had more than 6 readings in 12 months (presumed to be a sign of erroneous records); or they had insufficient English to understand (which makes the study culturally more coherent but less relevant).
- *Dwelling exclusions*. Excluded if dwelling was less than two years old or electricity and gas meters more than 5 m apart. Only 40.2% of customers who agreed in principle to take part passed the site survey and 69.3% of these had a successful installation (combined = 28%, overall = 2.5% of sampling base). Few flats could be included.
- *Geographic range*. The homes are all in London and south-east England. Recruitment was easier in the south-east because people were at home more often. Could represent bias towards older/retired households but Greys were not over-represented. Could also mean more owner occupiers and/or households with young children and/or single wage earner but more than one adult.

³ The stratification was based on fuel purchased from EDF (electricity or dual fuel) and a combination of estimated annual consumption (low/mid/high) and Ocean indirect demographics: (a) prepayment; (b) fuel poor (*area* likely to spend $\geq 10\%$ of take-home salary on fuel) merged with low consumption (= not high); (c) green (on a green tariff or have a very high propensity to be a green customer as determined by a lifestyle code); (d) grey (a group ≥ 55 , determined by a lifestyle code, an OAPP special condition code or Nectar data); (e) high consumption (Acorn 1 to 14 or 24 to 36).

Capabilities on project:
Building Engineering

B2 E.ON quality assessment

B2.1 Overview

The capacity of the E.ON trials to find significant effects is reduced by noise introduced by the design and method as a whole, although this would be partly balanced by most of the identified biases tending to increase the likelihood of a significant finding.

The net effect is a modest overall reduction in confidence in the null results in non-smart meter groups (less so for the RTD intervention) but each intervention was tested (individually or in combination) in at least two trial groups and two customer subgroups. Confidence is further reduced by factors specific to implementation of the interventions, in the sense that better implementation of the same type of intervention might have reduced consumption. On balance, there is reasonable confidence in these null findings. Nevertheless, the possibility cannot be excluded that different implementation of the same type of intervention would be more successful.

Confidence in the null findings for electricity consumption in the smart meter groups, especially the fuel poor group, is reduced a little by noise and bias but there is a reasonable chance that different implementation of the same type of intervention would be more successful.

Overall there is good reason to doubt the significant effects for HU_{DF} customers. These customers started with high levels of consumption and at least some of the reduction noticed is likely due to 'regression to the mean'. There is reasonable confidence in the other significant effects in smart meter groups.

There is an issue with applying the findings to the population at large because the stratification is difficult to interpret and may be difficult to replicate in modelling what would happen in a national roll-out.

B2.2 Sample bias

The control group had business-as-usual services and was selected to be appropriate to each individual trial group.⁴ The Cameo demographic profile was said to be "comparable" between trial cells⁵ and controls. There were very few significant differences between trial cells on survey variables; these few can probably be discounted because of the large number of statistical tests carried out, although significance was unlikely in many cases because of the small numbers per cell. This reduces but does not eliminate concern over the lack of multivariate analysis.

Sample retention to the end of the trial was reasonable, in the range 77-82% of consumers being retained across trial groups but slightly higher (88%) for controls.

Customers were excluded from the sample if: living in new builds with no consumption history; fewer than two actual meter reads in last 12 months; in current dispute over a bill; subject to an outstanding warrant for non-payment; or a prepayment customer. These are logical exclusions but could create a bias to less mobile households and those with a more positive view of their energy supplier, hence more likely to take action on energy efficiency issues. They

⁴ E.ON added graphical bill data to all customer bills as part of business as usual, during the early part of the trial (commencing December 2007, complete by end of March 2008). This was withheld from the control group so as to maintain a control for the trial group that had graphical bill data as the intervention. This added a difference between the control group and all trials other than the "Historic information" groups. This reduces clarity over intervention effects but not in a major way because there is consistently little impact of historic information.

⁵ Trial cells was the early terminology used by E.ON. It generally corresponds to trial groups but the terminology is retained here to avoid possible errors of interpretation.

Capabilities on project:
Building Engineering

might also introduce bias to households with someone at home during the day, hence more interest in saving energy.

The introductory letters sent to trial participants did not emphasise the energy-saving potential of the interventions applicable to the household. They instead emphasised limited availability, accurate bills and, where applicable, more frequent bills and the RTD. This was probably good for recruitment and may have reduced any bias towards more energy-aware households being recruited (although this might have been counterbalanced by the letters being rather wordy, even the shorter forms to some extent, making it doubtful that householders would read them in full).

Customers were selected into the non-smart meter trials and not invited to opt out but they could do so at any stage – this reduces selection bias relative to opting in but does not eliminate it. For the clip-on RTD, there were known opt-outs where customers declined to have the device fitted and unknown opt-outs where they did not use it. Customers were invited to have the RTD, described as being worth about £80 and an engineer then cold-called to fit it unless customer proactively declined. The high price of the device might encourage some customers but make others more suspicious of E.ON's motives. Overall, there might have been some selection bias, with customers accepting the RTD more interested in reducing energy demand. While such bias might generally be expected to make the interventions more likely to succeed, it is also possible that those most interested in energy savings had already taken action and therefore had less potential for further action.

For the additional bill data and advice interventions (with or without a smart meter), any opt-out might have been passive (i.e. customers do not look at the information). Any bias in these interventions might reasonably be viewed as representing the effectiveness of the intervention rather than sampling bias.

Customers were invited by personalised mail to apply to have a smart meter installed (all applications accepted if technically feasible). They were not told they would be part of a trial or have their consumption monitored. Letters instead "outlined the features of having smart meters installed" and the benefit of accurate billing. Selected customers were invited to have an RTD and an engineer then cold-called to fit it unless the customer proactively declined. The possible bias is complex and may have varied between customer strata (overall conversion rate was 22% but lower for HU_{DF} and higher for other strata). This has the potential to create a bias for all smart meter trial groups to be composed of customers more interested in energy saving, compared to non-smart meter groups and the control group. There could be a small additional bias in the same direction for groups given an RTD.

Typically, electricity-only trial cells had more candidate customers for the offer of a smart meter and no additional marketing was needed. Some other cells required telephone calls to follow up the initial letters to achieve the target numbers (up to 8 calls in the most difficult cells but only 2 or 3 in others). It is not specified which cells were which. This could mean that electricity-only cells and other "easier groups" were drawn from a population of more "enthusiastic" householders but the "more difficult" groups were drawn from a population of less enthusiastic householders, selectively sampled to over-represent enthusiastic householders. The overall effect could be similar levels of enthusiasm in the final sample.

B2.3 Intervention quality issues

Reviewing each of the interventions, it is now possible to see ways in which each could have been improved. The implication of this is that some interventions were successful even with their limitations, and greater energy savings might be achieved with improved interventions. In other cases, interventions were not successful but might have been if designed and executed differently.

In keeping with the effects on energy consumption, the problems were greater with the advice and additional bill data interventions. In both cases, the extra material was provided with customers' bills but in neither case were customers told in advance that their bills would be changing. There are weaknesses in the delivery and content of

Capabilities on project:
Building Engineering

the advice sheets that would reduce the likelihood of recipients (a) reading them (at all or in part), (b) understanding them, (c) finding them applicable to their household at that particular time, (d) deciding to take action and (e) knowing what action is most likely to save energy. E.g. bill inserts tend to be ignored; seasonal advice was not necessarily seasonally appropriate. The historic comparison information was on back of bill, at the foot of the "small print" column.

The delivery was better where new equipment was involved: installers were trained to demonstrate the functions of the equipment (although not to give energy efficiency advice). Training included getting the customer to turn a light or appliance on and off to see the display change, but it is not known how often this was actually done. Customers were also shown how to view historic information of the RTDs, using a demonstration unit that the installers carried with them.

This support would have been important because the RTD user guides are probably too complex for many users. The guides do not take users direct to how to display the most relevant information, or how to use this information. The guide to the mains RTD also does not have high quality visual presentation, which could undermine confidence in it.

The cost per kWh remained set at an initial default price for the clip-on RTD if it was not updated by customer, which could also have undermined confidence in the device.⁶

Gas smart meter installation was more likely to be aborted than electric (1,685 vs 211, 17% vs 2%). This implies that some dual fuel homes had only the electric meter installed and not the gas. The RTD would not have shown gas data for the affected homes.

B2.3 Noise in data

Several aspects of the E.ON trials potentially increase or decrease noise in the data, which would increase or decrease confidence in the significant differences or non-significant differences respectively.

- *Fuel type.* There is no record of whether the electricity-only customers had gas supplied by another company, hence this cannot be taken into account in the analysis.
- *Geographic control.* Most of the homes are in the English midlands but a few as far away as Devon, Kent or Yorkshire, hence the locations of trial and control homes could be different and/or within-group variance could be increased or decreased depending on how widely dispersed the homes are in each group. The variance would arise from a combination of weather/climatic differences (although the homes are probably sufficiently well clustered for this not to be a major concern) and local extraneous changes.⁷
- *External factors.* Consumption will be affected by factors extraneous to the project such as the economy, energy price changes, changing outdoor temperatures, credit crunch, press coverage and national campaigns. It is assumed that the effects of such factors are felt equally across all groups. This assumption probably has to be made but it is potentially flawed, e.g. because these factors create a pressure to reduce energy use in the control group, thus making the baseline change more difficult to exceed in the trial groups, or because trial participants who would have made investments in energy saving became concerned about spending capital.
- *Economy 7.* Not all smart meter Economy 7 control homes used an Economy 7 product (i.e. storage heaters) whereas the trial homes all did. Economy 7 meters might also be present in other groups.

⁶ This might also have been true of the mains RTD used with smart meters, and RTDs in other suppliers' trials.

⁷ E.g. a local authority or newspaper running a campaign or a major employer making redundancies. Sociodemographic variance needs to be considered separately because it arises within small areas.

Capabilities on project:
Building Engineering

- *Survey participation.* Some participants (including controls) took part in customer surveys early in the trial period, which could have influenced their subsequent behaviour. The control group and non-smart meter trial groups had the survey about six months earlier than the smart meter trial groups.

B2.4 Wider applicability of E.ON's findings

Several aspects of the trials reduce confidence in applying the findings more widely to the GB population, without entailing a bias in any obvious direction. It is also possible that these same aspects would either increase or decrease noise in the data, which would increase confidence in the significant or non-significant differences respectively.

- *Stratification.* Owing to limited information on the sociodemographics of the E.ON customer base, there was no attempt to make the groups representative of E.ON customers. 'Fuel poor' households were defined at postcode level, hence many households would not have been fuel poor (and some not 'fuel poor' households might have been fuel poor).
- *Customer exclusions.* Customers were excluded from the sample if not on standard, Age Concern, Age UK or green tariffs, or participating in other trials being run by E.ON. This is in addition to the biasing exclusions noted above.
- *Dwelling exclusions.* Excluded if dwelling less than two years old or if the meter was "unusual or complex" or had recently been exchanged.
- *Geographic range.* Most of the homes are in the English midlands.
- *Geographic targeting.* For the smart meter strata that had better recruitment responses, calls prioritised geographical availability to optimise the use of installers' time. This would create a geographic bias and reduced geographic variability compared to difficult cells, and a different geographic basis to the control groups.
- *Clip-on RTD installation.* The RTDs could not be fitted in all homes – there might therefore be dwelling bias in addition to any acceptance bias. Reasons for installation failure are not documented at household level.

Capabilities on project:
Building Engineering

B3 Scottish Power quality assessment

B3.1 Overview

Overall the metadata tend to give confidence in the null findings for Phases 1 and 2. The implication is that the three component interventions independently would also have had no effect.

It remains possible that reduction in demand could be brought about by more effective delivery and content of advice and consumption data, and/or greater take-up of the RTD. In this context, if those who used the RTD (at all or throughout) can be identified, further analysis of this group might be worthwhile.

The reductions in electricity consumption in Phase 3 are subject to several concerns. Apart from a range of sources of bias, the possibility of Hawthorne effects cannot be discounted, i.e. occupants can change their behaviour because they know they are being studied and not because of the intervention itself. This effect tends to be more prominent at the start of a study and, as each Wave of intervention was only 3 months, it is difficult to know whether the impact arises from the intervention itself or the Hawthorne effect. While prepayment customers might have been subject to the same influences, they might have had less scope to save energy and might reasonably have felt less inclined to do so.

The variation over time in Phase 3 outcomes might be seasonal as suggested by Scottish Power. It could also be at least partly due to smoothing of seasonal variation in the baseline and control group, as a result of data interpolation (as demonstrated by E.ON). Alternatively, customers might have lost momentum between waves because of time gap or the target could have become more difficult because they had already tried twice. It is also possible that the calculation of customer targets over-compensates for summer.

It is also uncertain whether a reduction in demand could have been brought about by challenge alone, without financial reward (other than the cost of the energy saved), or a greater effect could have been achieved with a differently structured schedule of incentives, set out in advance for customers.

The issues for gas are similar to those for electricity, the major difference being that the Phase 3 effect on gas consumption emerged gradually over the summer, rather than the sharp change seen at the start for electricity consumption. This is less likely to be due to the Hawthorne effect but is more likely to be a seasonal smoothing effect. It could alternatively be that the main gas intervention (the gas pledge) was introduced in August and its impact on heating would initially be minimal but increasing through the autumn. The variation over time might be seasonal as suggested by Scottish Power but this does not fit with the different seasonality compared to electricity consumption, with customers concentrating on gas consumption outside the heating season.

Overall the metadata leave little confidence in the significant reduction in consumption found in Phase 3. However, this finding shows (as a minimum) that the credit households had the capacity to reduce consumption if the motivation to do so was sufficient.

The lack of significant reductions in the prepayment groups is in stark contrast – the prepayment customers differed obviously in the means of payment (therefore meter type) but also in having lower initial consumption and a range of biases in different directions. On balance, the non-significant finding is credible in comparison with the credit groups. It is worth noting, however, that the percentage of customers meeting the challenge appears to tell a different story and the prepayment group might have had a different response to the financial incentive, rather than no response.

Capabilities on project:
Building Engineering

B3.2 Layered design

Layering was a poor experimental design because it makes it difficult to exclude the Hawthorne effect, since the study is divided into a series of short intervention periods, or to separate out the effects of different interventions, although, in hindsight, it may be a realistic way of representing the way that actual roll-out could occur.

It is not known whether the final financial incentive would have been equally effective without the earlier "priming" of ineffective interventions. Means and opportunity have to be added to the financial (and possibly environmental) motivation, and the earlier interventions might have provided these. However, the fact that effects were similar in all three credit test groups suggests that the incentive alone was the critical factor.

B3.3 Sample bias

Opting in to a trial may create a bias to customers more interested in energy saving (for financial or other reasons), this bias increasing with the demands of participation but decreasing with the value of the intervention as perceived by the customer. While such bias might generally be expected to make the interventions more likely to succeed, it is also possible that those most interested in energy savings had already taken action and therefore had less potential for further action. These influences may at least partly balance each other.

Scottish Power reduced such bias by approaching both the control group and test group participants (other than TG3) as though business as usual (they were not aware that they were in a trial). Within each trial group, all households were offered the same interventions, irrespective of social circumstance, payment method, meter type or payment history. This reduces (but does not eliminate) the risk of self-selection bias and also provides a control for it: the fact that TG1 and TG3 did not differ in their consumption changes reduces concerns generally about opt-in bias.⁸

Other aspects of the trial would tend to increase bias towards more energy-aware/literate households.

- Households were allocated to groups before first contact. Trial and control groups are implied to be comparable on Mosaic group but no actual tests were reported.
- The initial invitation to receive the planned interventions was sensibly composed but perhaps introduced too many new concepts at once: "Do you want to reduce your energy bills? ... we may be able to help ... join the growing number of our customers who can see exactly how much energy they're using – and how much it's costing day-to-day, minute-to-minute ... simple ... new gas and electricity meters ... also receive an energy saving monitor ... neat, easy to use device shows ... spending on energy ... no catches ... no cost ... After a few months, we will ask you about how useful you have found the energy saving monitor".
- Customers were excluded if, during the past two years, they had not continuously been supplied gas and electricity by Scottish Power, customer details associated with the premises had changed, or if there had been meter exchange or replacement. Credit customers were additionally excluded if they had been on prepayment in the past two years, if they were not on a single tariff or if they did not have valid actual meter reads over last four cyclic visits for both gas and electricity. Prepayment customers were excluded if they did not have an electricity prepayment key meter (installed no later than Dec 06), were not paying back debt for electricity (with or without debt for gas) or were not linked to a list of meter readings and date of readings from each payment transaction for Jan 06 to date of selection. If prepayment customers had a credit meter for gas, there additionally had to be successful capture of valid actual cyclic reads. These are logical exclusions but could create a bias to less mobile households and those with a more positive view of Scottish Power (hence the credibility of the company's advice

⁸ TG3 did not experience Phase 1 and were the only householders in the study asked to volunteer to have their meters upgraded. They were told that doing so would provide better visibility of energy use.

Capabilities on project:
Building Engineering

and take action on energy efficiency issues). They might also introduce bias to households with someone at home during the day, hence more interest in saving energy.

Retention of households in the sample was good, varying between groups from 83% to 90% of consumers retained over the period February 2009 - February 2010. Attrition was due to change of energy supplier for electricity and/or gas, tenancy or credit/prepayment status, or "by householder request" (e.g. customer did not wish to receive non-bill-related correspondence). This might further bias the sample in relation to mobility, ease of changing supplier and changing financial circumstances. However, a check for this does help to allay concerns: Control Group 1 maintained a similar time course in consumption after drop-outs, compared with a representative sample of non-participant households.

Prepayment groups had lower pretrial consumption than credit groups, by about 50% for electricity and 25% for gas. It could be argued that prepayment groups may be already at or close to the limit of what they could do to reduce their energy consumption without significant financial investment. Alternatively, prepayment households might already experience regular intervention by having to know when to top up and by how much, to last for how long.

B3.4 Intervention quality issues

Reviewing each of the interventions, it is now possible to see ways in which each could have been improved. The implication of this is that some interventions were successful even with their limitations, and greater energy savings might be achieved with improved interventions. In other cases, interventions were not successful but might have been if designed and executed differently.

Energy advice

Some of the literature provided to customers contains potentially confusing or misleading statements, which would tend to reduce the likelihood of the intervention being effective.

- Advice suggests to some customers that turning down their thermostat would reduce their heating bills by 10% (which is possible though not guaranteed) whilst other customers are told that the same action would reduce their energy bills by 10% (which is very unlikely).
- The clip-on information specifies a comfortable temperature range of 20-25°C, which might suggest to customers that they should not reduce their thermostat setting below 20°C, and that 25°C is acceptable from an energy perspective. The information also gives an energy cost example of boiling a kettle, which is a very low saving and could give the impression that energy is cheap and reducing use is not worth the effort.
- Green Challenge Wave 1 leaflet.
 - Starts with terms & conditions, which might have discouraged reading the rest. Then says the challenge should be easy (encouragement to some, removing the challenge element to others).
 - Two of the main six suggestions (low energy bulbs and cavity wall insulation) require a purchase, with no information on where to do it or what it might cost or what subsidies are available.
 - Chart of progression from no-cost to "more extensive" measures (which range in cost by two orders of magnitude) makes assumptions about what households had already done. Some of the no-cost measures depend on having or acquiring certain equipment (e.g. a shower, a washing machine with a cool wash programme). Others have to be interpreted or the logic might not be obvious, e.g. does "the amount of water you need" in a kettle include the water needed in an hour's time, or what does it mean to "overstock" a fridge and why is it a problem?
 - Some of the advice has risks, which householders might spot immediately or discover later, undermining all the advice. E.g. draughtproofing windows, door and floors too thoroughly can increase the condensation risk (or worse, if there is an unflued combustion appliance).

Capabilities on project:
Building Engineering

- Low-flow shower heads might not deliver an adequate flow as retrofit and/or householders might need to take a longer shower to compensate for reduced flow so there might not be actual energy saving for the outlay – a shower head is cheap but taps (and having them fitted) can be expensive.
- Reducing draughts is not the main point of floor insulation. Misspelling of "draught" does not give a professional image to the advice.
- Green Challenge Wave 2 leaflet.
 - Concentrates on a narrower range of behaviours – lighting, TVs and fridges, with a ready reckoner on savings potential – which is probably a good approach but some people might find the associated data tables hard to follow.
 - Final advice then focused on heating, not the electricity demand that was the main point of the challenge.
- Green Challenge Wave 2 mid-term update postcard. Encouragement to turn off appliances and hang washing outside (annoying if there is no secure, dry area outside, especially in early spring).
- Green Challenge Wave 2 conclusion. Given the timing ("With summer approaching"), the focus on heating is odd and, at times, confusing: what does turning the heating down actually mean – the boiler thermostat, the room/radiator thermostats or the heating period? The advice on thermostats seems to misunderstand what a thermostat is.
- Letter sent July 2009, between Green Challenge Wave 2 and Wave 3. Emphasis on turning down the thermostat (which should have no effect in summer) could disengage users.

Additional bill data

- Consumption comparison (year-on-year) letters were sent on only two occasions (June & November 2007), which limits their potential effect.

Demonstration of equipment

- All devices were installed by trained operators who subsequently explained how the devices worked, using a scripted demonstration. Scottish Power reports a low level of enquiries after installation. This should increase the likelihood of the intervention being effective and is particularly important given the relative complexity of the user guides for the devices used.

RTDs

- Approximately 40% of credit and 25% prepayment households could not accommodate the clip (mainly due to inaccessibility of meter tail).
- A letter and fresh batteries were sent during Phase 1 to encourage use of the clip-on but it is not known how many changed the batteries (either without detaching the clip or replacing it correctly afterwards).
- In approximately 23% of dual-fuel homes, the gas and electricity meters could not both be upgraded (most often because the gas meter could not be replaced). The RTD would also not have shown gas data for the affected homes. In effect, some TG1 homes would have been more like TG2 for gas consumption.
- Customers had questions about the running cost of the mains RTD and whether it should be switched off at night. Scottish Power suggested that it was good practice to switch off all non-essential appliances at night. This is right from an energy perspective but could cause problems if display settings are lost.
- User guides were probably too complex for many users – they do not take users direct to how to display the most relevant information, or how to use this information.
- Green Challenge communications did not mention RTDs specifically, although some did suggest that customers keep an eye on their meters.

Capabilities on project:
Building Engineering

B3.5 Noise in the data

Several aspects of the trials potentially increase or decrease noise in the data, which would increase or decrease confidence in the significant differences or non-significant differences respectively.

- *Geographic control.* The sample covers a narrow geographic range and is likely sufficiently well clustered that weather and seasonal correction are not necessary. Also reduces variance due to local extraneous changes.⁹
- *Trial group top ups.* Meter replacement was explained to householders as routine replacement, hence homes that had already recently had a meter replacement had to be excluded from Phase 2. Combined with failure to gain access to the property, this meant that sample numbers had dropped more than could be accepted. Therefore TG1 and TG2 were topped up with new participants during Phase 2.
- *Survey participation.* Some participants (including controls) took part in customer surveys early in the trial period, which could have influenced their subsequent behaviour.
- *External factors.* Consumption will be affected by factors extraneous to the project such as the economy, energy price changes, changing outdoor temperatures, credit crunch, press coverage and national campaigns. It is assumed that the effects of such factors are felt equally across all groups. This assumption probably has to be made but it is potentially flawed, e.g. because these factors create a pressure to reduce energy use in the control group, thus making the baseline change more difficult to exceed in the trial groups, or because trial participants who would have made investments in energy saving became concerned about spending capital.

B3.6 Wider applicability of Scottish Power's findings

Two aspects of the trials reduce confidence in applying the findings more widely to the GB population, without entailing a bias in any obvious direction. It is also possible that these same aspects would either increase or decrease noise in the data, which would increase confidence in the significant or non-significant differences respectively.

- *Dwelling exclusions.* Excluded if dwelling less than two years old.
- *Geographic range.* The sample was designed to be demographically representative of the Scottish Power customer base but covered a small geographic area.

⁹ E.g. a local authority or newspaper running a campaign or a major employer making redundancies. Sociodemographic variance needs to be considered separately because it arises within small areas.

Capabilities on project:
Building Engineering

B4 SSE Quality assessment

B4.1 Overview

The design and method of SSE's trials have the potential at least to increase noise in the data and, in some cases, bias – sometimes towards a significant effect and sometimes away (or in a direction that cannot be predicted). Some of these issues could be resolved in further analysis, thus allowing greater certainty in both significant and non-significant effects.

Looking at the metadata for each individual intervention, there is substantial scope for better implementation of the same type of intervention to reduce consumption.

On balance, there is reasonable confidence in SSE's null and significant findings but there is a reasonable chance that different implementation of the same type of interventions could have been more successful.

B4.2 Sample bias

Opting into the trial (or not opting out) has the potential to create a bias towards customers who are more interested in energy saving (for financial or other reasons), this bias increasing with the demands of participation but decreasing with the value of the intervention as perceived by the customer. While such bias might generally be expected to make the interventions more likely to succeed, it is also possible that those most interested in energy savings had already taken action and therefore had less potential for further action. These influences may at least partly balance each other. Opting in varied between trial groups, as described in Section 3 of the main report.

Households were allocated to groups before first contact but those customers aware of the trial were not told at recruitment stage the specific technology they would receive.

The wording of the initial letter to customers was generally a bit awkward in style and use of the jargon "trial" could be confusing or offputting for some customers. Who is on trial? The letter opens with "save the planet" motive, which is not something that appeals to all people; one form of letter then moves on to use "new technology" as a motive. A questionnaire had to be completed in the earlier phases of recruitment, which put off some potential recruits. Any demand on time is likely to be offputting and in this case there was likely to be at least one question the householder could not answer and because of ambiguity (e.g. "Is there a television in use ..."). The smart meter letter is inconsistent, stating that a limited number of smart meters is available but SSE wants to give one to every customer who wants one. This all tends to create bias towards more informed and interested customers.

The recruitment success rate is low (34-56%) except for the control group (93%), meaning that those who agreed to take part in trial groups (especially the Aware and Committed groups) are more likely to be interested in energy saving. In contrast, those interested in energy saving might have been less likely to join the control groups.

Pretrial consumption was similar across groups and therefore a major potential source of bias was not present.

Customers were excluded if they had not been with SSE for at least two years, with regular readings over that period. This could create a bias to households that are less mobile, have someone at home during the day (increasing energy demand) and more favourably disposed towards their supplier, which could mean they are more likely to take action on energy efficiency issues).

Retention in the sample was about 65% overall (said to be the expected level for business as usual, but worse than for other trials). There was a higher attrition rate in non-smart meter groups (including overall control group). This could tend to leave the final group more interested in energy saving, particularly the non-smart meter groups.

Capabilities on project:
Building Engineering

B4.3 Intervention quality issues

Reviewing each of the interventions, it is now possible to see ways in which each could have been improved. The implication of this is that some interventions were successful even with their limitations, and greater energy savings might be achieved with improved interventions. In other cases, interventions were not successful but might have been if designed and executed differently.

Advice

- The advice booklet offered a lot of useful information – possibly too much for households to take in or find the parts most relevant to them. The reader has to get to page 13 before finding advice on a likely savings action, such as installing efficient bulbs. However it was nicely laid out with margin tabs to find different sections. Unpredictable impact overall but could increase within-group variance by stretching the difference between those customers who are interested and those uninterested in energy.
- There was also a customer helpline and offers of free insulation. This is a sensible provision but it is not known how many households benefited from it.

Additional bill data

- The advance letter about graphs shows exactly what the customer will get, and the graphs themselves were big, in colour and on a separate page. Nevertheless, over two-thirds of recipients did not recall their bills having changed.
- Benchmark figures could be well out without reference to size of household and size and age of home, making it either easy or impossible for customers to try to get down to the benchmark level if it is lower than their own consumption. There is also a risk that benchmarking will cause those below their benchmark figure to relax and increase consumption, offsetting the efforts of others to reduce consumption (this would show in a change of variance rather than mean but benchmarking did not affect the variance).

Smart meter

- Manual readings were kept up to produce normal quarterly bills. This could cause customers to take the smart meter less seriously.

RTDs

- The customer had to install the clip-on device used in non-smart meter groups; the mains device used with smart meters was installed by the smart meter installer but s/he provided no explanation of the device. In both cases, the cost per kWh remained set at an initial default price if it was not updated by the customer, which could also have undermined confidence in the device.
- Two types of RTD were used with smart meters with different geographical distribution (ecoMeter and Home Energy Monitor – HEM). They had similar functions except the traffic light signals (representing current power load) were provided only with the ecoMeter. Also, while the ecoMeter has been found to be reliable in other trials, the HEM appears to have been problematic in this study. The type of RTD was not coded in the database, and therefore cannot be included in the analysis as a confounder, but is largely represented by the geographic variables.
- The user guides are designed for someone who bought the device and wants to use it – they were probably too complex to be an optimal part of an intervention. The guided do not take users directly to how to display the most relevant information, or how to use this information. The HEM guide is somewhat easier and does suggest how to use it to save energy. The ecoMeter guide also does not have high quality visual presentation, which could undermine confidence in it.

Capabilities on project:
Building Engineering

- *Smart meters.* For about 25% of dual fuel trials, only the electricity meter was installed and not the gas. The RTD would also not have shown gas data for the affected homes. These homes were counted as being dual fuel but with missing gas data.

Web information

- Customers received a letter saying the website was available but otherwise information about the website was provided through the website itself. The explanation of the graphs is in small font.
- Did not include real time feedback (delay of a day in data being available).
- “Energy Viewer” (personalised billing information) had a lot of options for displaying consumption data but no connection to advice on savings.
- Required customers to provide an email address but only about half did so (it is not clear whether the remainder were still counted as being in the trial; if they were, this potentially increases variance in the data).
- Low uptake: 1,428 were invited to use the web site. 733 applied, 372 logged in (182 more than once).
- Customers not involved in the trial were also able to access the website for general advice. This reduces the ability to detect impact of intervention but probably not in a major way.

Incentive to shift

- Initial letter to customers is confusing as it appears to assume the customer is already on a TOU tariff, before the shift incentive is explained.
- Includes non-consumption incentives (i.e. to pay on time or by direct debit). It is not clear whether this was common to all customers; if it was not, it could detract from the consumption incentive if it looks easier to get discounts other than by shifting load.
- Designed to for bills to break even for average users but eliminated the possibility of losing money. This is probably essential to recruit for the trial but makes a positive impact less likely and the findings less widely applicable.
- The use of “shadow bills” to indicate the financial impact of being on the TOU tariff might have been unnecessarily complicated and discouraged engagement, especially as shadow bills were sometimes missed (because of smart meter communications failure).
- Graphs of half-hourly energy use are probably too detailed for most people to take in. One bar per period (hourly consumption rate) plus peak half hour in each period would be easier to follow.

B4.4 Noise in data

Several aspects of the trials potentially increase or decrease noise in the data, which would increase or decrease confidence in the significant differences or non-significant differences respectively.

- *Geographic control.* The geographic spread of the sample is very wide. Season/weather corrections will deal with some aspects of this but not the impact of local extraneous changes.¹⁰
- *Stratification.* For electricity-only customers, there is no record of whether gas was supplied by another company.
- *Other data exclusions.* Data cleaning removed 16% of households from trials. This was evenly distributed across Mosaic groups but it is not clear whether it was biased by trial group or response to interventions.

¹⁰ E.g. a local authority or newspaper running a campaign or a major employer making redundancies. Sociodemographic variance needs to be considered separately because it arises within small areas.

Capabilities on project:
Building Engineering

- *External factors.* Consumption will be affected by factors extraneous to the project such as the economy, energy price changes, changing outdoor temperatures, credit crunch, press coverage and national campaigns. It is assumed that the effects of such factors are felt equally across all groups. This assumption probably has to be made but it is potentially flawed, e.g. because these factors create a pressure to reduce energy use in the control group, thus making the baseline change more difficult to exceed in the trial groups, or because trial participants who would have made investments in energy saving became concerned about spending capital.
- *Survey participation.* Some participants took part in customer surveys early in the trial period, which could have influenced their subsequent behaviour.

B4.5 Wider applicability of SSE's findings

Several aspects of the SSE trials reduce confidence in applying the findings more widely to the GB population, without entailing a bias in any obvious direction. It is also possible that these same aspects would either increase or decrease noise in the data, which would increase confidence in the significant or non-significant differences respectively.

- *Exclusions.* Households were excluded if they were on Economy 7 or another off-peak tariff, or if the home was less than 2 years old.
- *Geographic targeting.* Recruitment for smart meter groups targeted areas close to depots.
- *Match to population.* Not geographically or demographically matched but a wide geographic spread and could probably be grossed from equal numbers to actual distribution by Mosaic group. Larger than expected percentage of home owners, larger homes, high income, direct debit customers and use of energy-efficient lights.