Building Engineering

Ofgem

June 2011

Energy Demand Research Project: Final Analysis

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

c:\documents and settings\rossd1\my documents\edrp\public report\edrp public report - v2.2.docx

Acknowledgements

The final analysis presented in this report has been undertaken by a consortium led by AECOM. The authors wish to thank the following participating partners.

- The statistical analysis team which comprised the following (in alphabetical order of organisation).
  a) Reza Tolouei, AECOM
  b) Lorna Hamilton, the Building Research Establishment (BRE)
  c) Neil Spencer, the Statistical Services and Consultancy Unit (SSCU) at the University of Hertfordshire
  d) Eleanor Allan and Polly Grove, the Statistical Services Centre (SSC) at the University of Reading
- Sarah Darby of the Environmental Change Institute (ECI) at the University of Oxford for her advice on consumer behaviour related to energy use.
- Tom Lelyveld, AECOM and John Parsons, BEAMA for their input on the practical and technical issues associated with EDRP.

Significant input to this final analysis has also been provided by Ofgem's technical advisors to EDRP. In particular, we appreciate the assistance of the following.

- Joshua Thumim, Will Anderson, Tom Wilcox and Vicki White, the Centre for Sustainable Energy (CSE)
- David Shipworth, University College London (UCL)

In producing this report, we have been helped throughout by the four energy suppliers and their academic advisors. We acknowledge that in addition to these named individuals there have been many others working on EDRP within the supplier companies, equipment manufacturers and local communities.

- Martin Bell and Andrew Jones, EDF and Lee Stapleton, Science & Technology Policy Research (SPRU), the University of Sussex
- Andrew Brook and Steve Russell, E.ON and Andrew Wright, Hui-Wen Hsiao and Katherine Irvine, Institute of Energy and Sustainable Development (IESD), De Montfort University
- Graham Smith and Keith Thomson, Scottish Power
- Ian Lansley, Richard Westoby and Sai Rajagopal, Scottish and Southern Electricity and Carlos Barahona and Cathy Garlick, the Statistical Services Centre (SSC) at the University of Reading

Finally, we acknowledge the support of the funders of this final analysis.

- Kate Smith, Sarah Samuel and Harry Barton, Office of Gas and Electricity Markets (Ofgem)
- Michael Harrison and Pau Castells, Department of Energy and Climate Change (DECC)

This report was commissioned by Ofgem on behalf of DECC and any queries about the report or the Energy Demand Research Project should be addressed to DECC on 0300 060 4000 or at correspondence@decc.gsi.gov.uk.

Table of Contents

1 Executive Summary

This section summarises the overall findings and conclusions from the Energy Demand Research Project (EDRP). EDRP was a major project in Great Britain to test consumers' responses to different forms of information about their energy use. Four energy suppliers each conducted trials of the impacts of various interventions (individually or in combination) between 2007 and 2010. The interventions used were primarily directed at reducing domestic energy consumption, with a minority focused on shifting energy use from periods of peak demand. The project involved over 60,000 households, including 18,000 with smart meters. Measures were generally applied at household level but one supplier also tested action at community level.

The final analysis collates and builds on suppliers' findings and additional analysis of the data on energy consumption and consumer feedback, together with a review of the wider literature to identify the interventions that have proved most effective in reducing consumption, and key messages about how such interventions can best be delivered. The report provides a unique source of information particularly pertinent to the forthcoming national roll-out of smart meters. It sets out the quantitative savings that were achieved through a range of interventions enabled by smart meters, demonstrating the potential of smart metering as an enabling platform for measures to influence consumer behaviour.

1.1 Introduction

EDRP was launched in July 2007 and has been managed by Ofgem on behalf of DECC. The trials were undertaken by four energy supply companies: EDF Energy Customers Plc, E.ON UK Plc, Scottish Power Energy Retail Ltd and SSE Energy Supply Ltd (EDF, E.ON, Scottish Power and SSE). The Government allocated £9.75 million to the trials, match-funded by the energy suppliers taking part.

This report presents a final analysis of the findings from EDRP. The project was designed to help understand better how domestic consumers react to improved information about their energy consumption over the long term. This included investigating the impact of measures to reduce energy consumption and, in some cases, to shift energy demand from periods of peak demand.

The analysis presented in this report has been undertaken in the context of the Government's proposed roll-out of smart meters – hence there is a greater focus on the smart meter trials. The trials started prior to the Government's decision to roll out smart meters and so were designed to trial a range of different interventions. Nevertheless, much useful information has been acquired in support of the plans.

The trials have been made up of different combinations of measures and explored the responses of around 60,000 different households. The trials began in 2007 and finished towards the end of 2010.

This report presents the following.

- Findings reported by the energy suppliers on changes in energy consumption, resulting from the introduction of the interventions, and consumer feedback on the interventions.
- AECOM's evaluation of the energy suppliers' findings.
- AECOM's additional analysis of the energy suppliers' energy consumption data and consumer feedback. This builds on the work done by the suppliers to extend and clarify the impact of the various interventions tested.
- An assessment of the findings of the trials in the context of the academic and professional literature on trials of similar interventions (including a detailed literature review).
- The key practical and technical issues identified by the energy suppliers associated with the installation and operation of smart meters in EDRP and possible implications for the national roll-out of smart meters.

Capabilities on project:
Building Engineering

1.2 Summary of the trial design

The trials were constructed through a call for tenders to energy suppliers to bid for matched government funding to conduct trials on energy demand reduction. The brief was not prescriptive of the size, content, design, methods or form of analysis of the trials. This process resulted in commissioning of four essentially independent trials being conducted and analysed by the energy suppliers (EDF, E.ON, Scottish Power and SSE) and their academic advisors.

The energy suppliers each divided their trials into a number of trial groups to test the impact of different interventions. The interventions included the following (assessed either individually or in combination with each other).

- Energy efficiency advice.
- Historic energy consumption information (such as comparison of energy consumption with earlier periods).
- Benchmarking of the customer’s consumption against the consumption of comparable households.
- Customer engagement using targets (commitment to reduce consumption).
- Smart electricity and gas meters.
- Real-time display (RTD) devices that show energy use (including audible usage reduction alarms).
- Control of heating and hot water integrated with RTD.
- Financial incentives (including variable tariffs) to either reduce consumption or shift energy demand from periods of peak demand.
- Other digital media for delivering information (web, TV).

Table 1.1 provides a summary of the number of households who participated in the trials and the sub-set that had smart meters. The number of drop-outs during the trials ranged from 5% to 36% depending on the energy supplier. The analysis sections provide further details of the number of households included in the analysis for each of the different interventions and in-trial years.

Table 1.1 Number of households participating in EDRP

Energy supplier	*Total number of households*	*Households with smart meters*
EDF	1979	1879
E.ON	28450	8055
Scottish Power	3028	1330
SSE	27887	7106
Total	61344	18370

In addition, SSE assessed the effect of community engagement on behaviour and electricity demand reduction. This was undertaken in three villages: one each in England, Scotland and Wales. Each community had the same target and incentive: a £20,000 community project prize for achieving an average 10% reduction in electricity consumption over a three month period compared with the same three month period in 2007-8. Household-level interventions were provided, e.g. smart meters, RTDs (clip-on and linked to smart meters) and web access were common interventions deployed in the three communities. In addition, communities were given free rein to pursue other interventions and activities to achieve their targets (in individual dwellings, by insulation, new appliances, etc. or at community level).

1.3 Behaviour change

This report uses a simple theoretical framework based on the *means*, *motive* and *opportunity* for householders to change their behaviour (i.e. for householders to reduce energy demand, they must know what to do, have a reason for doing it and have the resources to do it). The wide range of behaviours that affect domestic energy demand may be characterised in terms of "*opportunity*" – the time required and the cost associated with them, and sometimes also the required space in the home. Therefore, a simple division of energy efficiency measures into "behaviour change" and "installation" does not adequately represent the range of options for saving energy. Insulation, for example, is not a behaviour, but installing insulation is a behaviour. Turning down a thermostat is a (no-cost) behaviour but it requires a thermostat to be installed.

The literature review conducted for this project shows that most changes seen in trials of energy demand reduction incentives tend to be – in the short term at least – those that require little investment of time or money. The EDRP customer surveys document a wide range of behaviours that householders used to reduce energy demand. While the trials do not allow changes in behaviour to be definitively tied to specific supplier interventions, the resulting changes in energy demand have been demonstrated more clearly than in past GB studies and – in some cases – for the first time.

1.4 Energy demand reduction seen in EDRP

Below is a summary of the changes in energy demand seen in the different interventions trialled in EDRP. The summary also takes account of a thorough review of findings from trials of similar interventions in the wider literature.[1]

1.4.1 Interventions without smart meters

With two exceptions, there was no significant reduction in energy consumption when the intervention did not include a smart meter.

The exceptions are interventions either using clip-on real-time displays (RTDs) of electricity consumption or "benchmarking" each customer's consumption against typical consumption in comparable households. In these two cases it was only SSE that found a significant reduction (in electricity consumption only, not gas consumption), and the effect was small (around 1% savings).[2]

The other trials found no statistically significant effect of RTDs, energy efficiency advice (on paper or online), historic feedback (on paper or online), self-reading of meters or financial incentives to save energy in the absence of smart meters.

[1] The findings presented in this executive summary are statistically significant. In interpreting the findings from EDRP we have considered the quality of the interventions, their delivery, the data and the design of the trials. Those interventions that did not show statistically significant savings might nevertheless be able to save energy in some customer groups and/or if implemented differently.

[2] Only the SSE trial included benchmarking. Although a small effect, this is one of the clearest pieces of evidence for an effect of benchmarking (the literature is generally positive but allows no quantification). There is concern that those who consume less than the benchmark amount may start to consume more; therefore, careful consideration needs to be given to how benchmarking might usefully be deployed in the population in general.

1.4.2 Interventions with smart meters

In contrast, interventions using smart meters were successful more frequently and with larger percentage savings in energy consumption. This may be explained partly by some aspect of receiving the smart meter (e.g. interaction with the installer or the positive image of getting new technology) but also the different options that were available once a smart meter was installed, e.g. more sophisticated RTDs (fitted by an installer), and more frequent and accurate historic feedback and billing.

Real-time displays

RTDs provide live data on energy consumption (kW and cost) and usually other information such as CO_2 emissions and energy consumption over specified periods. Some RTDs have audible alarms or visual signals to alert the customer to high consumption. The combination of smart meters and RTDs consistently resulted in energy savings of around 3% but with some higher and lower savings, depending on fuel, customer group and period.

In the case of electricity consumption, providing an RTD is the more important factor: savings were generally 2-4% higher than with a smart meter only (with a full range of 0-11% for some periods and customer groups) and these effects were persistent to the end of the trial. The percentage savings were greatest in the EDF trial and this may be because the accompanying advice was more effective. Only the Scottish Power trial showed no positive effect of RTDs with smart meters and this may be related to the fact that the meter replacement was presented as a routine replacement rather than as a smart meter or part of a research trial.

In the case of gas consumption, the smart meter itself (e.g. the information provided on consumption and cost) or some aspect of the experience of getting a smart meter appears to be a positive mechanism, resulting in savings of around 3%. E.ON found that these effects were persistent into the first quarter of the second in-trial year (i.e. for 15 months) and for one or two further quarters in some groups. The literature and other EDRP findings indicate that this effect may require support over time from other interventions (e.g. advice or billing information) to be sustained for longer periods.

The RTD findings are consistent with the literature. The effect of smart meters in isolation from other measures on gas consumption has not been investigated before but is in keeping with theoretical considerations, that real-time feedback is more relevant to electricity consumption than to gas. Applications of gas (e.g. heating and hot water) tend to be subject to more occasional adjustments having long-term effects (e.g. changing a thermostat setting in response to getting a smart meter or RTD).

The EDF survey data showed that customers expected, and could have benefited from, more engagement and instruction during installation of smart meters and RTDs. More generally, the written instructions may sometimes have been unclear or too complex, especially for people in the trial context (who were sent the device, rather than making an active decision to buy it). Hence support from the installers may be particularly important. This raises the possibility that greater savings might have been achieved with different procedures at this early stage.

In addition, EDRP provided the following evidence on the relevance of different aspects of the RTD, evidence that could be used in pursuing higher savings. Customer surveys about RTDs showed that cost information was used and valued more than unit (kW) information, and electricity information more than gas (this may be due in part to the order in which button pushes accessed different types of information, electricity generally being the default display). Displays of CO_2 emissions were generally not widely noticed or used or perceived as useful. Displays of temperature data were generally rated positively and may have been particularly useful in the early stages in responding to advice to reduce thermostat settings.

This is consistent with the literature, which additionally shows that portability of the RTD is a benefit (at least initially) and that appliance-specific feedback can have additional effects on energy savings (the latter was not tested in the

EDRP). The literature also identifies two key points that were not explored further in EDRP, but which are relevant to engaging consumers about RTDs:

- householders may find RTDs more useful in confirming savings after attempts to reduce consumption, rather than using an RTD to initiate savings;
- RTDs can be used to check that everything has been switched off before going to bed or leaving the house.

The mains powered RTDs (i.e. those connected to a smart meter), which displayed both gas and electricity consumption and tended to have more sophisticated functions, were consistently (across trials) more likely than clip-on RTDs to be fitted, retained, used and rated positively. Having an RTD also tended to result in more positive perception of the smart meter. The SSE survey found that reasons for not using the RTD differed between clip-on displays, not connected to a smart meter (where the functionality of the device itself was the dominant reason) and mains RTDs (where the usefulness of the information provided was dominant).

The audible alarm of high consumption in EDF's trial caused no incremental reduction in consumption and attracted only negative response in the customer survey. A "traffic lights" visual signal of consumption level, in contrast, was often the most positively rated feature. This is consistent with the limited evidence in the published literature.

There is nothing in the literature on the effect of an RTD with integral heating controller so the EDF trial of this device (combined with provision of advice) was the first test of its kind. There was no reduction of electricity consumption (data on change in gas consumption were not available). The customer survey revealed a positive response to the device but this may be because customers valued it mainly as a heating controller and did not use it so much as an RTD to view energy consumption information.

Energy efficiency advice and historic feedback

EDRP used generic written advice (not personalised to the customer) – mainly on paper but also via the web, a dedicated TV-based web page and RTDs. Advice was sent at varying frequencies and in varying amounts and styles. Historic feedback was principally a graphical comparison of consumption in the current bill/statement period and the same period the previous year. The EDRP findings for generic advice and historic feedback are consistent with the literature insofar as an effect of these interventions was not always seen and, when it was seen, the reduction in consumption was up to 5%. This was important to confirm in the UK context, given the previous dependence on evidence from elsewhere.

The particular combination of advice and historic feedback on consumption that EDF deployed (along with smart meters) reduced electricity consumption by 2.3% overall[3] in the first in-trial year.[4] The effect was persistent into the second in-trial year (4.0% saving). The E.ON trial found some weaker evidence of savings due to advice, in combination with smart meters and monthly feedback on consumption, but it is difficult to quantify because of the particular schedule of interventions. The other trials detected no effects of advice, historic feedback or the combination.

The effects observed by EDF occurred in spite of a survey finding low customer engagement with the material provided. This perhaps gives a clue as to why the EDF trial was effective: information was provided in simple, short statements, over a period of time – minimal but well presented and easy to absorb a little each month. The SSE advice booklet, in contrast, was comprehensive but required more effort from householders and was provided once

[3] When dual fuel and electricity-only customers were considered separately, the reduction was significant only for dual fuel customers (4.6% saving) and not for electricity-only customers (0.9% saving). However, in the second in-trial year, savings were significant for both dual fuel (2.4%) and electricity-only customers (5.0%).

[4] Gas consumption could not be assessed.

only, at the start of the trial. There was greater engagement with the SSE booklet if an RTD was also provided, so some of the effect attributed to the RTD may have been due to the RTD prompting interest in the advice and/or the advice helping householders to use the feedback from the RTD.

Similarly, the historic feedback was more obvious to customers in the EDF trial (large, in colour and on separate sheets) than that provided by E.ON. The SSE survey data did show that customers found the information on bills more useful if they also had advice, which further helps to explain the effect on consumption found by EDF. The survey also showed that many customers found the bill data and RTDs to be complementary, with a value in providing both.

The message is that advice should be provided as a fundamental requirement, and historic feedback can be useful, but the details of delivery, and combination with other interventions (e.g. smart meters and real-time feedback), are critical. EDRP has taken a step forward in understanding these details but there is more work to do and the optimum approach is likely to vary between customer segments, and to change over time as more energy knowledge becomes commonplace and people become more familiar with their consumption levels.

Incentives to reduce consumption

EDRP found no reliable or persistent effect of either financial incentives to reduce energy consumption or general statements of commitment to reduce consumption.

The literature provides little substantive evidence on financial incentives to meet a consumption reduction target except for the general (and obvious) point that sufficient incentive will prompt people to reduce consumption, but only for as long as the incentive is kept in place. Three EDRP trials employed financial incentives to reduce consumption but only Scottish Power saw reductions in consumption when the incentives were applied – only in the case of credit customers with smart meters and only for short periods. The Hawthorne effect[5] is a sufficient explanation of the Scottish Power findings. There are also concerns in the literature that using the financial motive in this way could focus householders' attention on financial savings and that this could reduce the chances of seeing long-term savings because other motives to reduce consumption are suppressed by the financial motive.

Similarly, having householders make a commitment to reduce consumption (without a specific target or reward) did not have any detectable effect on consumption. There is no directly comparable evidence in the literature but, in short-term trials, there is tentative evidence of savings prompted by realistic but stretching targets (without a reward), combined with frequent feedback.

Web-based interventions

EDRP used web-based services to provide advice, billing information and historic feedback (delayed by only a day but not real-time feedback), but without any effect on consumption.

The literature shows the potential benefits of online services to help consumers reduce energy demand, but also shows that the potential is rarely realised. EDRP also found this, with neither of the suppliers that used web-based interventions (EDF and SSE) seeing any energy savings as a result. The trials also showed that a major reason for failure is likely to be lack of engagement with the web sites, not necessarily a lack of effect among those who did use the sites.

---

[5] People may change their behaviour merely because they know they are being observed or tested, regardless of any specific attempts to change their behaviour. This "Hawthorne effect" tends to be short-term but can be reinstated by regularly changing the intervention, which is what happened in the Scottish Power trials.

Evidence on the provision of online real-time feedback is limited and, as yet, inconclusive in the literature. It was not tested in EDRP but surveys suggested that the online provision of feedback would have been more effective if real-time data had been provided.

1.4.3 Overview of changes in energy demand

A range of different smart meter-enabled interventions were demonstrated to deliver energy consumption reductions, and smart metering was demonstrated generally as a necessary enabling platform for behaviour change measures. While the savings were sometimes small in percentage terms, the absolute savings scaled up to national level would be substantial.

- The positive savings from smart meters depended on providing consumers with appropriate additional interventions, as discussed above.
- The provision of an RTD was particularly important in achieving savings in electricity consumption. Gas savings could be achieved through installation of a smart meter without further intervention, although evidence of persistence was not as strong as for electricity savings with RTDs.
- Electricity savings can be promoted through provision of advice and historic feedback on consumption but they cannot be relied upon on their own: one EDRP trial found significant savings with these two interventions in combination with smart meters and another found significant savings in combination with benchmarking against the consumption of a peer group (without smart meters).
- Financial incentives and commitment to reduce consumption, in contrast had either no effect or a very short-term effect.
- Delivery of information through the web or customers' TVs was also not successful in reducing consumption.
- Savings were generally persistent where the trial was long enough to test this, especially electricity savings from the combination of RTDs and smart meters. In contrast, any savings from financial incentives rapidly dissipated when the incentive was withdrawn.

However, the impact of an intervention depends on the detail of deployment: how a particular intervention is delivered and how it is combined with other interventions. Savings are not guaranteed simply by implementing a particular type of intervention and the following points need to be considered.

- A smart meter can provide key data and an RTD can relay that information. However, further information, advice and prompts are likely to be required if the impact is to be maximised.
- RTDs will have less impact if customers are relied upon to fit them: a significant proportion will simply not be fitted. Furthermore, where they are fitted, guidance needs to cover how to use the information that RTDs provide, not just how to access the information.
- *Consumers need to know what to do*: what *means* should be deployed to save energy. Some consumers already know (at least in part), others need further information. This information can be delivered through generic written material (on paper or online, before, during or after smart metering installation), verbal advice as part of the installer visit and by consumers being encouraged to experiment with an RTD to see the savings that could be achieved from a particular end-use of energy.
- *Quality matters*: information needs to be clear, easily seen amongst other material sent by suppliers, and presented in an attractive way. It also needs to be relevant and timely (e.g. appropriate to the season) and kept up to date as the options for action change (e.g. because of new technology or incentives). The design of RTDs and the explanation of how to use them are similarly essential to effective customer engagement, satisfaction and savings.

- *Quantity also matters*: a balance needs to be struck between providing sufficient information and avoiding information overload. For example, regular small nuggets of information appear to be more effective than a single delivery of comprehensive information (to provide information in manageable amounts and to maintain behaviour change prompts over an extended period).
- *People and households are not all the same*: the literature suggests that the more closely an intervention can be tailored to particular households or individuals, the more effective it is likely to be. Possible options range from full energy audits to a few selected pieces of advice, offered during smart meter installation. There is also potential to use the web more effectively: although web (and TV) interventions were unsuccessful in EDRP, information on the web could work for consumers more engaged with such forms of information and particularly if they receive tailored information (including real-time feedback and online audits).
- Quality, quantity and tailoring of interventions are relevant to all points in the customer journey: from engaging with an intervention at all (e.g. reading advice or installing an RTD), to the initial impact of the intervention and sustaining actions over a longer period.

Community engagement can also be an effective tool, making use of social networks and social capital, and moving social norms away from acceptance of energy wastage. It may, however, require a higher initial investment and will not necessarily work in all localities. Local support from a combination of experts and peers can help consumers to understand what to do, appreciate reasons for taking action (reasons that make sense to them personally) and provide the resources (time, space and money) necessary to take action: means, motive and opportunity.

1.5 Load shifting

Two trials (EDF and SSE) tested time-of-use (TOU) tariffs for electricity (i.e. tariffs that vary with time of day and sometimes season) in combination with smart meters and other interventions (advice, historic and real-time feedback, and incentives to reduce overall consumption). These trials showed effects on shifting load from the peak period, with bigger shifts at weekends than on weekdays. Estimates of the magnitude of shifting effect vary with trial but were up to 10%. The EDF trial showed that the effect is stronger with smaller households (1 or 2 people), thus providing a clear focus for where such interventions should be targeted. The effect was weaker in the SSE trial and this may be because awareness of the intervention was limited and it was seen as overly complex.

Neither of the TOU tariff trials involved any automation of energy-consuming appliances to facilitate load shifting. No data were gathered during the trials to provide evidence on what appliances or behaviours were responsible for the observed shifting.

The literature shows that time of use tariffs can also bring about reductions in total energy consumption. However, the evidence is almost exclusively from studies in hot regions (where the dominant energy demand is for air conditioning) and cold regions with electric heating. The limited evidence from the UK suggests small reductions (3% or less) in overall electricity demand and no such effect was detected in EDRP.

1.6 Population segment effects

Across all the trials, there was limited evidence of how different population segments were affected by the interventions. However, the following themes could be seen in the findings.

- Smaller households were more likely to save energy overall and to shift consumption from the evening peak period.
- E.ON's 'fuel poor' (FP) and not 'fuel poor' (NFP) groups are difficult to characterise – they differed in the fuel poverty index for their postcode and it is not known how many households were actually fuel poor in each group.

Nevertheless, the FP group is likely to be generally less affluent, with more pressure on energy costs. A generally more positive response to interventions in the FP group most likely signifies their greater motivation to save money.

- SSE found significant variation in energy consumption reduction with demographic group (based on the Mosaic classification) but the analysis did not permit identification of specific Mosaic groups that had a higher or lower propensity to reduce consumption. In no case did Mosaic group have a significant interaction with a trial intervention, i.e. the effects of interventions did not vary with demographic category.
- SSE also found variation in consumption reduction with postcode but again the analysis did not permit identification of specific locations that had a higher or lower propensity to reduce consumption. Also, savings did not depend on the more systematic geographic variables included in the analysis so the details of any location effect are difficult to define. In no case did location have a significant interaction with a trial intervention, i.e. the intervention effects did not vary with location.
- There were also some differences in savings between customers who purchased only electricity from the supplier responsible for the trial and those who also purchased gas from that supplier. These differences are difficult to explain with any certainty but they were not due to electricity-only customers not having gas heating because most of them did.
- Only SSE included prepayment smart meters (for electricity only) in its trial, with one trial group being given smart meters and an RTD. This group made savings similar to those in the credit customer groups.
- Scottish Power made more extensive investigations of prepayment customers but not with smart meters. These customers did not make savings relative to the control group under any interventions.
- Prepayment customers tend to have lower consumption anyway and therefore less scope to reduce. However, they may be more practiced at monitoring consumption, more aware of costs and have greater (financial) motivation to save energy. With these counterbalancing factors, savings appear to be possible (as shown by SSE) but not always achieved (as shown by Scottish Power). Logically, efforts to help prepayment customers reduce consumption might focus on means and opportunity, the motive already being present, but this was not explored in the trials.

1.7 Practical and technical issues

The primary purpose of EDRP was to investigate consumer behaviour and it was not set up as a technology or roll-out trial. The equipment used was what was either readily available or could be developed for use in the time available. Nevertheless the experiences of the installation and use of the equipment from EDRP provide valuable practical lessons for mass roll-out of smart metering.

Overall, the equipment employed worked well. Many of the practical problems encountered were reported as due to the ‘pilot’ nature of this project (e.g. equipment faults, data communication losses between the household and supplier, and managing the large volume of energy meter data received). Indeed, given the subsequent advances in technology, a number of the issues identified would not be encountered if starting EDRP today. The remaining issues identified are already informing work under the smart metering programme. It is also worth noting that some of the issues identified (access to properties, billing errors) would be encountered in “business as usual” meter replacement activity, whether smart or not, although over a longer period and therefore more easily managed.

A significant issue will be the customer interaction with the smart metering equipment (including RTDs). All trials reported some customers having difficulties in understanding the new equipment provided. This is not simply about the design of equipment, although that has been improving: care is needed in how RTDs are ‘sold’ and explained to the wide cross-section of population such that they know how to access and use the data that the display provides, and take an interest in doing so.

1.8 Further evidence development

Measuring changes in energy consumption and the associated household behaviour is a complex business. Carrying out research to attribute changes to particular experimental interventions is more complex still and this is reflected in uncertainties that remain on the subject of how to change householder behaviour. The literature shows some engagement of consumers with behaviour changes that require relatively little time and money to implement. If substantial savings are to be realised across wide segments of the population there is a need for consumers to move along a pathway of behaviour change, so that changes requiring more investment of time and money also occur more frequently.

In particular, the specific combinations of interventions that are most effective in achieving reductions in energy demand, and the details of how and when they should be implemented, are only partly understood. It is becoming clear that electricity and gas consumption are not affected in the same way but the details are only now starting to emerge. The way in which different segments of the population can be engaged, and how they will respond to interventions, also merits further investigation. The key questions concern getting consumers' attention, motivating them to take action and providing them with the necessary knowledge and resources.

The development of knowledge will need to be partly around the design of equipment and the associated user guidance – to establish the ideal balance of complex functionality and simple, attractive products. Equally, the media of communication could be enhanced, looking at approaches ranging from optimising the use of web-based approaches to person-to-person spread of knowledge and motivation through communities. The use of community-based approaches has been shown to be effective but, to be viable for application at national level, further research and analysis is needed to determine whether there are "key ingredients" for more cost-effective versions of the community approaches used to date and to provide good practice guidance.

Time-varying tariffs may play a role in managing energy demand, in controlling total and peak-time consumption. The optimum tariff levels and ratios, and the role of advice and technology in supporting behaviour change, are as yet poorly understood.

Some questions may be answered by further analysis of the EDRP data. More broadly, if specific issues are identified, more in-depth processing of the energy or survey data is certainly possible. Based on the experience of EDRP, much useful knowledge could also be gained from monitoring and evaluation of the smart meter roll-out. In particular, there is great potential for achieving energy savings through the installation process but the optimum level of interaction between households, installers and (if involved) energy advisors needs to be established.

There is also a need to understand better the impact of energy behaviour change interventions in the context of other policy initiatives targeting household energy use. For example, the Green Deal, Feed-in Tariff (FIT), Renewable Heat Incentive (RHI), Energy Company Obligation (ECO) and the Community Energy Saving Programme (CESP) will change the context in which behavioural interventions are played out. Changing the physical form and fabric of the property (Green Deal), the supply of electricity and heat (FITs and RHI), and the social, legal and financial context (CESP and Green Deal) will change the means, motive and opportunity of householders to engage in behaviour change. This, in turn, will impact on the efficacy of different packages of behavioural interventions in different ways. The capacity to evaluate the impact of existing behavioural change packages, and to develop and test new packages of behavioural interventions on a continuing basis will be important. In addition, there is a possibility that smart meters themselves, independently of any effects they have on energy demand, will be a useful tool for evaluating the impact of other policy initiatives.

2 Introduction

This report presents the findings from the Energy Demand Research Project (EDRP). EDRP was originally designed to help understand better how domestic consumers react to improved information about their energy consumption over the long term.

The interventions have included the following (assessed either individually or in combination with each other):

- energy efficiency advice;
- historic energy consumption information;
- benchmarking of the customer's consumption against the consumption of comparable households;
- customer engagement using commitment to reduce consumption;
- smart electricity and gas meters;
- real-time display (RTDs) devices which show energy use;
- control of heating and hot water integrated with an RTD;
- other digital media for delivering information (web, TV);
- financial incentives (including variable tariffs) to either reduce consumption or shift energy demand from periods of peak demand.

The idea for EDRP was established in 2006 when the then Department of Trade and Industry was interested in trialing smart metering and the Department of the Environment, Food and Rural Affairs was interested in trialing broader interventions to influence behaviour change. The trials were constructed through a call for tenders to energy suppliers to bid for matched government funding to conduct trials on energy demand reduction. The Government allocated £9.75 million to fund the trials. The brief was not prescriptive of the size, content, design, methods or form of analysis of the trials. This process resulted in commissioning of four essentially independent trials, conducted by the four successful energy suppliers (EDF, E.ON, Scottish Power and SSE) and their academic advisors.

EDRP was launched in July 2007. The energy suppliers carried out trials across Great Britain, incorporating different combinations of measures, and explored the responses of around 60,000 different households. The majority of the trials began in late 2007 or early 2008 and the trials finished towards the end of 2010[6]. AECOM was contracted in the autumn of 2010 to conduct an independent review of the trials and further analysis of the data collected.

This report presents AECOM's final analysis of the findings from EDRP. This analysis has been undertaken in the context of the Government's proposed roll-out of smart meters – hence there is a greater focus on the smart meter trials. The trials started prior to the Government's decision to roll out smart meters and were not specifically designed to feed into its implementation plans. Nevertheless, much useful information has been acquired in support of the plans.

The key components of this final analysis are as follows.

- An assessment of the energy suppliers' own findings.
- Further analysis of the trial data collected. This particularly included the energy consumption data collected for each supplier's trials, consumer feedback on their experience of the various measures and feedback from the energy suppliers on the practical issues associated with designing, installing and operating the various measures, including smart meters.
- Contextualise the findings with respect to the wider literature.

---

[6] EDF has continued with its heating controller trial group into early 2011 (see Section 3). This report presents analysis of data collected until the end of August 2010.

The contents of the remaining sections of the report are outlined below.

- Section 3 provides a summary of the trials undertaken by the four energy suppliers.
- Sections 4 to 7 present further analysis of the four energy suppliers' energy consumption data and consumer survey data. The principal purpose of this additional analysis is to build on the work done by the suppliers to extend and clarify the impact of the various interventions tested in EDRP.
- Section 8 puts the findings of the trials in the context of the academic and professional literature on trials of similar interventions.
- Section 9 brings together the key practical and technical issues identified by the energy suppliers associated with the installation and operation of smart meters in EDRP and considers possible implications for the national roll-out of smart meters.
- Section 10 presents the key conclusions from this study.

A glossary of terms and abbreviations is provided at the end of this document. Additional material is provided in appendices, in separate documents.

- Appendix A summarises the findings reported by the energy suppliers themselves (on changes in energy consumption resulting from the introduction of the interventions and on customer survey responses).
- Appendix B is a quality assessment of the final findings (i.e. following AECOM's analysis).
- Appendix C is a literature review of the impact (on energy consumption and the associated householder behaviour) of the interventions trialled in EDRP.
- Appendix D reports on SSE's community trials.
- Appendix E provides further analysis tables to supplement Sections 4 and 5.

3 Trial Design

Each of the four trials assessed multiple approaches to reducing energy consumption. This chapter describes the following for each of the four trials.

- The different interventions (measures to reduce energy consumption) assessed by each trial.
- The different trial groups within each trial. Each trial group comprised and assessed a different set of energy reduction measures.
- The methods of recruitment of trial participants.

The data collected (consumption and survey data) and the data processing and analysis are described in Sections 4 to 7, alongside the findings of the analysis.

3.1 EDF

3.1.1 Interventions

Non-smart meter interventions

EDF tested only one intervention that did not include a smart meter, the Read-Reduce-Reward (RRR) scheme for dual fuel customers, which combined the following interventions.

- Energy efficiency advice: monthly tips sent by post.
- Incentive to reduce consumption: reward for year-on-year reduction in consumption.[7]
- Customer engagement: customer reads meters and provides the readings to EDF.
- Web information: personalised consumption history available online.

Smart meter interventions

EDF tested the following interventions in trial groups provided with smart meters.

- Smart meter.
- Accurate billing and no meter reading visits.
- Additional bill data: graphs on monthly summaries (not bills) showing current period and historic energy consumption, cost and CO_2 emissions.
- Energy efficiency advice: monthly tips sent by post, on same sheet as additional bill data.
- Energy efficiency advice: monthly tips sent to RTD, TV or online.
- Real time display (“Wall Panel”): mains RTD with touch screen showing current electricity and gas use, cost (current month and per hour), CO_2 emissions, historic data and messages from EDF.
- Real time display (“basic display”): mains RTD with one-line display; five buttons to toggle between electricity use/cost on each tariff rate (current and per day, week and month), CO_2 emissions and tariff rates.
- Usage reduction alert (URA): RTD set up with audible alarm if consumption exceeds predefined daily level.
- Heating controller (HEC): RTD incorporating a controller for space heating and hot water.

[7] Rewards were provided as Nectar points (typical value of £1 per 200 points). A year-on-year reduction in energy use was rewarded with 1000 points per fuel. In addition, 250 points were offered per fuel per quarter for providing meter readings online (200 points for phoning in readings). This scheme is no longer available.

- Time of use tariff (TOUT): incentive to shift from peak period consumption.
- TV information: personalised consumption history available via a TV Freeview box.
- Web information: personalised consumption history available online.

The TOU tariff intervention was based on electricity tariffs varying with time of day. The peak period was 16:30-19:30, night period was 23:00-06:00 and off-peak period was 06:00-16:30 and 19:30-23:00 (the same times for both GMT and BST). The peak tariff was 161-169% of the off-peak tariff and the night tariff was 56-65%, depending on region. The off-peak tariff was between 8.41 and 9.03 pence per unit (excluding VAT).[8]

3.1.2 Trial Groups

The different trial groups are detailed in Table 3.1. The trial group numbers and names are referred to throughout this report in discussing the EDF trial.

- As noted above, there is only one non-smart trial group (RRR – Read-Reduce-Reward).
- All the other groups included smart meters and more accurate billing, including the control group. These comprise seven trial groups. The only subdivision is between households with EDF electricity-only accounts (who received electricity smart meters only) and households with EDF dual fuel accounts (who received electricity and gas smart meters).

3.1.3 Recruitment

Sampling frame and exclusions

EDF recruited from its customer base in London and the southeast of England. Initially customers were required to have 4 meter readings in the 12 months prior to the start of the trial to be eligible to join. This criterion was relaxed early in the recruitment process (but after the monthly advice and control groups had been recruited) to a minimum of 2 meter readings due to the shortage of customers fulfilling the original specification.

The following universal exclusions were applied.

- Customers on the Mailing Preference Service.
- EDF staff.
- Customers who had taken “Green products” (special tariffs or consumption reduction products).
- Customers with medical equipment in their homes.
- Customers with more than 6 readings in the 12 months.
- Customers for whom EDF did not hold a telephone number.
- Customers below a lower survival band of estimated annual energy consumption. This band is defined as 1,500 kWh estimated annual electricity consumption and 7,500 kWh estimated annual gas consumption.
- Economy 7 customers.

---

[8] These tariffs were used for trial purposes only and are no longer available.

Table 3.1 EDF trial groups

Trial group name	Interventions	Fuel type	Number installed
Paper$_{EO}$	• Smart meter + accurate billing • Additional bill data • Energy efficiency advice: sent by post	Electricity only	200
Paper$_{DF}$		Dual fuel	186
Wall Panel$_{EO}$	• Smart meter + accurate billing • RTD on wall (“Wall Panel”) • Energy efficiency advice: sent to RTD	Electricity only	170
Wall Panel$_{DF}$		Dual fuel	200
TV$_{EO}$	• Smart meter + accurate billing • TV information • Energy efficiency advice: sent to TV	Electricity only	97
TV$_{DF}$		Dual fuel	53
Usage Reduction Alert (URA)	• Smart meter + accurate billing • RTD on wall (“Wall Panel”) • Usage reduction alert • Energy efficiency advice: sent to RTD	Electricity only	200
Time of Use Tariff (TOUT)	• Smart meter + accurate billing • RTD (“basic display”) • Time of use tariff • Energy efficiency advice: sent by post[1]	Electricity only	194
Heating Controller (HEC)	• Smart meter + accurate billing • RTD on wall (“Wall Panel”) • Heating controller integrated with RTD • Energy efficiency advice: sent to RTD	Dual fuel	156
Control$_{EO}$	• Smart meter + accurate billing	Electricity only	187
Control$_{DF}$		Dual fuel	189
Read-Reduce-Reward (RRR)	• Energy efficiency advice • Incentive to reduce consumption • Customer engagement • Web information	Dual fuel	100
Web	• Smart meter + accurate billing • Web information • Energy efficiency advice: monthly tips online	Dual fuel	100
		Total	1979

[1] The intervention did not specifically include providing customers with energy efficiency advice. However, there was some general advice within the material sent to the customers about the intervention.

Sampling was stratified according to estimated annual energy consumption, with households divided into low, medium and high estimated annual consumption. Following this, further stratification divided potential recruits by the following EDF customer variables.

- Prepayment users: premises had a prepayment meter installed.
- Fuel poor: customers who are defined by a model that predicts the likelihood of a geographical area spending 10% of more of their take home salary on fuel bills.
- Green: customers on a green tariff or have a very high propensity to be a green customer based on a lifestyle code.
- Grey: customers of 55 years or older.
- High consumption: customers corresponding to ACORN demographics classes 1 to 14 or 24 to 36.
- Low consumption: the residual customers not in the other groups.

EDF amalgamated fuel poor and low consumption customers because they appeared to overlap somewhat and it was deemed desirable from a statistical point of view to limit the size of the sampling matrix.

A random number generator was used to allocate potential households to the trial groups, constrained by this stratification. Customer lists were prepared for each of the trial groups and supplied to an agency to undertake the recruitment. The sampling frames for each trial group were therefore randomised, with the following exceptions.

- Only medium baseline consumption households were recruited for the ‘time of use’ variable tariff trial.
- There were few high consumption households left on the lists, especially in London, towards the end of the recruitment for the TV and heating controller trial groups. These trial groups were predominantly recruited in the south-east.

Recruitment methods

The first step in the recruitment process was a phone call from the recruitment agency. All households in the EDF trial were offered the opportunity to join the trial to which they had been provisionally allocated and all were given the choice of whether or not to participate. All were aware that they would be participating in a trial, including control group households.

The recruitment agency sought to persuade households to join the trial. At first, the agency approached this task with a pro-environmental message but this was not successful so, after the first month, they adopted a four stage approach, promoting the benefits as follows: 1) saving money, 2) accurate bills, 3) saving energy, and 4) saving carbon. There were however important differences across the groups.

- The control group was offered a £50 shopping voucher and accurate bills but were not told that they would be able to save money or energy (only the very first customers – around ten recruited households – were not offered accurate bills).
- Each of the interventions was promoted on its own potential merits.
- The trial of the RTD with excess consumption alarm was difficult to recruit to because customers did not want an alarm.
- The variable tariff trial group was difficult to recruit to because customers did not understand the principle of load-shifting and did not believe that the company would want to help them save money. EDF had to provide the recruitment agency with a special training session specifically on this intervention.
- Across all groups, recruitment was easier in the south-east because people were at home more often when the agency called.

- Different groups were recruited, and installed, at different times. For those groups recruited at the beginning of the study, the time between recruitment and installation was about 6 weeks. This was eventually reduced to 2 weeks because of the initially high drop-out rate between recruitment and installation. The control group was recruited at the beginning of the study.

Households that agreed to join a trial were sent written terms and conditions to sign. These terms and conditions were simplified after the first month as the original documents were putting off 90% of the households.

The same installation teams were used for all trial groups except the heating controller trial group as these installations needed boiler experience. There was minimal interaction between installation teams and customers – the latter were given a booklet but were not given direct advice or instructions.

Table 3.1 (above) shows the installation rates for EDF's trial. Table 3.2 illustrates the timetable for the trial.

Table 3.2 Recruitment, installation and trial periods of EDF's trial

Trial group name	*Recruitment*	*Installation*	*Trial*
Paper	Oct 2007 – Dec 2008	Jan 2008 – Aug 2009	Jan 2009 – Sep 2010
Wall Panel	Oct 2007 – Dec 2008	Dec 2007 – Dec 2009	Jan 2009 – Sep 2010
TV	Oct 2007 – Sep 2009	May 2008 – Sep 2009	Sep 2009 – Sep 2010
Usage Reduction Alert (URA)	Oct 2007 – Dec 2008	Dec 2007 – May 2009	Jan 2009 – Sep 2010
Time of Use Tariff (TOUT)	Oct 2007 – Dec 2008	Jun 2008 – Dec 2009	Jan 2009 – Sep 2010
Heating controller (HEC)	Oct 2007 – Mar 2009	Feb 2008 – Aug 2009	Mar 2009 – Mar 2011
Control	Oct 2007 – Dec 2008	Aug 2008 – Apr 2009	Jan 2009 – Sep 2010
Web	Oct 2007 – Dec 2008	Jul 2008 – Jan 2009	Jan 2009 – Sep 2010

3.2 E.ON

3.2.1 Interventions

Non-smart meter interventions

E.ON's trial tested the effects of four interventions.

- Additional bill data: graphs on quarterly bills showing historic energy consumption information.
- Energy efficiency advice: monthly tips sent by post.
- Real time display: clip-on RTD showing current electricity use, cost, CO_2 emissions and historic data.
- Customer engagement: monthly request for customer to read meters and provide the reading to E.ON, so that E.ON could provide accurate bills (referred to as the "Hawthorne" group)[9].

[9] Designating this as the Hawthorne group is misleading: self-reading meters (and providing accurate bills) potentially achieve more than simply communicating to customers that they are in a trial.

Smart meter interventions

E.ON's trial tested the effects of six interventions.

- Smart meter.
- Accurate billing and no meter reading visits.
- Monthly bills.
- Additional bill data: graphs on monthly bills showing historic energy consumption information.
- Energy efficiency advice: monthly tips sent by post.
- Real time display: mains RTD showing current electricity and gas use, cost, CO_2 emissions and historic data, plus a "traffic light" indicator of current consumption.

3.2.2 Trial Groups

The different trial groups are detailed in Table 3.3. There are nine main trial groups which are then differentiated depending on stratification of the sample by customer type. The different categories and fuel/payment types are defined further in the recruitment section.

The control groups were combined together, in different ways, in the E.ON analysis to provide the best demographic match between controls and trial groups.

The non-smart meter trial groups did not have a full factorial design but did combine the advice and RTD interventions in one trial group. In effect, the additional bill data became part of the intervention in all trial groups during the early part of the trial, when this became part of standard billing procedure (except in the control group). The customer engagement intervention was withdrawn before trial completion.

Similarly, the smart meter trial groups did not have a full factorial design but added monthly bills, then advice, then RTDs to create groups with an increasing number of interventions.

3.2.3 Recruitment

Sampling frame and exclusions

E.ON used its customer base in the Midlands (25%) and East Midlands (75%) to recruit to its trial. The East Midlands has a legacy of a large number of households with Economy 7 (E7) meters, not all of which are used for storage heating (i.e. not all E7 metered households used an E7 product).

The sampling frame was limited to customers on standard, Age Concern or green tariffs. The following universal exclusions were applied.

- Customers with unusual or complex meters.
- Recent new builds with no consumption history and those with fewer than 2 actual meter reads in the last 12 months.
- Customers in other trials being run by E.ON.
- Customers with a recent meter exchange.
- Customers currently disputing a bill or where a warrant for non-payment was outstanding.

Customers with prepayment meters were excluded from the smart trial groups.

Table 3.3 E.ON trial groups and installation rates

Trial group	Trial Group Code	Interventions	Stratum	Fuel/payment type	Number installed	Group total
Control	C1	Control	FP	Std EO	1,358	5,022
	C2		FP	E7 EO	1,122	
	C3		FP	Std DF	1,555	
	C4		FP	E7 DF	987	
	C5		NFP	Std EO	681	2,609
	C6		NFP	E7 EO	458	
	C7		NFP	Mixed DF	1,470	
	C8		HUDF	Mixed DF		2,443
	C9		E7	E7 EO	606	1,271
	C10		E7	E7 DF	665	
	C11		NFP E7	Mixed DF		2,251
Non-smart metered trials	TG1-FP	• Additional bill data	FP	Std	754	1,525
	TG1-NFP		NFP	Mixed	771	
	TG2-FP	• Additional bill data • Energy efficiency advice	FP	Std	733	1,510
	TG2-NFP		NFP	Mixed	777	
	TG3-FP	• Self-reading of meters • Accurate billing • Additional bill data	FP	Std	743	1,507
	TG3-NFP		NFP	Mixed	764	
	TG4-FP	• Additional bill data • RTD	FP	Std	546	1,101
	TG4-NFP		NFP	Mixed	555	
	TG6-FP	• Additional bill data • Energy Efficiency Advice • RTD	FP	Std	592	1,156
	TG6-NFP		NFP	Mixed	564	
Smart metered trials	TG5-FP$_{EO}$	• Smart Meter • Accurate billing • Additional bill data	FP	Mixed EO	782	2,639
	TG5-FP		FP	Mixed DF		
	TG5-NFP		NFP	Mixed DF	719	
	TG5-HU$_{DF}$		HUDF	Mixed DF	696	
	TG5-E7		E7	E7 EO	442	
	TG7-FP$_{EO}$	• Smart Meter • Accurate billing • Monthly Bill • Additional bill data	FP	Mixed EO	733	1,436
	TG7-FP		FP	Mixed DF		
	TG7-HU$_{DF}$		HUDF	Mixed DF	703	
	TG8-FP$_{EO}$	• Smart Meter • Accurate billing • Monthly Bill • Additional bill data • Energy Efficiency Advice	FP	Mixed EO	743	1,456
	TG8-FP		FP	Mixed DF		
	TG8-HU$_{DF}$		HUDF	Mixed DF	713	
	TG9-FP	• Smart Meter • Accurate billing • Monthly Bill • Additional bill data • Energy Efficiency Advice • RTD	FP	Mixed DF	706	2,524
	TG9-NFP		NFP	Mixed DF	697	
	TG9-HU$_{DF}$		HUDF	Mixed DF	710	
	TG9-E7		E7	E7 EO	411	
					Total	28,450

EO = Electricity only, DF = Dual fuel, HU = High baseline use, FP = 'Fuel poor', NFP = 'Not fuel poor'
Std = Standard tariff, E7 = Economy 7, Mixed = Standard and E7 tariff.

Recruitment was randomised with no demographic stratification. However the groups themselves were stratified according to a four-fold categorisation of ‘fuel poor’ (FP), ‘not fuel poor’ (NFP), high baseline use dual fuel customers (HU_{DF}) and E7. These were defined as follows.

- FP households were defined using CSE’s fuel poverty indicator. This indicator provides an estimate of the proportion of households in any output area that are fuel poor. E.ON sorted its households by postcode sector, identified the highest level of fuel poverty among the output areas in each postcode sector, ranked all postcode sectors according to these values, then selected the 25% of postcode sectors with the highest ranks as ‘fuel poor’. The NFP households were defined as the households in the remaining postcode sectors. This is a highly inaccurate method for identifying fuel poverty and most of the FP households will in fact not be in fuel poverty. For consistency with E.ON’s analysis, we have continued to use these trial codes though we note that they are not accurate descriptions.
- HU_{DF} households were defined as those consuming more than 7,000 kWh of electricity per year (around double the Ofgem domestic average). This was the only group where any consumption threshold was applied; a full range of consumption levels will be present in all other groups.
- Households in the E7 group all have E7 meters. However these meters are also present in other groups.

E.ON also divided smart meter groups according to whether households had an E.ON electricity-only account (who received electricity smart meters only) and households with E.ON dual fuel accounts (who received electricity and gas smart meters). Dual fuel households were included in all trial groups except the E7 groups.

E.ON undertook CAMEO demographic profiling of its trial and control groups. Although there were prepayment customers in the non-smart trial groups, there were none in the control groups (or, as expected, the smart trial groups). As a result, E.ON removed the prepayment customers from the sample and developed a more complex specification of 11 control groups such that each household could be in more than one control group.

At this point the classification of households in the E7 groups was also reviewed. Households were only retained in the two E7 smart meter trial groups if they used an E7 product as well as an E7 meter, i.e. they were users of Economy 7 heating. However this exclusion was not applied to the E7 control groups which continued to include households with E7 meters but not E7 products.

Recruitment methods

Recruitment took place between April 2007 and October 2007 for the non-smart meter trial groups, between February 2008 and August 2008 for the smart meter trial groups, and between April 2007 and April 2008 for the controls.

Households in the non-smart meter trial groups were sent interventions as if this was business as usual. They were not invited to become part of a trial, nor were they aware that they were participating in a trial. However the households receiving the clip-on in-home display required an installer visit. These households were sent a letter informing them that the device was being given to them and that an installer would visit. Installers visited households without making prior appointments until trial quotas were filled. Consequently, although recruitment rates were 100% for the non-smart meter trial groups that received postal interventions only (graphs on bills and energy efficiency advice), they were lower for the non-smart meter households receiving the clip-on display, in part because access was not always possible or granted and in part because not all homes were suitable for the device (though reasons for installation failure were not logged by household).

All the smart trial groups required the active opt-in of households. Letters were sent to households describing the benefits of the particular combination of interventions they were being offered. These letters stressed the limited availability of the offer and encouraged customers to call a freephone number to sign up. The letter did not tell customers that they would be part of a trial. They subsequently supplemented this method with telephone calls to households that had received letters but not replied. As with the letters, these calls sought to persuade customers of the benefits of the interventions being offered. Among dual fuel households, 22% of booked smart meter installation appointments were gained through customers phoning in response to the letter; 78% were gained through the company phoning the customer (figures are not available for electricity-only households).

Towards the end of the recruitment process, E.ON also used the reason of meter recertification as a means of persuading customers to agree to smart meter installation. This was not used as a recruitment filter but only as an additional argument to use with households that happened to have old meters.

All E.ON control households were selected without the knowledge of the households and received no trial interventions. Table 3.3 includes the installation rates for E.ON's trial. Table 3.4 shows the progress of the trial.

Table 3.4 Recruitment, installation and trial periods of E.ON's trial

Trial Group		*Recruitment*	*Installation*	*Trial*
Control	All	Apr 2007 – Mar 2008	n/a	Oct 2009 – Jul 2010
Non-smart meter trials	TG1	Apr 2007 – Sept 2007	Apr 2007 – Sept 2007	Oct 2007 – Sep 2009
	TG2			Oct 2007 – Jul 2010
	TG3			Oct 2007 – Sep 2009
	TG4			Oct 2007 – Sep 2009
	TG6			Oct 2007 – Jul 2010
Smart meter trials	All	Feb 2008 – Aug 2008	Feb 2008 – Aug 2008	Aug 2008 – Jul 2010

The monthly energy efficiency tips were deployed in both non-smart and smart trials. However, the intervention was accidentally withdrawn for 12 months between August 2008 and July 2009. This means that the non-smart meter advice trials (TG2 & TG6), which started in October 2007, had an interruption in the intervention. The smart meter advice trials (TG8 & TG9), which started in August 2008, only received the full set of inventions measures a year later than intended.

3.3 Scottish Power

Scottish Power's trial was unique in EDRP in that the interventions were introduced in phases over time.

3.3.1 Interventions

Phase 1 (12 months)

Phase 1 of the trial implemented interventions without smart meters.

- Additional bill data: two mailings of year-on-year consumption history (number and footprint graphic) sent separately to actual bills.
- Energy efficiency advice: seasonal energy advice.
- Real time display: clip-on RTD showing current electricity use, cost, CO_2 emissions and historic data.

Phase 2 (7 months)

The following were implemented in Phase 2.

- Smart meter. The groups provided with smart meters continued to have meter readers visiting and received estimated bills when readings were not available.
- Reconfiguration of prepayment meters to provide higher granularity data by separately recording consumption in five periods of the day.
- Additional bill data: advanced consumption history (more granular for credit customers, meter reading information gathered from customer transactions for prepayment customers) sent separate to actual bills.
- Energy efficiency advice: seasonal energy advice.
- Real time display: clip-on RTD retained by prepayment customers, showing current electricity use, cost, CO_2 emissions and historic data.
- Real time display: mains RTD for credit customers showing current electricity and gas use, cost, CO_2 emissions and historic data, plus a "traffic light" indicator of current consumption.

Phase 3 (15 months)

The main change was the introduction of the "Green Challenge" financial incentives to reduce consumption in addition to the existing interventions in each of the trial groups. The challenge was spread over four "Waves", combining three main elements.

- Incentive to reduce consumption: reward for quarter-on-quarter reduction in electricity consumption (all Waves).
- Time of use: incentive to shift from peak period electricity consumption (in Wave 2).
- Customer engagement: pledge to reduce gas consumption (in Wave 3).

In more detail, the Green Challenge conditions were as follows

- Wave1 (Nov 08 - Jan 09). £10 voucher for using less electricity than quarterly target, £5 for just missing.
- Wave 2 (Feb - May 09). Main challenge was a £10 voucher for using less electricity than predicted. £5 to reduce peak time (early evening) consumption.
- Wave 3 (Aug - Nov 09). Pre-challenge letter to TG1 and TG3 to encourage use of RTD. Letter about customer's average monthly spend in Jan-Jun 09. Main challenge was a £10 voucher for using less electricity than target. Gas Savings Pledge introduced "to make customers take action on their gas use in advance of the winter months approaching". The pledge was to: make regular checks on hot water and heating timer, adjusting according to "seasonality"; switch off boiler if away (only if there was no risk of "adverse temperatures"); and take a Scottish Power energy efficiency survey[10]. Those who made the pledge were entered into a prize draw for a £500 voucher (with five £100 prizes for runners up), to purchase "low energy rated goods". The response rate to the pledge was 20%, which was said to be high for a Scottish Power mailing programme.
- Wave 4 (Nov 09 - Jan 10). Main challenge was a £10 voucher for using less electricity than target, "noting that anyone with two or three challenges worth of sustained reduction in energy use would find the fourth challenge the toughest yet". Customers also sent a free thermometer and information on how to use it to save energy.

Targets were based on customer's individual consumption over past three months. Typical consumption upturn/downturn (based on seasonal variation for Scottish Power customers) was then used to calculate an estimated reading over next three months, with a 5% reduction to generate a 'challenge' figure.

[10] The survey was an online tool, accessible by all Scottish Power customers.

3.3.2 Trial Groups

The different trial groups are detailed below, showing which interventions were included in which phase.

Trial Group 1 (TG1) - Credit households

- Phase 1: Additional bill data; energy efficiency advice; RTD (clip-on).
- Phase 2: Smart meter; additional bill data; energy efficiency advice, RTD (mains).
- Phase 3: As Phase 2 plus Green Challenge programme.

Trial Group 2 (TG2) - Credit households

- Phase 1: Additional bill data; energy efficiency advice.
- Phase 2: Smart meter; additional bill data; energy efficiency advice.
- Phase 3: As Phase 2 plus Green Challenge programme.

Trial Group 3 (TG3) - Credit households

- Phase 1: Not yet recruited into the study.
- Phase 2: As TG1 but recruited part way through the phase.
- Phase 3: As TG1.

Trial Group 4 (TG4) - Prepayment households

- Phase 1: Additional bill data; energy efficiency advice; RTD (clip-on).
- Phase 2: Reconfigured prepayment meters; additional bill data; energy efficiency advice; RTD (clip-on).
- Phase 3: As Phase 2 plus Green Challenge programme

Trial Group 5 (TG5) - Prepayment households

- Phase 1: Additional bill data; energy efficiency advice.
- Phase 2: Reconfigured prepayment meters; additional bill data; energy efficiency advice.
- Phase 3: As Phase 2 plus Green Challenge programme.

Control Group 1 (CG1) - Credit households

This group received business as usual Scottish Power services.

Control Group 2 (CG2) - Prepayment households

This group received business as usual Scottish Power services.

3.3.3 Recruitment

Sampling frame and exclusions

Control and trial groups were sampled from Scottish Power's customer base in the Greater Glasgow and Lanarkshire areas.

For the credit households, the following criteria had to be met.

- Both electricity and gas supplied by Scottish Power, and not paid for by prepayment, for more than 2 years.

- Same customer details for premises for last 2 years.
- Single rate tariff.
- No meter exchange or replacement in past 2 years.
- A history of valid actual meter reads over the last four cyclic visits for both electricity and gas.

For the prepayment households, the following criteria had to be met.

- Customer has prepayment key meter for electricity (gas could be either credit or prepayment meter).
- Both electricity and gas supplied by Scottish Power for more than 2 years.
- Meter reads and dates for each payment transaction from Jan 2006 to present.
- Date of key meter installation no later than Dec 2006.
- Same customer details at premises for last 2 years.
- Single rate tariff.
- For gas credit meters, a history of valid actual cyclic reads.
- Customers paying back debt for electricity, not necessarily for gas.

Selected households were ordered by level of annual consumption and sampling to all groups was balanced by this ranking. There was no demographic stratification but analysis with the Mosaic classification indicated a close match between trial and control groups.

Recruitment methods

Households participating in Scottish Power's trials were treated, as far as possible, as 'business as usual' customers. They did not know that they were participating in a trial and they had no opportunity to accept or refuse postal interventions. However the installation of the clip-on displays required both suitable wiring at the meter and customer consent.

- In Trial Group 1, the RTD (clip-on) was offered to customers during a 'normal' cyclic meter read visit, without pre-warning. Of the dwellings visited, 42% were deemed to be unsuitable for a clip-on display and in 30% customers refused to have a display.
- In Trial Group 4, a letter was sent to prepayment customers offering them a RTD (clip-on). This letter was itself an intervention as it included 'common sense energy advice'. Customers did not have to reply as all were visited by installers and offered the device; 22% of customers refused it.

Smart meter installations in Phase 2 were treated as normal meter replacements. Households in Trial Groups 1 and 3 also received a mains RTD at this point. For households in Trial Group 1, this was presented as part of the package of their upgrade. For households in Trial Group 3, who were joining the trials at this point and had not had the prior experience of the clip-on meter, Scottish Power made a more active offer. They were sent a letter offering them the smart meter (including RTD) and highlighting its benefits. However, as only around 6% of customers responded to this letter, the company then proceeded to telephone customers who had been sent the letter, explaining the benefits of the meter and encouraging them to consider the offer – though the company states that they did not actively persuade customers to sign up.

Despite the 'business as usual' approach, many of the credit households recruited for Phase 1 did not go forward to Phase 2. The principal reasons for this were (a) they could not be contacted, and (b) they refused the upgrade or could not be upgraded. As this depleted the trial groups, which were also depleted from customers leaving the trials due to changes in supplier, Scottish Power 'topped up' these groups with new households. This means that, strictly, Trial Groups 1 and 2 are subdivided into subgroups which did and did not experience Phase 1.

Details of participant numbers in each trial group are shown in Table 3.5. Table 3.6 shows the progress of the trial.

Table 3.5 Scottish Power installation rates

		TG1	*TG2*	*TG3*	*TG4*	*TG5*	*CG1*	*CG2*	*Total*
Phase 1	Start of Phase	395	400	0	408	400	400	400	2,403
Phase 2	Top up of participants	100	118	317	0	0	60	30	
	Start of Phase	288	285	277	291	277	333	316	2,067
Phase 3	Start of Phase	275	261	250	249	250	314	294	1,946

Table 3.6 Recruitment, installation and trial periods of Scottish Power's trial

Trial Group	*Recruitment*	*Installation*	*Trial Phase*
TG1	Apr 2007 – May 2007 Topped up Jul/Aug 2008 for Phase 2	Apr 2007 – May 2007 for Phase 1 Apr 2008 – Aug 2008 for Phase 2	Phase 1: Apr 2007 – Mar 2008 Phase 2: Apr 2008 – Oct 2008 Phase 3: Nov 2008 – Jan 2010 The exception is TG3 which was not included in Phase 1 and began Phase 2 in June 2008.
TG2	Apr 2007 – May 2007 Topped up Jul/Aug 2008 for Phase 2	Apr 2007 – May 2007 for Phase 1 Apr 2008 – Aug 2008 for Phase 2	
TG3	Jun 2008 – Jul 2008	Jun 2008 – Aug 2008	
TG4	Apr 2007 – May 2007	Apr 2007 – May 2007 Reconfigured meters Aug 2008 for Phase 2	
TG5	Apr 2007 – May 2007	Apr 2007 – May 2007 Reconfigured meters Aug 2008 for Phase 2	
CG1	Apr 2007 – May 2007	n/a	
CG2	Apr 2007 – May 2007	n/a	

3.4 SSE

Only the interventions with individual households are covered here. SSE's community trials are described in Appendix D.

3.4.1 Interventions

Non-smart meter interventions

SSE tested the effects of four interventions.

- Additional bill data: graphs on quarterly bills showing historic energy consumption information.
- Energy efficiency advice: an advice booklet sent by post and/or provided online.
- Real time display: clip-on RTD showing current electricity use, cost, CO_2 emissions and historic data.
- Benchmarking: customer's consumption compared with that of other households of comparable demographic group in the region.

Smart meter intervention

SSE tested the effects of the following interventions applied in homes provided with a smart meter.

- Smart meter.
- Monthly bills.
- Additional bill data: graphs on monthly bills showing historic and half-hourly energy consumption information.
- Energy efficiency advice: an advice booklet sent by post and/or provided online.
- Real time display: mains RTD showing current electricity and gas use, cost, CO_2 emissions and historic data, plus (on one of the two models of RTD used) a "traffic light" indicator of current consumption.
- Time of use tariff: incentive to shift from peak period consumption (incentive applied to electricity consumption only).
- Incentive to reduce consumption: reward of 5% of the bill (equivalent to the amount of VAT paid) for a 10% year-on-year reduction in consumption (incentive applied to electricity consumption only).
- Web information: personalised consumption history available online.

The incentive to shift intervention was based on electricity tariffs varying with time of day, season and day of the week (weekday vs weekend). The peak period was 16:00-19:00, night period was 00:30-07:30 and off-peak period was 07:30-16:00 and 19:00-00:30. Low Season was March-October and the off-peak tariff was between 10.29 and 10.88 pence per unit (excluding VAT), varying with region. The peak tariff was 180-190% of the off-peak tariff (for both weekdays and weekends) and the night tariff was 50-60%. High Season was November-February and the off-peak tariff was between 10.87 and 11.46 pence per unit. The peak tariff was 180% of the off-peak tariff at weekends and 210% on weekdays; the night tariff was 50-60%.

3.4.2 Trial Groups

Compared to the other trials, SSE's assessed more combinations of interventions, which are shown in Tables 3.7 and 3.8. The groups are listed in an order to emphasise the partly factorial design, rather than the SSE trial number order. In addition, Trial Group 32 was a control group receiving business as usual services.

Beyond these groups, customers were stratified according to whether they were Aware, Unaware or Committed to the trials (including the control group). In some trial conditions, only the Aware and Committed groups were used.

- Committed: customers were aware of the trial and, in a signed statement to SSE, have said that they were committed to reducing energy.
- Aware: customers were aware of the trial.
- Unaware: customers were unaware that they were participating in a trial.

All the trial groups referred to above pay by credit. In addition, Trial Group 31 (not shown in the tables) had a prepayment smart meter and wirelessly linked RTD powered from the mains. Trial Group 33 was the associated control group with standard prepayment meters and received business as usual services. These two groups were not stratified by their level of awareness.

Table 3.7 SSE non-smart trial groups

Trial Group	*Booklet*	*Clip-on RTD*	*Quarterly bills with graphs*	*Benchmarking*
4	✓			
3		✓		
1	✓	✓		
5	✓		✓	
2	✓	✓	✓	
6	✓		✓	✓

Table 3.8 SSE credit smart trial groups (all included smart meters and accurate billing)

Trial Group[1]	*Booklet*	*Monthly bills with graphs*	*Incentive to shift*	*Incentive to reduce*	*RTD*	*Web information*
15						
16		✓	✓			
17		✓		✓		
18		✓	✓	✓		
27					✓	
28			✓		✓	
29				✓	✓	
30			✓	✓	✓	
11	✓	✓				
12	✓	✓	✓			
13	✓	✓		✓		
14	✓	✓	✓	✓		
23	✓	✓			✓	
24	✓	✓	✓		✓	
25	✓	✓		✓	✓	
26	✓	✓	✓	✓	✓	
10	✓	✓				✓
7	✓	✓	✓			✓
8	✓	✓		✓		✓
9	✓	✓	✓	✓		✓
22	✓	✓			✓	✓
19	✓	✓	✓		✓	✓
20	✓	✓		✓	✓	✓
21	✓	✓	✓	✓	✓	✓

[1] The groups are listed in an order to emphasise the partly factorial design, rather than the SSE trial number order. Trial Group 15 only had smart meters and none of the intervention measures itemised in the table.

Capabilities on project:
Building Engineering

3.4.3 Recruitment

Sampling frame and exclusions

SSE used its national customer base as its sampling frame. An initial 1.4 million customers were selected as potential candidates on the basis of (a) having single rate electricity tariffs and (b) having good consumption records (4 actual reads over the past 2 years).

Prospective candidate lists for each trial were prepared from this customer base using a dual stratification of Mosaic group (a classification of households into different demographic groups) and fuel mix (electricity only or dual fuel). The selection according to Mosaic group was designed to ensure that all trial and control groups had similar demographic profiles, as far as possible. However, as each Mosaic class was given equal weight in this profiling, the groups are not directly representative of the GB population, which has an uneven Mosaic profile.

Payment type was also used as a selection variable, not for stratification but to distinguish credit and prepayment trial groups. The prepayment trials were not subject to this Mosaic stratification because the demographic profile of prepayment households is known to be highly skewed.

The actual selection of households for each of the trial groups was iterative with some trial groups selected before others. Technology-dependent projects (those with smart meters) were recruited in areas close to selected depots where installers had been trained.

Recruitment methods

SSE used different recruitment methods for each of its subgroups of Unaware, Aware and Committed customers. Initially the approach was as follows.

- Unaware customers were sent non-smart interventions, including the clip-on display, without prior warning or consent. The installation of smart meters took the same form as a standard meter change and supplementary interventions followed, again without prior notice.
- Aware customers were actively recruited to a trial but not told about the specific technology they would get. Customers were sent a letter with a recruitment questionnaire.
- Committed customers were sent a letter and questionnaire, plus a postcard, pledging to try and save energy, which they could sign and return.

The initial recruitment rates for the Aware and Committed groups were between 3% and 10%. A sample of customers who had received letters was telephoned to identify their reasons for not responding. The recruitment survey was identified as a key reason for this refusal rate with the length and complexity of the questionnaire and the income question deterring people. “A very large number” of customers also said they were simply not interested. Other reasons for refusal included being elderly, multi-occupancy households, little energy usage, moving house or supplier, and changing meter type.

Various changes were made to the recruitment methods to address these problems.

- Withdrawing the recruitment questionnaire.
- Changing the correspondence, including the addition of a two-week deadline in a revised letter for prospective Committed trialists.
- Systematically following up letters with telephone calls.
- Latterly, accepting a statement of commitment from customers by telephone rather than a written commitment.

Some of the households on the Aware prospective lists were telephoned and asked if they would commit to saving energy, boosting the Committed group recruitment. But most of the Aware households were not asked about their level of commitment. Households on the Committed prospective lists were switched to the Aware groups if they chose not to make the commitment.

Table 3.9 summarises the number of installations for SSE's trial. Table 3.10 illustrates the timetable for this trial.

Table 3.9 The number of households with different meter types

Trials	*Installed*
Non-smart meters	15,239
Credit smart meters	6,670
Prepayment smart meters	436
Controls	5,542
Total	27,887

Table 3.10 Recruitment, installation and trial periods of SSE's trial

<table>
<tr><th colspan="2">Trial Group Number</th><th>Recruitment</th><th>Installation</th><th>Trial</th></tr>
<tr><td rowspan="2">Non-smart meter trials</td><td>1,2,3</td><td>Apr 2007 – Oct 2007</td><td>Apr 2007 – Oct 2007</td><td>Nov 2007 – Sep 2010</td></tr>
<tr><td>4,5,6</td><td>Apr 2007 – Oct 2007</td><td>Apr 2007 – Oct 2007</td><td>Dec 2007 – Sep 2010</td></tr>
<tr><td>Smart meter trials</td><td>17-31</td><td>Apr 2007 – Sep 2008</td><td>Apr 2007 – Sep 2008</td><td>Nov 2008 – Sep 2010[1]</td></tr>
<tr><td>Controls</td><td>32,33</td><td>Apr 2007 – Oct 2007</td><td>Apr 2007 – Oct 2007</td><td>Nov 2007 – Sep 2010</td></tr>
</table>

[1] For trials with either incentive to reduce or incentive to shift, Sep 2009 – Sep 2010.

4 Analysis of EDF data

This section presents AECOM's analysis of both the energy consumption data and the consumer survey data from EDF. The summary at the end brings together findings from EDF's and AECOM's analysis. Appendix B1 summarises the issues that need to be considered, related to the research design and execution, when interpreting or applying the findings.

4.1 Energy Consumption Data

4.1.1 Introduction

The original analysis by EDF did not use any pretrial energy consumption data as a baseline to assess the intervention measures: this was the single greatest concern for the findings reported by EDF. We are grateful to EDF for supplying the baseline data for electricity consumption. This allowed us to undertake new analysis to assess the change between pretrial and in-trial energy consumption. As pretrial data were only available for electricity, no further analysis of gas consumption was undertaken.

The approach taken was to assess change from a whole year of pretrial readings to a whole year of in-trial readings. We would have preferred to analyse the change in electricity consumption between pretrial and in-trial periods at different times of the year as the impact of the interventions may vary according to the time of year and time from start of intervention. However, the pretrial data are based on manually read non-smart meter data. As noted in Section 3.1, the criteria for recruiting homes only required them to have two readings in the 12 month period up to the start of the trial. Given the variation in electricity consumption between seasons, it was considered too inaccurate to use this information to interpolate monthly or even quarterly pretrial electricity consumption levels.

On behalf of Ofgem, the Centre for Sustainable Energy (CSE) has developed a database to collate the energy suppliers' data as well as carrying out validation checks. CSE provided AECOM with appropriately aggregated and formatted data for use in its analysis. The new EDF pretrial data were sent to CSE to be prepared in a similar way to the other data.

The electricity data were provided in three formats:

- raw energy consumption data;
- weather-adjusted energy consumption data: takes into account variations in weather conditions for the same day between different years;
- weather-adjusted and time-corrected energy consumption data: takes into account of weather variations both between years and for different days within a given year).

The weather-adjusted data were considered more useful for the analysis than the raw data. The analyses compared electricity consumption across different years (i.e. between pretrial and in-trial time periods). This correction adjusted for the different weather conditions between years which may impact on electricity consumption. The time-correction factor was not used as the analysis was undertaken on annual electricity consumption. Previous work by Ofgem, CSE and the energy suppliers had investigated the robustness of the correction factors, which was therefore not investigated further in this work.

Two types of smart meter electricity consumption data were provided:

- half-hourly consumption data;
- actual remote meter readings (typically weekly).

The latter were used for most of the analysis. This was appropriate as the data were aggregated over a year-long period for the analysis – hence there was no value in using half-hourly data. Furthermore, where half-hourly data are missing, an estimate is needed for the consumption during the missing period. At the start of the trial, there were several quarter periods where high numbers of households had percentages of missing data which fell outside an acceptable threshold (greater than 10% of electricity consumption data were missing).

The electricity consumption data provided to AECOM were aggregated into quarterly periods, and these needed to be further aggregated into yearly periods for each household.

- The pretrial year comprised the four quarters before the quarter in which the meter was installed.
- The first in-trial year comprised the four quarters after the quarter in which the meter was installed.
- The quarter in which the meter was installed was omitted from the analysis as electricity usage in that quarter would include usage both before and after the trial conditions began.

Initial examination of the distributions of the weather-corrected electricity consumption data in the pretrial year and in-trial years showed that the data did not follow a normal distribution (i.e. a bell shaped curve). Similarly, the changes in electricity consumption from pretrial to in-trial years were not normally distributed (nor were log transformations of the data). In all cases, the distributions were discovered to be leptokurtic[11] and influenced by extreme deviations from the mean. A multivariate modelling strategy had been the planned approach to the analysis, in which it would have been possible to assess separately the impact of different interventions (and other variables, as identified in Chapter 3) on the change in electricity consumption. Although some departure from normality is acceptable if classical modelling techniques are to be used, the data here were substantially different from that which would usually be acceptable. Modelling approaches that avoid having to assume normality would require the investment of considerable additional time, which was not available.

As a result, the analysis focused on using ‘non-parametric’ methods, which make no assumption about the data distribution and are therefore suitable for robust analysis of the data in this trial. In keeping with this, medians are reported, rather than the usual arithmetic means, to indicate the average response to the interventions.

For analysis of consumption at different times of day (to examine load-shifting effects), half-hourly consumption data were aggregated to each tariff period.

4.1.2 Analysis methods

Changes in Electricity Consumption

There is a danger that households with unusually high or low consumption levels and changes in consumption will bias the calculation of summary statistics. Rather than use arbitrary methods to exclude households from the analyses, we overcame this by using the non-parametric Mann-Whitney U test to test differences between groups, and reporting median consumption and changes in consumption. The use of non-parametric techniques is advantageous for reasons beyond not having to assume that the data are normally distributed: such techniques are less sensitive to outliers than parametric techniques and, although care has been taken to avoid errors in the data preparation, it is impossible to be definitive that there are no figures that are larger or smaller than they should be. The robust analysis reported here reduces the impact of such problems on the results obtained.

[11] Leptokurtic distributions are pointed and narrow in the centre, rather than the rounded ‘normal’ distribution.

A series of analyses was undertaken comparing pairs of trial groups as in Table 4.1. The subscripts ‘EO’ and ‘DF’ refer to EDF customers who were ‘electricity-only’ and ‘dual fuel’ respectively. The trial groups are described in Section 3.1 (Table 3.1).

- For most of the analyses, trial groups are compared with their control groups.
- Analyses 11-13 instead assess differences between selected trial groups, in particular to compare the Wall Panel group with URA, TOUT and HEC respectively (each of which also had an RTD but with a different additional intervention in each case).
- Analyses 14-16 merged groups that had the same intervention and differed only in whether EDF supplied them with electricity only or both electricity and gas (dual fuel groups).

EDF’s own analysis did not consider the impact of whether the household was supplied with electricity only or gas and electricity. Furthermore, it solely compared trial groups with their control groups.

For each pair of trial groups listed in Table 4.1, the Mann-Whitney U test was used to compare groups on:

- pretrial electricity consumption;
- the change in electricity consumption from pretrial to the first in-trial year;
- the change in electricity consumption from pretrial to the second in-trial year.

Table 4.1 Trial groups compared in analysis of EDF electricity consumption

Analysis	Groups to be compared	
	(i)	(ii)
1	Paper_{EO}	Control_{EO}
2	Paper_{DF}	Control_{DF}
3	Wall Panel_{EO}	Control_{EO}
4	Wall Panel_{DF}	Control_{DF}
5	TV_{EO}	Control_{EO}
6	TV_{DF}	Control_{DF}
7	URA	Control_{EO}
8	TOUT	Control_{EO}
9	HEC	Control_{DF}
10	Web	Control_{DF}
11	URA	Wall Panel_{EO}
12	TOUT	Wall Panel_{EO}
13	HEC	Wall Panel_{DF}
14	Paper_{EO+DF}	Control_{EO+DF}
15	$\text{Wall Panel}_{EO+DF}$	Control_{EO+DF}
16	TV_{EO+DF}	Control_{EO+DF}

Changes in time of use of electricity

The main reason for the time of use tariff (TOUT) intervention was to see if having a price tariff that varied depending on time of day caused the proportion of electricity used in the peak period to be reduced. In order to examine whether this took place, trial group TOUT was compared with the trial group having just an RTD (Wall $Panel_{EO}$), and with the control group ($Control_{EO}$).

Only in-trial data could be used for this analysis because it requires data from smart meters, broken down by time of day. For each household, electricity consumption data were aggregated over successive 28 day periods for each of the four tariff periods. Furthermore, the data were separated by weekdays and weekends to investigate differences in behaviour between the two periods.

This analysis was based on using half-hourly electricity consumption data. As discussed earlier, there can be missing data (e.g. due to faults in data transmission between the smart meter and supplier) and estimates need to be made for the electricity consumption during that 28-day period. Where substantial amounts of data are missing, it is not advisable to make adjustments because of the high uncertainty surrounding the estimate. For the vast majority of households, there were no missing data. It was ascertained that making adjustments for situations when some data were missing (but limiting it to cases where less than 10% of data were missing to reduce uncertainties) would make only a marginal difference to the number of households available for analysis and therefore these households were not used.

For each 28-day period, the proportion of the electricity consumption that had been used during the peak period was calculated separately for weekdays and weekend. The distribution of the logarithm of these proportions was sufficiently normally distributed for a modelling exercise to be undertaken. As these data form a hierarchy of each household having a set of 28 day periods, it was considered appropriate to undertake multi-level modelling.

As well as including the three trial groups in the modelling process, a set of potentially confounding variables (i.e. variables that might affect the consumption level or pattern) was also included:

- electricity distribution area (London or rest of south-east England);
- ACORN category (a demographic measure);
- type of heating (electric fires/fan heaters; electric storage heating; gas central heating; gas fires/room heaters; paraffin/oil heaters; no heating);
- whether or not the household had a programmer to control the heating;
- number of people in the household (in three age categories: under 16, 16-64, and 65 and over).[12]

These variables were selected because they were available from information known by EDF about the household location and from the brief recruitment survey undertaken by EDF.

In order to be recruited for TOUT, a household had to have a ‘medium’ level of pretrial electricity consumption. However, households in Wall $Panel_{EO}$ and $Control_{EO}$ did not have such a restriction and information was not available to duplicate the filtering process used for TOUT. To assess whether it was inappropriate to compare groups, an analysis of the consumption patterns in $Control_{EO}$ was undertaken. Splitting the households in $Control_{EO}$ into equal thirds according to the pretrial consumption values that had been obtained, the proportion of electricity consumption in the peak tariff period was examined (using multi-level modelling) in order to see if it was affected by pretrial consumption level. For both the weekday and weekend peaks, the pretrial consumption level did not affect

[12] The age bands have been cited differently (by a year on either side) in different sources but these are the age bands in the EDF survey questionnaire itself.

the proportion of consumption falling in the peak period. The decision was therefore made that it was appropriate to proceed with the analysis without regard to the pretrial consumption limitation.

4.1.3 Results

Annual consumption

Summary statistics are presented in Appendix E1.[13]

- Median electricity consumption values for each trial group for the pretrial year, in-trial year 1 and in-trial year 2 are given in Table E1.1.
- The differences in electricity consumption between pretrial and in-trial year 1 and between pretrial and in-trial year 2 are given in Table E1.2.

The trial groups to be compared were first assessed using Mann-Whitney U tests to see if they differed in their electricity consumption during the pretrial year (see Table E1.3 for full details). None of the pairs of groups had significantly different consumption in the pretrial year. This means there is unlikely to be bias in the analysis of changes in consumption due to the pairs of groups starting from differing baseline conditions. Nevertheless, it is appropriate to take into account pretrial consumption in the analysis of effects of interventions as non-significant pretrial differences are still relevant in the analyses.

The trial group pairs (as shown in Table 4.1) were then compared to determine whether they differed in the change in consumption from the pretrial year to in-trial year 1. The results of the Mann-Whitney U tests are given in Table E1.4 and summarised in Table 4.2. The p-value represents the confidence in the finding (a value of less that 0.05 is generally accepted as an indication that we should be confident in the finding). It should be noted that the number of households with valid data differs from trial group to trial group and that, in particular, trial groups TV_{EO} and TOUT do not have sufficient data for a meaningful analysis to be done.

A similar analysis was also carried out to compare differences in energy consumption between pretrial and in-trial year 2. The results of the Mann-Whitney U tests are given in the Table E1.5 and summarised in Table 4.3. There are now more trial groups with insufficient households for a meaningful analysis to take place.

[13] The number of households with valid data differs between the pretrial year and in-trial years. Our comparison between pretrial and in-trial consumption used only those households with valid data in both comparison years. For this reason, the differences determined in the analysis should not be compared with the differences between medians in the Appendix E tables, which are based on all households with data in the particular year only. There are several reasons for differences in the number of households with valid data between years. For example, sufficient pretrial data were not available for all households in the trial groups and some households were recruited too late into the trial to have a complete second year of in-trial data. Households that left the study for whatever reason during the trial have been omitted from the data.

Table 4.2 Comparisons of changes in electricity consumption from pretrial to in-trial year 1

Group (i)	*Median % change in consumption*	*Group (ii)*	*Median % change in consumption*	*Difference in % change in consumption Group (i) - Group (ii)*	*p-value from Mann-Whitney*
$Paper_{EO}$	-3.0%	$Control_{EO}$	-2.1%	-0.9%	0.325
$Paper_{DF}$	-6.2%	$Control_{DF}$	-1.6%	-4.6%	0.025*
Wall $Panel_{EO}$	-9.3%	$Control_{EO}$	-2.1%	-7.2%	0.007**
Wall $Panel_{DF}$	-2.6%	$Control_{DF}$	-1.6%	-1.0%	0.980
TV_{EO}	-5.6%	$Control_{EO}$	-2.1%	-3.5%	0.383
TV_{DF}	-5.6%	$Control_{DF}$	-1.6%	-4.0%	0.618
URA	-9.4%	$Control_{EO}$	-2.1%	-7.3%	0.011*
TOUT	-14.9%	$Control_{EO}$	-2.1%	-12.8%	0.221
HEC	-3.3%	$Control_{DF}$	-1.6%	-1.7%	0.339
Web	-2.6%	$Control_{DF}$	-1.6%	-1.0%	0.322
URA	-9.4%	Wall $Panel_{EO}$	-9.3%	-0.1%	0.757
TOUT	-14.9%	Wall $Panel_{EO}$	-9.3%	-5.6%	0.693
HEC	-3.3%	Wall $Panel_{DF}$	-2.6%	-0.7%	0.464
$Paper_{EO+DF}$	-4.1%	$Control_{EO+DF}$	-1.8%	-2.30%	0.020*
Wall $Panel_{EO+DF}$	-5.8%	$Control_{EO+DF}$	-1.8%	-4.0%	0.047*
TV_{EO+DF}	-5.6%	$Control_{EO+DF}$	-1.8%	-3.8%	0.250

* p<0.05; ** p<0.01 Grey text means sample numbers were small.

Table 4.3 Comparisons of changes in electricity consumption from pretrial to in-trial year 2

Group (i)	*Median % change in consumption*	*Group (ii)*	*Median % change in consumption*	*Difference in % change in consumption Group (i) - Group (ii)*	*p-value from Mann-Whitney*
$Paper_{EO}$	-8.9%	$Control_{EO}$	-3.9%	-5.0%	0.045*
$Paper_{DF}$	-10.1%	$Control_{DF}$	-7.7%	-2.4%	0.029*
Wall $Panel_{EO}$	-14.9%	$Control_{EO}$	-3.9%	-11.0%	0.004**
Wall $Panel_{DF}$	-9.1%	$Control_{DF}$	-7.7%	-1.4%	0.871
TV_{EO}	n/a	$Control_{EO}$	-3.9%	n/a	n/a
TV_{DF}	n/a	$Control_{DF}$	-7.7%	n/a	n/a
URA	-15.2%	$Control_{EO}$	-3.9%	-11.3%	0.427
TOUT	-9.7%	$Control_{EO}$	-3.9%	-5.8%	0.886
HEC	n/a	$Control_{DF}$	-7.7%	n/a	n/a
Web	n/a	$Control_{DF}$	-7.7%	n/a	n/a
URA	-15.2%	Wall $Panel_{EO}$	-14.9%	-0.3%	0.531
TOUT	-9.7%	Wall $Panel_{EO}$	-14.9%	5.2%	0.210
HEC	n/a	Wall $Panel_{DF}$	-9.1%	n/a	n/a
$Paper_{EO+DF}$	-9.5%	$Control_{EO+DF}$	-5.5%	-4.0%	0.004**
Wall $Panel_{EO+DF}$	-11.3%	$Control_{EO+DF}$	-5.5%	-5.8%	0.061~
TV_{EO+DF}	n/a	$Control_{EO+DF}$	-5.5%	n/a	n/a

~ p<0.1; * p<0.05; ** p<0.01 Grey text means sample numbers were zero or small.

There are several group comparisons where, at the 5% level of significance ($p<0.05$), we have sufficient evidence to make the following claims.

- $Paper_{EO}$ (electricity-only customers with additional bill data and energy efficiency advice by post in addition to a smart meter) show greater reductions in consumption than $Control_{EO}$ (electricity-only customers with just a smart meter). This finding is significant only for in-trial year 2. It must be borne in mind that the $Paper_{EO}$ households involved in the pretrial to in-trial year 2 analysis are those that were recruited sufficiently early for them to have a complete set of data for in-trial year 2. It is possible that these households were recruited easily because they are inherently more committed to reducing electricity consumption than is typical for all households in $Paper_{EO}$ and this is why a significant effect is shown. It is also possible, given that the p-value for this comparison is 0.045, quite near the 0.05 cut-off for claiming significance at this level, that the significant finding is an anomaly. One would expect 1 in 20 (5%) of tests to give a significant result ($p<0.05$) even if there were no differences between any of the groups and it is important to bear this in mind when interpreting the results of a substantial number of analyses.
- $Paper_{DF}$ (dual fuel customers with additional bill data and energy efficiency advice by post in addition to a smart meter) show greater reductions in consumption than $Control_{DF}$ (dual fuel customers with just a smart meter). This finding is persistent across both trial years.
- $Paper_{EO+DF}$ (electricity-only and dual fuel customers with additional bill data and energy efficiency advice by post in addition to a smart meter) show greater reductions in consumption than $Control_{EO+DF}$ (electricity-only and dual fuel customers with just a smart meter). This finding is persistent across both trial years.
- Wall $Panel_{EO}$ (electricity-only customers with an RTD and energy efficiency advice by RTD in addition to a smart meter) show greater reductions in consumption than $Control_{EO}$ (electricity-only customers with just a smart meter). This finding is persistent across both trial years.
- Wall $Panel_{EO+DF}$ (electricity-only and dual fuel customers with an RTD and energy efficiency advice by RTD in addition to a smart meter) show greater reductions in consumption than $Control_{EO+DF}$ (electricity-only and dual fuel customers with just a smart meter). This finding is persistent across both trial years, although only at a 10% level of significance for in-trial year 2.
- URA (electricity-only customers with a real time display RTD, energy efficiency advice by RTD and usage reduction alert in addition to a smart meter) show greater reductions in consumption than $Control_{EO}$ (electricity-only customers with just a smart meter). This is only significant for in-trial year 1. This could be due to the very small number of households in URA for in-trial year 2 and thus insufficient data to identify a difference rather than a difference not existing (in fact, the difference in medians was greater in year 2).

EDF found the following significant differences in their analysis compared the control group (see Appendix A1):

- significantly lower in Paper (by 0.7%), Wall Panel (by 1.1%) and TOUT (by 1.5%);
- significantly higher in TV (by 8.4%), HEC (by 13.2%) and Web (by 14.1%);
- not significantly different in URA.

It seems that the differences not seen in the comparison with pretrial consumption are attributable to differences that existed before the interventions were introduced. The exception may be TOUT, where Table 4.2 shows a large difference in consumption that is not significant because of small samples size.

Time of Use Analysis

Table 4.4 shows the proportion of consumption that occurred in the weekday and weekend peak for each of the three trial groups. This suggests that, although there may be a difference in the proportion for the weekend, this may not exist for weekdays but this was subject to statistical testing, as described below. The differences reported in this section are significant at the 5% level of significance ($p<0.05$) unless stated otherwise. These differences are "predicted" by the model, using all levels of each independent variable rather than being based on the mean for each level. Hence, for example, the effect of number of people in the household is presented as though each extra person has the same effect.

Table 4.4 Summary statistics for proportion of consumption in peak periods

		TOUT	*Wall Panel$_{EO}$*	*Control$_{EO}$*
Proportion of consumption in weekday peak period	Households with valid data	170	141	135
	Mean	20%	20%	20%
	Minimum	12%	13%	13%
	Maximum	33%	34%	42%
Proportion of consumption in weekend peak period	Households with valid data	170	141	135
	Mean	18%	19%	19%
	Minimum	12%	11%	12%
	Maximum	32%	27%	31%

Weekdays

In the multilevel modelling analysis of the proportion of weekday electricity consumption in the peak period, the following effects were observed.

The finding of greatest relevance here is that it appears to become more difficult to manage electricity use out of the peak period as the number of people in the household increases. Differences between trial groups TOUT and Control$_{EO}$ depend on the number of people in the household aged 16 to 64. The results are presented in Figure 4.1, which shows the relative proportion of consumption in the peak period between TOUT and Control$_{EO}$. For example, a figure of 90% means that the proportion of consumption in the peak period in TOUT is only 90% of that in Control$_{EO}$. This percentage increases by 3.9% per additional person. The analysis indicates a benefit from using the time of use tariff when there are two or fewer people in the household in this age-range but a disbenefit when there are three or more householders in this age group. While it is unsurprising that the benefit depends on the number of people in this household, this could be key information for targeting time of use tariffs most effectively.

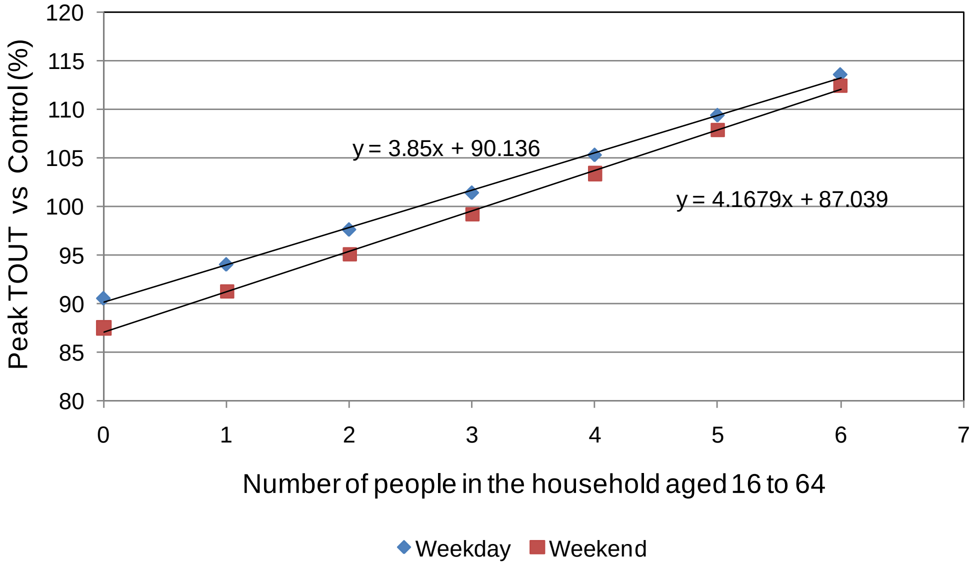
There is no overall significant difference between trial groups at the 5% level but the proportion of consumption in the peak period in TOUT is 95.6% of that in Wall Panel$_{EO}$ (effect significant at the 10% level).

In addition to the effects of trial group, significant effects were observed for other variables.

- For each additional person aged 16 to 64 in Wall Panel$_{EO}$ and TOUT, the proportion of consumption in the peak period increases by 4.3%. The interaction effect between the number of people aged 16 to 64 and Control$_{EO}$ is reported above.
- For each additional person aged under 16 in the household, the proportion of consumption in the peak period increases by 3.7%.
- Households in south-east England have a proportion of consumption in the peak period 5.4% higher than those in London.
- The type of heating used makes a difference to the proportion of consumption in the peak period. Those that have paraffin/oil heaters or no recorded heating have a larger proportion of their consumption in the peak period than households heated by gas or electricity. The effect for no recorded heating is significant at the 10% level of significance.

There was no significant effect of the number of people aged 65 and over, ACORN code or whether the household has a heating programmer.

Figure 4.1 Interaction of TOUT and $Control_{EO}$ with the number of people in the household aged 16 to 64[14]



Weekends

In the multilevel modelling analysis of the proportion of weekend electricity consumption in the peak period, the following effects were observed.

Again the finding of greatest relevance is that it appears to become more difficult to manage electricity use out of the peak period as the number of people in the household aged 16 to 64 increases. The results are presented in Figure 4.1, which shows the relative proportion of consumption in the peak period between TOUT and $Control_{EO}$. This percentage increases at 4.2% per additional person. The analysis indicates a benefit from using the time of use tariff when there are two or fewer people in the household in this age-range but a disbenefit when there are more than three householders in this age group. With three householders, the effect is neutral or very modestly positive.

There is no overall significant difference between trial groups Wall $Panel_{EO}$ and $Control_{EO}$ but TOUT has a proportion of consumption in the peak period 91.9% of that in Wall $Panel_{EO}$, indicating a stronger overall impact of TOUT at weekends than on weekdays ($p<0.05$).

In addition to the effects of trial group, significant effects were observed for other variables.

- For each additional person aged under 16 in the household, the proportion of consumption in the peak period increases by 2.1%.
- For each additional person aged 16 to 64 in trial groups Wall $Panel_{EO}$ and TOUT, the proportion of consumption in the peak period increases by 4.8%. The interaction effect between the number of people aged 16 to 64 and $Control_{EO}$ is reported above.

[14] There are fewer large households so the apparent negative impact of the tariff on households of more than three people may be an artefact of modelling based mainly on small households.

- There is an interaction between location (south-east England or London) and the type of heating used.
  - Households in south-east England with electric fires have a proportion of consumption in the peak period 79.6% higher than those in London.
  - Households in south-east England with gas fires/room heaters have a proportion of consumption in the peak period 20.3% less than those in London.
  - For other types of household heating, households in south-east England have a proportion of consumption in the peak period 6.4% higher than those in London.
  - Those that have paraffin/oil heaters or no recorded heating have a larger proportion of their consumption in the peak period than those households heated by gas or electricity.
  - Those households with electric storage heaters have a smaller proportion of their consumption in the peak period than those heated by gas.

There was no significant effect of the number of people aged 65 and over, ACORN code or whether the household has a heating programmer.

4.2 Survey Data

4.2.1 Rationale for analyses carried out

In the final consumer survey, EDF collected data from the majority of participants. The scope of the survey was restricted as a result of the priority of achieving a high response rate. Here we report on the responses obtained to a selection of the questions asked. The reasons behind the choice of questions to analyse are listed below.

- Questions were chosen to be analysed where further investigations could usefully add to those analyses already undertaken by EDF.
- Focus was put on understanding change in consumer behaviour and/or energy use.
- In order for the conclusions drawn from the analyses to be robust, the questions chosen should not be unduly open to interpretation by respondents (e.g. because of their wording or because the response choices given for the questions were ambiguous).
- It was important that questions chosen for analysis did not suffer from potential distortion by their position in the survey (e.g. because of questionnaire routing). Thus questions were only included where answers to earlier questions in the survey did not have the potential to distort responses.
- In order for effective analyses to be undertaken, sufficient numbers of respondents had to exist. This often meant merging electricity-only and dual fuel customers.

For the selected questions, the tables below show the responses given according to the trial group to which the respondent belonged. Trial groups are omitted from the analysis where the question did not apply to them (e.g. if the question was about RTDs and the group was not given RTDs).

Chi-square tests were used to test whether or not the responses to the questions are independent of the trial group to which the respondents belonged. In some circumstances, response categories have been combined in order to have sufficient number of respondents in the cells of the table for the assumptions behind the chi-square test to be valid. Missing data (including responses such as “don’t know” and “refuse to answer”) have been omitted except where they have a particular relevant interpretation.

In some cases, scale responses have been treated as interval scales for the purpose of further exploration of significant effects. The scale responses are generally not normally distributed so the analysis of variance (ANOVA) findings should be treated as indicative only, and a p-value is therefore not stated.

4.2.2 Results

Tabulations of the questions selected for analysis against trial group of respondent are shown below. The p-values that result from the chi-square tests are shown beneath the tables. Further comments are made where the p-value is less than 0.05 (indicating that there is sufficient evidence to reject the claim that the responses are independent of trial group at the 5% level of significance).

Table 4.5 When was the loft insulation installed or last topped up?

	Trial Group						
	$Paper_{EO+DF}$	*Wall $Panel_{EO+DF}$*	*URA*	*TOUT*	*HEC*	*$Control_{EO+DF}$*	*Total*
In the last two years	26.7%	28.9%	20.2%	25.3%	18.5%	17.2%	23.0%
Longer ago	73.3%	71.1%	79.8%	74.7%	81.5%	82.8%	77.0%
Total	100% (n=90)	100% (n=166)	100% (n=89)	100% (n=95)	100% (n=92)	100% (n=151)	100% (n=683)

p = 0.129

There appears to be some correspondence between a higher percentage of insulation during the past two years (i.e. the trial period) and two trial groups in which significant reductions in electricity consumption were seen (i.e. Paper and Wall Panel) and also TOUT, for which the impact on consumption could not be tested. However, this was not statistically significant. EDF presented findings on other actions (see Appendix A1.1) – cavity wall insulation, double glazing, draughtproofing and fitting a new boiler. These variables were not available in our database.

Table 4.6 Over the past two years, has the number of electrical appliances in use in your home decreased, increased or stayed the same?

	Trial Group						
	$Paper_{EO+DF}$	*Wall $Panel_{EO+DF}$*	*URA*	*TOUT*	*HEC*	*$Control_{EO+DF}$*	*Total*
Increased	15.5%	14.4%	21.9%	15.7%	22.7%	20.6%	18.2%
Stayed the same	78.2%	83.0%	71.9%	79.6%	75.3%	76.2%	77.8%
Decreased	6.4%	2.7%	6.1%	4.6%	2.1%	3.2%	4.0%
Total	100% (n=110)	100% (n=188)	100% (n=114)	100% (n=108)	100% (n=97)	100% (n=189)	100% (n=806)

Because of small numbers in the “Decreased” category, it was necessary to merge with the “Stayed the same” category to conduct a valid test. p = 0.309

Again there is a non-significant trend, for increase in number of appliances to be less likely in trial groups Paper, Wall Panel and TOUT.

Table 4.7 Do you have a visual display in your home, that tells you how much electricity or gas you are using?

	Trial Group					
	Wall Panel$_{EO}$	*Wall Panel$_{DF}$*	*URA*	*TOUT*	*HEC*	*Total*
No	45.8%	35.2%	50.9%	62.4%	36.1%	46.5%
Yes	54.2%	64.8%	49.1%	37.6%	63.9%	53.5%
Total	100% (n=83)	100% (n=105)	100% (n=114)	100% (n=109)	100% (n=97)	100% (n=508)

$p < 0.001$

The percentage being aware of the RTD was highest among dual fuel customers (Wall Panel$_{DF}$ and HEC – 65% and 64% respectively) and lowest where the most basic device was provided (TOUT – 38%). These extremes most likely account for the significant overall effect and the other two groups had an intermediate and similar percentage (Wall Panel$_{EO}$ and URA – 54% and 49% respectively).

Table 4.8 How useful do you find the visual display?

	Trial Group					
	Wall Panel$_{EO}$	*Wall Panel$_{DF}$*	*URA*	*TOUT*	*HEC*	*Total*
Not at all	7.1%	4.6%	6.0%	12.5%	0.0%	4.9%
Not very useful	23.8%	9.2%	12.0%	25.0%	4.8%	12.8%
Quite useful	38.1%	38.5%	38.0%	54.2%	30.6%	37.9%
Very useful	31.0%	47.7%	44.0%	8.3%	64.5%	44.4%
Total	100% (n=42)	100% (n=65)	100% (n=50)	100% (n=24)	100% (n=62)	100% (n=243)

Because of small numbers in the “Not at all” category, it was necessary to merge with the “Not very useful” category to conduct a valid test. $p < 0.001$

To explore this significant effect further, the response scale was numbered 1-4 and treated as an interval scale. The mean value of this scale rating was highest for HEC (3.6), lowest for TOUT (2.6) and intermediate for Wall Panel$_{EO}$ (2.9), Wall Panel$_{DF}$ (3.3) and URA (3.2). A one-way ANOVA supports the chi-square test in showing an overall effect and indicates that it is due mainly to the following differences:

HEC > Wall Panel$_{EO}$ and TOUT

TOUT < Wall Panel$_{DF}$ and URA and HEC.

In trial group HEC, the RTD included a controller for heating and hot water, so it may be this function that resulted in its higher ratings. TOUT had the most basic RTD and this seems to have been reflected in less positive ratings.

Table 4.9 To what extent do you agree or disagree that your smart meter technology has enabled you to plan or budget for your energy use in the home?

	Trial Group							*Total*
	Paper$_{EO+DF}$	*Wall Panel$_{EO+DF}$*	*TV$_{EO+DF}$*	*URA*	*TOUT*	*HEC*	*Control$_{EO+DF}$*	
Strongly agree	5.4%	26.8%	4.3%	18.3%	11.7%	28.1%	6.0%	15.0%
Agree	30.1%	42.5%	37.6%	43.1%	53.4%	47.9%	30.1%	40.2%
Disagree	46.2%	20.1%	37.6%	24.8%	25.2%	18.8%	37.3%	29.4%
Strongly disagree	18.3%	10.6%	20.4%	13.8%	9.7%	5.2%	26.5%	15.4%
Total	100% (n=93)	100% (n=179)	100% (n=93)	100% (n=109)	100% (n=103)	100% (n=96)	100% (n=166)	100% (n=839)

$p < 0.001$

To explore this significant effect further, the response scale was numbered 1-4 and treated as an interval scale (lower numbers meaning greater agreement). A one-way ANOVA supports the chi-square test in showing an overall effect and indicates that it is due mainly to a difference between groups that had an RTD and those that did not. Mean ratings for groups with an RTD were: HEC (2.0), Wall Panel$_{EO+DF}$ (2.2), TOUT (2.3) and URA (2.3). For the other groups, the means were: TV$_{EO+DF}$ (2.7), Paper$_{EO+DF}$ (2.8) and Control$_{EO+DF}$ (2.8). This is a clear indication that the RTD was a key element in the smart meter technology.

4.3 Summary of EDF and AECOM findings

4.3.1 Demand reduction

Electricity

Non-smart meter group

EDF found that consumption was higher in the RRR group, who read their own meters and could get a financial reward for reducing consumption. No baseline data were available to AECOM for this group and other data issues had been identified for this trial group (see Section A1.2), meaning that no clear conclusion can be drawn as to its effectiveness.

Smart meter groups

The EDF and AECOM analyses took quite different approaches to the data but both found significant reductions in energy use in the Paper and Wall Panel groups. The AECOM analysis, taking account of baseline consumption, estimated the percentage savings as 2.3% for Paper and 4.0% for Wall Panel.

Both of these trial groups had composite interventions:

- Paper combined additional consumption data with energy efficiency advice, (both sent with bill statements);
- Wall Panel combined an RTD showing both electricity and gas consumption data with energy efficiency advice delivered through the RTD.

Hence, the benefits cannot be attributed to any single factor; indeed it is likely to be the combinations that were effective, not a single ingredient.

In both cases, the effect depended on whether EDF supplied electricity only, or both electricity and gas to the household. The effect in the Paper group was greater for dual fuel customers and the Wall Panel effect was greater for electricity-only customers. No definitive explanation can be offered for this, but a number of possibilities may be considered:

- if a customer buys only electricity from EDF, a gas smart meter would not be fitted and the RTD would display only electricity data, hence focusing conservation efforts on electricity rather than gas;
- focusing on electricity might also make the RTD itself simpler to operate;
- dual fuel customers are more likely to see the advice as being relevant to them, because it is more likely that at least one specific piece of advice will apply;
- dual fuel customers have made a market choice to take both fuels from one supplier – specifically EDF – and might therefore be more energy-aware (or energy-cost-aware) and willing to look at advice from EDF;
- the two customer types differed in how easily they could be recruited, and hence in their motivation at the start of the trial.

These are *post hoc* explanations and should be tested further if the distinction between customer types becomes relevant to key decisions.

In the case of TV, HEC and Web, EDF found that consumption was higher in the trial groups than in the control group whereas we saw a non-significant reduction. Our analysis, allowing for baseline consumption is the more relevant but the analyses at least agree that no benefit was demonstrated for these interventions. This is not to say that it would be impossible to design effective interventions using the same basic approach.

In the case of URA, EDF found no significant difference between trial and control group whereas AECOM found a significant reduction in energy use (of 7.3%), taking into account the control group and baseline consumption. The latter result is more relevant and consistent with the finding for Wall Panel, since URA was identical to Wall Panel except that the alarm function had been fixed by EDF, prior to installation.

In the case of TOUT, EDF found a small but significant difference between trial and control group. AECOM found a large reduction in energy use, taking into account the control group and baseline consumption, but it was not statistically significant (or meaningful) because the sample size was too small for households where both in-trial and pretrial data were available. The effect on overall consumption is therefore unproven but plausible; further evidence on TOUT is presented below, under load shifting and survey findings.

Overall, EDF found that differences in energy consumption between trial and control groups were more clear-cut for smaller households (one or two people). This finding should be viewed cautiously because the analysis did not take account of baseline data, but it is in keeping with the load-shifting results (see below, Section 4.3.3).

Gas

Non-smart meter group

EDF found that consumption was 2.1% lower in the RRR group, who read their own meters and could get a financial reward for reducing consumption. No baseline gas data were available and other data issues had been identified for this trial group (see Section A1.2), meaning that no clear conclusion can be drawn as to its effectiveness.

Smart meter groups

In EDF's analysis, gas consumption was higher in all groups than in the control group and this was significant for TV (by 14.7%), HEC (by 13.2%) and Web (by 14.1%) but not for Paper or Wall Panel. No gas data were available for the other groups. No baseline data were available for gas consumption. Given the impact of controlling for baseline consumption in the analysis of electricity consumption, no further analysis was conducted using the gas consumption data and EDF's gas findings should not be seen as reliable.

4.3.2 Persistence of effects

Electricity

A second full year of in-trial data was available for enough households (that also had baseline data) to allow AECOM to conduct analysis for the Paper and Wall Panel groups only.

In the case of Paper, the reduction in demand remained significant in the second year, for electricity-only customers (5.0% saving, $p<0.05$), dual fuel customers (2.4% saving, $p<0.05$) and both customer types combined (4.0% saving, $p<0.01$).

In the case of Wall Panel, the reduction in demand remained significant in the second year, for electricity-only customers (11.0% saving, $p<0.01$) and marginally significant for both customer types combined (5.8% saving, $p<0.1$). It remained non-significant for dual fuel customers.

The persistence of other effects can be judged by looking at the variation of differences between the trial and control groups over time, using quarterly figures tabulated by EDF. Although this does not take into account the baseline consumption, it does indicate whether any initial effect (if present) remained over time. The observed trends over time confirm the persistence of effects for Paper and Wall Panel. For the less certain effects – URA and TOUT – any initial effect is eroded over the first few quarters.

Gas

Given that no initial impact could be determined with confidence, the persistence of effects cannot be examined but the difference between trial and control groups appears to have stabilised after the first in-trial year (based on quarterly figures tabulated by EDF). However, examination of quarterly figures created more concern over the data, with marked inexplicable peaks in some groups.

4.3.3 Load shifting

We analysed load shifting only for electricity because this was the specific target of the intervention and therefore the sample consisted of electricity-only customers.

At weekends, there was a significant ($p<0.05$) overall reduction in the percentage of consumption that occurred in the peak tariff period, compared with the Wall Panel group (relative peak time consumption in TOUT was 92% of that in Wall Panel).

Comparing TOUT with the control group, the overall difference was modified by a significant interaction effect, meaning that peak load was reduced in TOUT only for smaller households (based on the number of people in the age range 16-64) – for larger households, relative peak time consumption actually increased. Modelled relative peak time consumption in TOUT was 86% that in the control group with nobody aged 16-64 in the household (this would be almost exclusively pensioner households), 90% with one person and 94% with two people. The tipping point is three people (98%).

On weekdays, the load-shifting effect is similar but weaker. There was a marginally significant ($p<0.1$) overall reduction in the percentage of consumption that occurred in the peak tariff period, compared with the Wall Panel group (relative peak time consumption in TOUT was 96% of that in Wall Panel).

Comparing TOUT with the control group, there is same significant interaction as was seen for weekends. Modelled relative peak time consumption in TOUT was 89% that in the control group with nobody aged 16-64 in the household, 93% with one person and 97% with two people. The tipping point is three people (101%).

4.3.4 Implications of the survey findings

Combining the insights from EDF's and AECOM's analysis of survey data, the following observations can be made.

EDF reports the following key points about motivation to change, from the customer research as a whole.

1. Although the trials have prompted some behavioural change, some respondents were happy to benefit from no more meter readers or estimated bills and nothing else.
2. Positive attitude and intentions were not always translated into less consumption or lower expenditure.
3. Core triggers to uptake are emerging: *cost* savings, desire for *control* and less *hassle*.
    - The cost savings motivation is most directly met by TOUT.
    - Control was particularly strong on all interventions with an RTD.
    - The immediacy of feedback in Wall Panel, HEC and TOUT was preferred to retrospective information in the Web, TV and Paper groups.

Customers expected, and could have benefited from, more engagement and instruction during installation of equipment.

Whether using an RTD or TV to access information, only the first or second screens in the access sequence tend to be used, so the most useful information should be put on those screens. This would have particularly impacted the TOUT intervention, where the basic RTD required multiple button-pushes to access data.

Looking at specific interventions, the following points can be made.

- RRR customers were not aware they were part of trial but reading their own meter appears to have raised some awareness about costs and consumption.
- Paper group customers indicated very little engagement with the information provided and yet it appears to have been enough to bring about a reduction in consumption.[15] This is perhaps an indication that the amount of engagement is not a good guide to whether effective action will be taken. Responding to a few pieces of the advice is all that is needed.
- In groups with the Wall Panel type RTD, customers varied in their awareness of the device, from 49% for URA, through 60% for Wall Panel to 64% for HEC. Splitting the Wall Panel group, awareness was 65% for dual fuel customers but only 54% for electricity-only customers. So awareness was greatest among dual fuel customers (in HEC or Wall Panel) but this was not associated with greater reductions in consumption.
- The percentage aware of the RTD dropped to 38% for TOUT, and the low-tech design (used only for TOUT) was a disappointment to some householders. It appears that the tariff effect could have been better supported by an alternative RTD.
- Many of the customers who were aware of the RTD were frequently accessing information from their RTD (39% daily or several times a week) though mainly limited to one householder (the bill payer).

[15] The delivery of the Paper intervention was less reliable in the second in-trial year and this may have unduly influenced the survey responses, since the survey was at the end of the trial. It is possible that engagement had been greater in the first in-trial year.

- Based on ratings of usefulness of RTDs, HEC was the best, with 95% rating it “quite” or “very” useful and TOUT the least useful (63%), with Wall Panel (79%) and URA (82%) intermediate.
- The URA intervention did not appear to be sufficiently well understood and any effects in this group are probably due to the RTD, with the alarm tending to *reduce* its effectiveness.
- The limited use that customers made of the TV and Web interventions means that their effects on consumption cannot be fairly judged until more effective implementation is achieved, making the interventions less dependent on the existing technology in the home, more easily accessible, more focused on the key information that users want to see, and with better linkage between consumption data (real-time and historic), advice and access to external financial and technical support.

Records of specific actions taken were difficult to interpret with great confidence or attribute to specific interventions but, in two cases of energy-saving action (loft and cavity wall insulation) the three groups with a (non-significantly) higher than average percentage taking the action correspond with the groups where there was a significant difference in electricity consumption between trial and control group (Paper, Wall Panel and TOUT). Wall Panel was also in the top three for draughtproofing, as were TV and HEC. These are the less expensive actions, with financial support often available, which might therefore be carried out quickly. In the case of double glazing and getting a new boiler – more expensive actions – HEC stood out (along with TV in the case of double glazing). There is also a non-significant trend for the increase in number of appliances to be less likely in trial groups Paper, Wall Panel and TOUT.

The survey question “To what extent do you agree or disagree that your Smart meter technology has enabled you to plan or budget for your energy use in the home?” offers an overall rating of the interventions. On this rating, groups that had an RTD were viewed significantly more positively than those that did not.

Focus group evaluations of the interventions placed Wall Panel, HEC and TOUT in the top three, RRR and URA in the middle two and Paper, Web and TV in the bottom three. In the case of HEC, the high position (and high ratings in some survey questions) may be due to the controller function rather than consumption feedback. These ratings are not fully reflected in the consumption findings but the ratings are derived from small focus groups which consist of different people for each intervention.

Across all the survey evidence, there is a common theme of the RTD being a key element in the interventions – most clearly in HEC (combined with a controller for heating and hot water) and least in TOUT where the device was the most basic and not the main feature of the intervention. This is in spite of the fact that a majority of those who had been given an RTD were not aware of having it. Hence there is also scope for RTDs to have greater impact, simply by ensuring that more households who receive one retain a greater awareness of its existence and actually make use of it. This would come partly from using well designed devices and partly through the installation/delivery and support processes that are put in place.

5 Analysis of E.ON data

This section presents AECOM's analysis of both the energy consumption data and the consumer survey data from E.ON. The summary at the end brings together overall findings from E.ON's analysis and AECOM's. Appendix B2 summarises the issues that need to be considered, related to the research design and execution, when interpreting or applying the findings.

5.1 Energy Consumption Data

5.1.1 Introduction

The analyses focus on the trials in which a smart meter was installed, since these are most relevant to the roll-out of smart meters and also the trials for which E.ON's analysis showed most potential to learn more by further analysis.

Data were provided in three formats for electricity and two formats for gas:

- raw energy consumption data (electricity and gas);
- weather-adjusted energy consumption data: to take into account variations in weather conditions for the same day between different years (electricity only);
- weather-adjusted and time-corrected energy consumption data: to take into account of weather variations both between years and for different days within a given year (electricity and gas).

The weather-adjusted data were considered more useful for the electricity analysis than the raw data. The analyses compared electricity consumption across different years (i.e. between pretrial and in-trial time periods). This correction adjusted for the different weather conditions between years which may impact on electricity consumption. Time correction was not used as E.ON reports that anomalies in the time-correction process appeared to lead to inconsistent data. As a result of this, raw meter readings had to be used for the analysis of gas consumption. This had implications for comparisons between the trial and control groups as differences in weather conditions could potentially confound the analyses. The approach taken to address this is presented at the end of this introduction.

Furthermore, the following approach was taken.

- Pretrial meter reads were typically available for each household at 3 to 6 month intervals (as noted in Section 3.2, the criteria for recruiting homes only required them to have two readings in the 12 month period up to the start of the trial). Given this frequency of the pretrial data, the approach to the analysis was to assess change from a whole year of pretrial readings to a whole year of in-trial readings. Given the variation in electricity consumption between seasons, it was considered too inaccurate to use this information to interpolate monthly or even quarterly pretrial electricity consumption levels.
- For the smart meter interventions, it was appropriate to use the meter readings automatically transmitted at monthly intervals rather than using the more frequent (half-hourly) consumption data. The latter dataset typically has missing data in each quarter which would need to be estimated to determine total consumption.

The electricity consumption data provided had been aggregated into quarterly periods and these were now further aggregated into annual consumption data.

- For each household, the pretrial year comprised the four quarters before the quarter in which the smart meter was installed.
- The first in-trial year comprised the four quarters after the quarter in which the meter was installed or the trial began (whichever was the later).
- The quarter in which the meter was installed was omitted as energy usage in that quarter would include usage both before and after the trial conditions began.

Similarly, the gas consumption data were aggregated into yearly periods. The approach taken was to have common pretrial and in-trial years for all trial and control groups in each analysis. Thus, while weather varied between years, it varied similarly for trial and control groups (especially given that the participants were nearly all from the same region, the English midlands). This mitigates the possible confounding effect of not using weather-corrected data: if this approach had not been taken, differences between groups might have been confounded by differences in weather conditions between the groups.

- For each household, the pretrial year comprised the period 01/04/2007 to 31/3/2008.
- For trial groups, the first in-trial year comprised the four quarters after the quarter in which the meter was installed or the trial began (whatever was the later). In order to maintain the required common in-trial dates, we restricted analysis to the households that had their first in-trial year as 01/10.2008 – 30/09/2009. This comprised most of the households.
- Data from a second entire in-trial year were only available for analysis for a relatively small percentage of households. Thus, in order to examine changes over time in differences between groups, we considered further full in-trial years by using rolling totals. This means that each successive year of data was shifted forward by one quarter, rather than a full year, and therefore overlapped the previous year by three quarters. Three such additional years were examined, by rolling forward from the first in-trial year by one, two and three quarters.

When interpreting the rolling totals, it is important to recall that the energy advice element of TG8 and TG9 was not introduced at the start of the trial period (see Section 3.2.3). In effect, this means that advice was being delivered only during the final quarter of the first in-trial year but during the whole of the in-trial year rolled forward by three quarters (“in-trial year 1 plus 3 quarters”).

Examination of the energy consumption data in the pretrial and in-trial years showed that they did not follow a normal distribution (i.e. a bell shaped curve). Similarly, the changes in energy consumption from pretrial to in-trial years were not normally distributed (nor were log transformations of the data). In all cases, the distributions were leptokurtic and influenced by extreme deviations from the mean. Similar to the EDF analysis, the E.ON analysis focused on using robust non-parametric methods suitable for the data in this trial. In keeping with this, medians are reported, rather than the usual arithmetic means, to indicate the average response to the interventions.

5.1.2 Analysis methods

Because of the data distributions, as noted above, there is a risk that households with unusually high or low consumption levels and changes in consumption will bias the calculation of summary statistics. Rather than use arbitrary methods to exclude such households from the analyses, we overcome this by adopting a non-parametric statistical test (the Mann-Whitney U test) and reporting median values for consumption and changes in consumption. Non-parametric techniques are less sensitive to outliers than parametric techniques and, although care has been taken to avoid errors in the data preparation, it is impossible to be definitive that there are no figures that are larger or smaller than they should be. The robust analysis reported here reduces the impact of such problems on the results obtained.

A series of analyses was undertaken comparing pairs of trial groups as in Table 5.1.[16] These trial group numbers were defined in Section 3.2 (Table 3.3). Households that left the study for whatever reason during the trial have been omitted from the analysis.

---

[16] The possibility of combining customer strata into a single analysis was also considered. This would have been reasonable if multivariate analysis had been possible but a simple combination of groups was not considered viable. The main reason is that the E.ON analysis had shown (and our analysis confirmed) that the strata had responded differently to the interventions. It would therefore be misleading to report an overall effect.

Table 5.1 Trial groups compared in analysis of E.ON electricity and gas consumption

Analysis	*Groups to be compared for electricity*	
E1	TG5-FP_{EO}	C1+C2
E2	TG5-FP	C3+C4
E3	TG5-NFP	C7
E4	TG5-HU_{DF}	C8
E5	TG5-E7	C2+C6+C9
E6	TG7-FP_{EO}	C1+C2
E7	TG7-FP	C3+C4
E8	TG7-HU_{DF}	C8
E9	TG8-FP_{EO}	C1+C2
E10	TG8-FP	C3+C4
E11	TG8-HU_{DF}	C8
E12	TG9-FP	C3+C4
E13	TG9-NFP	C7
E14	TG9-HU_{DF}	C8
E15	TG9-E7	C2+C6+C9
E16	TG7-FP_{EO}	TG5-FP_{EO}
E17	TG8-FP_{EO}	TG5-FP_{EO}
E18	TG7-FP	TG5-FP
E19	TG8-FP	TG5-FP
E20	TG9-FP	TG5-FP
E21	TG9-NFP	TG5-NFP
E22	TG7-HU_{DF}	TG5-HU_{DF}
E23	TG8-HU_{DF}	TG5-HU_{DF}
E24	TG9-HU_{DF}	TG5-HU_{DF}
E25	TG9-E7	TG5-E7
E26	TG8-FP_{EO}	TG7-FP_{EO}
E27	TG8-FP	TG7-FP
E28	TG9-FP	TG7-FP
E29	TG8-HU_{DF}	TG7-HU_{DF}
E30	TG9-HU_{DF}	TG7-HU_{DF}
E31	TG9-FP	TG8-FP
E32	TG9-HU_{DF}	TG8-HU_{DF}

Analysis	*Groups to be compared for gas*	
G1	TG5-FP	C2+C4
G2	TG5-NFP	C7
G3	TG5-HU_{DF}	C8
G4	TG7-FP	C2+C4
G5	TG7-HU_{DF}	C8
G6	TG8-FP	C2+C4
G7	TG8-HU_{DF}	C8
G8	TG9-FP	C2+C4
G9	TG9-NFP	C7
G10	TG9-HU_{DF}	C8
G11	TG7-FP	TG5-FP
G12	TG8-FP	TG5-FP
G13	TG9-FP	TG5-FP
G14	TG9-NFP	TG5-NFP
G15	TG7-HU_{DF}	TG5-HU_{DF}
G16	TG8-HU_{DF}	TG5-HU_{DF}
G17	TG9-HU_{DF}	TG5-HU_{DF}
G18	TG8-FP	TG7-FP
G19	TG9-FP	TG7-FP
G20	TG8-HU_{DF}	TG7-HU_{DF}
G21	TG9-HU_{DF}	TG7-HU_{DF}
G22	TG9-FP	TG8-FP
G23	TG9-HU_{DF}	TG8-HU_{DF}

For approximately one half of the electricity analyses (E1-E15), the consumption in the trial group is compared against the non-smart control group. The other analyses are comparisons against smart meter interventions because, given that the Government's policy is for all homes to have smart meters, these comparisons provide information on what additional actions would help to reduce energy demand.

- E16-E25: The comparison group is a smart meter only.
- E26-E30: The comparison group is a smart meter plus monthly bills.
- E31-E32: The comparison group is a smart meter plus monthly bills plus energy efficiency advice.

Similarly, for approximately one half of the gas analyses (G1-G10), the consumption in the trial groups is compared against the non-smart control groups. The other analyses are comparisons against smart meter interventions.

- G11-G17: The comparison group is a smart meter only.
- G18-G21: The comparison group is a smart meter plus monthly bills.
- G22-G23: The comparison group is a smart meter plus monthly bills plus energy efficiency advice.

For each pair of trial groups, the Mann-Whitney U test was use to compare the groups listed in Table 5.1, on:

- pretrial electricity/gas consumption;
- the change in electricity/gas consumption from pretrial to the first in-trial year;
- the change in electricity/gas consumption from pretrial to the first in-trial year plus one, two and three quarters.

Using non-parametric statistics limits the options for multivariate analysis, i.e. taking into account variables other than the interventions. However, two possible confounders were identified that could be used to split the sample for independent analysis with sufficient numbers in each group (at least for some trials): geographic location and whether the household had participated in a survey before the end of the trial.

- For the electricity consumption analyses, location was defined as being according to the distribution area, with households in two areas (coded "B" and "E") encompassing almost the entire dataset.
- For the gas consumption analyses, location was again defined according to the distribution area, with "East Midlands" and "West Midlands" encompassing almost the entire dataset.

5.1.3 Electricity consumption results

Summary statistics are presented in Appendix E2.[17]

- Median electricity consumption values for the pretrial and in-trial years for each trial group are given in Tables E2.1-E2.5.
- The differences in electricity consumption between pretrial and in-trial years are given in Tables E2.6-E2.9.

In preliminary analysis, data were split by location and survey participation, for the pretrial year and in-trial year 1 only. Where the sample sizes are sufficient, the results are similar to those seen in the whole sample. To test differences directly, the change in consumption from pretrial to the first in-trial year was compared between the two locations and between the two survey participation groups; this was done independently for each trial and control group, using the Mann-Whitney U test. Only one difference was significant at the 5% level of significance, out of 37 comparisons; this provides confidence in analysing all households together, as below, rather than separately according to location or survey participation.

The trial groups to be compared were then tested for differences in electricity consumption during the pretrial year (see Table E2.10 for details). Several of the comparison pairs have pretrial data that differ significantly from each other. As many of these are comparisons between trial and control groups, this casts doubt on whether the creation of the control groups has been successful, confirming the need to control for baseline consumption in the analysis.

---

[17] The number of households with valid data differs between the pretrial year and in-trial years. Our comparison between pretrial and in-trial consumption used only those households with valid data in both comparison years. For this reason, the differences determined in the analysis should not be compared with the differences between medians in the Appendix E tables, which are based on all households with data in the particular year only. There are several reasons for differences in the number of households with valid data between years. For example, sufficient pretrial data were not available for all households in the trial groups and some households were recruited too late into the trial to have a complete second year of in-trial data. Households that left the study for whatever reason during the trial have been omitted from the data.

The concern is greatest for the high use dual fuel (HU_{DF}) group, since the pretrial consumption is much higher in the trial groups than the control groups; while statistical controls will mathematically adjust for this, they do not eliminate the possibility that the trial groups had greater behavioural potential to reduce consumption and/or greater risk of regression to the mean.[18]

The trial groups to be compared were then assessed to see if they had similar changes in consumption from the pretrial year to in-trial years. The table of results from the Mann-Whitney U tests are shown in Appendix E (Tables E2.11 – E2.14) and summarised below.

Comparisons with non-smart meter control groups

To make it more readable, we have reported the analysis below, using the following strata as defined in Section 3.2.

- 'Fuel poor' (electricity only) – FP_{EO}
- 'Fuel poor' – FP
- Not 'fuel poor' – NFP
- High use (dual fuel) – HU_{DF}
- Economy 7 – E7

Comparisons E1-E15 are presented in Table 5.2 for changes from pretrial to the first in-trial year, with all significant differences identified. Redefining trial periods (see Section 5.1.1) has allowed additional significant effects to be seen, in the FP and E7 trial groups, where there were only non-significant trends in the E.ON analysis. Figures 5.1 to 5.5 show electricity consumption for each of the customer strata, over the four trial periods (the first in-trial year and the first in-trial year rolled forward by one, two and three quarters). Each bar represents the change from the baseline year, adjusted for change in the control group.

Table 5.2 Median changes in electricity consumption (pretrial vs in-trial year 1)[1]

	Control	*TG5: Smart meter*		*TG7: Smart meter & monthly bills*		*TG8: Smart meter & monthly bills & energy advice*		*TG9: Smart meter & monthly bills & energy advice & RTD*	
Stratum	*Change against pretrial*	*Change against pretrial*	*Adjusted for control*	*Change against pretrial*	*Adjusted for control*	*Change against pretrial*	*Adjusted for control*	*Change against pretrial*	*Adjusted for control*
FP_{EO}	-5.9%	-3.6%	2.3%**	-3.8%	2.1%~	-5.6%	0.3%		
FP	-5.2%	-4.1%	1.1%	-4.0%	1.2%	-2.7%	2.5%	-6.9%	-1.7%*
NFP	-4.9%	-3.9%	1.0%					-6.2%	-1.3%
HU_{DF}	-4.2%	-6.1%	-1.9%**	-7.6%	-3.4%**	-6.6%	-2.4%**	-8.1%	-3.9%**
E7	-7.0%	-6.6%	0.4%					-9.9%	-2.9%**

**p<0.01; ~p<0.1

[1] Blank cells are where there were no trial groups

[18] Regression to the mean refers to the statistical phenomenon whereby values above (or below) the mean are likely to decrease (increase) if measurements of the same group are made a second time.

Figure 5.1 Changes over time in electricity consumption for 'fuel poor' (electricity only)

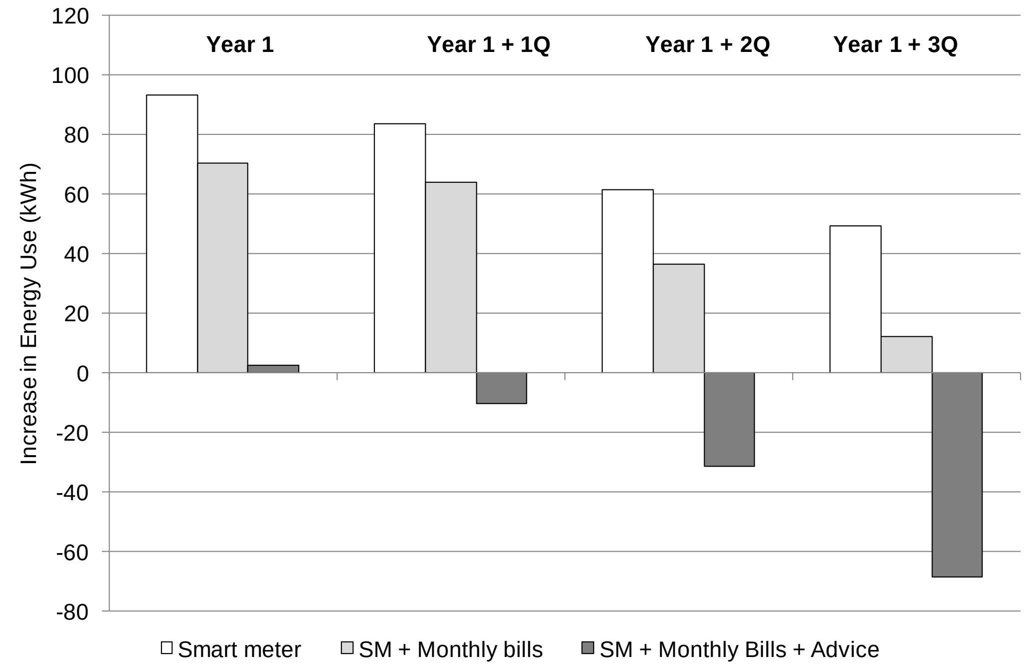


Figure 5.2 Changes over time in electricity consumption for 'fuel poor'

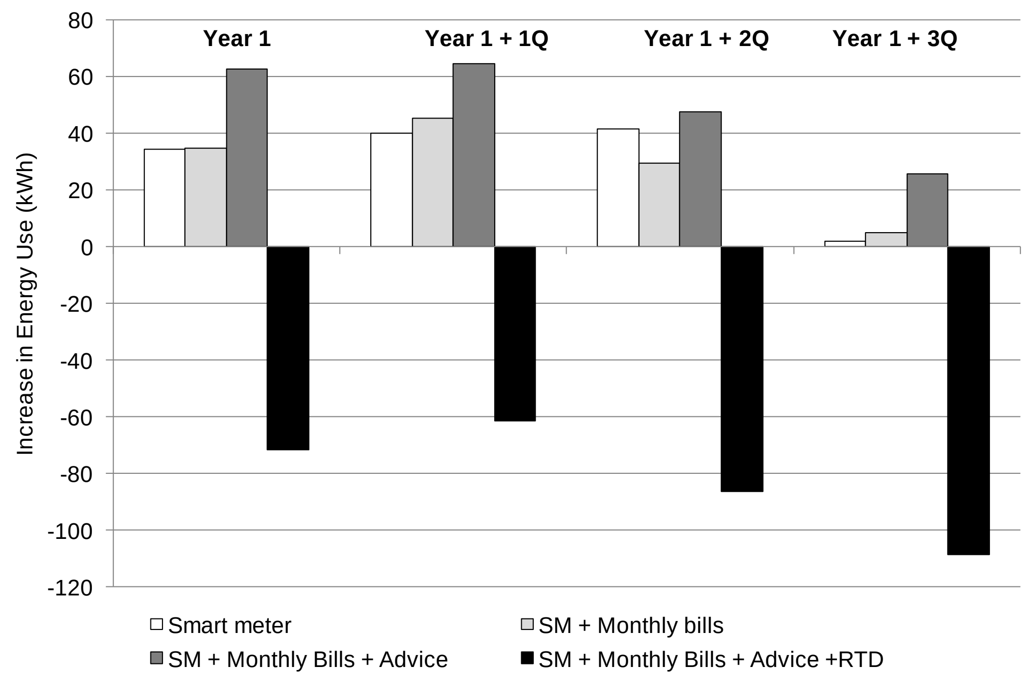


Figure 5.3 Changes over time in electricity consumption for not 'fuel poor'

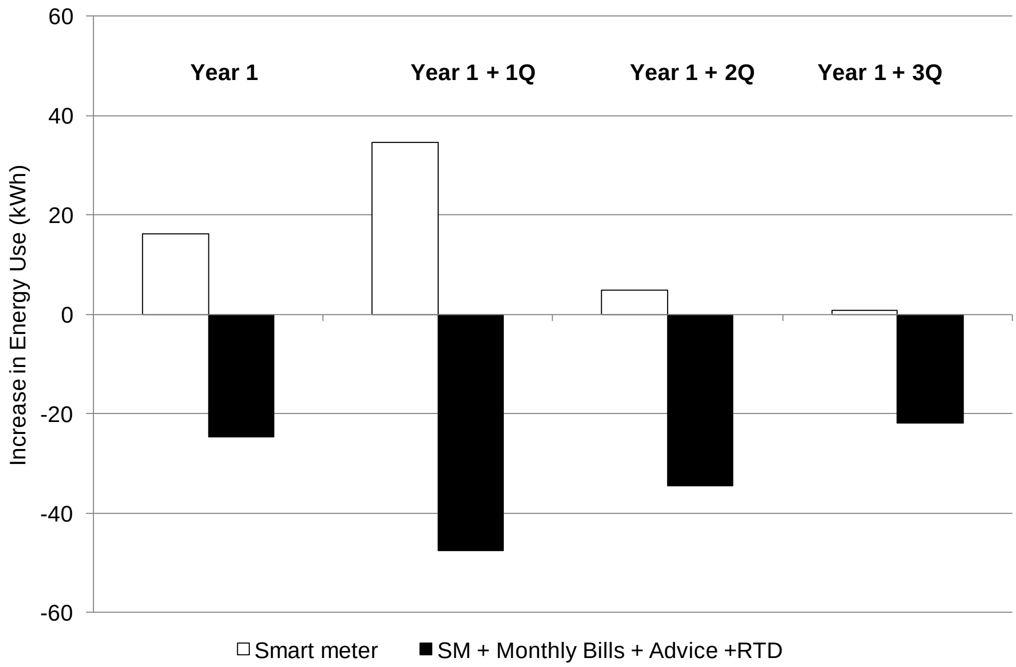


Figure 5.4 Changes over time in electricity consumption for high use dual fuel

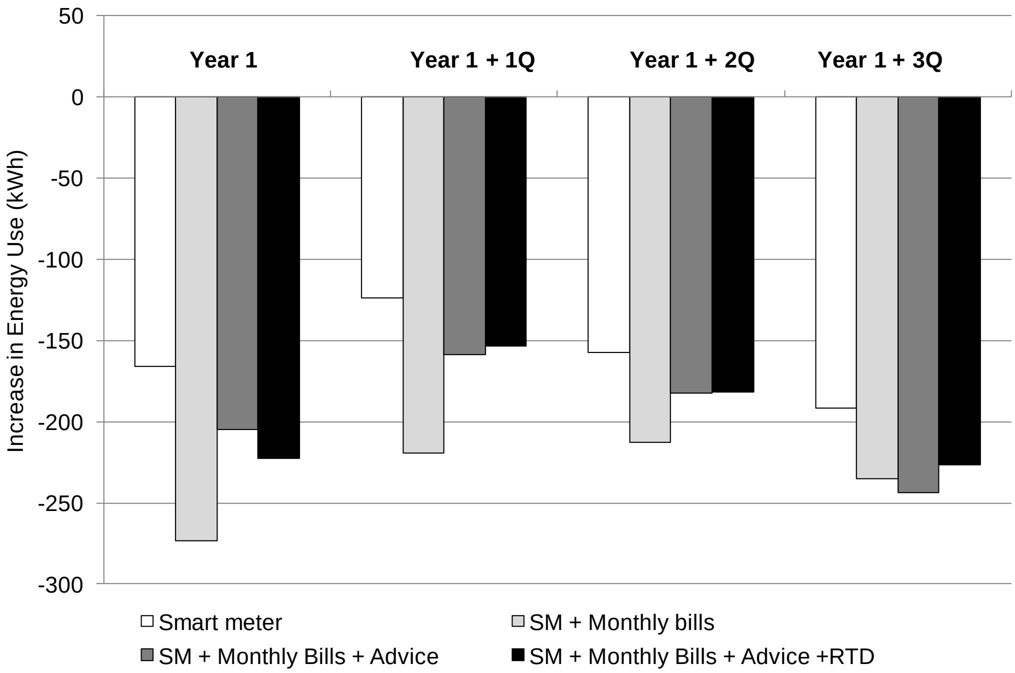
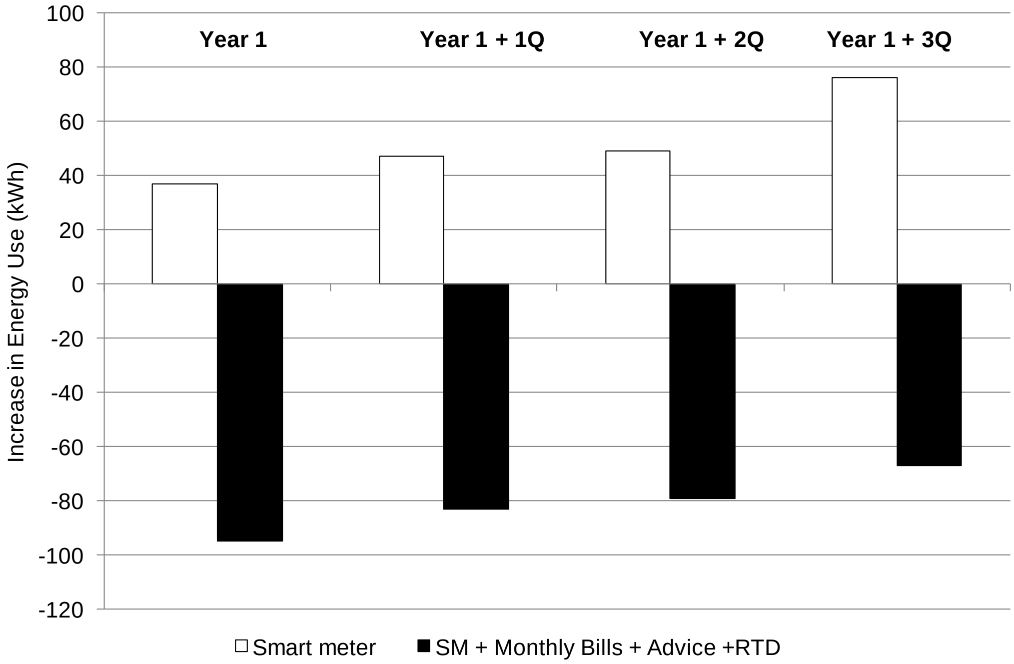


Figure 5.5 Changes over time in electricity consumption for Economy 7



In year 1, the reduction in consumption is significantly greater in the trial group than in the control group for all HU_{DF} groups. It is also significantly greater in TG9 for all strata except NFP (where the non-significant trend is in the same direction). The opposite effect occurs for FP_{EO} in TG5 but in this case the pretrial consumption was significantly higher in the control group.

For all the differences that are significant ($p<0.05$) in year 1, the difference remains significant through to year 1 plus 3 quarters (except for the increase in consumption for FP_{EO} in TG5, which reaches only the $p<0.1$ level in the final year).

As noted above, the HU_{DF} trial groups had significantly higher consumption pretrial than their control groups, by more than 1000 kWh. The changes observed in these groups could therefore be, at least in part, due to "regression to the mean". The other two significant reductions in consumption are not affected by this concern – in TG9, i.e. smart meter trial groups with RTDs. Just as important, in the absence of an RTD, consumption typically increased in all strata other than HU_{DF}.

Comparisons with smart meter control groups

The remaining comparisons are between trial groups, effectively using smart meter trial groups as the control groups, and investigating the impact of adding additional measures. These comparisons are shown in Table 5.3.

All the significant differences represent a reduction in consumption with the addition of an intervention. Incremental effects of interventions are seen mainly for FP trial groups, where there is a significant difference between TG9 and all three comparison trial groups. There is also a significant difference between TG8 and TG5 but note that (a) pretrial consumption was significantly higher in TG8 and (b) advice was not actually introduced until the final quarter of the period, so the intervention in TG8 differed little from that in TG7. There is also a significant difference for E7 between TG9 and TG5, the only comparison that could be made for E7. It can also be noted that the effects in the HU_{DF} stratum did not differ significantly across the interventions, again suggesting that the effects may be attributed to regression to the mean.

The persistence of these effects across rolling years was as follows.

Comparison with smart meter only

- The effects in TG9 – for FP and E7 – remain significant through to year 1 plus 3 quarters. In the case of FP, the effect may even be getting stronger over time.
- The effect in TG8 – for FP_{EO} – drops from $p<0.05$ in year 1 to $p<0.1$ in year 1 plus 1 quarter but then returns to $p<0.05$ in the remaining two periods.
- These trends may reflect the fact that TG8 and TG9 were not receiving advice in all four quarters until the final year but, in general, advice did not appear to be having a significant effect on consumption.

Comparison with smart meter plus monthly bills

- The effect in TG9 – for FP – remains significant until in-trial year 1 plus 1 quarter but then drops to $p<0.1$.

Comparison with smart meter, monthly bills and energy advice

- The effect in TG9 – for FP – remains significant through to in-trial year 1 plus 3 quarters.

The final year may be seen as most relevant to evaluating the effect of energy advice, since advice was delivered in all four quarters in that year. Where an RTD was also provided (TG9) any effect of advice is masked. The HU_{DF} groups did not differ from each other in any year. Hence, the effect of advice can be assessed only in the FP groups, by comparing TG8 with the non-smart control group and the other smart meter groups. In all four periods, TG8 did not differ significantly from the control group or TG7. Across all four periods, reduction in consumption in TG8 was consistently greater than in TG5 for FP_{EO} and consistently less that TG9 in FP. Hence, there is no clear evidence for an effect of advice.

Table 5.3 Median changes in electricity consumption (pretrial vs in-trial year 1) for smart meter 'controls'[1]

		'Control'	TG7: Smart meter & monthly bills		TG8: Smart meter & monthly bills & energy advice		TG9: Smart meter & monthly bills & energy advice & RTD	
'Control'	Stratum	Change against pretrial	Change against pretrial	Adjusted for 'control'	Change against pretrial	Adjusted for 'control'	Change against pretrial	Adjusted for 'control'
TG5: Smart meter only	FP_{EO}	-3.6%	-3.8%	-0.2%	-5.6%	-2.0%*		
	FP	-4.1%	-4.0%	0.1%	-2.7%	1.4%	-6.9%	-2.8%**
	NFP	-3.9%					-6.2%	-2.3%
	HU_{DF}	-6.1%	-7.6%	-1.5%~	-6.6%	-0.5%	-8.1%	-2.0%
	E7	-6.6%					-9.9%	-3.3%**
TG7	FP_{EO}	-3.8%			-5.6%	-1.8%		
	FP	-4.0%			-2.7%	1.3%	-6.9%	-2.9%*
	HU_{DF}	-7.6%			-6.6%	1.0%	-8.1%	-0.5%
TG8	FP	-2.7%					-6.9%	-4.2%**
	HU_{DF}	-6.6%					-8.1%	-1.5%

**p<0.01; *p<0.05; ~p<0.1

[1] Blank cells are where there were no trial groups

5.1.4 Gas consumption results

Presentation of the findings follows the same pattern as for the electricity findings. The trial groups to be compared are shown in Table 5.1. Summary statistics are presented in Appendix E2.

In preliminary analysis, data were split by location and survey participation, for the pretrial year and in-trial year 1 only. Where the sample sizes are sufficient, the results are similar to those seen in the whole sample. To test differences directly, the change in consumption from pretrial to the first in-trial year was compared between the two locations and between the two survey participation groups; this was done independently for each trial and control group, using the Mann-Whitney U test. Similarly to the electricity analysis, only two tests are significant at the 5% level so we can feel confident in analysing the whole sample, as below, rather than separately according to location or survey participation.

Descriptive statistics for pretrial and in-trial year 1 are presented in Tables E2.15-E2.23[19]. The results of Mann-Whitney U tests comparing trial and control groups are shown in Table E2.24 for comparison of baseline consumption and Tables E2.25-B2.28 for comparison of differences between in-trial and pretrial years. To make the findings more readable, again we have presented the main results in a series of summary tables below. All tests of difference were performed using Mann-Whitney U tests.

The trial groups to be compared were first tested for differences in gas consumption during the pretrial year (see Table E2.24 for details). Several of the comparison pairs differ significantly from each other. As many of these are comparisons between trial and control groups, this casts doubt on whether the creation of the control groups has been successful, confirming the need to control for baseline consumption in the analysis. The concern is greatest for HU_{DF}, since the pretrial consumption is much higher in the trial groups than the control groups; while statistical controls will mathematically adjust for this, they do not eliminate the possibility that the trial groups had greater behavioural potential to reduce consumption and/or greater risk of regression to the mean.

Comparisons with non-smart meter control groups

Table 5.4 presents the results for comparisons G1-G10 for changes from pretrial to the first in-trial year. All were different at the 5% level of significance ($p<0.05$) and some at the 1% level. Figures 5.6 to 5.8 show gas consumption for each of the customer strata, over the four trial periods (the first in-trial year, the first in-trial year plus 1 quarter, the first in-trial year plus 2 quarters and the first in-trial year plus 3 quarters).

Table 5.4 Median changes in gas consumption (pretrial vs in-trial year 1)[1]

	Control	*TG5: Smart meter*		*TG7: Smart meter & monthly bills*		*TG8: Smart meter & monthly bills & energy advice*		*TG9: Smart meter & monthly bills & energy advice & RTD*	
Stratum	*Change against pretrial*	*Change against pretrial*	*Adjusted for control*	*Change against pretrial*	*Adjusted for control*	*Change against pretrial*	*Adjusted for control*	*Change against pretrial*	*Adjusted for control*
FP	-0.7%	-5.1%	-4.4%**	-7.4%	-6.7%**	-7.9%	-7.2%**	-5.3%	-4.6%**
NFP	-1.3%	-4.9%	-3.6%**					-6.2%	-4.9%**
HU_{DF}	-1.2%	-3.5%	-2.3%*	-3.7%	-2.5%**	-3.6%	-2.4%**	-3.4%	-2.2%**

**p<0.01; *p<0.05

[1] Blank cells are where there were no trial groups

[19] In some of these tables, the number of households with valid data is too small in some groups for analyses to be undertaken appropriately but the results are included (in grey text) for completeness.

Figure 5.6 Changes over time in gas consumption for 'fuel poor'

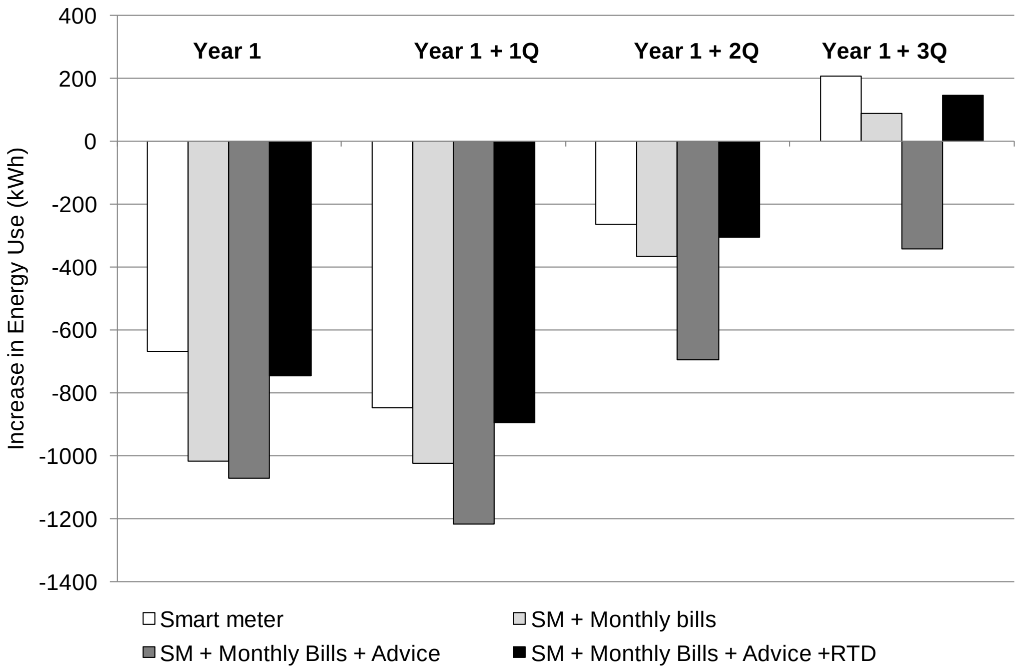


Figure 5.7 Changes over time in gas consumption for not 'fuel poor'

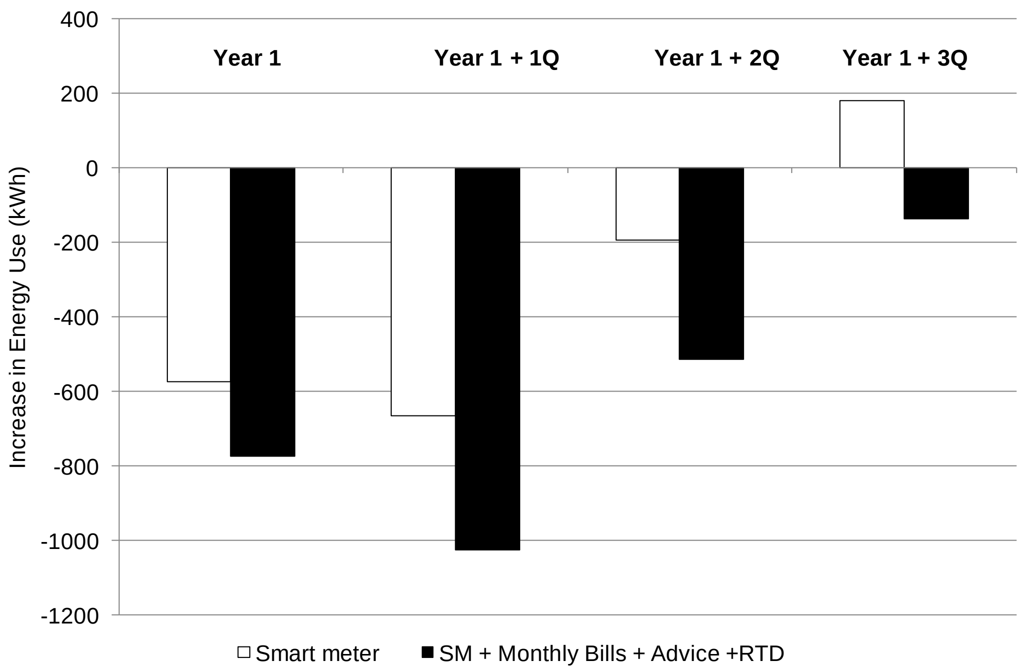
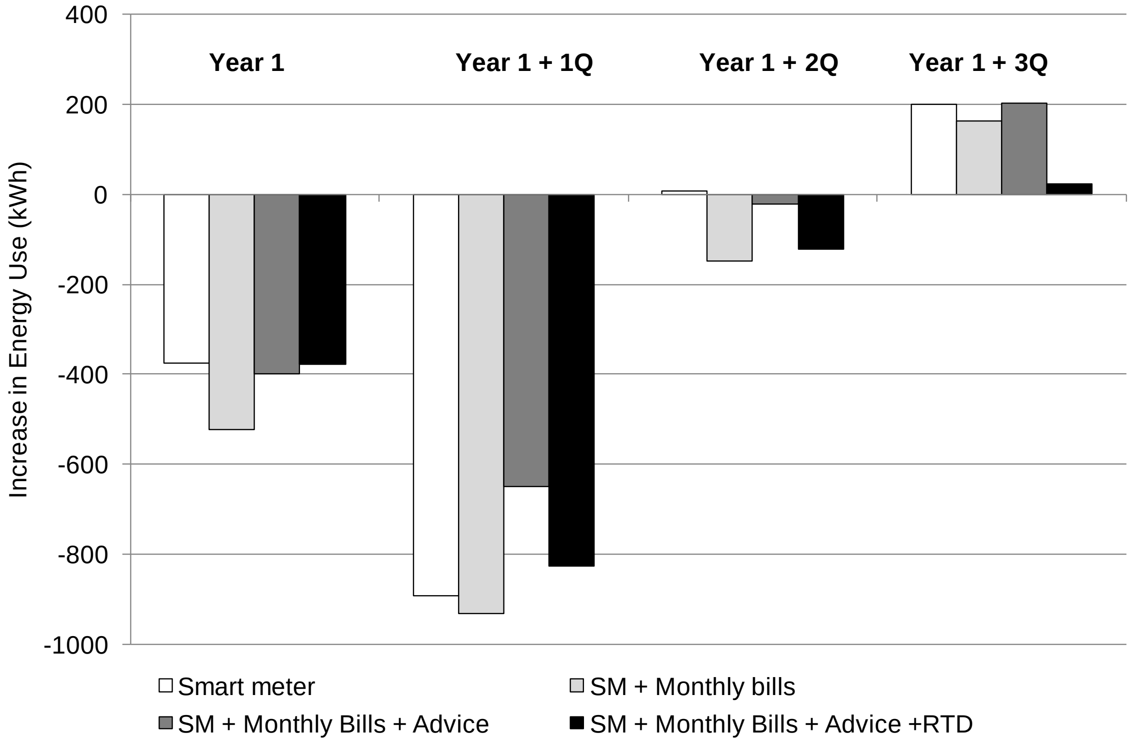


Figure 5.8 Changes over time in gas consumption for high use dual fuel



Reductions in gas consumption persisted to in-trial year 1 plus 1 quarter and, in some cases to in-trial year 1 plus 2 quarters (with a trend for some strengthening of the effect in in-trial year 1 plus 1 quarter). The only effect that persisted through to in-trial year 1 plus 3 quarters was in TG8 (smart meter, monthly bills and energy advice) for FP, which had the greatest median percentage difference in in-trial year 1.

However, as noted above, the HU_{DF} trial groups had significantly higher consumption pretrial, by more than 2000 kWh. The changes observed in these groups could therefore be, at least in part, due to “regression to the mean”. The NFP trial groups with RTDs also had higher pretrial consumption than its control group but regression to the mean is less likely as an explanation here, for two reasons: (a) the pretrial difference was smaller (less than 800 kWh) and (b) the effect is greater in TG8 (where there was not a significant pretrial difference).

Comparisons with smart meter control groups

The remaining comparisons are between trial groups, effectively using smart meter trial groups as the control groups, and investigating the impact of additional measures. These comparisons are shown in Table 5.5. The only case of an incremental reduction was between TG9 and TG5 for NFP and this persisted through to in-trial year 1 plus 1 quarter. The opposite effect occurs in the comparison of TG9 and TG8 for FP: there is a relative increase in consumption (and this effect persists through to in-trial year 1 plus 3 quarters although dipping just below the 5% level of significance during in-trial year 1 plus 2 quarters).

Using the same approach as for electricity consumption, FP groups were compared to determine whether an effect of energy advice emerged as this intervention progressively affected more of the in-trial years. Again, no such effect could be seen.

Overall, the findings suggest that gas consumption is reduced following the introduction of a smart meter, without any additional intervention.

As with the electricity findings, redefining trial periods has allowed additional significant effects to be seen, in the NFP and HU_{DF} strata, where there were only non-significant trends in the E.ON analysis. The fact that effects generally waned over time would explain why E.ON did not see the effects over the full in-trial period.

Table 5.5 Median changes in gas consumption (pretrial vs in-trial year 1) for smart meter 'controls'[1]

		Control	*TG7: Smart meter & monthly bills*		*TG8: Smart meter & monthly bills & energy advice*		*TG9: Smart meter & monthly bills & energy advice & RTD*	
'Control'	*Customer stratum*	*Change against pretrial*	*Change against pretrial*	*Adjusted for control*	*Change against pretrial*	*Adjusted for control*	*Change against pretrial*	*Adjusted for control*
TG5: Smart meter only	FP	-5.1%	-7.4%	-2.3%	-7.9%	-2.8%~	-5.3%	-0.2%
	NFP	-4.9%					-6.2%	-1.3%*
	HU_{DF}	-3.5%	-3.7%	-0.2%~	-3.6%	-0.1%	-3.4%	0.1%
TG7	FP	-7.4%			-7.9%	-0.5%	-5.3%	2.1%
	HU_{DF}	-3.7%			-3.6%	0.1%	-3.4%	0.3%
TG8	FP	-7.9%					-5.3%	2.6%**
	HU_{DF}	-3.6%					-3.4%	0.2%

**p<0.01; *p<0.05 ~p<0.1

[1] Blank cells are where there were no trial groups

5.2 Survey Data

5.2.1 Rationale for analyses carried out

Through consumer surveys, E.ON collected a considerable amount of data. Here we report on the responses obtained for a selection of the questions asked. The reasons behind the choice of questions to analyse are listed below.

- Questions were chosen to be analysed where further investigations could usefully add to those analyses already undertaken by the energy supplier and reported on elsewhere.
- Focus was put on understanding change in consumer behaviour and/or energy use. Thus, questions relevant to this goal were selected for analysis.
- In order for the conclusions drawn from the analyses to be robust, the questions chosen could not be open to interpretation by respondents because of their wording and the response choices given for the questions had to be unambiguous.
- It was important that questions chosen for analysis did not suffer from potential distortion by their position in the survey. Thus questions were only included where answers to earlier questions in the survey did not have the potential to influence or distort responses.
- In order for effective analyses to be undertaken, sufficient numbers of respondents had to exist.

For the selected questions, the tables below show summaries of the responses given. Trial groups are omitted from the analysis where the question did not apply to them (e.g. if the question was about RTDs and the group was not given RTDs).

Many of the tables show the responses given according to the trial group to which the respondent belonged. For these, chi-square tests of independence have been carried out to consider whether or not the responses to the questions are independent of the trial group to which the respondents belonged. In some circumstances, response categories have been combined in order to have sufficient number of respondents in the cells of the table for the assumptions behind the chi-square test to be valid. Missing data (including the responses equivalent to "don't know" and "refuse to answer") have been omitted. The p-values that result from the chi-square tests are shown beneath the tables.

Where the p-value from the chi-square test is less than 0.05 (indicating that there is sufficient evidence to reject the claim that the responses are independent of trial group at the 5% level of significance), the differences between the responses given by the different trial groups were examined and are reported on below.

5.2.2 Results

Actions taken by households

Respondents were asked whether they had taken (and/or were still taking) a range of actions that could reduce their energy consumption. The percentage taking each action was first compared in the 'fuel poor' (FP) and not 'fuel poor' (NFP) strata, since they were represented in all trial groups.

The general form of the first set of relevant questions was "Still thinking about the past 2 years, have you […] at all to reduce your household energy consumption?". Table 5.6 shows the percentage answering "Yes" and the p-value associated with the chi-square test for each action. The follow-up question, if the action had been taken at all, was "And do you still […]?". Table 5.7 shows the equivalent figures relating to this question.

Whether or not the difference between trial groups was significant, the percentage having taken (or still taking) the action was consistently higher in the smart meter groups receiving advice (with or without an RTD). The groups least likely to have taken (or be taking) action were those receiving extra billing information but not advice or an RTD (with or without a smart meter) and the control group.

These differences are also reflected in analysis of the total number of actions taken, which differed significantly between trial groups ($p=0.002$ and $p<0.001$ for actions taken and actions still being taken, respectively)[20]. This is represented by cumulative percentages in Table 5.8 for actions taken and Table 5.9 for actions still being taken. Tables 5.10 and 5.11 show the p-values from Mann-Whitney U tests, comparing each trial group with each other trial group, for actions taken and actions still being taken, respectively.

There were sufficient survey responses to undertake a separate analysis of the high user dual fuel (HU_{DF}) stratum, for the smart meter and control groups only. The pattern of differences was similar to the FP/NFP analysis but there was a significant effect for only one action ("filled the kettle less") – see Tables 5.12 and 5.13 for actions taken and actions still being taken, respectively. The total number of actions taken did not differ significantly between trial groups ($p=0.266$ and $p=0.312$ for actions taken and actions still being taken, respectively).[21] This is represented by cumulative percentages in Table 5.14 for actions taken and Table 5.15 for actions still being taken. Because the overall effect of trial group was not significant, pairs were not compared using Mann-Whitney U tests.

[20] Using Kruskal-Wallis nonparametric analysis of variance to test for differences between groups.

[21] Using Kruskal-Wallis nonparametric analysis of variance to test for differences between groups.

Table 5.6 Percentage of each trial group that had taken each action (FP and NFP strata)

	Trial group									
Action (p-value)	*1*	*2*	*4*	*5*	*6*	*7*	*8*	*9*	*10*	*Total*
Turned the thermostat down (p=0.044)	64.7%	**74.0%**	68.9%	**73.8%**	67.0%	66.7%	**76.5%**	**73.6%**	67.0%	70.2%
Reduced the amount of time your heating is on (p=0.024)	66.2%	72.5%	**74.4%**	72.2%	69.1%	65.0%	**76.1%**	**77.1%**	67.0%	70.8%
Had showers instead of baths (p=0.001)	62.7%	69.0%	71.1%	72.2%	71.8%	71.1%	**76.9%**	**81.1%**	66.5%	71.2%
Fitted a hot water cylinder jacket (p=0.581)	31.8%	35.0%	**38.9%**	33.9%	**38.8%**	30.1%	34.8%	32.3%	35.8%	34.5%
Filled the kettle less (p=0.169)	73.1%	75.0%	74.4%	77.4%	75.5%	72.8%	**78.9%**	**81.6%**	71.0%	75.3%
Put lids on pans when cooking (p=0.028)	65.7%	73.5%	71.1%	**76.6%**	**77.1%**	71.5%	**76.9%**	**79.6%**	70.1%	73.4%
Used lights less – by this we mean switching lights off when you leave a room or using fewer lights in rooms (p=0.538)	80.1%	77.0%	81.7%	76.2%	82.4%	78.5%	80.6%	**84.1%**	78.9%	79.7%
Sample size	201	200	180	248	188	246	247	201	355	2066

Shading = higher than the total %, Bold = Higher by 3%.

Table 5.7 Percentage of each trial group still taking each action (FP and NFP strata)

	Trial group									
Action (p-value)	*1*	*2*	*4*	*5*	*6*	*7*	*8*	*9*	*10*	*Total*
Turn the thermostat down (p=0.111)	62.7%	70.5%	68.3%	**72.6%**	64.9%	65.0%	**72.9%**	70.6%	64.2%	67.9%
Reduce the amount of time your heating is on (p=0.006)	63.7%	69.5%	**73.9%**	69.8%	67.0%	63.8%	**74.5%**	**75.6%**	62.8%	68.5%
Have showers instead of baths (p=0.001)	61.2%	66.0%	70.0%	71.8%	70.7%	69.5%	**75.3%**	**79.1%**	64.5%	69.6%
Fill the kettle less (p=0.035)	70.6%	72.5%	73.3%	76.2%	74.5%	71.1%	**77.3%**	**80.6%**	67.3%	73.3%
Put lids on pans when cooking (p=0.018)	64.2%	71.0%	70.6%	**75.4%**	**75.5%**	70.3%	**75.3%**	**78.1%**	67.0%	71.7%
Use lights less – by this we mean switching lights off when you leave a room or using fewer lights in rooms (p=0.362)	77.6%	73.5%	80.0%	74.6%	**81.4%**	76.8%	78.5%	**82.1%**	75.5%	77.5%
Sample size	201	200	180	248	188	246	247	201	355	2066

Shading = higher than the total %, Bold = Higher by 3%.

Table 5.8 Sum of actions taken: cumulative percentages (FP and NFP strata)

Sum	Trial group 1	2	4	5	6	7	8	9	10	Overall
0	5.5%	5.5%	3.9%	4.0%	4.3%	5.7%	3.2%	0.5%	5.1%	4.3%
1	11.4%	7.0%	6.1%	5.6%	8.0%	8.5%	4.0%	2.0%	7.6%	6.7%
2	14.9%	13.0%	8.9%	10.1%	10.6%	13.0%	6.5%	6.5%	15.2%	11.2%
3	22.4%	18.0%	17.8%	19.0%	20.7%	24.0%	16.6%	16.9%	23.7%	20.2%
4	42.8%	32.0%	35.0%	33.5%	33.0%	41.1%	29.6%	27.4%	40.6%	35.4%
5	67.7%	61.0%	62.2%	57.7%	54.8%	63.4%	54.7%	53.2%	63.1%	59.9%
6	91.0%	87.5%	85.6%	87.9%	86.7%	88.6%	84.6%	84.1%	88.5%	87.3%
7	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%
Sample size	201	200	180	248	188	246	247	201	355	2066

Table 5.9 Sum of actions still taken: cumulative percentages (FP and NFP strata)

Sum	Trial group 1	2	4	5	6	7	8	9	10	Overall
0	8.5%	8.5%	5.6%	5.2%	6.4%	7.3%	4.5%	2.5%	8.5%	6.4%
1	14.4%	10.0%	7.8%	8.1%	10.1%	11.0%	6.1%	5.0%	11.5%	9.4%
2	17.4%	16.0%	11.1%	13.3%	13.8%	15.0%	10.5%	10.9%	20.6%	14.7%
3	29.9%	24.0%	23.9%	23.4%	25.5%	29.7%	21.1%	19.4%	31.8%	25.8%
4	50.2%	42.0%	45.6%	41.9%	41.0%	46.7%	38.1%	33.3%	51.0%	43.8%
5	79.6%	76.5%	70.0%	67.7%	69.1%	73.6%	66.0%	62.7%	75.2%	71.3%
6	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%
Sample size	201	200	180	248	188	246	247	201	355	2066

Table 5.10 Sum of actions taken: p-values for differences between pairs of groups (FP and NFP strata)

Trial Number	Trial group 1	2	4	5	6	7	8	9	10
1	-								
2	0.050*	-							
4	0.069~	0.926	-						
5	0.022*	0.793	0.763	-					
6	0.019*	0.590	0.596	0.767	-				
7	0.501	0.198	0.223	0.107	0.086~	-			
8	0.001**	0.188	0.185	0.260	0.467	0.007**	-		
9	<0.001**	0.099~	0.104	0.146	0.294	0.003**	0.713	-	
10	0.416	0.191	0.215	0.092~	0.075~	0.953	0.004**	0.002**	-

~p<0.1, *p<0.05, **p<0.01 (Shading also denotes p<0.01 – because of the large number of independent tests, this more conservative criterion should be used.)

Table 5.11 Sum of actions still taken: p-values for differences between pairs of groups (FP and NFP strata)

Trial Number	*Trial group*								
	1	*2*	*4*	*5*	*6*	*7*	*8*	*9*	*10*
1	-								
2	0.146	-							
4	0.056~	0.594	-						
5	0.010*	0.251	0.620	-					
6	0.029*	0.388	0.813	0.806	-				
7	0.294	0.696	0.325	0.115	0.226	-			
8	0.001**	0.051~	0.209	0.420	0.323	0.017*	-		
9	<0.001**	0.006**	0.043*	0.100	0.077~	0.002**	0.369	-	
10	0.851	0.176	0.052~	0.008**	0.031	0.325	<0.001**	<0.001**	-

~p<0.1, *p<0.05, **p<0.01 (Shading also denotes p<0.01 – because of the large number of independent tests, this more conservative criterion should be used.)

Table 5.12 Percentage of each trial group that had taken each action (HU_{DF})

	Trial group					
Action (p-value)	*5*	*7*	*8*	*9*	*10*	*Total*
Turned the thermostat down (p=0.730)	68.0%	74.0%	75.0%	**76.0%**	72.0%	73.0%
Reduced the amount of time your heating is on (p=0.710)	68.0%	69.0%	70.0%	**76.0%**	68.0%	70.2%
Had showers instead of baths (p=0.189)	(74.0%	74.0%	**85.0%**	73.0%	72.0%	75.6%
Fitted a hot water cylinder jacket (p=0.754)	31.0%	31.0%	34.0%	**38.0%**	37.0%	34.2%
Filled the kettle less (p=0.010)	64.0%	70.0%	**83.0%**	**81.0%**	77.0%	75.0%
Put lids on pans when cooking (p=0.212)	78.0%	66.0%	78.0%	78.0%	76.0%	75.2%
Used lights less – by this we mean switching lights off when you leave a room or using fewer lights in rooms (p=0.884)	78.0%	77.0%	**81.0%**	79.0%	75.0%	78.0%
Sample size	100	100	100	100	100	500

Shading = higher than the total %, Bold = Higher by 3%.

Table 5.13 Percentage of each trial group still taking each action (HU_{DF})

	Trial group					
Action (p-value)	*5*	*7*	*8*	*9*	*10*	*Total*
Turn the thermostat down (p=0.683)	64.0%	71.0%	71.0%	**73.0%**	71.0%	70.0%
Reduce the amount of time your heating is on (p=0.452)	64.0%	67.0%	68.0%	**76.0%**	68.0%	68.6%
Have showers instead of baths (p=0.311)	71.0%	72.0%	**82.0%**	71.0%	71.0%	73.4%
Fill the kettle less (p=0.010)	60.0%	68.0%	**80.0%**	**78.0%**	75.0%	72.2%
Put lids on pans when cooking (p=0.242)	74.0%	64.0%	**76.0%**	**77.0%**	74.0%	73.0%
Use lights less – by this we mean switching lights off when you leave a room or using fewer lights in rooms (p=0.943)	74.0%	75.0%	78.0%	78.0%	75.0%	76.0%
Sample size	100	100	100	100	100	500

Shading = higher than the total %, Bold = Higher by 3%.

Table 5.14 Sum of actions taken: cumulative percentages (HU_{DF})

Sum	Trial group 5	7	8	9	10	Overall
0	8.0%	5.0%	2.0%	6.0%	1.0%	4.4%
1	11.0%	7.0%	3.0%	8.0%	4.0%	6.6%
2	14.0%	14.0%	3.0%	9.0%	9.0%	9.8%
3	25.0%	26.0%	13.0%	16.0%	20.0%	20.0%
4	40.0%	40.0%	33.0%	28.0%	38.0%	35.8%
5	56.0%	61.0%	57.0%	51.0%	62.0%	57.4%
6	85.0%	86.0%	83.0%	81.0%	89.0%	84.8%
7	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%
Sample size	100	100	100	100	100	500

Table 5.15 Sum of actions still taken: cumulative percentages (HU_{DF})

Sum	Trial group 5	7	8	9	10	Overall
0	12.0%	7.0%	4.0%	8.0%	3.0%	6.8%
1	15.0%	11.0%	6.0%	11.0%	7.0%	10.0%
2	19.0%	16.0%	7.0%	12.0%	12.0%	13.2%
3	33.0%	32.0%	21.0%	19.0%	26.0%	26.2%
4	48.0%	46.0%	40.0%	32.0%	48.0%	42.8%
5	66.0%	71.0%	67.0%	65.0%	70.0%	67.8%
6	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%
Sample size	100	100	100	100	100	500

The findings from this section of the survey must be taken with some caution, given that:

- it is not always clear how respondents might have interpreted the questions;
- not all actions would have been available to all respondents (e.g. taking more showers is an option only if there is a shower);
- some actions need to be repeated (e.g. putting lids on cooking pans), others are one-off or infrequent changes (e.g. heating thermostat or timer settings);
- responses required recall of a two-year period;
- each question allows a wide range of frequency or intensity of action within the answer “yes”.

The overall pattern is more meaningful than any one question; the overall pattern tends to reinforce the findings for energy consumption, while adding to the doubt over the effects in the HU_{DF} stratum being genuine.

Awareness of billing changes

Only 66% of those who had been receiving monthly bills were aware of it. This should probably be compared with 81% of the other groups saying they received quarterly bills (84% quarterly or less frequently), rather than an optimistic 100%. Nevertheless, it limits the potential for the billing interventions to have any effect. The quality of data did not allow further exploration of the billing interventions.

Response to energy efficiency advice

Four trial groups received (with their bills) advice on reducing energy use. TG2 had advice only, TG6 also received a clip-on RTD. TG8 received a smart meter and monthly bills with historic feedback and TG9 in addition received a mains RTD displaying both gas and electricity data.

Overall response

There was a significant difference between trial groups in the percentage of respondents who recalled receiving the advice. With or without a smart meter, overall 57% of respondents recalled receiving the advice but there was a distinct difference between non-smart meter groups – 48% if an RTD was provided, 65% if not. It appears that one intervention may interfere with recall of another.

Among those who recalled receiving the advice, ratings of how useful it had been did not differ significantly between trial groups (p=0.890). Across all groups, 8.8% found the advice “Not at all useful”, 20.7% “Not very useful”, 50.9% “Quite useful” and 19.6% “Very useful”.

Response to specific aspects of the advice

Trial groups did not differ significantly in their rating of the usefulness of specific aspects of the advice. The percentage responding in each category and the overall ratings, across groups, are shown in Table 5.16. Advice on heating, lighting and appliances were most highly rated. Advice on saving energy while cooking was least highly rated. Hot water and insulation advice were intermediate.

Table 5.16 Ratings of the usefulness of different aspects of advice

Aspect of advice	*Not at all useful*	*Not very useful*	*Quite useful*	*Very useful*	*Mean rating[22] (sample size)*
Insulation (e.g. cavity wall and loft insulation, draught exclusion)	15.8%	18.6%	45.7%	19.9%	2.7 (543)
Heating (e.g. servicing boiler, turning thermostat down, fitting thermostatic controls)	11.2%	16.7%	51.7%	20.4%	2.8 (544)
Hot water (e.g. immersion heater, fitting cylinder jacket, advice on showers and baths)	12.8%	20.0%	48.9%	18.3%	2.7 (530)
Cooking (e.g. not overfilling kettles, putting lids on pans, using the microwave and grill)	16.5%	20.7%	44.9%	17.9%	2.6 (521)
Appliances (e.g. energy efficient ratings, filling the freezer)	11.9%	18.4%	48.6%	21.2%	2.7 (523)
Lighting (e.g. energy saving bulbs, nightlights)	10.4%	19.7%	48.1%	21.7%	2.8 (538)

These differences may be attributable to some combination of the advice itself (amount, clarity and quality), the household’s prior awareness of the information provided, and relevance to the household (whether the advice relates to actions the household has not already taken, can take and is prepared to consider taking). Interpretation of the findings is therefore problematic but they may be taken to indicate that:

- advice on heating, lighting and appliances is the easiest place to start;
- more work is needed on delivering advice relating to cooking (perhaps because cooking involves decisions that distract from, override or are unrelated to energy use).

[22] Mean ratings were calculated by assigning values of 1-4 to the ratings, the value increasing with usefulness.

Response to RTDs

Three trial groups were provided with an RTD – TG4 had a clip-on RTD only, TG6 had a clip-on RTD plus advice and TG9 has a mains RTD displaying both gas and electricity data, plus smart meter, advice and monthly bills with historic feedback. Responses from these three groups were compared.

Overall response

Only 48.0% of respondents in TG9 were aware they had an RTD but this was non-significantly better than TG4 (40.7%) and TG6 (42.0%).

Among those who knew they had an RTD, a similar percentage of each trial group looked at it every day (51-54%). However, the percentage who looked at it less often (rather than "Never" or "Don't know") increased with complexity of intervention: TG4=14.3%, TG6=27.4%, TG9=37.6%. Therefore, the percentage who ever looked at it increased with complexity of intervention: TG4=68.6%, TG6=79.5%, TG9=88.2%. This effect of trial group was significant ($p<0.001$). Multiplying these figures by the proportions who were even aware of the RTD, the overall percentages who ever looked at the RTD become: TG4=28%, TG6=33%, TG9=42%.

TG9 was also more likely to report finding the RTD useful ($p=0.013$) – see Table 5.17.

Table 5.17 Ratings of the usefulness of the RTD – percentage in each category

	Trial group			Total
	4	6	9	
Not at all useful	18.2%	14.5%	4.1%	9.4%
Not very useful	6.1%	7.2%	8.1%	7.5%
Quite useful	25.8%	37.7%	33.1%	32.6%
Very useful	50.0%	40.6%	54.7%	50.5%
Sample size	66	69	172	307

Response to specific aspects of the display

Respondents were asked about the usefulness of different aspects of the display, using the general form of question "How useful do you find the following aspects of your visual display, in helping you to manage your energy use – […]?" Because the RTDs were different, there were also differences in the questions between TG9 and TGs 4 and 6. These questions were asked only to those who said they were aware of the particular feature, so they should be taken as a measure of users' opinions, not the opinions of the whole sample. The view might be taken that those who did not know about the feature found it not useful; taking this into account would tend to increase the difference between aspects of the displays because those with the higher ratings also tended to have larger numbers of respondents.

Mean ratings were calculated by assigning values of 1-4 to the ratings, the value increasing with usefulness.

Ratings from TG9 are shown in Table 5.18. The "traffic lights" were rated the most useful feature, current electricity usage in kW the least. Intermediate were display of electricity usage in pence per hour, kWh over defined periods and gas consumption over defined periods. There was no difference in overall ratings of numeric and graphic display of information.

TG4 and TG6 did not differ in their ratings of each display feature so the combined ratings are shown in Table 5.19. Respondents rated cost and temperature information the most useful, the least useful being greenhouse gas emissions and the usage alarm feature. Intermediate were kW and humidity displays.

Table 5.18 Ratings of usefulness of display features (TG9)

Aspect of display	*Not at all useful*	*Not very useful*	*Quite useful*	*Very useful*	*Mean rating (sample size)*
Real time electricity usage in pence per hour	6.9%	17.0%	34.6%	41.5%	3.2 (159)
Real time electricity usage in units (kW)	8.8%	23.3%	39.0%	28.9%	3.0 (159)
Traffic light indicators	4.8%	9.6%	24.6%	61.1%	3.5 (167)
Information about electricity consumption across periods of time (e.g. per day/week/month/year)	6.9%	11.3%	38.8%	43.1%	3.2 (160)
Information about gas consumption across periods of time (e.g. per day/week/month/year)	7.9%	9.3%	38.6%	44.3%	3.2 (140)
Numerical energy use or cost information	9.1%	18.8%	39.0%	33.1%	3.0 (154)
Graphical energy use or cost information	5.3%	18.0%	42.0%	34.7%	3.0 (150)

Table 5.19 Ratings of usefulness of display features (TG4/6)

Aspect of display	*Not at all useful*	*Not very useful*	*Quite useful*	*Very useful*	*Mean rating (sample size)*
Real time electricity usage in pence per hour	16.9%	8.5%	35.4%	39.2%	2.8 (130)
Real time electricity usage in units (kW)	16.4%	11.7%	39.1%	32.8%	2.7 (128)
Display of greenhouse gas emissions	32.3%	26.0%	21.9%	19.8%	2.2 (96)
Temperature display	15.0%	9.2%	39.2%	36.7%	2.8 (120)
Humidity display	22.1%	15.0%	34.5%	28.3%	2.6 (113)
Alarm when cost per hour exceeds set limit	36.0%	12.8%	25.6%	25.6%	2.4 (86)

The two types of RTD can be compared on two aspects of the display – real time electricity usage in pence per hour and real time electricity usage in kW. In each case, respondents in TG9 gave the higher ratings. Another comparison that can be made is the highest rating each respondent gives to any aspect of the display. The concept behind this comparison is that, if one aspect is useful, the display is useful. Whereas 91% of TG9 found at least one aspect of the display useful or very useful (and 3% found no aspect useful), the equivalent figures for TG4/6 were 83% and 13%.

5.3 Summary of E.ON and AECOM findings

5.3.1 Demand reduction

Electricity

Non-smart meter groups

None of the non-smart meter interventions had a significant impact on electricity consumption.

Smart meter groups

The E.ON and AECOM findings did not differ dramatically but the AECOM analysis – allowing for the non-normal distribution of data and defining a rolling set of trial years – revealed some additional significant effects.

Significant reductions in consumption were seen in all high use dual fuel (HU_{DF}) groups and in all strata in TG9 except the not 'fuel poor' (NFP), for which there was a non-significant reduction.

Findings in the HU_{DF} groups are most likely to be due to regression to the mean rather than the effect of any particular intervention, compounded by a mismatch between pretrial consumption (which was far higher in the HU_{DF} trial groups than the HU_{DF} control groups).

Hence, the effect in TG9 (i.e. provision of an RTD in addition to advice and monthly bills) appears to be the key ingredient. Equally importantly, there were (mainly non-significant) increases in consumption in groups without an RTD. Since there was not a smart meter trial group with RTDs but without advice or monthly bills, it is uncertain whether RTDs would have had the same impact without these supporting interventions.

Gas

Non-smart meter groups

None of the non-smart meter interventions had a significant impact on gas consumption.

Smart meter groups

As for electricity, the E.ON and AECOM gas findings did not differ dramatically but the AECOM analysis – allowing for the non-normal distribution of data and defining a rolling set of trial years – revealed some additional significant effects.

Reductions in gas consumption were significant in every intervention involving a smart meter.

Although the HU_{DF} groups were defined by high electricity consumption, they also had high gas consumption – higher than other strata and higher in the HU_{DF} trial groups than the HU_{DF} control groups. Findings in the HU_{DF} groups are therefore most likely to be due to regression to the mean rather than the effect of any particular intervention.

Seasonal smoothing effects

The E.ON findings also demonstrated an important issue with comparing data between smart and non-smart meters. The smart meter groups show greater seasonal extremes than the non-smart control groups for both electricity and gas, with significantly lower consumption than the control group in almost every spring/summer quarter in every group, sometimes outweighing higher consumption in autumn/winter quarters.

This may be entirely due to “smoothing” of the non-smart data because of the need to interpolate infrequent readings over each quarter. If so, then analysis at sub-annual level would be valid only when comparing smart meter trials with smart-meter-only groups treated as control groups. In the current assessment, it is assumed that smoothing is a sufficient account of this seasonal variation but this needs to be better understood for any future trials.

5.3.2 Persistence of effects

Electricity

The significant reductions in consumption persisted through to in-trial year 1 plus 3 quarters. In the ‘fuel poor’ (FP) groups, savings if anything increased over time.

Gas

Reductions in gas consumption persisted to in-trial year 1 plus 1 quarter and, in some cases to in-trial year 1 plus 2 quarters (with a trend for some strengthening of the effect in in-trial year 1 plus 1 quarter). The only effect that persisted through to in-trial year 1 plus 3 quarters was in TG8 (smart meter, monthly bills and energy advice) for FP, which had the greatest median percentage difference in in-trial year 1.

5.3.3 Implications of the survey findings

Differences in survey response are consistent with the analysis of consumption data, the RTD being more likely to be used in TG9 than in the two non-smart meter groups (TG4 and TG6), and rated as more useful. Since the groups differed in both the type of RTD and the context of other interventions, the reasons for the differences in response cannot be stated with certainty. However, the differences in the RTDs themselves are the most likely explanation.

Reports of actions taken further indicate the benefit of RTDs but also suggest that energy advice was affecting behaviour by the end of the trial, perhaps too late to be seen in the consumption data. This was more distinct in the smart meter groups than in the non-smart meter groups. The additional bill data did not show any positive effect and may even have had a negative effect. Behavioural changes were more weakly evidenced in the HU_{DF} groups, again supporting the conclusion that changes in consumption in these groups were artefactual.

Across all the interventions, there was scope for more households to have become and remained aware of the information or technology provided, and to have engaged with it on a regular basis. Hence, there is potential for greater impact on consumption – with or without smart meters.

The survey also provides evidence relating to the optimum design of RTDs, supporting findings in the literature that cost data and simple graphics are key features. In contrast, CO_2 emissions and the usage alarm were not widely used or highly rated features. One surprise is perhaps the high ratings given to the temperature display of the clip-on RTD but this might have been particularly useful when interpreting advice about room temperature.

6 Analysis of Scottish Power Data

This section presents AECOM's analysis of the energy consumption data and the consumer survey data from Scottish Power. The summary at the end brings together findings from Scottish Power's analysis and AECOM's. Appendix B3 summarises the issues that need to be considered, related to the research design and execution, when interpreting or applying the findings.

6.1 Energy Consumption Data

6.1.1 Introduction

Scottish Power carried out a range of analyses to investigate the impact of its interventions on energy consumption. Those analyses were based on "Test Points" – 90-day periods at strategic points in the trial. In the course of reviewing the findings from all the suppliers, it became apparent that use of such a short comparison period is problematic, especially when comparing data from smart and non-smart meters. Therefore, our analysis uses as close to a full year of data as the trial design allows.[23]

The analysis by Scottish Power used the ratio of consumption in the trial period to consumption in the same months in the pretrial period whereas other suppliers used the difference in consumption between periods. The pros and cons of using the ratio were discussed with Scottish Power but we are convinced that it is appropriate to use the difference (and this approach has been seen universally in the literature review where baseline data were available). Scottish Power's reasons for using ratios were (a) this better represented differential opportunities to save energy among high and low users; (b) financial incentives were based on percentage and (c) the ease of achieving absolute savings levels would vary with season. However, using ratios means that saving a kWh is given different value for different customers and reductions from a high baseline are less likely to be detected. It is unlikely that customers would have thought about the ratio basis of the targets, especially since none asked Scottish Power how the target was calculated.

Furthermore, in our analysis a constant baseline period was maintained, using only pretrial data. Scottish Power had compared Phases 2 and 3 with Phase 1, which looks at cumulative rather than overall effects – this could fail to detect effects if they emerge in two stages. The Phases and trial groups are described in Section 3.3.

Scottish Power found no effect of any intervention in Phases 1 and 2 of its trial. In Phase 3, there were significant savings in all three groups of credit customers but in neither group of prepayment customers. We did not expect this to change dramatically in the new analysis and we therefore focussed on comparing trial groups with each other, rather than with the control groups. Credit and prepayment customers were included in separate analyses since the Scottish Power analysis had shown that they have quite different responses to the intervention. The main point of the AECOM analysis was to check directly whether adding an RTD to other levels of intervention had any additional impact on consumption.

Data were provided in three formats:

- raw energy consumption data;
- weather-adjusted energy consumption data: takes into account variations in weather conditions for the same day between different years;

---

[23] We would have preferred to analyse the change in energy consumption between pretrial and in-trial periods at different times of the year as the impact of the interventions may vary according to the time of year and time from start of intervention (and, in Phase 3, with the details of the intervention). However, the pretrial data are derived from non-smart meters and, as shown in Section 2.3, households were selected that only had a minimum of two readings per year. Given the variation in consumption between seasons, it was considered too inaccurate to use this information to interpolate monthly or even quarterly pretrial energy consumption levels for comparison with smart meter data.

- weather-adjusted and time-corrected energy consumption data: takes into account of weather variations both between years and for different days within a given year).

The weather- and time-corrected data, as used by Scottish Power, were used for the analysis. This makes some allowance for comparison over years that have different weather and for use of periods less than 12 months. Previous work by Ofgem, CSE and the energy suppliers had investigated the robustness of the correction factors and this was therefore not investigated further in this work. Data were provided in the form of quarterly aggregate figures – this was appropriate to the analysis periods used.

For the smart meter interventions, in the quarterly aggregates, it was appropriate to use the meter readings automatically transmitted at monthly intervals rather than using the more frequent (half-hourly) consumption data. The latter dataset typically has missing data in each quarter which need to be estimated to determine total consumption.

Outliers were excluded using the same approach as Scottish Power, i.e. changes greater than 50% between pretrial and in-trial period were excluded. Initial examination of the resulting distributions of the changes in corrected energy consumption showed that they were sufficiently close to a normal distribution (i.e. a bell shaped curve) for parametric statistical techniques to be used. A multivariate modelling strategy had earlier been planned but the introduction of survey variables would have unacceptably reduced the sample sizes and other possible confounders were not available in the database. Hence bivariate statistical methods were used.

The same baseline year was used for each Phase (Apr 06 – Mar 07) but different quarters within that year were used to match the quarters analysed in each Phase. Also, the households included in each baseline period were those for which data were available for the Phase being analysed. For both these reasons, the baseline figures differ between Phases. The means for the baseline period are shown in Table 6.1, along with each figure normalised to CG1=100, to make comparison easier. The trial groups are described in Section 3.3.2.

Table 6.1 Baseline consumption figures used for each Phase

	Electricity consumption						*Gas consumption*					
	kWh			*Normalised*			*MWh*			*Normalised*		
Phase:	*1*	*2*	*3*	*1*	*2*	*3*	*1*	*2*	*3*	*1*	*2*	*3*
TG1	4223	4486	4264	103	101	101	24.07	31.84	24.34	101	102	102
TG2	3818	4279	4019	93	97	95	23.11	29.57	23.03	97	95	97
TG3		4313	4088		97	97		29.99	23.81		96	100
CG1	4107	4431	4231	100	100	100	23.90	31.29	23.81	100	100	100
TG4		4356			98		21.25	31.14	21.69	89	100	91
TG5		4123			93		21.00	35.30	21.37	88	113	90
CG2		4481			101		19.45	25.82	19.37	81	83	81

6.1.2 Phase 1

Comparisons made

Phase 1 lasted for 12 months but there was a staggered start, the households that received an RTD being supplied with the equipment over the first two months. Therefore, for the analysis, the quarterly data were aggregated over a 9-month period (i.e. missing the first quarter). The periods compared were pretrial (Jul 06 – Mar 07) with in-trial (Jul 07 – Mar 08).

The questions posed were (a) has any of the interventions caused a change in consumption that differs from the applicable control group and (b) does an intervention with RTD cause a different effect to an intervention without RTD? Hence the two credit trial groups (TG1 with RTD, TG2 without RTD) were compared with each other and with the control group CG1, using analysis of variance (ANOVA) to determine whether the three groups differed from each other at all. Where the ANOVA showed a significant overall effect, it was followed up by using Tukey multiple comparison tests to determine which particular groups differ from each other. Similarly, for the prepayment groups, TG4, TG5 and CG2 were compared (using gas data only – insufficient data were available for electricity). Note that all of the groups in Phase 1 used a non-smart meter.

Results

Figures for changes from pretrial (mean change, sample size and standard deviation) are shown in Table 6.2. The effect of trial group on electricity consumption was not significant ($p>0.1$) for the credit groups. This means that the changes between the pretrial and in-trial period did not differ significantly between any pairs of groups tested.

The effect of trial group on gas consumption was significant ($p=0.002$) for the credit groups but not for prepayment groups ($p>0.1$). Multiple comparisons showed that the reduction in consumption was *greater* in the control group than in TG1 ($p=0.034$) and TG2 ($p=0.001$).

The analysis confirms the Scottish Power findings that the Phase 1 interventions did not result in any reduction in energy consumption. In the case of gas consumption, the significant effect shows the opposite of the expected difference, i.e. the reduction is greater in the control group.

These results are not accounted for by differences in baseline consumption: TG1 had baseline electricity and gas consumption slightly higher than CG1 whereas TG2 had slightly lower consumption. Both TG4 and TG5 had higher baseline consumption than CG2.

Table 6.2 Changes in energy consumption (kWh) between pretrial and Phase 1[24]

	Electricity			*Gas*		
Group	*Mean*	*N*	*S.D.*	*Mean*	*N*	*S.D.*
TG1	-101	204	620	-1207	194	2912
TG2	-60	217	482	-907	209	2911
CG1	-78	130	450	-2000	203	3605
TG4				-955	182	2687
TG5				-931	176	2650
CG2				-1453	168	3007

6.1.3 Phase 2

Comparisons made

Phase 2 lasted for 7 months but there was a staggered start: the trial group (TG1-5) households received a smart meter or had prepayment meters reconfigured over the first five months. Therefore, it was considered valid to analyse data for only one in-trial quarter (Jul 08 – Sept 08), using only households for which there are smart or reconfigured meter data throughout the quarter, and comparing with a pretrial period of Jul 06 – Sept 06. Because of the short period analysed, it was not considered appropriate to compare smart trial group data with non-smart

[24] A negative value means a reduction in consumption.

control group data. Although the trial groups had non-smart pretrial data, this was consistent across groups and therefore of lesser concern.

The question posed was whether an intervention with an RTD causes a different effect to an intervention without an RTD. Hence the two established credit trial groups (TG1 and TG2) were compared with each other using a t-test. TG3 was not used in the analysis because it was in the process of joining the trial during Phase 2 and therefore differed from TG2 in ways other than the provision of an RTD. Similarly, for the prepayment groups, TG4 and TG5 were compared. TG1 had an upgraded RTD at this stage whereas TG4 retained the existing device.

Results

The effect of trial group was not significant ($p>0.1$) for electricity or gas consumption, for either credit or prepayment groups. The analysis confirms the Scottish Power findings that the Phase 2 interventions did not result in any reduction in energy consumption, although the non-significant trend was the same in all four comparisons, i.e. the group with an RTD made savings relative to the group without. Figures for changes from baseline are shown in Table 6.3.

These results are not accounted for by differences in baseline consumption: TG1 had slightly higher baseline electricity and gas consumption than TG2. TG4 had slightly higher baseline electricity consumption than TG5. These differences would tend to make it more likely for the reduction to be higher in the groups with an RTD. Only for gas consumption in TG4 and TG5 is the trend reversed.

Table 6.3 Changes in energy consumption (kWh) between pretrial and Phase 2[25]

	Electricity			*Gas*		
Group	*Mean*	*N*	*S.D.*	*Mean*	*N*	*S.D.*
TG1	-103	167	1013	-7763	88	6427
TG2	-84	134	715	-6334	75	6964
TG4	9	36	866	-1844	127	5824
TG5	39	39	977	-1224	142	6795

6.1.4 Phase 3

Comparisons made

Phase 3 lasted for 15 months, which gave the opportunity to compare a full year of data with the available baseline year (Apr 06 – Mar 07), meaning that all trial and control groups could be included. The in-trial period selected was the whole of 2009, the period from the mid-term update in Wave 1 to the postcard update in Wave 4.

The questions posed were (a) does the series of financial incentives in Phase 3 result in a reduction in energy consumption, (b) does an intervention with an RTD cause a different effect to an intervention without an RTD and (c) does this depend on the prior experience of Phase 1? Hence the analysis compared credit trial groups TG1, TG2 and TG3 with each other and with CG1, using analysis of variance (ANOVA) with Tukey multiple comparison tests to determine which (if any) groups differ from each other. Similarly, for the prepayment groups, TG4, TG5 and CG2 were compared (using gas data only - insufficient data were available for electricity).

[25] A negative value means a reduction in consumption.

Results

Figures for changes from the pretrial period are shown in Table 6.4.

The effect of trial group on electricity consumption was not significant ($p>0.1$) for the credit groups. This means that the changes between pretrial and in-trial periods did not differ significantly between any pairs of groups. This conflicts with the Scottish Power findings that the reduction in electricity use was greater in the trial groups in three or four out of the five Test Points during Phase 3. This is most likely to be due to the loss of intervention effect that Scottish Power saw in the middle of Phase 3.[26] In any case, it emphasises the risk of looking at periods of less than a year and seeing temporary or artefactual effects.

The results are not accounted for by differences in baseline consumption, which is similar in all four groups (slightly higher in TG1, slightly lower in TG2 and TG3 than in the control group).

The effect of trial group on gas consumption was significant ($p=0.001$) for the credit groups but not for prepayment groups ($p>0.1$). Multiple comparisons showed that the reduction in consumption was *less* in the control group than in TG1 ($p<0.001$), TG2 ($p=0.031$) and TG3 ($p=0.006$) but TG1, TG2 and TG3 did not differ from each other ($p>0.1$). This is consistent with the Scottish Power findings, showing an overall saving of gas by all three credit trial groups but no significant additional effect of including an RTD. Concerns remain that this can be accounted for by the Hawthorne effect rather than a genuine effect of the interventions.

The significant effect in the credit groups is not accounted for by differences in baseline consumption: TG1 had slightly higher baseline electricity and gas consumption than the control group, TG2 had slightly lower consumption and TG3 the same. The non-significant effect in the prepayment groups are not accounted for by differences in baseline consumption, which is higher in the trial groups than the control group.

Table 6.4 Changes in energy consumption (kWh) between baseline and Phase 3[27]

	Electricity			*Gas*		
Group	*Mean*	*N*	*S.D.*	*Mean*	*N*	*S.D.*
TG1	-337	190	756	-3047	101	3343
TG2	-243	188	641	-2418	99	2755
TG3	-307	211	689	-2670	107	3463
CG1	-176	60	490	-987	61	3098
TG4				-1305	30	2572
TG5				-1596	43	3544
CG2				-1492	33	3905

[26] It might also arise from using a different baseline period or testing differences rather than ratios, but this did not change the other findings.

[27] A negative value means a reduction in consumption.

6.2 Survey data

6.2.1 Rationale for analyses carried out

Scottish Power carried out three main surveys, as described in Appendix A3. Here we report on the responses obtained to a selection of the questions asked. The reasons behind the choice of questions to analyse are listed below.

- Questions were chosen to be analysed where further investigations could usefully add to those analyses already undertaken by Scottish Power.
- Focus was put on understanding change in consumer behaviour and/or energy use.
- In order for the conclusions drawn from the analyses to be robust, the questions chosen should not be unduly open to interpretation by respondents (e.g. because of their wording or because the response choices given for the questions were ambiguous).
- It was important that questions chosen for analysis did not suffer from potential distortion by their position in the survey (e.g. because of questionnaire routing). Thus questions were only included where answers to earlier questions in the survey did not have the potential to distort responses.
- In order for effective analyses to be undertaken, sufficient numbers of respondents had to exist.

For the selected questions, the tables below show the responses given according to the trial group to which the respondent belonged. Trial groups are omitted from the analysis where the question did not apply to them (e.g. if the question was about RTDs and the group was not given RTDs).

Chi-square tests have been carried out to test whether or not the responses to the questions are independent of the trial group to which the respondents belonged. In some circumstances, response categories have been combined in order to have sufficient number of respondents in the cells of the table for the assumptions behind the chi-square test to be valid. Missing data (including responses such as “don’t know” and “refuse to answer”) have been omitted except where they have a particular relevant interpretation.

6.2.2 Results

Checking energy use

Respondents were asked how often someone in the household checks “how much gas and electricity or energy you use”. The frequency of checking does not itself represent intention to reduce energy use, so the response categories “frequently” and “occasionally” were combined and contrasted with “never” and “only when the bill comes in”. This also increased cell counts and thus allowed chi-square tests to be conducted. A follow-up question asks whether anything has happened to change how often checks are made. Together, these questions provide some insight into the impact of the trial interventions. The percentages checking frequently or occasionally are shown in Table 6.5.

Table 6.5 Percentage checking frequently or occasionally

	Wave 1		Wave 2		Wave 3	
Customer type:	*Credit*	*Prepay.*	*Credit*	*Prepay.*	*Credit*	*Prepay.*
Control group (CG1/CG2)	28	39	29	62	31	54
Trial group without RTD (TG2/TG5)	41	43	33	65	53	72
Original trial group with RTD (TG1/TG4)	47	63	72	82	70	70
New trial group with RTD (TG3)	-	-	59	-	71	-

In Wave 1, checking is more frequent if an RTD had been provided and among prepayment customers. The 41% figure for TG2 creates exceptions to this, but it appears to have been a temporary effect: the effect of RTD and customer type become more distinct in Wave 2 and the percentage of “checkers” has also increased, particularly in the prepayment groups. The difference between TG1 and TG3 also suggests that it may take a time to develop a habit of checking: this difference has disappeared in Wave 3 and checking in TG4 has also dropped back to the same level as TG1 and TG3, as trial participation “matures”. Checking in TG2 and TG5 (trial groups without RTDs) has also increased by Wave 3 (probably in response to the Green Challenge), narrowing the gap with RTD credit groups and eliminating the gap with the RTD prepayment group.

As shown in Table 6.6, in Wave 1, a higher percentage report something having happened if an RTD had been provided. In Wave 2, only TG3 stands out, having recently joined the trial. By Wave 3, the percentages for TG2 and TG5 have increased, perhaps meaning the Green Challenge was more salient to households that had not benefited from an RTD.

Table 6.6 Percentage reporting that something had happened to change the frequency of checking[28]

	Wave 1		Wave 2		Wave 3	
Customer type:	*Credit*	*Prepay.*	*Credit*	*Prepay.*	*Credit*	*Prepay.*
Control group (CG1/CG2)	5	10	27	20	15	20
Trial group without RTD (TG2/TG5)	17	11	28	19	44	30
Original trial group with RTD (TG1/TG4)	57	30	34	21	31	17
New trial group with RTD (TG3)	-	-	56	-	42	-

Response to RTD

Overall response

Recall of installation

In the short term, the fact that an RTD had been installed was recalled to a similar extent across trial groups – 87% in Wave 1 and, for the ecoMeter installation in Wave 2, 87% in TG1 and 94% in TG3. EcoMeter installation recall held up in Wave 3 (86% in TG1, 92% in TG3) but recall of the clip-on RTD fell to 74% in TG4. The clip-on and mains RTDS also differed in the percentage that were still fitted and working at Wave 3: 58% for the clip-on and 77% for the mains RTD.

---

[28] This usually meant *increasing* the frequency of checking.

Ease of use

Reported ease of changing RTD settings does not differ between groups. Overall, 32% found it difficult or very difficult, 60% easy or very easy and 9% neither. While the net response is positive, there is clearly scope for improvement.

Looking at the display

Respondents were asked how often they or anyone else in the household looked at the display. To allow statistical comparisons, the number of response categories has been reduced to three:

- daily (several times per day / once a day);
- less often (every 2-3 days / once a week / occasionally);
- rarely/uncertain/never (only when we get a bill / never / don't know / unsure / on receipt of a letter from Scottish Power, encouraging its use).

The percentages in each category are shown in Table 6.7. In TG4, the frequency of looking at the display increases markedly across waves; in TG1 it increases but not so obviously and in TG3 it decreases. This does not mirror the change in consumption in Phase 3, supporting the contention that the financial incentives were responsible, not the RTD.

Table 6.7 Percentage showing how often the household looked at the RTD

	TG1			*TG3*			*TG4*		
	W1	*W2*	*W3*	*W1*	*W2*	*W3*	*W1*	*W2*	*W3*
Daily	49	38	47	-	47	28	25	38	50
Less often	4	42	31	-	42	46	22	22	39
Rarely/uncertain/never	47	21	22	-	11	26	54	40	11

Specific RTD functions

Recall of functions

In Wave 1, both TG1 (credit) and TG4 (prepayment) had the same type of clip-on RTD and participants in each group had similar recall of specific functions of the device. The display feature most often recalled first was consumption in units (47%), followed by consumption in cost (29%). When prompted further, recall rose and the figures are shown in Table 6.8.

TG4 retained the same RTD in Wave 2 and the figures for percentage recall of features kept to a similar pattern, although at a lower overall level. With the higher specification RTD in Wave 2, TG3 had slightly higher recall of functions than TG1; this is perhaps because TG3 was getting an RTD for the first time and suffered less from confusion with the clip-on device. TG1 and TG3 had a similar pattern of recall, units of electricity again being the most frequent first mentioned (61%). Once first mentions and all other mentions were taken into account, the percentages of respondents recalling each function were as in Table 6.8.

In Wave 3, TG4 was not asked the question. Responses did not differ significantly between TG1 and TG3 and the combined percentages are shown in Table 6.8. The figures kept to a similar pattern, although at a lower overall level, as Wave 2.

Table 6.8 Percentage who recall specific functions of the RTD

	Wave 1	*Wave 2*			*Wave 3*
	TG1/4	*TG4*	*TG1*	*TG3*	*TG1/3*
Electricity (units)	80	72	76	73	59
Electricity (cost)	71	46	59	63	52
Gas (units)	-	-	23	38	26
Gas (cost)	-	-	-	-	22
CO_2 emissions	32	9	16	16	17
'Traffic lights'	-	-	27	48	33
Usage alarm	29	-	-	-	-
Temperature/humidity	53	14	7	9	9

Overall, electricity units and cost are clearly best recalled. Gas usage appears less salient, perhaps because there are fewer gas appliances and they are less relevant for regular checking. Awareness of the CO_2 function is low and falling. The usage alarm has a similar initial level of awareness to CO_2 but is not recorded in the database in Waves 2 and 3. The traffic lights function also has a low level of awareness but this may be an artefact of the question format since the traffic lights reflect electricity usage rather than being an entirely separate function. The high initial awareness of the temperature function is interesting – it may be that it is salient initially, as people consider adjusting their thermostat but its usefulness is then largely exhausted. There is a decline in recall of other functions too, but not as great a decline as for temperature.

Use of functions

Respondents reported how frequently they used each function of their RTD: frequently (F), occasionally (O) or never (N). At Wave 1, TG1 and TG4 did not differ in their reports of using any individual function. The most frequently used were electricity units (F=42, O=20) and cost (F=38, O=20), followed by the temperature and humidity function (F=22, O=17). Other functions (greenhouse gas emissions and the usage alarm) had percentages below 20 for frequently and occasionally combined.

Because the question at Wave 2 was asked only if the respondent recalled the function, in most cases there are insufficient responses to compare trial groups or display functions. A new variable was therefore created, which was the frequency of the most frequently used function. Waves 1 and 2 are compared in Table 6.9, using this variable. Frequency of use had dropped in Wave 2 and was lower in both waves among prepayment customers using the simpler clip-on RTD.

Table 6.9 Frequency of use of any function of the RTD[29]

	TG1			*TG3*			*TG4*		
	F	*O*	*N*	*F*	*O*	*N*	*F*	*O*	*N*
Wave 1	59	16	26				43	25	32
Wave 2	41	47	13	43	45	12	39	29	32

[29] Percentages of respondents who used any one (or more) of the display functions frequently, occasionally or never.

Relevance of functions

Respondents rated the relevance of each RTD function on a scale of 1 to 5, where 1 means not at all relevant and 5 means very relevant. Table 6.10 shows mean ratings for each trial group and survey Wave. Where differences are reported below, they are significant differences as determined by chi-square tests on the frequencies of each response category.

In Wave 1, TG1 and TG4 do not differ in their ratings of the relevance of each function. Display of electricity usage (in either kW or cost) is rated more highly than CO_2 or the alarm function. Ratings of the temperature display are not recorded.

In Wave 2, the pattern is the same for TG4 (temperature is now reported and falls in the middle of the rank order). In TG1 and TG3, ratings are similar for all functions (electricity and gas usage in units and cost, and the traffic light indicator) except temperature and the CO_2 display, which had lower ratings.

In Wave 3, the same question is asked but its context has changed, making it less directly related to saving energy. In addition, the traffic lights are not described so specifically, being included in a single question about “the alarm or warning light”. Hence, comparisons are less meaningful but temperature moves up the rank order (to joint top with electricity cost in TG4). CO_2 remains at lowest rank.

There is little overall difference between trial groups but a trend for ratings to be highest in TG1 in Waves 1 and 2. The highest rating given to any feature was also analysed on the basis that, if one feature is relevant, the RTD is relevant. This also showed a trend for higher ratings in TG1, which is almost eliminated by Wave 3. Trends over time can largely be accounted for by the introduction of the mains RTD in Wave 2 but there is also a small trend for ratings to improve where the RTD remains the same.

Table 6.10 Mean ratings of relevance of each RTD function

<table>
<tr><th rowspan="2"></th><th colspan="2">Wave 1</th><th colspan="3">Wave 2</th><th colspan="3">Wave 3</th></tr>
<tr><th>TG1</th><th>TG4</th><th>TG4</th><th>TG1</th><th>TG3</th><th>TG4</th><th>TG1</th><th>TG3</th></tr>
<tr><td>Electricity (units)</td><td>4.0</td><td>3.7</td><td>4.1</td><td>4.2</td><td>3.9</td><td>3.7</td><td>4.3</td><td>4.1</td></tr>
<tr><td>Electricity (cost)</td><td>4.0</td><td>4.0</td><td>4.2</td><td>4.4</td><td>4.0</td><td>4.2</td><td>4.4</td><td>4.3</td></tr>
<tr><td>Gas (units)</td><td>-</td><td>-</td><td>-</td><td>4.2</td><td>3.9</td><td>-</td><td>-</td><td>-</td></tr>
<tr><td>Gas (cost)</td><td>-</td><td>-</td><td>-</td><td>4.2</td><td>3.9</td><td>-</td><td>-</td><td>-</td></tr>
<tr><td>CO_2 emissions</td><td>3.3</td><td>3.1</td><td>3.1</td><td>3.8</td><td>3.4</td><td>3.6</td><td>3.5</td><td>3.7</td></tr>
<tr><td>‘Traffic lights’</td><td>-</td><td>-</td><td>-</td><td>4.4</td><td>4.1</td><td rowspan="2">3.5</td><td rowspan="2">4.2</td><td rowspan="2">4.2</td></tr>
<tr><td>Usage alarm</td><td>3.5</td><td>3.3</td><td>3.3</td><td>-</td><td>-</td></tr>
<tr><td>Temperature/humidity</td><td>-</td><td>-</td><td>3.7</td><td>3.9</td><td>3.5</td><td>4.2</td><td>3.8</td><td>3.9</td></tr>
<tr><td>Highest rating of any feature</td><td>4.3</td><td>4.0</td><td>4.3</td><td>4.5</td><td>4.3</td><td>4.5</td><td>4.5</td><td>4.4</td></tr>
</table>

6.3 Summary of Scottish Power and AECOM findings

6.3.1 Demand reduction

Electricity

Non-smart meter groups

None of the non-smart meter interventions had a significant effect on electricity consumption – this applies to all the findings for Phase 1 and the prepayment customer findings in Phases 2 and 3. Both Scottish Power and AECOM found this.

Smart meter groups

The smart meter interventions had no effect in Phase 2, i.e. smart meter plus letters giving time of day breakdown of consumption, with or without an RTD being provided. Both Scottish Power and AECOM found this.

In Phase 3, Scottish Power found significant reductions in all three credit customer groups (TG1 and TG3 with an RTD, TG2 without) during three of the four Waves of the Green Challenge financial incentive, each of which lasted three months. These three Waves had an incentive for electricity consumption reduction only, the other Wave having an additional gas-related incentive. AECOM found no effect over the central 12 months of Phase 3 as a whole although the trend was in the same direction as that observed by Scottish Power. The Scottish Power findings relate to each individual Wave and therefore give a more fine grain analysis but the control group and baseline non-smart meter data are less reliable over the shorter period. On balance, it is reasonable to concur with the Scottish Power conclusion that there was a temporary effect of the three Waves that focused on reducing electricity consumption. The major concern with the finding is that it could be due to the Hawthorne effect rather than being a true effect of the intervention. However, the finding shows as a minimum that the credit households had the capacity to reduce consumption if the motivation to do so was sufficient.

It is also debatable whether these are best viewed as smart meter interventions, given the efforts made to prevent customers from knowing they had a smart meter. They are better seen as tests of RTDs and financial incentives: although the smart meter technology was being tested, the experience of installation and understanding of the meter were not. At no point did the energy consumption of the groups with RTDs differ significantly from the groups without RTDs.

Gas

Non-smart meter groups

None of the non-smart meter interventions had a significant effect on gas consumption – this applies to all the findings for Phase 1 and the prepayment customer findings in Phase 2 and 3. Both Scottish Power and AECOM found this. The only exception was that AECOM found that consumption by credit customers decreased *more* in the control groups than in the trial groups in Phase 1.

Smart meter groups

The smart meter interventions had no effect in Phase 2, i.e. smart meter plus letters giving time of day breakdown of consumption, with or without an RTD being provided. Both Scottish Power and AECOM found this.

In Phase 3, Scottish Power found significant reductions in all three credit customer groups during the third of four Waves of the Green Challenge financial incentive, each of which lasted three months, and during a three-month break in use of incentives. The third Wave was the only one that had a gas-related incentive. AECOM found a significant reduction over the central 12 months of Phase 3 as a whole. The Scottish Power findings relate to each individual Wave and therefore give a more fine grain analysis but the control group and baseline non-smart meter data are less reliable over the shorter period. On balance, it is reasonable to concur with the Scottish Power conclusion that there was a temporary effect of the Wave that had a gas-related incentive, but also the non-incentivised period preceding it. As for electricity, the major concern is that the finding could be due to the Hawthorne effect rather than being a true effect of the intervention.

As with electricity, the interventions are better seen as tests of RTDs and financial incentives than tests of smart meters. At no point did the groups with RTDs differ significantly from the groups without RTDs.

6.3.2 Implications of the survey findings

Energy-saving measures

At both Wave 2 and Wave 3 surveys (representing Phases 2 and 3), there was no significant difference between groups in the percentage reporting that they had taken new measures to save energy (or increased an activity) during the previous 12 months. There was also no difference in the mean number of new or increased activities. There was a tendency for the control groups to have taken actions earlier in the trial than the trial groups, resulting in similar numbers overall by the end of the trial. This is in keeping with the generally null or negative effect of the interventions on energy consumption.

Advice

Although 37% said the mailings were telling them nothing new, 66% would like to receive more of the same information, and qualitative research indicated that people wanted to be kept regularly informed on energy matters and relating to energy use (in a personal rather than generic way), as a reminder. The implication is that repetition of advice is not seen as a problem.

Use of RTDs

The clip-on RTD offered in Phase 1 was not accepted by 30% of credit customers and 22% of prepayment customers. Of those installed in the initial credit trial group (TG1), only 42% were still operational at Phase 2, when installers fitted smart meters. In the prepayment group, 58% of survey respondents said they were still fitted and working during the Wave 3 interventions (towards the end of the study). Overall, this limits the potential of the RTD to have any effect on consumption.

The main-powered RTD fitted in Phase 2 for credit customers fared better, 77% being reported as fitted and working at Wave 3. Recall of the RTD being fitted was also higher in Wave 3 for the mains device (86-92%) than for the clip-on (74%). Over the three Phases and three trial groups, between 11% and 54% of those who knew they had the RTD rarely or never checked it. Figures were generally better for the mains device (11-26%) than the clip-on (11-54%) and general improved through the trial (47-54% in Wave 1, 11-26% in Wave 3). Nevertheless, overall there was limited potential for the RTD to affect consumption. Part of the problem appears to have been that 32% of users (across all groups that had an RTD) found it difficult or very difficult to change the settings.

A more positive point is that RTDs did appear to make a difference to customers checking their energy use. Initially, prepayment customers were more likely to be checking their energy use, other than when they receive a bill. In the course of the trial, credit customers who received an RTD came up to the same level of checking as prepayment customers. In addition, receiving an RTD was associated with participants noting that something had happened to change their frequency of checking. This effect was seen, even though not all recipients used (or even accepted) the RTD.

RTD features

Not all the display features were equally noticed, used or valued. Units of electricity was the feature most likely to be mentioned first, followed by cost of electricity but this is likely to be because of the set-up of the devices. In questions about the use or relevance of the display features, electricity cost was seen as more important. On all measures, CO_2 emissions and (when available) gas consumption and an audible consumption alarm scored poorly. Temperature display was rated more highly than might have been anticipated, perhaps most so at the start, as a one-off response to advice about thermostat settings. The overall “winner” was the traffic lights display on the mains RTD, providing a simple visual signal about consumption rate.

Intervention communications

Customer mailings were an important element in the Scottish Power trial and their quality was monitored and improved through the course of the trial. Recall and comprehension during Phase 3 was the highest achieved in the project (and similar among credit and prepayment customers – the latter had recalled the material less well in earlier Waves). It was these later letters that offered financial benefits so it may be this, rather than changes in style, that was responsible. Self-rated comprehension of various aspects of the mailings did not differ between trial groups.

7 Analysis of SSE Data

This section presents AECOM's analysis of SSE's energy consumption data and consumer survey data. The summary at the end brings together findings from SSE's analysis and AECOM's analysis. Appendix B4 summarises the issues that need to be considered, related to the research design and execution, when interpreting or applying the findings.

7.1 Energy Consumption Data

7.1.1 Introduction

The analysis reported by SSE is comprehensive in the sense that it covers all interventions and revealed a coherent set of significant effects. However, because of the unbalanced factorial design, some of the effects are difficult to interpret. The analysis presented here seeks to provide further insight into the findings, particularly the two incentive interventions (incentive to reduce, incentive to shift), which could be examined only in the smart meter groups.

Data preparation was different to that of the other three energy suppliers.

- For SSE, we used a data-set provided directly by SSE's analysis team. This was the same data-set used by SSE in its own analysis.
- For the other suppliers, the data-set was provided by the Centre of Sustainable Energy (CSE). Whilst the data should have been the same as used in the energy supplier's own analysis, there will have been some differences in the way that the data were prepared for analysis (e.g. selection of households with valid data, choice of pretrial and in-trial data).

SSE used weather-adjusted and time-corrected electricity and gas consumption data to take account of weather variations, both between years and for different days within a given year. Previous work by Ofgem, CSE and the energy suppliers had investigated the robustness of the correction factors and this was therefore not investigated in this work.

The analysis of changes in energy consumption was based on manual readings for non-smart meters (in-trial data for non-smart meter trial and control groups and all pretrial data). At least two meter readings were required for each household during the pretrial period and two meter readings during the trial. For smart meters, the energy consumption during the trial was based on automatically transmitted meter readings (typically sent at daily intervals).

Households were removed where the difference between pretrial and in-trial data was:

- greater than 5,000 kWh for electricity consumption analysis;
- greater than 10,000 kWh for gas consumption analysis.

The time of use analysis was based on electricity smart meter readings only. It was necessary to do this analysis on the half-hourly consumption data to be able to determine the consumption at different times of day. Where half-hourly data are missing, an estimate is needed for the consumption during the missing period. An acceptable threshold was judged to be when less than 10% of consumption data were missing.

7.1.2 Electricity demand reduction

Main analysis

The main factorial analysis was undertaken to investigate further the effects of the different components of the smart meter interventions. This included only trial groups 7 to 14 and 19 to 26 as they formed a balanced factorial design (all combinations of four of the interventions: incentive to shift, incentive to reduce, RTD and web information) for households that additionally all had a smart meter, energy advice booklet and monthly bills with consumption history information. In total, 3308 households were included.

The dependent variable was the change in consumption (pretrial to in-trial) for each household. A number of ‘confounder’ variables were added to the model to account for variability in the response due to factors other than the four interventions:

- region within England, Scotland or Wales (16 regions in total);
- postcode (first 2 letters of postcode);
- grid supply point (gsp);
- Mosaic demographic classification group;
- awareness of the trial (Aware/Unaware/Committed – see Section 3.4);
- fuel type (electric smart meter only, gas smart meter only, electric and gas smart meters).[30]

Table 7.1 shows a summary of the impact of the smart meter intervention measures. No main effects of interventions were significant at the 5% level but Web information was significant at the 10% level. The trend, however, was for less reduction in consumption where information was provided using the web. No two-way interactions between the four interventions were found to be significant at the 10% level once adjusted for all other effects in the model. The model summary is presented in Table 7.2. The adjusted means (least square means) for the main effects are presented in Table 7.3.

Table 7.1 Summary of changes in the mean by trial group

	Interventions active in the trial group								
Trial group	*Incentive to shift*	*Incentive to reduce*	*RTD*	*Web information*	*Number of households*	*Pretrial mean (kWh)*	*In-trial mean (kWh)*	*Change in mean (kWh)*	*Percentage savings (%)*
11					410	3883	3592	-292	7.5
12	✓				180	3942	3657	-285	7.2
13		✓			184	4062	3822	-241	5.9
14	✓	✓			193	3874	3553	-321	8.3
23			✓		251	3920	3555	-365	9.3
24	✓		✓		135	3971	3680	-291	7.3
25		✓	✓		161	4117	3959	-457	11.1
26	✓	✓	✓		139	4014	3790	-225	5.6
10				✓	349	4104	3889	-216	5.3
7	✓			✓	196	4168	3895	-274	6.6
8		✓		✓	195	3805	3528	-277	7.3
9	✓	✓		✓	195	3907	3647	-260	6.7
22			✓	✓	254	3954	3777	-177	4.5
19	✓		✓	✓	153	4078	3743	-334	8.2
20		✓	✓	✓	152	3889	3542	-347	8.9
21	✓	✓	✓	✓	161	3949	3633	-316	8.0

[30] We had wished this variable to reflect whether households had an electricity smart meter and/or a gas smart meter. However, this was not possible with the data-set available. Instead this variable classifies households into electricity-only, gas only, or dual fuel according to the data that remained after cleaning. The electricity-only and dual fuel cases were included in the analysis.

Table 7.2 Model summary

Source	*DF*	*Sum of squares*	*Mean square*	*F*	*p*
Region	4	2519596	629899	0.86	0.4889
Postcode	7	11956105	1708015	2.32	0.0230*
Gsp	1	2115648	2115648	2.88	0.0898~
Mosaic	10	13219704	1321970	1.80	0.0556~
Awareness	2	2361795	1180897	1.61	0.2007
Fuel type	1	140348	140348	0.19	0.6621
Incentive to shift	1	342408	342408	0.47	0.4949
Incentive to reduce	1	175588	175588	0.24	0.6250
Web information	1	2596008	2596008	3.53	0.0603~
RTD	1	909633	909633	1.24	0.2660

*p<0.05, ~p<0.10

Table 7.3 Adjusted mean change for main effects of interest

	Mean	*Standard error*
No incentive to shift	-537	96.6
Incentive to shift	-514	97.5
No incentive to reduce	-517	95.6
Incentive to reduce	-533	98.5
No RTD	-554	97.1
RTD	-497	96.5
No web information	-508	96.6
Web information	-542	97.1

The effect of postcode was significant at the 5% level but Region and Grid supply point were not, indicating a geographic variation that was not systematic as to which area of the country was represented. The effect of Mosaic group was marginally significant (i.e. at the 10% level), indicating some demographic variation in electricity savings but this variation was independent of trial interventions.

The non-significant trends were for reductions in consumption to be greater where the incentive to reduce was provided, but smaller where an RTD or the incentive to shift was provided. Percentage savings in each cell of the factorial design are shown in Table 7.4. The effect of the incentive to reduce appears to have been more positive in the absence of the incentive to shift, except where there is no feedback via either an RTD or the web. This kind of complex interaction, suggesting interference between interventions, could not be detected by the analysis. The RTD effect was positive in 6 out of 8 cells but this was not sufficient to create an overall positive effect.

Table 7.4 Percentage savings in each cell of the analysis

Incentive to reduce	Incentive to shift	RTD		No RTD	
		Web information	No web information	Web information	No web information
Yes	Yes	8.0	5.6	6.7	8.3
	No	8.9	11.1	7.3	5.9
No	Yes	8.2	7.3	6.6	7.2
	No	4.5	9.3	5.3	7.5

This main analysis could not include all the trial cells. Therefore, a series of supplementary analyses was conducted. Note that each set of results relates only to the analysis of that specific set of trial groups; no comparisons can be made across sets.

Effect of smart meter – with and without RTD

This comprised two sets of trial groups:

- with no other intervention [TGs 15 & 27 vs 32 & 3];
- with booklet and graphs on bills [TGs 11 & 23 vs 2 & 5].

Table 7.5 shows the interventions for these trial groups. There was only one household in TG15, and so this group was deleted from the analysis. The results presented here therefore relate only to TGs 32 (Control group), 2, 3, 5, 11 and 27. In total there were 8875 households in this set.

Table 7.5 Summary of trial groups

Trial group	*Smart meter*	*Booklet*	*Quarterly bills with graphs*	*Monthly bills with graphs*	*Incentive to shift*	*Incentive to reduce*	*RTD*	*Web information*
32								
3							✓	
5		✓	✓					
2		✓	✓				✓	
15	✓							
27	✓						✓	
11	✓	✓		✓				
23	✓	✓		✓			✓	

The results are presented in Tables 7.6 - 7.9. This analysis means that groups with a smart meter had greater reductions in consumption than groups without a smart meter if they also had the advice booklet and additional data on their bills (monthly for smart meter groups, quarterly for others) – with or without an RTD. The difference made by the smart meter was marginal if the advice booklet and additional data on bills were not provided, even though an RTD was provided.

Table 7.6 Summary of changes in the mean by trial group

Trial group	*Number of households*	*Pretrial mean (kWh)*	*In-trial mean (kWh)*	*Change in mean (kWh)*	*Percentage savings (%)*
32	3502	3915	3749	-166	4.23
3	1203	4016	3829	-187	4.66
5	1939	3855	3692	-163	4.24
2	1297	4027	3842	-184	4.58
27	273	3879	3567	-312	8.03
11	410	3883	3592	-292	7.51
23	251	3920	3555	-365	9.31

Table 7.7 Model summary

Source	*DF*	*Sum of squares*	*Mean square*	*F*	*p*
Region	14	10720107	765722	1.18	0.2865
Postcode	143	92705809	648292	1.00	0.5005
Gsp	12	7439311	619943	0.95	0.4933
Mosaic	11	15245499	1385954	2.13	0.0156
Awareness	2	2212261	1106131	1.70	0.1831
Fuel type	1	2043805	2043805	3.14	0.0765
Trial group	6	10822609	1803768	2.77	0.0109

Table 7.8 Adjusted mean change

Trial group	*Mean change*	*Standard error*
32	-141	28.0
3	-165	32.3
5	-113	29.4
2	-163	32.0
27	-266	59.5
11	-257	54.4
23	-323	62.3

Table 7.9 Specific comparisons of interest

Trial groups	*Estimate*[31]	*Standard error*	*t*	*p*
3 vs 27	-101	58.8	1.72	0.0858
2&5 vs 11&23	-152	45.9	3.31	0.0009

Effect of booklet & graphs with smart meter

This comprised two sets of trial groups:

- effect of booklet with graphs and incentives [TGs 12-14 vs 16-18];
- effect of booklet and graphs with RTD and incentives [TGs 23-26 vs 27-30].

Table 7.10 shows the interventions for these trial groups. The results are presented in Tables 7.11 - 7.14. This analysis means that the advice booklet and additional graphical data on monthly bills did not affect consumption in the smart meter groups. In total there were 2511 households in this set.

[31] This refers to the estimate of difference between the trial groups, based on the statistical model. Refer to the text for interpretation of positive and negative values.

Capabilities on project:
Building Engineering

Table 7.10 Summary of trial groups

Trial group	*Smart meter*	*Booklet*	*Quarterly bills with graphs*	*Monthly bills with graphs*	*Incentive to shift*	*Incentive to reduce*	*RTD*	*Web information*
16	✓			✓	✓			
17	✓			✓		✓		
18	✓			✓	✓	✓		
27	✓						✓	
28	✓				✓		✓	
29	✓					✓	✓	
30	✓				✓	✓	✓	
12	✓	✓		✓	✓			
13	✓	✓		✓		✓		
14	✓	✓		✓	✓	✓		
23	✓	✓		✓			✓	
24	✓	✓		✓	✓		✓	
25	✓	✓		✓		✓	✓	
26	✓	✓		✓	✓	✓	✓	

Table 7.11 Summary of changes in the mean by trial group

Trial group	*Number of households*	*Pretrial mean (kWh)*	*In-trial mean (kWh)*	*Change in mean (kWh)*	*Percent savings (%)*
16	176	3970	3771	-199	5.01
17	188	4002	3612	-391	9.76
18	184	1076	3676	-401	9.83
27	273	3879	3567	-312	8.03
28	161	4271	3843	-428	10.02
29	138	3924	3613	-311	7.93
30	148	3994	3595	-399	9.99
12	180	3942	3657	-285	7.24
13	184	4062	3822	-241	5.93
14	193	3874	3553	-321	8.29
23	251	3920	3555	-365	9.31
24	135	3971	3680	-291	7.33
25	161	4117	3659	-457	11.11
26	139	4014	3790	-225	5.60

Table 7.12 Model summary

Source	*DF*	*Sum of squares*	*Mean square*	*F*	*p*
Region	4	1410911	352728	0.52	0.7221
Postcode	5	7261111	1452222	2.14	0.0586
Gsp	1	768912	768912	1.13	0.2878
Mosaic	11	21360085	21360085	2.86	0.0010
Awareness	2	791958	791958	0.58	0.5587
Fuel type	1	900738	900738	1.32	0.2499
Trial	13	13350657	13350657	1.51	0.1058

Table 7.13 Adjusted mean change

Trial group	*Mean*	*Standard error*
16	-227	150
17	-422	149
18	-418	150
27	-316	141
28	-422	150
29	-345	154
30	-429	152
12	-303	147
13	-253	150
14	-352	148
23	-363	144
24	-318	154
25	-476	151
26	-230	154

Table 7.14 Specific comparisons of interest

Trial groups	*Estimate*	*Standard error*	*t*	*p*
12-14 vs 16-18	-53.3	49.9	1.07	0.2858
23-26 vs 27-30	-31.6	45.7	0.69	0.4891
12-14 & 16-18 vs 23-30	-33.3	34.8	0.96	0.3383

Effect of booklet and RTD, without smart meter

This comprised one set of trial groups, with factorial analysis [TGs 32, 1, 3 & 4]. Table 7.15 shows the interventions for these trial groups. The results are shown in Tables 7.16 to 7.18. This analysis means that the advice booklet did not affect consumption in the non-smart meter groups but the RTD had a small (1%) and marginally significant effect ($p<0.1$), which did not depend on the availability of the booklet. The results presented here are based on 9406 households.

Capabilities on project:
Building Engineering

Table 7.15 Summary of trial groups

Trial Group	Smart meter	Booklet	Quarterly bills with graphs	Monthly bills with graphs	Incentive to shift	Incentive to reduce	RTD	Web information
32								
4		✓						
3							✓	
1		✓					✓	

Table 7.16 Summary of changes in the mean by trial group

Trial group	Number of households	Pretrial mean (kWh)	In-trial mean (kWh)	Change in mean (kWh)	Percent savings (%)
32	3502	3915	3749	-166	4.23
4	3455	3888	3719	-169	4.35
3	1203	4016	3829	-187	4.66
1	1246	4073	3846	-228	5.60

Table 7.17 Model summary

Source	DF	Sum of squares	Mean square	F	p
Region	14	9701938	692996	1.09	0.3584
Postcode	146	103792692	710909	1.12	0.1531
Gsp	12	3897843	324820	0.51	0.9085
Mosaic	11	6179408	561764	0.89	0.5537
Awareness	2	2589152	1294576	2.04	0.1299
Fuel type	1	3990408	3990408	6.29	0.0121
Booklet	1	6759	6759	0.01	0.9178
RTD	1	1969239	1969239	3.10	0.0781
Booklet*RTD	1	430046	430046	0.68	0.4103

Table 7.18 Adjusted mean change

Booklet	RTD	Mean	Standard error
No	-	-200	24.5
Yes	-	-205	23.4
-	No	-184	23.3
-	Yes	-220	25.0
No	No	-189	26.3
No	Yes	-210	30.7
Yes	No	-179	24.7
Yes	Yes	-231	29.5

Effect of graphs and RTD, added to booklet, without smart meter

This comprised one set of trial groups, with factorial analysis [TGs 1, 2, 4 & 5]. Table 7.19 shows the interventions for these trial groups. The results are shown in Tables 7.20 to 7.22. This analysis means that the additional data on bills did not affect consumption in the non-smart meter groups that had the booklet. However, the RTD did have a small (1%) significant effect (p<0.05), which did not depend on the availability of the graphs. The results presented here are based on 7937 households.

Table 7.19 Summary of trial groups

Trial group	*Smart meter*	*Booklet*	*Quarterly bills with graphs*	*Monthly bills with graphs*	*Incentive to shift*	*Incentive to reduce*	*RTD*	*Web information*
4		✓						
1		✓					✓	
5		✓	✓					
2		✓	✓				✓	

Table 7.20 Summary of changes in the mean by trial group

Trial group	*Number of households*	*Pretrial mean (kWh)*	*In-trial mean (kWh)*	*Change in mean (kWh)*	*Percent savings (%)*
4	3455	3888	3719	-169	4.35
1	1246	4073	3846	-228	5.60
5	1939	3855	3692	-163	4.24
2	1297	4027	3842	-184	4.58

Table 7.21 Model summary

Source	*DF*	*Sum of squares*	*Mean square*	*F*	*p*
Region	14	13783485	984535	1.59	0.0741
Postcode	140	76704777	547891	0.88	0.8322
Gsp	12	5424759	452063	0.73	0.7238
Mosaic	11	8935472	812315	1.31	0.2110
Awareness	2	4250856	2125428	3.43	0.0325
Fuel type	1	1610699	1610699	2.60	1.1070
Graphs	1	1302563	1302563	2.10	0.1472
RTD	1	3434975	3434975	5.54	0.0186
Graphs*RTD	1	26664	26664	0.04	0.8357

Table 7.22 Adjusted mean change

Graphs	*RTD*	*Mean*	*Standard error*
No	-	-220	23.7
Yes	-	-194	25.3
-	No	-183	23.7
-	Yes	-230	25.3
No	No	-198	25.0
No	Yes	-241	29.7
Yes	No	-168	28.0
Yes	Yes	-220	30.5

Effect of benchmarking

This comprised one set of trial groups, with factorial analysis [TGs 5 & 6]. Table 7.23 shows the interventions for these trial groups. The results are shown in Tables 7.24 to 7.26. This analysis means that the addition of benchmarking comparisons on bills in the non-smart meter groups that had the booklet but no other intervention did have a significant effect ($p<0.05$) but the effect was small (savings of about 1%). The results presented here are based on 3841 households.

Table 7.23 Summary of trial groups

Trial group	*Smart meter*	*Booklet*	*Quarterly bills with graphs*	*Monthly bills with graphs*	*Incentive to shift*	*Incentive to reduce*	*RTD*	*Web information*
5		✓	✓					
6		✓	✓[32]					

Table 7.24 Summary of changes in the mean by trial group

Trial group	*Number of households*	*Pretrial mean (kWh)*	*In-trial mean (kWh)*	*Change in mean (kWh)*	*Percent savings (%)*
5	1939	3855	3692	-163	4.24
6	1902	3932	3724	-208	5.30

Table 7.25 Model summary

Source	*DF*	*Sum of squares*	*Mean square*	*F*	*p*
Region	14	7970655	569333	0.91	0.5478
Postcode	131	88237130	673566	1.08	0.2651
Gsp	11	10550163	959106	1.53	0.1128
Mosaic	11	16137018	1467001	2.34	0.0071
Awareness	2	5338463	2669231	4.26	0.0141
Fuel type	1	1234228	1234228	1.97	0.1604
Trial group	1	3154338	3154338	5.04	0.0248

[32] With benchmarking.

Table 7.26 Adjusted mean change

Trial group	*Mean*	*Standard error*
5	-123	36.4
6	-186	36.2

7.1.3 Gas demand reduction

Main analysis

A factorial analysis was undertaken similar to that of electricity demand reduction. In total, 1164 households were included. The same confounding variables were included except that gas zone replaced grid supply point. A summary of changes in consumption is shown in Table 7.27.

No main effects of interventions were significant at either the 5% or 10% level. However, there was a significant interaction effect at the 5% level between the incentive to shift and web information, and a marginally significant interaction effect (at the 10% level) between the incentive to reduce and the RTD, once adjusted for all other effects in the model. The model summary is presented in Table 7.28. The adjusted means (least square means) for the main effects are presented in Table 7.29.

Percentage savings in each cell of the factorial design are shown in Table 7.30. The first interaction means that savings were greater if participants had both the incentive to shift and web information, or neither, compared with those that had only web information. The second interaction means that the incentive to reduce was effective only when an RTD was provided. The significant differences are identified in Table 7.31.

Table 7.27 Summary of changes in the mean by trial group

Trial group	*Interventions active in the trial group*				*Number of households*	*Pretrial mean (kWh)*	*In-trial mean (kWh)*	*Change in mean (kWh)*	*Percentage savings (%)*
	Incentive to shift	*Incentive to reduce*	*RTD*	*Web information*					
11					134	17691	15576	-2115	12.0
12	✓				65	17811	16010	-1801	10.1
13		✓			64	19394	17284	-2110	10.9
14	✓	✓			70	17941	15838	-2103	11.7
23			✓		84	17341	15556	-1785	10.3
24	✓		✓		48	18046	16600	-1446	8.0
25		✓	✓		49	18006	15509	-2497	13.9
26	✓	✓	✓		53	16162	13645	-2517	15.6
10				✓	136	17793	16018	-1775	10.0
7	✓			✓	66	18761	16435	-2326	12.4
8		✓		✓	73	17959	16295	-1664	9.3
9	✓	✓		✓	70	19300	17081	-2219	11.5
22			✓	✓	89	18095	16377	-1718	9.5
19	✓		✓	✓	45	16200	14463	-1737	10.1
20		✓	✓	✓	55	18040	16417	-1623	9.0
21	✓	✓	✓	✓	63	17897	15201	-2696	15.1

Table 7.28 Model summary

Source	*DF*	*Sum of squares*	*Mean square*	*F*	*p*
Postcode	4	6466280	16165702	2.11	0.0778~
Gas zone	1	1985490	1985490	0.26	0.6110
Mosaic	10	199437888	19943789	2.6	0.0040**
Awareness	2	9713007	4856503	0.63	0.5311
Fuel type	1	29764970	29764970	3.88	0.0491*
Incentive to shift	1	13072577	13072577	1.70	0.1920
Incentive to reduce	1	15937749	15937749	2.08	0.1497
Web information	1	7986630	7986630	1.04	0.3077
RTD	1	498691	498691	0.07	0.7987
Incentive to shift*web information	1	32379064	32379065	4.22	0.0401*
Incentive to reduce*RTD	1	22448938	22448938	2.93	0.0874~

**p<0.01, *p<0.05, ~p<0.10

Table 7.29 Adjusted mean change for main effects of interest

Incentive to shift	*Incentive to reduce*	*Web information*	*RTD*	*Mean*	*Standard error*
No	-	-	-	-2197	344
Yes	-	-	-	-2424	356
-	No	-	-	-2159	345
-	Yes	-	-	-2462	356
-	-	No	-	-2364	353
-	-	Yes	-	-2257	345
-	-	-	No	-2313	351
-	-	-	Yes	-2308	347
No	-	No	-	-2423	365
No	-	Yes	-	-1970	356
Yes	-	No	-	-2303	382
Yes	-	Yes	-	-2544	374
-	No	-	No	-2305	359
-	No	-	Yes	-2012	364
-	Yes	-	No	-2321	378
-	Yes	-	Yes	-2603	378

Table 7.30 Percentage savings in each cell of the analysis

Incentive to reduce	*Incentive to shift*	*RTD*		*No RTD*	
		Web information	*No web information*	*Web information*	*No web information*
Yes	Yes	15.1	15.6	11.5	11.7
	No	9.0	13.9	9.3	10.9
No	Yes	10.7	8.0	12.4	10.1
	No	9.5	10.3	10.0	12.0

Table 7.31 Specific comparisons of interest

Comparison	*Estimate*	*Standard error*	*t*	*p*
Difference between the levels of incentive to shift when web information is not provided	121	248	0.49	0.627
Difference between the levels of incentive to shift when web information is provided	-574	243	-2.37	0.018*
Difference between the levels of web information when incentive to shift is absent	454	214	2.12	0.034*
Difference between the levels of web information when incentive to shift is present	-241	257	-0.94	0.349
Difference between the levels of incentive to reduce when RTD is provided	-16	229	-0.07	0.944
Difference between the levels of incentive to reduce when RTD is not provided	-591	264	-2.24	0.025*
Difference between the levels of RTD when incentive to reduce is absent	293	220	1.33	0.184
Difference between the levels of RTD when incentive to reduce is present	-282	254	-1.11	0.267

*$p<0.05$

This main analysis could not include all the trial cells. Therefore, a series of supplementary analyses was conducted.

Effect of smart meter – with and without RTD

This comprised two sets of trial groups:

- with no other intervention [TGs 15 & 27 vs 32 & 3];
- with booklet and graphs on bills [TGs 11 & 23 vs 2 & 5].

Table 7.32 shows the interventions for these trial groups. There were no households in TG15, and so this group was deleted from the analysis. In total there were 3395 households in this set.

The results are presented in Tables 7.33 - 7.36. This analysis means that groups with a smart meter had greater reductions in consumption than groups without a smart meter if they also had the advice booklet and additional data on their bills (monthly for smart meter groups, quarterly for others) – relative savings were 6.7% with or without an RTD. The difference made by the smart meter (6.4%) is also significant if the advice booklet and additional data on bills were not provided (an RTD was also provided in both the smart meter and the non-smart meter group).

Table 7.32 Summary of trial groups

Trial group	*Smart meter*	*Booklet*	*Quarterly bills with graphs*	*Monthly bills with graphs*	*Incentive to shift*	*Incentive to reduce*	*RTD*	*Web information*
32								
3							✓	
5		✓	✓					
2		✓	✓				✓	
15	✓							
27	✓						✓	
11	✓	✓		✓				
23	✓	✓		✓			✓	

Table 7.33 Summary of changes in the mean by trial group

Trial group	*Number of households*	*Pretrial mean (kWh)*	*In-trial mean (kWh)*	*Change in mean (kWh)*	*Percent savings (%)*
32	1317	17812	16690	-1122	6.30
3	488	18103	17519	-583	3.22
5	767	17723	16796	-928	5.24
2	485	17944	17298	-646	3.60
27	120	17941	16220	-1721	9.59
11	134	17691	15576	-2115	11.96
23	84	17341	15556	-1785	10.29

Table 7.34 Model summary

Source	*DF*	*Sum of squares*	*Mean square*	*F*	*p*
Region	15	63378202	4225213	0.51	0.9344
Postcode	129	1144287359	8870445	1.08	0.2566
Gas Zone	13	95420827	7340064	0.89	0.5591
Mosaic	11	127415595	11583236	1.41	0.1607
Awareness	2	3904007	1952004	0.24	0.7884
Fuel type	1	14306249	14306249	1.74	0.1870
Trial group	6	181544006	30257334	3.68	0.0012

Table 7.35 Adjusted mean change

Trial group	*Mean*	*Standard error*
32	-906	173
3	-358	198
5	-712	182
2	-551	194
27	-1316	346
11	-1652	349
23	-1327	394

Table 7.36 Specific comparisons of interest

Trial groups	*Estimate*	*Standard error*	*t*	*p*
3 vs 27	-957	333	2.87	0.004
11&23 vs 2&5	-858	293	2.92	0.004

Effect of booklet & graphs with smart meter

This comprised two sets of trial groups:

- effect of booklet with graphs and incentives [TGs 12-14 vs 16-18];
- effect of booklet and graphs with RTD and incentives [TGs 23-26 vs 27-30].

Table 7.37 shows the interventions for these trial groups. In total there were 920 households in this set. The results are presented in Tables 7.38 - 7.41. This analysis means that the advice booklet and additional data on monthly bills did not affect consumption in the smart meter groups.

Capabilities on project:
Building Engineering

Table 7.37 Summary of trial groups

Trial group	*Smart meter*	*Booklet*	*Quarterly bills with graphs*	*Monthly bills with graphs*	*Incentive to shift*	*Incentive to reduce*	*RTD*	*Web information*
16	✓			✓	✓			
17	✓			✓		✓		
18	✓			✓	✓	✓		
27	✓						✓	
28	✓				✓		✓	
29	✓					✓	✓	
30	✓				✓	✓	✓	
12	✓	✓		✓	✓			
13	✓	✓		✓		✓		
14	✓	✓		✓	✓	✓		
23	✓	✓		✓			✓	
24	✓	✓		✓	✓		✓	
25	✓	✓		✓		✓	✓	
26	✓	✓		✓	✓	✓	✓	

Table 7.38 Summary of changes in the mean by trial group

Trial group	*Number of households*	*Pretrial mean (kWh)*	*In-trial mean (kWh)*	*Change in mean (kWh)*	*Percent savings (%)*
16	64	18956	16636	-2321	12.24
17	64	18061	16905	-1156	6.40
18	64	17475	15526	-1948	11.15
27	120	17941	16220	-1721	9.59
28	54	17838	15935	-1903	10.67
29	58	16397	14166	-2231	13.61
30	63	17663	15793	-1870	10.59
12	65	17811	16010	-1801	10.11
13	64	19394	17284	-2110	10.88
14	70	17941	15838	-2103	11.72
23	84	17341	15556	-1785	10.29
24	48	18046	16600	-1446	8.01
25	49	18006	15509	-2497	13.87
26	53	16162	13645	-2517	15.57

Table 7.39 Model summary

Source	*DF*	*Sum of squares*	*Mean square*	*F*	*p*
Region	3	47710564	15903521	1.89	0.130
Postcode	1	4265875	4265875	0.51	0.477
Gas Zone	1	659612	659612	0.08	0.780
Mosaic	11	75833786	6893981	0.82	0.622
Awareness	2	12554421	6277211	0.75	0.475
Fuel type	1	46803280	46803280	5.56	0.019
Trial	13	90040057	6926158	0.82	0.637

Table 7.40 Adjusted mean change

Trial group	*Mean*	*Standard error*
16	-1964	708
17	-909	704
18	-1647	715
27	-1537	614
28	-1726	726
29	-1951	715
30	-1501	715
12	-1573	711
13	-1820	707
14	-1836	700
23	-1658	662
24	-1145	741
25	-2207	741
26	-2137	730

Table 7.41 Specific comparisons of interest

Trial groups	*Estimate*	*Standard error*	*t*	*p*
12-14 vs 16-18	-237	297	-0.80	0.424
23-26 vs 27-30	-108	266	-0.41	0.685
12-14 & 16-18 vs 23-30	109	205	0.53	0.596

Effect of booklet and RTD, without smart meter

This comprised one set of trial groups, with factorial analysis [TGs 32, 1, 3 & 4]. Table 7.42 shows the interventions for these trial groups. The results are shown in Tables 7.43 to 7.45. This analysis means that the advice booklet did not affect consumption in the non-smart meter groups but the RTD had a significant effect (in the opposite of the expected direction), which did not depend on the availability of the booklet. The results presented here are based on 3506 households.

Table 7.42 Summary of trial groups

Trial Group	*Smart meter*	*Booklet*	*Quarterly bills with graphs*	*Monthly bills with graphs*	*Incentive to shift*	*Incentive to reduce*	*RTD*	*Web information*
32								
4		✓						
3							✓	
1		✓					✓	

Table 7.43 Summary of changes in the mean by trial group

Trial group	*Number of households*	*Pretrial mean (kWh)*	*In-trial mean (kWh)*	*Change in mean (kWh)*	*Percent savings (%)*
32	1317	17812	16690	-1122	6.30
4	1172	17352	16580	-772	4.45
3	488	18103	17519	-583	3.22
1	529	17760	17067	-693	3.90

Table 7.44 Model summary

Source	*DF*	*Sum of squares*	*Mean square*	*F*	*p*
Region	15	154702381	10313492	1.29	0.202
Postcode	134	1200310506	8957541	1.12	0.173
Gas Zone	13	100527337	7732872	0.96	0.484
Mosaic	11	149068190	13551654	1.69	0.069
Awareness	2	22970889	11485444	1.43	0.239
Fuel type	1	81934043	81934043	10.22	0.001
Booklet	1	17511349	17511349	2.18	0.140
RTD	1	86693127	86693127	10.81	0.001
Booklet*RTD	1	25427804	25427804	3.17	0.075

Table 7.45 Adjusted mean change

Booklet	*RTD*	*Mean*	*Standard error*
No	-	-392	159
Yes	-	-292	155
-	No	-549	152
-	Yes	-135	164
No	No	-703	167
No	Yes	-82	193
Yes	No	-394	163
Yes	Yes	-189	186

Effect of graphs and RTD, added to booklet, without smart meter

This comprised one set of trial groups with factorial analysis [TGs 1, 2, 4 & 5]. Table 7.46 shows the interventions for these trial groups. The results are shown in Tables 7.47 to 7.49. This analysis means that neither the additional data on bills nor the RTD affected consumption in the non-smart meter groups that had the booklet. The results presented here are based on 2953 households.

Table 7.46 Summary of trial groups

Trial group	*Smart meter*	*Booklet*	*Quarterly bills with graphs*	*Monthly bills with graphs*	*Incentive to shift*	*Incentive to reduce*	*RTD*	*Web information*
4		✓						
1		✓					✓	
5		✓	✓					
2		✓	✓				✓	

Table 7.47 Summary of changes in the mean by trial group

Trial group	*Number of households*	*Pretrial mean (kWh)*	*In-trial mean (kWh)*	*Change in mean (kWh)*	*Percent savings (%)*
1	529	17760	17067	-693	3.90
2	485	17944	17298	-646	3.60
4	1172	17352	16580	-772	4.45
5	767	17723	16796	-928	5.24

Table 7.48 Model summary

Source	*DF*	*Sum of squares*	*Mean square*	*F*	*p*
Region	14	160781506	11484393	1.46	0.117
Postcode	131	1166454835	8904235	1.13	0.146
Gas Zone	10	38588709	3858871	0.49	0.897
Mosaic	11	173615010	15783183	2.01	0.024
Awareness	2	33763324	16881662	2.15	0.117
Fuel type	1	21770385	21770385	2.77	0.096
Graphs	1	941443	971443	0.12	0.725
RTD	1	18488800	18488800	2.35	0.125
Graphs*RTD	1	261869	261869	0.03	0.855

Table 7.49 Adjusted mean change

Graphs	*RTD*	*Mean*	*Standard error*
No	-	-504	164
Yes	-	-553	169
-	No	-623	162
-	Yes	-433	171
No	No	-608	171
No	Yes	-400	193
Yes	No	-638	183
Yes	Yes	-497	196

Effect of benchmarking

This comprised one set of trial groups with factorial analysis [TGs 5 & 6]. Table 7.50 shows the interventions for these trial groups. The results are shown in Tables 7.51 to 7.53. This analysis means that the addition of benchmarking comparisons on bills in the non-smart meter groups that had the booklet but no other intervention did not have a significant effect. The results presented here are based on 1558 households.

Table 7.50 Summary of trial groups

Trial group	*Smart meter*	*Booklet*	*Quarterly bills with graphs*	*Monthly bills with graphs*	*Incentive to shift*	*Incentive to reduce*	*RTD*	*Web information*
5		✓	✓					
6		✓	✓[33]					

Table 7.51 Summary of changes in the mean by trial group

Trial group	*Number of households*	*Pretrial mean (kWh)*	*In-trial mean (kWh)*	*Change in mean (kWh)*	*Percent savings (%)*
5	767	17723	16796	-928	5.24
6	791	17603	16570	-1033	5.87

Table 7.52 Model summary

Source	*DF*	*Sum of squares*	*Mean square*	*F*	*p*
Region	14	17813979	1272427	0.16	1.000
Postcode	114	992394214	8705212	1.12	0.198
Gas Zone	9	42158473	4684275	0.60	0.797
Mosaic	11	65807673	5982516	0.77	0.674
Awareness	2	1777915	888957	0.11	0.892
Fuel type	1	414351	414351	0.05	0.818
Trial group	1	12863781	12863781	1.65	0.199

[33] With benchmarking.

Table 7.53 Adjusted mean change

Trial group	*Mean*	*Standard error*
5	-958	243
6	-1131	243

7.1.4 Load shifting

Introduction

SSE's incentive to shift intervention was based on electricity tariffs varying with time of day, season and day of the week (weekday vs weekend). The original analysis by SSE examined consumption separately each tariff period. Here we report analysis of the proportion of electricity consumption that falls in the peak period; the main hypothesis is that this proportion should be lower when the incentive to reduce is in effect. Simple descriptive summaries for consumption during the peak period, total consumption, and the percentage of total consumption that was in the peak period are shown in Tables 7.55 and 7.56. In each case, the second group in each pair was the one that had the incentive to shift.

Table 7.55 Descriptive statistics (high season)

	Weekday kWh					*Weekend kWh*					
Trial Group	*Peak*	*SD peak*	*Total*	*SD total*	*% peak*	*Peak*	*SD peak*	*Total*	*SD total*	*% peak*	*n*
27	221.8	188.9	1148.9	1173.6	20.0	93.8	78.2	494.7	487.3	19.5	233
28	210.3	127.8	1084.8	685.9	19.6	95.1	59.9	484.6	302.9	19.6	124
29	212.3	149.3	1077.9	648.7	19.6	92.3	57.9	465.0	272.9	19.8	121
30	207.7	133.7	1109.5	750.4	19.2	92.2	59.9	490.0	320.0	19.0	127
11	218.7	145.9	1124.4	803.1	19.8	93.2	65.1	485.1	350.0	19.3	305
12	206.6	126.4	1129.1	707.9	18.7	90.1	55.9	494.3	297.9	18.4	162
13	238.6	151.8	1228.2	845.6	20.0	103.2	69.7	526.7	365.9	19.8	153
14	199.5	119.0	1041.7	624.0	19.2	87.7	57.6	456.5	282.8	19.0	160
23	214.7	142.2	1107.2	807.9	20.0	92.3	64.5	475.8	350.3	19.6	206
24	217.8	147.1	1152.1	876.9	19.6	92.4	65.7	491.3	371.2	19.1	103
25	201.8	169.6	1078.4	931.0	19.3	86.8	71.7	462.2	388.0	19.1	126
26	227.5	153.8	1204.2	825.4	19.3	98.1	73.3	517.3	370.7	18.7	121
10	227.8	146.5	1197.3	781.7	19.6	99.6	64.3	516.9	333.3	19.7	283
7	217.8	124.0	1167.0	722.7	19.3	96.2	58.7	510.4	317.5	19.0	173
8	203.7	136.5	1038.6	721.3	20.0	89.8	60.8	455.0	314.3	20.0	162
9	216.0	143.5	1106.3	744.0	19.8	93.6	62.1	482.8	316.1	19.3	177
22	203.6	135.9	1074.1	703.2	19.1	88.2	59.0	461.5	294.4	19.0	210
19	213.2	126.4	1134.9	707.8	19.2	93.0	52.2	486.7	278.5	19.3	131
20	203.4	130.9	1072.6	681.6	19.1	85.7	58.2	463.8	303.7	18.4	137
21	214.9	180.8	1102.5	962.3	19.7	90.3	71.9	474.6	394.6	19.0	140

Table 7.56 Descriptive statistics (low season)

Trial Group	Weekday kWh					Weekend kWh					n
	Peak	*SD peak*	*Total*	*SD total*	*% Peak*	*Peak*	*SD peak*	*Total*	*SD total*	*% Peak*	
27	365.4	277.5	1979.2	1696.8	18.7	138.6	105.6	794.4	641.8	17.6	233
28	362.1	246.6	1979.0	1303.7	18.2	146.4	105.1	829.3	540.9	17.4	124
29	362.9	226.0	1944.6	1118.4	18.6	139.7	83.1	787.8	461.6	17.8	121
30	350.3	194.7	1934.1	1021.9	18.2	138.9	83.4	800.9	453.0	17.3	127
11	359.2	223.7	1913.5	1173.3	18.8	135.8	92.1	773.6	492.3	17.4	305
12	346.3	185.7	1962.7	1105.3	18.1	132.5	75.5	801.3	455.0	16.7	162
13	386.9	228.5	2072.6	1264.7	18.9	149.7	97.0	834.2	520.1	17.9	153
14	334.5	169.8	1823.5	883.2	18.4	129.6	74.1	744.9	389.4	17.2	160
23	355.8	218.1	1913.8	1222.4	18.9	139.1	95.0	784.0	516.4	17.7	206
24	356.0	202.0	1982.8	1160.7	18.4	132.1	75.0	781.5	439.7	17.2	103
25	331.8	223.4	1845.1	1135.6	18.1	130.6	89.3	746.2	442.7	17.3	126
26	382.8	254.3	2109.4	1419.2	18.1	145.6	99.7	850.0	557.1	16.9	121
10	371.4	228.9	1996.8	1180.8	18.7	144.8	90.5	808.2	476.1	17.9	283
7	363.6	204.2	2008.6	1165.4	18.3	142.0	87.4	825.7	494.9	17.2	173
8	347.8	214.8	1837.1	1071.6	18.9	134.6	84.1	750.5	449.0	17.8	162
9	363.8	234.8	2006.3	1334.2	18.4	142.5	92.9	815.9	530.8	17.4	177
22	330.5	197.1	1838.4	1032.0	17.8	130.4	83.6	747.0	420.3	17.1	210
19	361.4	200.8	2011.1	1164.0	18.3	140.8	79.7	810.2	453.3	17.5	131
20	327.1	196.9	1792.1	999.2	18.2	124.5	77.4	722.5	405.8	16.9	137
21	365.1	252.4	1954.2	1545.1	18.9	137.3	94.6	792.0	609.1	17.4	140

This analysis necessarily excludes non-smart meter groups because data were not available with a time-of-day breakdown. Smart meter trial groups were further excluded where there was not a comparable pair of trial groups, having the same interventions except that one group had the incentive to shift and the other did not. Trial numbers 15 and 17 were also omitted from the analysis because small sample sizes (1 and 3 respectively). In addition, households with a zero reading, of which there were 2, and households with missing data for the mosaic variable were similarly omitted from the analysis set (12 households). Trial number 31 (prepayment customers) was omitted because there was no relevant comparison group. A total of 3653 households were included in this analysis after these omissions.

The data analysed here are percentages and, in many cases, when analysing percentage data it is necessary to apply a transformation to the data in order to fulfil the assumptions underlying the analysis method. Appropriate transformations were investigated on all response variables of interest but they did not improve the model or change the overall conclusions. Hence the untransformed percentage data were analysed.

A factorial repeated measurements analysis was carried out to generate separate models for weekdays and weekends. In each case, the low and high season were regarded as the “repeated measurements”. In order to undertake a repeated measurements analysis, mixed models were fitted to the data with households as a random effect. This induced a compound symmetry structure on the repeated measurements – i.e. constant variance for the two seasons, and a correlation between them. Fixed effects were included for interventions (Booklet, Bills with graphs, Incentive to shift, Incentive to reduce, RTD and Web information). The model included the same confounding variables as the analysis of total consumption (Section 7.1.2).

It was not possible to fit all two-factor interactions in the model – some were aliased with other effects and some were not estimable. All two-factor interactions that were possible to estimate have been included. The interactions of season with all these main effects and two-factor interactions were also included.

Results reported here are the results of the hypothesis tests from the mixed model. Least squares means for all the main effects for interventions are presented irrespective of whether or not the effect was significant. However for the interactions, least squares means are only presented for those that were significant.

Analysis of weekdays

The model summary is presented in Table 7.57. Adjusted means (least square means) are presented in Table 7.58. The difference between the two seasons is by far the dominant effect in this analysis, with the percentage of usage in the peak period being higher in the high season compared to the low season. The percentage of consumption that falls in the peak period is reduced by the incentive to shift but by only a small amount – from 19.8% to 19.5%. The effect of the incentive to shift was greater in the absence of an RTD or in the absence of web information, suggesting some kind of interference effect if too many interventions are in effect at the same time.

Table 7.57 Model summary

Effect[34]	*Num DF*	*Den DF*	*F*	*p*
Region	4	3638	1.07	0.3721
Postcode	7	3638	2.53	0.0133*
Gsp	1	3638	1.55	0.2137
Mosaic	11	3638	3.61	<0.0001**
Awareness	2	3638	1.65	0.1929
Fuel type	1	3638	0.93	0.3343
Booklet	1	3638	1.09	0.2968
Graphs on bills	1	3638	0.52	0.4701
Incentive to shift	1	3638	5.05	0.0247*
Incentive to reduce	1	3638	0.09	0.7685
RTD	1	3638	0.75	0.3861
Web information	1	3638	0.09	0.7705
Incentive to shift × RTD	1	3638	3.98	0.0462*
Incentive to shift × Web information	1	3638	4.42	0.0357*
Season	1	3638	481.09	<0.0001**

*p<0.05, **p<0.01

[34] Type 2 tests of fixed effects.

Table 7.58 Adjusted mean percentage of consumption in peak period

Effect	Season	Booklet	Graphs on bills	Incentive to shift	Incentive to reduce	RTD	Web information	Mean	Standard Error
Booklet	-	No	-	-	-	-	-	19.81	0.57
	-	Yes	-	-	-	-	-	19.48	0.58
Graphs on bills	-	-	No	-	-	-	-	19.51	0.61
	-	-	Yes	-	-	-	-	19.79	0.57
Incentive to shift	-	-	-	No	-	-	-	19.81	0.56
	-	-	-	Yes	-	-	-	19.49	0.56
Incentive to reduce	-	-	-	-	No	-	-	19.60	0.56
	-	-	-	-	Yes	-	-	19.70	0.57
RTD	-	-	-	-	-	No	-	19.71	0.57
	-	-	-	-	-	Yes	-	19.59	0.55
Web information	-	-	-	-	-	-	No	19.65	0.56
	-	-	-	-	-	-	Yes	19.65	0.56
Incentive to shift x RTD	-	-	-	No	-	No	-	20.01	0.57
	-	-	-	No	-	Yes	-	19.61	0.56
	-	-	-	Yes	-	No	-	19.40	0.58
	-	-	-	Yes	-	Yes	-	19.57	0.56
Incentive to shift x Web information	-	-	-	No	-	-	No	19.96	0.56
	-	-	-	No	-	-	Yes	19.66	0.57
	-	-	-	Yes	-	-	No	19.33	0.57
	-	-	-	Yes	-	-	Yes	19.64	0.58
Season	High	-	-	-	-	-	-	20.17	0.56
	Low	-	-	-	-	-	-	19.13	0.56

Analysis of weekends

The model summary is presented in Table 7.59. Adjusted means (least square means) are presented in Table 7.60. As for weekdays, the difference between the two seasons is by far the dominant effect in this analysis, with the percentage of usage in the peak period being higher in the high season compared to the low season. The percentage of consumption that falls in the peak period is reduced by the incentive to shift but by only a small amount – from 19.4% to 18.9%. The effect of the incentive to shift was greater in the absence of an RTD, suggesting some kind of interference effect if too many interventions are in effect at the same time.

Two other significant interactions have no obvious explanation: consumption in the peak period was higher if there was web information but no RTD, or if the incentive to reduce was in effect without an RTD.

Table 7.59 Model summary

Effect[35]	Num DF	Den DF	F	p
Region	4	3638	1.56	0.1823
Postcode	7	3638	1.40	0.2008
Gsp	1	3638	0.66	0.4159
Mosaic	11	3638	3.58	<0.0001**
Awareness	2	3638	1.75	0.1737
Fuel type	1	3638	1.21	0.2708
Booklet	1	3638	3.75	0.0530~
Graphs on bills	1	3638	0.79	0.3732
Incentive to shift	1	3638	11.01	0.0009**
Incentive to reduce	1	3638	0.01	0.9217
RTD	1	3638	1.34	0.2475
Web information	1	3638	0.94	0.3334
Incentive to shift × RTD	1	3638	5.13	0.0236*
Incentive to reduce × RTD	1	3638	5.41	0.0201*
RTD × Web information	1	3638	4.36	0.0369*
Season	1	3638	1521.94	<0.0001**

*p<0.05, **p<0.01, ~p<0.10

[35] Type 2 tests of fixed effects.

Table 7.60 Adjusted mean % of total usage used at peak times

Effect	Season	Booklet	Graphs on bills	Incentive to shift	Incentive to reduce	RTD	Web information	Mean	Standard Error
Booklet	-	No	-	-	-	-	-	19.44	0.52
	-	Yes	-	-	-	-	-	18.88	0.53
Graphs on bills	-	-	No	-	-	-	-	19.01	0.55
	-	-	Yes	-	-	-	-	19.31	0.51
Incentive to shift	-	-	-	No	-	-	-	19.39	0.50
	-	-	-	Yes	-	-	-	18.94	0.51
Incentive to reduce	-	-	-	-	No	-	-	19.11	0.50
	-	-	-	-	Yes	-	-	19.21	0.52
RTD	-	-	-	-	-	No	-	19.24	0.51
	-	-	-	-	-	Yes	-	19.08	0.50
Web information	-	-	-	-	-	-	No	19.11	0.50
	-	-	-	-	-	-	Yes	19.22	0.51
Incentive to shift x RTD	-	-	-	No	-	No	-	19.61	0.52
	-	-	-	No	-	Yes	-	19.16	0.51
	-	-	-	Yes	-	No	-	18.86	0.53
	-	-	-	Yes	-	Yes	-	19.01	0.51
Incentive to reduce x RTD	-	-	-	-	No	No	-	19.02	0.52
	-	-	-	-	No	Yes	-	19.20	0.50
	-	-	-	-	Yes	No	-	19.45	0.54
	-	-	-	-	Yes	Yes	-	18.96	0.52
RTD x Web information	-	-	-	-	-	No	No	19.03	0.52
	-	-	-	-	-	No	Yes	19.44	0.52
	-	-	-	-	-	Yes	No	19.18	0.51
	-	-	-	-	-	Yes	Yes	18.99	0.52
Season	High	-	-	-	-	-	-	20.10	0.50
	Low	-	-	-	-	-	-	18.22	0.50

7.2 Summary of SSE and AECOM findings[36]

7.2.1 Demand reduction

Electricity

Non-smart meter groups

The analysis by SSE, comparing all trial groups in a single multivariate analysis, found no effect of any intervention in the non-smart meter groups except a small significant effect ($p<0.05$) of providing a clip-on RTD. Reduction in consumption relative to the control group was 1% and this was confirmed in the AECOM analysis that compared selected trial groups.

[36] SSE's community trials are considered separately, in Appendix D.

The AECOM analysis also confirmed no effect of the advice booklet or graphs on bills in the absence of a smart meter. An additional small significant effect ($p<0.05$) was seen, comparing the trial group with the booklet, graphs on bills and benchmarking information with the most similar group (which had the booklet and graphs on bills but not benchmarking information). The additional savings attributable to benchmarking were 1%.

Smart meter groups

The analysis by SSE, comparing all trial groups in a single multivariate analysis, found a significant effect of "meter type" (a combination of smart vs non-smart "standard" meter and whether or not an RTD was provided). This meant that savings relative to the control group were 2.5% with a smart meter, 3.6% with a smart meter and RTD and 2.9% with a prepayment smart meter and RTD.

Since there was only one prepayment group, SSE's results for this group were taken as definitive and AECOM conducted some further analysis of the other smart meter groups. The significant effect of providing smart meters was confirmed when comparing three pairs of trial groups that differed only in the provision of a smart meter (relative savings of 3.3-4.7%). The largest effect (4.7%) was with the booklet, graphs on bills and RTD also provided in both smart meter and non-smart meter groups. Comparing two groups with smart meters, all having the booklet and graphs on bills, savings were1.8% greater in the group that also had an RTD.

The absence of effects of other interventions was also confirmed by the AECOM analysis.

Gas

Non-smart meter groups

The analysis by SSE, comparing all trial groups in a single multivariate analysis, found no effect of any intervention in the non-smart meter groups. This was confirmed in the AECOM analysis that compared selected trial groups. The significant reductions that had been seen for electricity consumption (of RTDs and benchmarking) were, in fact, non-significantly reversed.

Smart meter groups

The analysis by SSE, comparing all trial groups in a single multivariate analysis, found a significant effect of "meter type" (a combination of smart vs non-smart "standard" meter and whether or not an RTD was provided). This meant that savings relative to the control group were 3.0% with a smart meter and 3.2% with a smart meter and RTD. Relative savings with an electricity prepayment smart meter and RTD (0.3%) were not statistically significant.

The analysis by AECOM confirmed these effects and the non-significant other main effects. It also found a significant interaction between the incentive to shift and provision of information via the web, and a marginally significant interaction between incentive to reduce and provision of an RTD.

The first interaction means that savings were greater if participants had both the incentive to shift and web information (12.4%)[37] or neither (11.8%), compared with those that had web information without the incentive to shift (9.5%). Savings in groups with the incentive to shift but not web information (11.4%) were intermediate and did not differ from the other groups. There is no clear interpretation of this.

The second interaction means that the incentive to reduce was effective only when an RTD was provided (3.8% relative savings due to the incentive to reduce, with an RTD also provided, -0.3% without an RTD). There was also a

[37] Percentages shown are the average for the trial groups having the interventions indicated, regardless of other interventions that were in effect. The figures are not adjusted for savings in the control group so only the relative magnitudes should be considered.

positive effect of the RTD on savings, when the incentive to reduce was in effect (2.6%), which was reversed without the incentive to reduce (-1.5%), but neither difference was significant when tested separately.

Effects of non-intervention variables

Testing for differences across all trial groups, SSE found a significant effect of awareness of the trial on electricity demand reduction (relative to baseline), percentage savings being higher for the Aware (6.4%) and Committed (6.1%) groups than the Unaware group (4.0%). This was not seen in the AECOM analysis, which can be attributed to the smaller subsamples analysed.

The Committed group might have been more committed before the trial (which could mean either they were more likely to take action or they had already taken a lot of action and there is not much more they can do). Or they could have made a minimal, nominal commitment, based on the wording on the card. The latter is more likely in the case of interventions that offered the customer some material benefit (e.g. an RTD or a discount on their bill). Hence the Commitment may not mean the same thing for each trial group but this did not show in any significant interactions.

Postcode (for electricity only) and Mosaic group (for gas and electricity) also had significant effects in the SSE analysis but Grid distribution point (an industry geographic code) did not. These effects were also found in the AECOM analysis, which also included geographic region in the analysis, finding no significant effect for this additional variable. This emphasises the need to control for location and demographic variables but the type of analysis used did not permit identification of specific locations or Mosaic groups that had a higher propensity to reduce consumption. Since postcode was significant but not the more geographically systematic variables, the details of any location effect are difficult to define.

In no case did one of these variables have a significant interaction with a trial intervention, i.e. the intervention effects did not vary with awareness, location or demographic category.

7.2.2 Load shifting

SSE's analysis of smart meter data by tariff period and high/low season showed an effect of Incentive, and possibly Awareness, on consumption. Consumption is higher at night in high season for customers on the "Shift" incentive (for electricity) or "Shift + Reduce" incentive (for gas). The percentage shift from peak to night electricity usage is estimated as 8.5-10.1%, based on peak season consumption at night (overall shift from peak is not estimated).

AECOM's analysis confirmed that the percentage of consumption that falls in the peak period is reduced by the incentive to shift but by only a small amount – from 19.8% to 19.5% on weekdays and from 19.4% to 18.9% at weekends. The weekday effect of the incentive to shift was greater in the absence of an RTD and in the absence of web information, suggesting some kind of interference effect if too many interventions are in effect at the same time. The interaction with RTD was also seen at weekends.

7.2.3 Implications of the survey findings

Reasons for joining the trial

Householders in Projects 3 and 4[38] were asked what had motivated them to join the trial. The most common single reason was saving money (41% and 42% in Projects 3 and 4 respectively), followed by helping the environment

[38] Breakdown of survey findings was based on SSE's original "Project" classification (described in Appendix A4.1) rather than the final Trial Groups. Project 3 consisted of all Trial Groups with smart meters but not RTDs whereas Project 4 had both. Both Projects had a range of other interventions.

(21% and 17%). Project 4, with RTDs installed, were more likely to say the reason was to obtain detailed information about energy consumption (12% vs 5%) and to identify excessive energy use (6% vs 3%). These questions were asked more than two years after the decision so they should not necessarily be seen as reliable, but they do give an indication of what issues were uppermost in respondents' minds, which could be useful in presenting changes to consumers in future.

Specific actions taken

The survey recorded actions taken during the past two years (i.e. the trial period): boiler replacement, loft insulation installed or topped up, cavity wall insulation installed, double glazing, draughtproofing of doors and replacement of "standard bulbs" with energy saving bulbs. The percentage of households that had taken each action varied widely between groups but with no obvious pattern and not baselined against capacity to act (i.e. whether the change was physically possible and not already made).

Energy advice booklet

Credit customers with smart meters were asked whether they recalled receiving the advice booklet and a similar percentage (80%) did, whether or not they had also received an RTD. However, customers with an RTD were more likely to still have the booklet (59% vs 79%), still refer to it (8% vs 15%) and say it was quite or very useful (78% vs 86%). Having an RTD appears to have brought about greater interest in the advice booklet.

The most highly rated specific aspects of information were "Tells you which appliances are high energy users" (30%), "Tips on saving money/energy - unplug appliances" (8%) and information on insulation (4%). Reasons for not finding the booklet useful are equally important in guiding future design of advice; among those who had received information but not found it useful, the most common reasons were: "Tells you nothing new" (44%), "It's common sense" (21%), "Couldn't be bothered to read it" (11%), "Couldn't use less / already save energy" (6%), "Too much information to take in" (6%). The general theme is that the information was redundant or too much for the user to work through.

Information on bills

In an early survey of households receiving additional bill data (graphs on bills), only 32% recalled something different about the bill and 27% recalled what it was. Asked directly about the graphs, 48% said they had looked at the graphs in detail and 38% that it made them think about energy use. This limits the potential of the graphs to affect consumption. Nevertheless, for those who were aware of the graphs, the final survey provides some evidence on their role in customers' thinking.

Final survey respondents rated the usefulness of information on electricity bill and statements "in helping you to reduce your energy consumption". The percentage responding quite or very useful was higher in the non-smart meter groups that had extra information on their bills in addition to energy advice than in the one that had advice only (60% vs 34%). The figures reported do not allow this comparison for smart meter groups. Oddly, the percentage for non-smart meter groups with RTDs but not extra bill information was intermediate, at 50%.

Where the extra information was provided, the items on the bill rated as most useful changed from units used to the graphs or comparison with previous bills.

Householders who had a visual display of some kind were asked to compare billing information with information obtained from the display. Responses did not vary greatly with the other trial variables except that, unsurprisingly, prepayment customers were less likely to look at a bill at all. Overall, 30% thought the bill information added to the visual display because either it was easier to read or it was possible to sit and study it. A further 27% thought the bill reinforces the display information. The remaining 42% either did not look at information on the bill (12%) or did not think it provided anything more useful than the display information (30%).

Website

Overall, only 50% of respondents in smart meter groups were aware of the website, only 9% used it (and only 20% of website users used it once a month or, in the case of one user, more often). This clearly limits the capacity of the intervention to have any impact.

Smart meters

Respondents were generally satisfied with the explanation provided at the time of installation about how their smart meter worked – 68% were very or somewhat satisfied in groups that did not also have an RTD but this rose to 81% and 83% for credit and prepayment customers respectively with an RTD. Differences in recall of the smart meter information booklet were more distinct between these three groups: 41%, 73% and 85% respectively. The RTD appears to have had a positive impact on awareness of the smart meter.

Householders generally reported more positive changes in attitudes and awareness in the groups that received smart meters, particularly if they also had an RTD.

Prepayment customers were asked to compare their smart meter with their old meter – 57% though the smart meter was better, 11% that the old meter was better. Also, 56% thought the smart meter had been quite or very useful in helping them to save money on their energy bills and the same percentage would be quite or very likely to recommend it to their family and friends.

RTDs

In an early survey of households sent the clip-on RTD, 83% of those responding to the survey had received the RTD and 46% had received and fitted it (53% of those who received it). Only 31% were still using it. This limits the potential of the RTD to affect consumption.

Of those who had installed the RTD, most (86%) thought it had been easy to do but 10% thought the instructions were not clear. Overall, 17% of those who had received the RTD thought it complicated, or thought the instructions were not clear. This further reduces the potential of the RTD to affect consumption and emphasises the benefit of RTDs being fitted and explained by an installer.

The final survey suggests that the RTD provided with smart meters was viewed far more frequently than the clip-on device provided with non-smart meters: 37% vs 26% every day, 62% vs 41% at least once a week. This is important in terms of limiting the capacity of the intervention to have any effect on consumption and this is reinforced by ratings of usefulness. All respondents who had an RTD were asked about its usefulness, regardless of whether they were using it or not. Again the groups with smart meters were more likely to say the RTD was quite or very useful: 75% of credit customers and 70% of prepayment customers with smart meters vs 50% of those without.

Regarding specific aspects of the display, cost information was rated more highly than power/energy information. Where available, the traffic light display was seen as the most useful feature – 32% of credit customers and 42% of prepayment customers rated it the most useful feature and 64% of each rated it useful.

Turning to the reasons why the RTD was useful, responses were similar in each group. The percentages agreeing with a set of possible reasons were: “It provides reassurance that I’m doing the right things to reduce my energy” (73%), “It gives me more control over my energy usage” (68%), “It helps me to encourage other members of the household" (49%), "It means I have more influence on my bills" (63%), "I would miss it if it was taken away" (62%). This could be incorporated into material used to encourage uptake of RTDs.

Of those who had an RTD but were no longer using it, 42% non-smart meter groups and 32% of credit customers with smart meters had never looked at it whereas 9% and 24% respectively had used in for a few months, 49% and 43% for shorter periods. In the smart meter groups, the dominant reasons for not using the RTD related to the usefulness of the information provided – in general or after a period of use. In non-smart meter groups, the issue is more often the functionality of the device itself (except that the greater complexity of the smart meter RTD perhaps lead to a greater problem with knowing "what the buttons do"). This emphasises the need to explain how to use the information provided by RTDs linked to smart meters, not just how to access the information.

RTDs were most often located in the kitchen: 29% of clip-on RTDs and a higher percentage of RTDs provided with smart meters – 52% for credit customers (Project 4) and 47% in the prepayment trial group. Other common locations were the lounge (13% overall), hall (11%) and in a drawer/cupboard (8%). This suggests where householders find it most useful to keep the RTD, which could be incorporated into user guidance.

Incentive to reduce demand

While 72% of those in the "Incentive to reduce" groups were aware of the incentive; 37% of them never met the target and a further 9% did not know whether they had ever met the target (46% in total); 14% always or mostly met the target and a total of 45% met it more than once. The target may have been too challenging (or progress too difficult to monitor) to have any effect.

Incentive to shift

With 83% of those in the "Incentive to shift" groups aware of the incentive, this is a high level of awareness relative to other interventions, although only 75% were aware and had some understanding of how the incentive worked. Of those who were aware, 40% thought they had shifted consumption and made savings, 33% that they had shifted but without making savings and 28% that they had not shifted consumption.

This intervention was generally perceived as complex and this may be worse when it is combined with the incentive to reduce because cost saving is confused with energy saving.

Customers who had been in neither incentive group were asked "How much cheaper than the peak daytime tariff would the night tariff have to be to encourage you to move some of your consumption?" Group average figures ranged from 19-32% (the mean of the group averages was 25%).

8 EDRP Findings in the Context of the Literature

8.1 Introduction

8.1.1 Structure of the chapter

This section sets the EDRP findings in the context of the wider literature on interventions to change energy-related behaviour in homes. In doing so, it draws on the findings by the suppliers themselves (from Appendix A), further analysis by AECOM (in Sections 4-7) and the review of the literature in Appendix C.[39]

The EDRP trials employed a range of interventions that have been tested in other studies too, in a variety of forms, with different experimental approaches and with varying degrees of success.[40] The interventions are reviewed one at a time in the following subsections, grouped as follows.

- Energy efficiency advice (on paper or TV/web-based).
- Historic feedback: additional bill data (on paper or TV/web-based), such as comparison of energy consumption with earlier periods, and monthly bills.
- Benchmarking the customer's consumption against a peer group of comparable households.
- Customer engagement using commitment to reduce consumption.
- Smart meters (sometimes including accurate bills and no calls by meter readers).
- Real-time display (RTD) devices (including usage reduction alert).
- Heating controller integrated with RTD.
- Time of use tariff/incentive.
- Incentive to reduce consumption.
- Overview of web-based interventions.

Some of the interventions were tested both with and without smart meters, others only with smart meters and one intervention (benchmarking) only without smart meters. With two exceptions, there was no significant reduction in energy consumption when the intervention did not include a smart meter. The exceptions are clip-on RTDs and benchmarking; in these cases it was only SSE that found a significant reduction (in electricity consumption, not gas consumption), and the effect was small. In contrast, interventions using smart meters were successful more frequently and with larger percentage savings (particularly the combination of smart meters and RTDs, but also smart meters with no other intervention). These findings are detailed in the following subsections.

In each subsection, first the key conclusions of the literature review are stated. These conclusions are focused on the kinds of interventions used in EDRP, and on the situation in the UK. This section does not cite the literature directly – readers should refer to Appendix C to view the details of the evidence. The conclusions should not be taken as applying generally to all contexts throughout the world. This focus is problematic insofar as much of the evidence comes from outside the UK and it has been necessary to make extrapolations and assumptions in order to draw conclusions about the UK. The related findings from EDRP for the particular intervention are then summarised. Finally, comments are made on the consistency of the findings and the overall conclusions for each intervention.

[39] A literature review is subject to revision as new work is published – this is always true but particularly relevant to the current review, where a conclusion may depend on one or two key findings. In this situation, one contrary new finding can change the picture dramatically, which one would not expect to happen if a conclusion is supported by a large number of existing findings.

[40] Other studies have tested interventions that were not part of EDRP (e.g. regulation, labelling, accreditation, subsidies, loans, grants, education, appeals, taxation and energy price structure). These are not reviewed in detail but it is important not to lose sight of them in the context of a complete package of options to maximise the energy-saving benefits of a smart meter roll-out.

8.1.2 Behaviour change

The science of behaviour change crosses boundaries of psychology, sociology, ergonomics and economics; it also interlinks with engineering and product design. For this reason, there have been multiple models of behaviour and behaviour change, each with different emphasis and often using similar terms with different meaning, or different terms with similar meaning. This report simplifies the theoretical framework to allow easy comparison of the roles of different types of intervention. The framework used is based on the *means*, *motive* and *opportunity* for change.

- The *means* is the technology (a characteristic of the building fabric or services) or behaviour that will lead to reduced energy use and/or carbon dioxide (CO_2) emissions. This includes the person making the change having knowledge about current consumption and the technology or behaviour that would reduce consumption.
- The *motive* is the reason why households will want to make the change.
- The *opportunity* is the resource (e.g. time, space or money) to make the change.

In other words, for householders to reduce energy demand, they must know what to do, have a reason for doing it and have the resources to do it.

Money appears twice in the framework – as both motive and an essential element in opportunity. But other motives are also important – both those that relate to environmental impact and those that do not (e.g. social influences or achieving a personal sense of control over energy use).

Table 8.1 shows the principal expected contributions of each type of intervention. It is clear that no intervention offers the full package and that combinations therefore need to be designed that will complement each other. This is evidenced in specific cases in the following sections but the table allows other possibilities to be identified, particularly where supplementary interventions need to be identified to support a main intervention.

Table 8.1 The principal means, motive and opportunity contributions of each intervention

	Means			*Motive*			
	Technical information	*Behavioural information*	*Technology*	*Environmental*	*Financial*	*Other*	*Opportunity*
Energy efficiency advice	**✓**	**✓**		✓	✓	✓	✓
Historic feedback	✓	✓		**✓**	**✓**		
Benchmarking				✓	✓	**✓**	
Customer engagement using targets				✓	✓	**✓**	
Smart meter	✓	✓		✓	✓	**✓**	
RTD	✓	✓		✓	**✓**	✓	✓
RTD with heating controller	✓	✓	**✓**	✓	**✓**	✓	✓
Time of use tariff/incentive				✓	**✓**		
Incentive to reduce consumption					**✓**		✓

Large ticks indicate the likely major elements, small ticks the other likely elements.

A wide range of behaviours affect domestic energy demand. These behaviours may be characterised in various ways but one that is helpful for current purposes is based on "*opportunity*": the resources required to execute or to change the behaviour. Behaviours vary in the time required and in the cost associated with them, as shown in

Table 8.2. This simple categorisation is for convenience only: both cost and time are continuous variables and some behaviours may even be described as “no cost”. It is also important to understand that cost and time may be perceived quite differently according to the amount of money and spare time a householder has, and the facilities available in the home. Turning down a thermostat costs nothing if there is already a thermostat; it costs more if there is a central heating system without a (functional) thermostat. It costs even more if there is no central heating system at all.

Table 8.2 Examples of behaviours classified according to these two resource demands

	Low cost	*High cost*
High time demand	Insulate loft	Insulate solid external walls
Low time demand	Reduce temperature setting on central heating room thermostat	Replace boiler with a more efficient model

The third key element of opportunity (space) is also relevant in some cases. The cost of installing a heat pump, for example, depends on the amount of land available, and the financial return on solar power depends on the size, pitch and orientation of the roof. Critically, it matters whether the householders own the space – i.e. whether they are owner occupiers.

Cost and time will also depend on the knowledge and skills of the householder. The cost of insulating a loft, for example, is higher if the householder is not physically able to do it him/herself (but paying someone else to do it will reduce the time demand). The time demand is greater if s/he has to learn how to do it, clear a lot of stored items out of the loft, then (because of inexperience) work at a slow pace to ensure everything is done correctly.

Hence, a simple division of energy efficiency measures into “behaviour” and “installation” does not adequately represent the options. Insulation is not a behaviour, but installing insulation is a behaviour. Turning down a thermostat is a behaviour but it requires the technology to be installed.

The literature review conducted for this project shows that most changes seen in trials of energy demand reduction incentives tend to be – in the short term at least – those that require little investment of time or money. The EDRP customer surveys document a wide range of behaviours that householders used to reduce energy demand. While they do not allow specific changes in behaviour to be definitively tied to specific supplier interventions, the resulting changes in energy demand have been demonstrated more clearly than in past UK studies and – in some cases – for the first time.

8.2 Energy efficiency advice

8.2.1 The literature

Principles

Advice is an essential element of the *means* for change. In short, if people do not know how to save energy, it is unlikely they will do it. But advice itself is not sufficient: it can provide a degree of *motive* and/or help people to create *opportunity* (by facilitating access to the energy-saving *means* that the advice relates to, e.g. advising loft insulation and then offering grants for installation) but these elements tend to come mainly from other sources. The success of advice as an intervention is therefore, logically, dependent on motive and opportunity being either already present or created by some other intervention. It is also dependent on the householder not already being familiar with the information provided.

Advice can be delivered at a range of levels, through various media, and in combination with one or more other interventions; there is no reason why each approach should be equally successful in the population as a whole or with particular individuals or groups. The EDRP trials used generic written advice, delivered directly to households, mainly on paper but also testing the impact of advice via the web, an RTD or the participant's TV. Advice was sometimes deployed on its own and sometimes in combination with other interventions.

Evidence

There is little evidence that energy demand is reduced by household-level generic advice (i.e. information delivered to households but not tailored to the particular household) on its own. This may be because the evidence base is very limited, few studies having delivered such advice without supplementary interventions. When used as a supplement to other interventions, generic advice can increase energy savings by perhaps 5%. Although again the evidence is thin, where it is not feasible to give personal verbal advice, delivering advice by means more imaginative than leaflets and booklets may achieve higher savings – for example, using acted scenarios or providing written advice in conjunction with home energy audits.

In summary, the limited evidence (little of it from the UK) suggests that advice may be necessary (unless the recipients are being told something they already know) but is rarely sufficient on its own to bring about reductions in energy demand.

In principle, the effectiveness of advice should depend on the extent to which the originator is trusted (trusted to be reliable as to the facts, competent to deliver, honest in expression and transparent in motive). Hence, even the best designed advice could fail if the target audience does not trust the source. This makes it problematic to judge effectiveness merely from a description of its content and medium. While generalisations about Government and/or energy suppliers not being trusted can inform interpretation of research findings, they are not definitive. Nevertheless, the challenge for suppliers is certainly twofold – to convince customers that their advice can be trusted and to motivate customers to read and apply it.

8.2.2 EDRP findings

Non-smart meter interventions

None of the trials showed a significant reduction in gas or electricity consumption by providing advice (on paper or on the web), on its own or in combination with other interventions, in the absence of a smart meter.

Smart meter interventions

Where a smart meter was provided, there was no trial of providing advice alone, in the absence of other interventions.

Where advice was provided along with other interventions, in most cases there was no significant effect of the combination. The factorial analysis that was possible with the SSE data showed no significant effect of advice or any interaction effect with other variables. The Scottish Power trial design did not allow the effects of advice to be assessed independently of other interventions but all the trial groups received advice and no significant savings were seen in any phase until financial incentives were offered.

In the E.ON trial, significant reductions in electricity consumption were seen in all high use dual fuel (HU_{DF}) groups, regardless of whether advice or any other intervention was provided in addition to the smart meter. Hence this does not indicate an effect of advice; also findings in the HU_{DF} groups generally are most likely to be due to regression to the mean rather than the effect of any particular intervention. Significant reductions in electricity consumption were also seen in two of the other three customer strata that had an RTD along with a smart meter, additional bill data and advice, but not in the same strata in the absence of an RTD. Again, advice could not be seen as the key component but it is uncertain whether RTDs would have had the same impact without advice also being given.

Reductions in gas consumption were significant in every E.ON intervention involving a smart meter, so no specific effect of advice can be imputed.

The evidence from EDF is more complicated to unravel but does show the most positive indication from any of the trials of the benefit of advice. EDF provided advice in all its seven smart meter interventions, although it was minimal in the “time of use tariff” (TOUT) group. Analysis of changes in gas consumption was not possible for any group because baseline data were not available. There were significant reductions in electricity consumption in three groups.

- Paper: Additional bill data + Energy efficiency advice sent by post. Savings of 2.3% overall in the first in-trial year but significant for dual fuel customers (4.6%), not electricity-only customers (0.9%).
- Wall Panel: Real time display (“Wall Panel”) + Energy efficiency advice (sent via the RTD). Savings of 4.0% overall in the first in-trial year but significant for electricity-only customers (7.2%), not dual fuel customers (1.0%).
- Usage Reduction Alert (URA): Real time display (“Wall Panel”) + Usage reduction alert + Energy efficiency advice (sent via the RTD). Savings of 7.3% overall in the first in-trial year for electricity-only customers. There were no dual fuel customers in this group.

In two other groups, savings were not significant.

- HEC: Real time display (“Wall Panel”) + Heating controller + Energy efficiency advice (sent via the RTD). Dual fuel customers only.
- Web: Web consumption information + Energy efficiency advice: monthly tips online. Dual fuel customers only.

Two other groups did not have sufficient numbers to allow analysis.

- TOUT: Real time display (“basic display”) + Time of use tariff + Minimal energy efficiency advice (sent by post). Electricity-only.
- TV: TV consumption information + Energy efficiency advice (sent via the TV). Dual fuel + electricity-only.

Taking these findings overall, savings among electricity-only customers in the Wall Panel and URA group can most logically be attributed to the RTDs provided, although the advice may well have played a facilitating role. These effects are therefore considered under the RTD heading, in Section 8.7.

The principal finding to be considered here is therefore for the Paper group.[41] It may have been advice or the additional data on bills that was more important, or the particular combination could have been critical to the success of this intervention.

No definitive explanation can be offered for the fact that the savings were greater for dual fuel customers but a number of possibilities may be considered:

- dual fuel customers are more likely to see the advice as being relevant to them, because it is more likely that at least one specific piece of advice will apply (this may require the further assumption that, where an electricity-

[41] The TV trial also yielded some useful survey findings. The limited use that customers made of TV means that its effects on consumption cannot be fairly judged until more effective implementation is achieved, making the interventions less dependent on the existing technology in the home, more easily accessible, more focused on the key information that users want to see, and with better linkage between advice, consumption data (real-time and historic) and access to external financial and technical support.

only customer has gas supplied by another company, the advice tends to be thought of as applying to EDF services);

- dual fuel customers have made a market choice to take both fuels from one supplier – specifically EDF – and might therefore be more energy-aware (or energy-cost-aware) and willing to look at advice from EDF;
- the two customer types differed in how easily they could be recruited, and hence in their motivation at the start of the trial.

These are *post hoc* explanations and should be tested further if the distinction between customer types becomes relevant to key decisions.

Persistence of effects

In the EDF Paper group, the reduction in electricity demand remained significant in the second in-trial year, for dual fuel customers (2.4% saving, $p<0.05$) and both customer types combined (4.0% saving, $p<0.01$), and became significant for electricity-only customers (5.0% saving, $p<0.05$).

In the E.ON trial, reductions in gas consumption persisted to in-trial year 1 plus 1 quarter and, in some cases to in-trial year 1 plus 2 quarters (with a trend for some strengthening of the effect in in-trial year 1 plus 1 quarter). The only effect that persisted through to in-trial year 1 plus 3 quarters was in TG8 (smart meter, monthly bills and energy advice) for FP. This may be taken as advice enhancing the persistence of the effect but the advice was not introduced until the final quarter of the first in-trial year. Hence it is more likely an effect of a new element of the intervention rekindling interest in energy savings.

Findings from customer surveys

EDF

- Core triggers to uptake are emerging: *cost* savings, desire for *control* and less *hassle*. The immediacy of Wall Panel, HEC and TOUT was preferred to retrospective information in the Paper and TV groups. Focus group evaluations of the interventions placed Paper and TV in the bottom three.
- Paper group customers indicated very little engagement with the information provided and yet it appears to have been enough to bring about a reduction in consumption.[42] This is perhaps an indication that the amount of engagement is not a good guide to whether effective action will be taken. Responding to a few pieces of the advice is all that is needed.
- Records of specific actions taken were difficult to interpret with great confidence but, in two cases of energy-saving action (loft and cavity wall insulation), Paper was one of three groups with a (non-significantly) higher than average percentage taking the action. TV was in the top three for draughtproofing. These are the less expensive actions, with financial support often available, which might therefore be carried out quickly. In the case of double glazing – a more expensive action – TV stood out (along with HEC). There is also a non-significant trend for increase in number of appliances to be less likely in the Paper group.
- When using the TV to access information, only the first or second screens in the access sequence tend to be used, so the most useful information should be put on those screens.

[42] The delivery of the Paper intervention was less reliable in the second in-trial year and this may have unduly influenced the survey responses, since the survey was at the end of the trial. It is possible that engagement had been greater in the first in-trial year.

E.ON

- Reports of actions taken suggest that energy advice was affecting behaviour by the end of the trial, perhaps too late to be seen in the consumption data. This was more distinct in the smart meter groups (which had delayed commencement of advice) than in the non-smart meter groups (which had a mid-trial break in delivery of advice).

Scottish Power

- Although 37% said the mailings were telling them nothing new, 66% would like to receive more of the same information, and qualitative research indicated that people wanted to be kept regularly informed on energy matters and relating to energy use (in a personal rather than generic way), as a reminder. The implication is that repetition of advice is not necessarily a problem.
- Customer mailings were an important element in the Scottish Power trial and their quality was monitored and improved through the course of the trial. Recall and comprehension during Phase 3 was the highest achieved in the project (and similar among credit and prepayment customers – the latter had recalled the material less well in earlier Phases). Since it was Phase 3 that introduced financial benefits, it may have been this (rather than changes in style) that was responsible. The implication is that advice may have been less well received earlier in the trial, when it was a more substantial component of the overall package of interventions.

SSE

- Credit customers with smart meters were asked whether they recalled receiving the advice booklet and a similar percentage (80%) did, whether or not they had also received an RTD. However, customers with an RTD were more likely to still have the booklet (59% vs 79%), still refer to it (8% vs 15%) and say it was quite or very useful (78% vs 86%). Having an RTD appears to have brought about greater interest in the advice booklet.
- The most highly rated specific aspects of information were “Tells you which appliances are high energy users“ (30%), “Tips on saving money/energy - unplug appliances” (8%) and information on insulation (4%).
- Reasons for not finding the booklet useful are equally important in guiding the future design of advice. Among those who had received information but not found it useful, the most common reasons were: “Tells you nothing new” (44%), “It’s common sense” (21%), “Couldn’t be bothered to read it” (11%), “Couldn’t use less / already save energy” (6%), “Too much information to take in” (6%). The general theme is that the information was redundant or too much for the user to work through.

8.2.3 Conclusions

The EDRP findings are consistent with the literature insofar as an effect of energy efficiency advice was not always seen and, when it was seen, the reduction in annual consumption was up to 5%. This was important to confirm in the UK context, given the previous dependence on evidence from elsewhere.

The particular combination of advice and historic feedback on consumption that EDF deployed (along with smart meters) reduced electricity consumption by 2.3% overall in the first in-trial year and this was statistically significant. When dual fuel and electricity-only customers were considered separately, the reduction was significant only for dual fuel customers (4.6% saving) and not for electricity-only customers (0.9% saving).[43] Furthermore, the effect was persistent into the second in-trial year for both customer types combined (4.0% saving) and dual fuel customers (2.4% saving), and became statistically significant for electricity-only customers (5.0% saving). The E.ON trial found some weaker evidence of savings due to advice, in combination with smart meters and historic feedback on consumption, but it is difficult to quantify because of the particular schedule of interventions.

[43] Changes in gas consumption could not be assessed because of lack of baseline data.

There was no evidence of any effect in any trial when smart meters were not provided. In other trials with smart meters, SSE showed no significant effect of advice and, whilst the Scottish Power trial design did not allow the effects of advice to be assessed independently of other interventions, the results also suggested no significant effect of advice.

The effects observed in the EDF trial were seen in spite of survey findings that there was low customer engagement with the EDF energy efficiency and consumption history material. This perhaps gives a clue as to why the EDF trial was effective: information was provided in simple, short statements, over a period of time – minimal but well presented and easy to absorb a little each month. The SSE advice booklet was more comprehensive but required more effort from householders. An interesting twist was that there was greater engagement with the SSE booklet if an RTD was provided, so some of the effect attributed to the RTD may have been due to the RTD prompting interest in the advice and, conversely, the advice helping householders to use the feedback from the RTD.

The message is that advice should be provided but the details of delivery (e.g. clarity, quantity of information, timing) and combination with other interventions, are critical. The credibility of advice can also be undermined if it carries risks (e.g. blanket advice to draughtproof homes, regardless of potential for increasing risk of condensation or build-up of toxic combustion gases) or its effectiveness is uncertain. EDRP has taken a step forward in understanding these details but there is more work to do and the optimum approach is likely to vary between customer segments and to change over time as more energy knowledge becomes commonplace.

8.3 Historic feedback

8.3.1 The literature

Principles

While advice plays a key role in helping people to understand how to reduce energy use, another important ingredient is feedback, i.e. for consumers to know how much energy they have been using (and, ideally, when and how they have been using it). Historic feedback means access to past consumption data (where the past could be yesterday or last year). There are many permutations of how such feedback may be provided – it can vary in frequency, duration, immediacy, content (kWh, cost, CO_2, etc.), breakdown (by time, space and appliance), medium (and details of the medium, such as aesthetics and simplicity of access or use), comparisons (historic or normative), and combination with other instruments. The standard quarterly billing procedure (often including estimates of consumption) or monthly identical payments (based on estimated annual consumption) do not give high grade feedback.

Benchmarking a customer’s consumption against a sample of “peer group” comparable homes is considered separately in Section 8.4 although it is a form of historic feedback, because the dynamics of how it operates are distinct.

Real-time feedback is also discussed later (Section 8.7). The distinction is important because, although there is a general finding that households take a positive view of feedback, it matters how detailed it is and how closely linked to specific actions, in time and in level of disaggregation. Logically, aggregated feedback (e.g. quarterly or annual consumption) is more relevant to one-off changes that have a persistent impact, such as installing insulation or upgrading a heating system. More fine-grain, real-time feedback is more relevant to routine behaviour and purchases of equipment used intermittently (e.g. washing machines, televisions). By extension, aggregated feedback may be more relevant to the fuel used for heating (most often gas) and real-time feedback to electricity.

EDRP tested several modifications to the information customers receive with (or about) their bills. Where a smart meter was not fitted, suppliers provided historic comparison data with bills or asked customers to read their own meters each month. With smart meters, a wider range of options was possible, including bills being more accurate and/or more frequent, and providing more detailed breakdown of consumption across the day or the year. The extra bill-related information was delivered variously on paper, via the web or through the customer's TV. RTDs also offer historic feedback but this is considered later, as part of the whole package of information offered by RTDs.

Such feedback serves mainly to enhance *motive* – specifically through raising consumers' awareness of their energy use and, potentially, indirectly through improving the relationship with the supplier. Consumers can also, in principle, use the information to provide the *means*, i.e. to identify technical or behavioural changes (e.g. buying a bigger TV or switching from baths to showers) that have contributed to a change in consumption. But this depends on having sufficient existing knowledge and motivation, in addition to time (even if only a little time is needed, it can represent an *opportunity* barrier).

The general assumption has been that feedback will allow a consumer to make better informed choices and, critically, that the consumer will choose to use less energy. However, consumers may instead decide that the energy costs are so low that it is not worth the effort to reduce consumption, and may even see that the cost is much lower than s/he previously thought, and it is therefore OK to use more energy. In any sample of households, it is possible that all three outcomes will occur in different households. Add to this the fact that people are not accustomed to getting historic comparison data and may be unsure what to do with it, and the uncertainty of impact increases. Furthermore, there are likely to be people who already think that energy costs too much and therefore perceive enhanced billing information as "rubbing it in" rather than being helpful.

Evidence

One clear characteristic of the effect of historic feedback is its variability. This variability is not surprising, given the wide range of forms and contexts of feedback.

Few past studies have quantified the effect of enhanced billing in isolation (which is logical from the perspective that billing in itself does not provide a complete *means-motive-opportunity* package). Furthermore, the evidence comes mainly from cold climates. Hence it is difficult to determine what a "normal" impact would be for the purpose of comparison with EDRP. Fortunately, a range of other studies provide evidence on different aspects of what could, in theory, be achieved with enhanced billing. These studies fall into four main groups – studies of:

- households given historic feedback on consumption as part of normal business but not with the bill;
- households given web-based access to historic feedback on consumption;
- households given historic feedback on consumption for research purposes;
- households reading their own meters.

Taking the evidence as a whole, enhanced billing has the potential to reduce energy consumption but its effectiveness will depend on the details of the enhancement, the match to what customers want and understand, and the extent to which they are motivated to absorb and use the new information. The savings achievable from enhanced routine billing (at intervals of a month or more) appear to vary with national context (although the evidence comes mainly from cold climates): internationally, savings have been estimated at 5.5% energy savings but this seems optimistic in relation to routine monthly billing in the UK, 2-3% being more likely.

More frequent feedback (e.g. weekly) may achieve greater savings, particularly if householders are focusing for a limited period on making energy savings. While this may be impractical through the billing system, it could become feasible with wider use of smart meters and households having online access to their consumption data.

The alternative of encouraging householders to read their own meters and keep track of consumption may well be more effective. It is difficult to draw definitive conclusions from the literature because various combinations of interventions are involved, in various countries. However, the savings are generally more consistent than for studies without customer meter readings. The potential savings are perhaps up to 5% for electricity and more for gas. The difference between electricity and gas is consistent with the different levels of feedback granularity that are relevant to each, as noted above. It is not unreasonable to say that reading one's own meter is an active engagement with energy use that could both focus the mind of the customer and produce useful feedback, even if the reading is not used for billing purposes.

There may also be indirect benefits of enhanced billing. It can improve householders' sense of control over their energy costs, potentially leading to improved relations with energy suppliers, which might itself be a facilitator for other interventions.

Enhanced billing is relatively cheap to implement and can be deployed on an opt-out basis rather than opt-in, and could therefore be the most cost-effective type of feedback intervention. However, the distinction needs to be made between active opt-out (e.g. where someone declines the offer of enhanced billing) and passive opt-out (e.g. where additional billing information is received but not read). There are significant barriers to be overcome in delivering information with bills in such a way that customers notice, read, understand and apply the information. Attempts to deliver the information by other means have focused on the web but these have not yet been developed to point at which energy conservation benefits can be demonstrated.

8.3.2 EDRP findings

Non-smart meter interventions

None of the trials showed a significant reduction in gas or electricity consumption by providing historic feedback (on paper), on its own or in combination with other interventions, in the absence of a smart meter.

Regarding self-reading of meters, only the EDF "read-reduce-reward" (RRR) intervention tested this: customers read their own meters and could get a financial reward for reducing consumption. No baseline data were available and other data issues had been identified for this trial group meaning that no clear conclusion can be drawn as to its effectiveness.

Smart meter interventions

Where a smart meter was provided, in most cases there was no effect of providing historic feedback (on paper or via the web or a TV), alone or in combination with other interventions. The factorial analysis that was possible with the SSE data showed no significant effect of historic feedback or any interaction effect with other variables.

The Scottish Power trial design did not allow the effects of historic feedback to be assessed independently of other interventions but all the trial groups received historic feedback and no significant savings were seen in any phase until financial incentives were offered. It might reasonably be said that the feedback sent to customers was an important element in the way the financial incentives operated but this cannot be proved or disproved from the data. In any case, the effects of the incentives themselves are in some doubt (see Section 8.10).

In the E.ON trial, significant reductions in electricity consumption were seen in all high user dual fuel (HU_{DF}) groups, regardless of whether historic feedback or any other intervention was provided in addition to the smart meter. Hence this does not indicate an effect of historic feedback; also findings in the HU_{DF} groups generally are most likely to be due to regression to the mean rather than the effect of any particular intervention. Significant reductions in electricity consumption were also seen in two of the other three customer strata that had an RTD along with a smart meter, historic feedback and advice, but not in the same strata in the absence of an RTD. Therefore, historic feedback could not be seen as the key component but it is uncertain whether RTDs would have had the same impact without historic feedback also being given.

Reductions in gas consumption were significant in every E.ON intervention involving a smart meter, so no specific effect of historic feedback can be imputed.

EDF combined historic feedback with energy efficiency advice in its Paper group, which was discussed in Section 8.2. In this group there were electricity savings of 2.3% overall in the first in-trial year but significant for dual fuel customers (4.6%), not electricity-only customers (0.9%). It may have been advice or the historic feedback that was more important, or the particular combination could have been critical to the success of this intervention. Analysis of changes in gas consumption was not possible because baseline data were not available.

Persistence of effects

In the EDF Paper group, the reduction in electricity demand remained significant in the second in-trial year, for dual fuel customers (2.4% saving, $p<0.05$) and both customer types combined (4.0% saving, $p<0.01$), and became significant for electricity-only customers (5.0% saving, $p<0.05$).

Findings from customer surveys

EDF

- Core triggers to uptake are emerging: *cost* savings, desire for *control* and less *hassle*. The immediacy of feedback from RTDs was preferred to retrospective information in the Paper and TV groups. Focus group evaluations of the interventions placed Paper and TV in the bottom three.
- RRR customers were not aware they were part of trial but reading their own meter appears to have raised some awareness about costs and consumption. Focus group evaluations of the interventions placed RRR in the middle two.
- Records of specific actions taken were difficult to interpret with great confidence but, in two cases of energy-saving action (loft and cavity wall insulation), Paper was one of three groups with a (non-significantly) higher than average percentage taking the action. These are the less expensive actions, with financial support often available, which might therefore be carried out quickly. There is also a non-significant trend for increase in number of appliances to be less likely in the Paper trial group.

Scottish Power

- Customer mailings were an important element in the Scottish Power trial and their quality was monitored and improved through the course of the trial. Recall and comprehension during Phase 3 was the highest achieved in the project (and similar among credit and prepayment customers – the latter had recalled the material less well in earlier Waves). Since it was Phase 3 that introduced financial benefits, it may have been this (rather than changes in style) that was responsible. The implication is that historic feedback may have been less well received earlier in the trial, when it was a more substantial component of the overall package of interventions.

SSE

- In an early survey of households receiving additional bill data (graphs on bills), only 32% recalled something different about the bill and 27% recalled what it was. Asked directly about the graphs, 48% said they had looked at the graphs in detail and 38% that it made them think about energy use. This limits the potential of the graphs to affect consumption. Nevertheless, for those who were aware of the graphs, the final survey provides some evidence on their role in customers’ thinking.
- Final survey respondents rated the usefulness of information on electricity bill and statements “in helping you to reduce your energy consumption”. The percentage responding quite or very useful was higher in the non-smart meter groups that had extra information on their bills (in addition to energy advice) than in the one that had advice only (60% vs 34%). The figures reported do not allow this comparison for smart meter groups.

- Where the extra information was provided, the items on the bill rated as most useful shifted from units used to the graphs or comparison with previous bills.
- Householders who had an RTD of some kind were asked to compare billing information with information obtained from the display. Responses did vary greatly with the other trial variables except that, unsurprisingly, prepayment customers were less likely to look at a bill at all. Overall, 30% thought the bill information added to the visual display because either it was easier to read or it was possible to sit and study it. A further 27% thought the bill reinforces the display information. The remaining 42% either did not look at information on the bill (12%) or did not think it provided anything more useful than the display information (30%).

8.3.3 Conclusions

The EDRP findings are consistent with the literature insofar as an effect of historic feedback was not always found and, when it was found, the impact on consumption was up to 5%. This was important to confirm in the UK context, given the previous dependence on evidence from elsewhere.

As noted in Section 8.2, the particular combination of advice and historic feedback on consumption that EDF deployed (along with smart meters) reduced electricity consumption by 2.3% overall in the first in-trial year but statistically significant only for dual fuel customers (4.6%), not electricity-only customers (0.9%).[44] Furthermore, the effect was persistent into the second in-trial year for both customers types combined (4.0% saving) and dual fuel customers (2.4%), and became significant for electricity-only customers (5.0%). The other trials found no evidence of savings due to historic feedback, with or without smart meters.

Effects were seen in the EDF trial in spite of survey findings that there was low customer engagement with the EDF material. The billing material was more obvious to customers than that provided by E.ON (large, in colour and on separate sheets) and, as noted in Section 8.2, the advice was more accessible than that provided by SSE. The SSE survey data did show that customers found the information on bills more useful if they also had advice, which helps to explain the effect on consumption found by EDF. This survey also showed that many customers found the bill data and RTDs to be complementary, with a value in providing both.

Historic feedback can be useful but the details of delivery, and combination with other interventions, are critical. EDRP has taken a step forward in understanding these details but there is more work to do and the optimum approach is likely to change over time as people become more familiar with their consumption levels, and to vary between customers, depending on their current understanding of energy use and interest in reducing consumption.

8.4 Benchmarking (comparative or normative feedback)

8.4.1 The literature

Principles

Benchmarking takes historic feedback a stage further, by showing customers how their consumption compares with other households that are in some way comparable. While typically associated with historic feedback, it could in principle be incorporated into real-time feedback.

Such comparative feedback may provoke competition, social comparison or peer pressure. In doing so, it shifts the mechanism of action towards more socially based *motive* but leaves some scope for environmental and financial motive. Only one of the EDRP trials used benchmarking, and not in isolation from other interventions.

---

[44] Gas consumption could not be assessed because of the lack of baseline data.

Implementation of benchmarking faces three substantial issues.

- A comparison group would ideally need to be of households of the same composition, in a similar financial position, in a similar dwelling in a similar climate – this difficult to achieve (especially with the limited information available to energy suppliers). Furthermore, this fact is sufficiently obvious that customers may be suspicious about the validity of their comparison group (perhaps more so for gas because it depends too much on the amount of time spent at home and the characteristics of the home).
- Households that are already below the benchmark values may find no reason to reduce consumption and might even increase consumption. This effect is perhaps best evidenced by success in eliminating it – there is evidence that negative impact may be overcome by use of simple normative messages or even minor changes in presentation such as the inclusion of “smiley” icons with low users’ bills.
- The more subtle issue is that people tend to place normative information low on their list of what they believe influences their behaviour and yet it can be high on the list of what actually influences their behaviour. Hence, normative messages need to overcome people’s belief that they are not influenced by them, which could lead to a tendency to ignore them.

Evidence

If these problems are well evidenced, there is less evidence on what works, since benchmarking tends to be combined with other interventions in a way that makes it difficult to isolate its effects. Both positive and negative effects, and no effect, have been observed and the evidence does not allow an estimate of likely savings to be calculated. In particular, there is some evidence from the USA of differences in impact, depending on the household’s political and environmental views – if confirmed, this would suggest it would be more efficient to target particular population segments with benchmarking data. One approach would be to base comparisons on groups that have something in common, e.g. because they have volunteered to join an online energy comparison group or local network.

8.4.2 EDRP findings

Non-smart meter interventions

Only SSE included benchmarking in its trial. There was a small but significant effect ($p<0.05$) on electricity consumption, comparing the trial group with advice, historic feedback and benchmarking information with the most similar group (which had the advice and historic feedback but not benchmarking information). The additional savings attributable to benchmarking were 1%. The same comparison for gas consumption did not show a significant effect.

Smart meter interventions

No smart meter interventions included benchmarking.

Persistence of effects

The persistence of the benchmarking effect could not be investigated because the in-trial period did not extend to a second year.

Findings from customer surveys

The SSE survey did not provide further insight into the benchmarking intervention.

8.4.3 Conclusions

Only the SSE trial used benchmarking. There was a small but significant effect on electricity consumption, comparing the trial group with advice, historic feedback and benchmarking information with the most similar group (which had the advice and historic feedback but not benchmarking information). Neither group had smart meters. The additional savings attributable to benchmarking were 1%.

Although a small effect, this is one of the clearest pieces of evidence for an effect of benchmarking, in the context of a literature that is generally positive but allows no quantification. There is concern in the literature that those who consume less than the benchmark amounts may start to consume more; therefore, careful consideration needs to be given to how benchmarking might usefully be deployed in the population in general.

8.5 Customer engagement using commitment to reduce consumption

8.5.1 The literature

Principles

The EDRP trials included some testing of a commitment to reduce consumption (without a specific target or financial incentive to fulfil the commitment). The principal route of impact for this kind of commitment would be to reinforce any existing motive to save, and possibly to introduce a social motive in terms of the relationship between the customer and the supplier. The impact of financial incentives to save is discussed separately in Section 8.10.

Evidence

There has been little theoretical impetus behind achieving savings through a general commitment to reduce consumption and, consequently, no clear evidence on what effect it is likely to have.

A few studies have given households an explicit savings target (without contingent financial reward, other than reduced energy costs). In most cases, the effect of the target cannot be isolated from other aspects of the intervention. In short-term trials, there is tentative evidence of a benefit of realistic but stretching targets, combined with frequent feedback, but there is insufficient evidence to quantify such an effect.

Similarly to benchmarking, commitments and target-setting might be most effective within groups that have something in common, e.g. because they have volunteered to join an online energy comparison group or local network.

8.5.2 EDRP findings

Non-smart meter interventions

The Scottish Power trial included a three-month period in which customers were invited to make a pledge to reduce gas consumption. This applied to prepayment customers with non-smart meters, who simultaneously had other interventions (principally financial incentives to reduce electricity use and, in one of the two trial groups, an RTD), so the impact of the pledge cannot independently be evaluated. There was no significant effect of the package of interventions on gas or electricity consumption.

SSE divided its sample into customers who were Unaware (not aware that they were in a trial), Aware (that they were in a trial) and Committed (aware and additionally made a commitment to reduce consumption, in writing or verbally by telephone). Testing for differences across all trial groups, SSE found a significant effect of awareness of the trial on electricity demand reduction (relative to baseline), this being higher for the Aware (6.4%) and Committed

(6.1%) groups than the Unaware group (4.0%). Commitment did not have an effect in addition to the effect of being aware of the trial. There was no interaction effect between awareness and any trial intervention, i.e. the intervention effects did not vary with awareness.

Smart meter interventions

The Scottish Power trial included a three-month period in which customers were invited to make a pledge to reduce gas consumption. This applied to credit customers with smart meters, who simultaneously had other interventions (principally financial incentives to reduce electricity use and, in two out of the three trial groups, an RTD), so the impact of the pledge cannot independently be evaluated. However, this was the trial period in which there was a reduction in gas consumption, not electricity consumption. The period of intervention was so short that the Hawthorne effect could explain any savings made.

Commitment to reduce consumption did not have a significant effect in the SSE trial, as noted above, under non-smart meter interventions.

Persistence of effects

The persistence of any commitment effect could not be investigated because the in-trial period did not extend to a second year.

Findings from customer surveys

The Scottish Power and SSE surveys did not provide further insight into the commitment interventions.

8.5.3 Conclusions

Commitment to reduce consumption did not have any reliable effect on consumption, with or without smart meters, other than an effect due to making customers aware of the trial. This is a new finding, with no comparable evidence in the literature.[45]

8.6 Smart meters

8.6.1 The literature

Principles

This section is about the impact of installing a smart meter without any other scheduled interventions. Merely having a smart meter installed is a minimal reason to expect any change in consumption: the main point is that the meter allows a range of other interventions, as described elsewhere in this review. The question here is whether the experience of getting a smart meter can itself influence energy use. If the experience is simply being told that a replacement meter has been installed, somewhere out of sight, the consequences should be small. But this is not necessarily the total experience.

In the course of the exercise, customers might experience any or all of:

- being told they are among the first to get the latest technology;
- renewed positive interaction with the supplier;

[45] In short-term trials, there is related tentative evidence of a benefit of realistic but stretching targets, combined with frequent feedback, but there is insufficient evidence to quantify such an effect.

- reassurance that the meter accommodation is now safer;
- a friendly or unfriendly installer;
- positive or negative attention from friends and neighbours;
- informal energy advice from the installer.

If the smart meter is easier to read than the meter it replaces, householders may read it more often, which could itself have a positive impact on energy demand (especially for gas, since one-time adjustments to heating or hot water boilers or controls, or new insulation, can have long-lasting effects). There may be merit in encouraging this as part of smart meter installation; it would be perhaps the least “high tech” application of smart meters but potentially one of the more effective in relation to energy demand reduction.

Any of these factors might have some effect on householders’ knowledge of *means* and/or their *motive* to save.

The EDRP trials did not set out to test installation effects but they arise through different approaches to recruitment and installation. The installation procedure, for example, ranged from trying to minimise any impression that the work was anything other than “business as usual” or promoting minimal interaction between installer and householder, to training installers to demonstrate the installed technology to householders.

Evidence

There is, as yet, insufficient evidence in the literature to evaluate the impact of the above factors.

8.6.2 EDRP findings

Non-smart meter interventions

This section is about smart meters so it may seem odd to report findings on non-smart meter interventions. However, it is worth making the point that EDRP tested various types of intervention without smart meters, with several variations in detail and in different combinations. Only one trial found any significant reductions in energy consumption, and then only a small effect (1% reduction in electricity use): SSE saw this in its tests of RTDs and of benchmarking. In general, interventions in the absence of a smart meter did not succeed. This may be explained partly by some aspect of receiving the smart meter (e.g. interaction with the installer or the positive image of getting new technology) but also the different options that were available once a smart meter was installed, e.g. more sophisticated RTDs (fitted by an installer), and more frequent and accurate historic feedback and billing.

Smart meter interventions

In contrast to interventions without a smart meter, a good number of interventions with a smart meter did succeed. This is the first key point about smart meters: they facilitate other interventions that would otherwise not be feasible. Other sections cover the detail of this; the remainder of this section is about whether smart meters, in the absence of other planned interventions, bring about a reduction in energy demand.

EDF used smart meters in the control group and there was no comparison with households that did not have a smart meter or any other intervention, so the effect of the smart meter itself cannot be determined.

Scottish Power also did not test the effect of smart meters without other interventions, and the design did not permit any comparison between groups that had identical interventions except for provision of a smart meter. The key point from this trial is perhaps that Scottish Power made strenuous efforts to prevent customers from knowing they had a smart meter – it was not presented to them as such, and bills continued to be based on business-as-usual meter readings. It is also debatable whether the interventions are best viewed as smart meter interventions; in that context, the generally null findings are more easily explained and have a greater value in the context of the whole of EDRP.

E.ON found that installing a smart meter, with no other intervention, had no effect on electricity consumption.[46] In contrast, reductions in gas consumption were significant in every intervention involving a smart meter, including the group having a smart meter alone (along with cessation of meter readings and an end to estimated bills). The range of gas savings was 2.3-4.4%.

SSE found that electricity savings attributable to having a smart meter were 2.5%. This is based on all comparisons between groups that did and did not have a smart meter: the simple effect of having a smart meter and no other intervention could not be tested with the available data. The 2.5% therefore represents the enhancement of other interventions by the smart meter. The level of enhancement varied with the combination of other interventions and reached 4.7% when combined with an advice booklet, graphs on bills and an RTD but the RTD provided with smart meters was of a higher specification than the one provided with non-smart meters. Without an RTD, the effect of having a smart meter was 3.3%.

Similarly, the gas savings attributable to having a smart meter were 3.0%. For gas consumption, there was not a significant effect of providing an RTD, so the effect was more clearly attributable to the smart meter itself.

Persistence of effects

Reductions in gas consumption in the E.ON trial persisted (in fact, slightly increased) to the first quarter of the second in-trial year (i.e. for 15 months) and for one or two further quarters in some groups. This suggests that the experience of getting the smart meter prompted some initial action (e.g. turning down a thermostat) but the effect may require support over time from other interventions (e.g. advice or billing information) to be sustained for longer periods.

The persistence of any effect could not be investigated in the SSE trial because the in-trial period did not extend to a second year.

Seasonal variation

The E.ON findings also demonstrated an important issue with comparing data between smart and non-smart meters. The smart meter groups show greater seasonal extremes than the non-smart control groups for both electricity and gas, with significantly lower consumption than the control group in almost every spring/summer quarter in every group, sometimes outweighing higher consumption in autumn/winter quarters.

This may be entirely due to “smoothing” of the non-smart data because of the interpolation procedure, with drift between reading and trial period. If so, then analysis at sub-annual level would be valid only when comparing smart meter trials with smart-meter-only groups treated as control groups. In the current assessment, it is assumed that smoothing is a sufficient account but this needs to be better understood for any future trials.

Findings from customer surveys

The customer surveys offer some insight into the installation process and perception of the smart meter among prepayment customers.

---

[46] Other than in the HU_{DF} groups, where any intervention had the same effect, most likely due to regression to the mean rather than the effect of any particular intervention.

EDF

- Customers expected, and could have benefited from, more engagement and instruction during installation of equipment.
- Although the trials have prompted some behavioural change, some respondents were happy to benefit from no more meter readers or estimated bills and nothing else.

SSE

- Respondents were generally satisfied with the explanation provided at the time of installation about how their smart meter worked – 68% were very or somewhat satisfied in groups that did not also have an RTD but this rose to 81% and 83% for credit and prepayment customers respectively with an RTD. Differences in recall of the smart meter information booklet were more distinct between these three groups: 41%, 73% and 85% respectively. It is uncertain whether the RTD made customers more aware of the smart meter and associated information, or customers thought of the RTD as being the smart meter.
- Prepayment customers were asked to compare their smart meter with their old meter – 57% thought the smart meter was better, 11% that the old meter was better. 56% thought the smart meter had been quite or very useful in helping them to save money on their energy bills and the same percentage would be quite or very likely to recommend it to their family and friends.

8.6.3 Conclusions

The literature does not provide direct evidence of the impact of installing a smart meter without any other scheduled interventions. Two of the EDRP trials (E.ON and SSE) provide the first evidence on this, showing that some aspect of the experience of getting a smart meter can itself prompt a reduction in energy consumption, particularly gas consumption (savings of around 3%). The clearer effect for gas consumption makes sense in the context that simple one-off changes (e.g. reducing a thermostat setting) can have big effects on gas demand.

Reductions in gas consumption in the E.ON trial persisted (in fact, slightly increased) to the first quarter of the second in-trial year (i.e. for 15 months) and for one or two further quarters in some groups. This suggests that the experience of getting the smart meter prompted some initial action (e.g. turning down a thermostat) but the effect may require support over time from other interventions (e.g. advice or billing information) to be sustained for longer periods.

The relevant feature of the customers’ experience of the smart meter installation cannot be stated with certainty but there was a clear difference between E.ON and SSE, which set out to explain to customers what they were receiving, and Scottish Power, which ran the trial as a blind trial (every effort was made to prevent customers knowing they had a smart meter, by presenting it as a routine meter replacement). While smart- and non-smart conditions could not be compared in the EDF trial, the survey data showed that (a) customers expected, and could have benefited from, more engagement and instruction during installation of equipment and (b) some respondents were happy to benefit from no more meter readers or estimated bills and nothing else.

8.7 Real-time display (RTD) devices

8.7.1 The literature

Principles

RTDs go a stage beyond historic feedback by showing the current rate of energy consumption, generally with the option of expressing this in kW, cost or CO_2 emissions. They often have other functions such as displaying historic consumption for various periods, temperature displays, alarms and simple visual signals to indicate high consumption rates. RTDs range from battery-powered devices that clip on to the live cable of an electricity meter (smart or not) to relatively sophisticated mains-powered devices, showing both electricity and gas consumption by using the signal from smart meters. The latter type is also capable of greater accuracy. The EDRP trials represent all these options, alone or in various combinations with other interventions.

As noted earlier, RTDs are probably most relevant to routine behaviour and purchases of equipment used intermittently (e.g. washing machines, televisions). Householders can check the power consumption or cost of using an item, at a point in time or over a short period, and make informed decisions such as whether to turn something off or down, or replace it. Consumers might see, for example, the jump in power when a kettle or cooker is switched on, and the actual relative demand of household appliances – in use and on standby – which can be very different from what they had assumed or been led to believe. Consumers can also use RTDs to confirm the benefit of changes they have made.

In this way, RTDs can support:

- *means* (by identifying what changes could be made) but only if the consumer understands how to do this, and has the knowledge to act upon observations;
- *motive* (by showing how much energy is being used) but only if the consumer already has sufficient motivation to want to check energy use (and possibly also by engaging other motives such as using or showing off new technology);
- *opportunity* (because they save householders time relative to taking a series of meter readings and calculating differences).[47]

Thus, RTDs offer all three elements of change but incompletely in each case and, logically, dependent on the extent to which households actually refer to them, particularly when they are first installed and there is most to learn. Generally, financial motivation is likely to be the strongest element.

As noted earlier, in any sample of households, the response to feedback may be choosing to use less energy but some consumers may decide that the energy costs are so low that it is not worth the effort, and others may even see that the cost is much lower than s/he previously thought, and it is therefore OK to use more energy.

Evidence

This review has drawn on extensive evidence from simple trials that compared households provided with an RTD with a control group that did not have real-time feedback, more complex trials that combined an RTD with other interventions and web sites that provided real-time feedback. Some evidence from studies of student accommodation is also used and two related types of real-time feedback are considered – usage alarms and thermometers. The evidence is based on both quantified energy savings and studies of user response to RTDs.

Energy demand

Most of the published research involves ‘clip-on’ RTDs, used without smart meters. One implication of this is that most studies have been of electricity consumption only, not gas. There are also some examples of studies using main-powered meters or real-time feedback via the web.

---

[47] It could be argued that this is theoretical because householders generally do not use their meters in this way. The alternative view is that the opportunity benefit accrues from the time required to make checks being shortened to the point where more householders do it at all. So, although householders then take more time for checking, this is because it is now seen as better value use of time.

Most of the research tends to focus on family homes that are owner-occupied. Many used relatively small samples and almost all were, to some extent, ‘opt-in’ trials or experiments. Opting in is likely to have some impact on the trial outcome. It is often assumed that people who opt in will be more “enthusiastic” about saving energy and will therefore reduce consumption more as a result of the additional feedback. There is a counter-argument that enthusiasts will have already done much of what the RTD indicates they should do, hence they have less capacity to reduce consumption further. Therefore, the question of the experimental impact of opting in becomes an empirical one; while the evidence is not definitive, there does not appear to be a strong overall influence of the take-up rate in opt-in trials. Either way, it may be that the optimum target is people who have not yet taken much interest in conserving energy but who could be motivated in the process of providing an RTD and informed how to use the device to fulfil their newfound motivation.

The impact of RTDs on energy demand is highly variable, depending on the design of the RTD itself, who is using it and how often. The evidence suggests that the uncertainty of impact can be reduced, and the positive impact increased, by combining provision of an RTD with other interventions – including energy efficiency advice but also something else. The “something else” needs to engage the household with using the RTD, for example through a savings target, a time-varying tariff or some other reason to self-monitor consumption.

This conclusion depends heavily on research outside the UK, little of it being even from temperate climates, so expected percentage savings in the UK are difficult to estimate. However, a base level effect of RTDs alone could be less than 3% electricity savings whereas supplementary interventions that increase engagement could double or triple the benefit. Far fewer studies have tested effects on gas consumption but they have generally shown a benefit; while it is not feasible to quantify it, savings tend to be of a similar order to those for electricity.

One caveat is that a small trial in California found a negative effect of providing an RTD to homes on a time of use tariff. This may be because feedback indicated how cheap electricity was during off-peak period, encouraging over-compensation for peak time savings.

There is little published evidence on the use of audible alarms or visible signals in tandem with RTDs. The limited evidence (generally qualitative) indicates that audible alarms tend not to be used but visible signals with clear meaning may be viewed positively and used as part of an energy-saving strategy. Examples in the literature are visual signals of high current usage (in relation to a baseline or target, similar to the ecoMeter “traffic lights”), the current price band (for households on variable tariffs), or to indicate that the heating was on or that the air conditioning could be turned off.

A thermometer may be a simple and effective feedback device but there is insufficient evidence to say whether it actually has any effect on energy demand.

Evidence on the provision of online real-time feedback is limited and, as yet, inconclusive.

It has been argued that targeting an RTD intervention on the households best equipped to use it (in terms of motivation and understanding) could multiply the benefit per targeted household, although the total benefit would necessarily be reduced and a wider range of households could benefit if given some additional support. In any case, careful attention to RTD clarity and reliability remains important.

RTD design and use

The literature provides useful insight into how RTDs are used, including the display features that people find most useful. This insight could be incorporated into future RTD design and guidance. The following points can be made about how RTDs are used.

- The first barrier to RTDs being effective is to have them actually installed, maintained in working order (e.g. by replacing batteries) and used, which makes the installation process and user guides critical to their success.

- Most of the energy savings from feedback programmes arise from simple changes in routine behaviours, rather than investments.
- A key use of RTDs is to identify specific uses of electricity so that related behaviour can be modified. For this reason, the option of appliance-specific monitoring is attractive to users.
- Users can experience a strong impact of seeing readings jump up and down as appliances are switched on or off. An RTD's effectiveness may depend on users getting to this point. Guidance needs to manage the associated risk that the visual impact of brief use of high wattage appliances may distract users from the potentially large consumption due to extended use of lower wattage appliances.
- The impact of RTDs is extended where participants go on to discuss the issues raised with their children.
- Being able to move around the home with an RTD is most useful when it is first acquired, working out the impact of different appliances. After this initial period, there is less concern about portability and the display tends to be left in a prominent position in the home.
- People will not reduce consumption for a particular application simply because the RTD shows that consumption is high: they have certain non-negotiable uses of energy.
- One specific strategy is to use the RTD to check whether everything is switched off at night. This may be the application that has the most persistent benefit, after all appliances have been checked and usage monitored over a period.
- While RTDs can be used to identify opportunities for savings, the more powerful use may be in confirming that actions to reduce energy demand have been successful: the RTD may be more effective in maintaining conservation behaviour than initiating it.
- External support may be needed but unavailable (or difficult to access), for example to inform purchase choices or to establish norms in relation to comparable households, or to obtain planning permission for more substantial installations.

The RTD and its instructions need to be kept simple; while more sophisticated devices may be imagined, it is more important that users have easy access to the aspects of the display they find most useful and are most likely to use (and that they can easily return to those aspects if they get lost in trying to access other information). The following points can also be noted.

- Cost information tends to be the preferred display information, and the most easily understood, but tracking energy use over time – as costs change – probably needs users to refer to kW and kWh, even if they do not fully understand these units.
- Both current and cumulative consumption figures have their different purposes.
- Few people value display of CO_2 emissions and, for any given household, they are of limited additional relevance because they are directly dependent on the energy used.
- Indications of consumption against a target level can be helpful if someone is trying to reduce consumption but can be a source of stress if money is tight – a constant reminder of money being spent.
- Although accurate numbers are important, a rapidly changing rate is better expressed as an analogue indicator.

But different market sectors have different preferences for RTD design and display features, giving priority variously to saving money, environmental benefit, acquiring and manipulating data and the 'new gadget' itself with technological and aesthetic appeal. Taking this a step further, it may be critical for installers to have a small set of options for how to present a device to different households.

The literature also gives some indication of who responds most positively to feedback:

- investment is more likely in higher income households and when moving into a new home;
- those who do invest tend to save most energy;
- higher savings are associated with higher levels of education or income; larger homes and households; strong environmental values; and younger people.

In short, RTDs show promising results in the short to medium term, perhaps more so with already motivated customers but sometimes with householders who previously showed little interest. However, the displays can also cause frustration in people who cannot use the information to improve their situation, and they can cause conflict in the household, for example when one member wants to make changes in thermostat settings or appliance usage and another does not. There are challenges ahead to widen the motivation to understand energy use, to use RTDs as part of wider programmes (e.g. community initiatives or retrofits), and to maintain interest, e.g. through time-varying tariffs.

8.7.2 EDRP findings

Non-smart meter interventions

All trials except EDF's tested the effects of clip-on RTDs with non-smart meters, providing feedback on electricity consumption but not gas. In combination with advice and/or additional bill data, or on their own, neither E.ON nor Scottish Power saw any significant effect on gas or electricity consumption. SSE also found no effect on gas consumption but did find a small significant effect ($p<0.05$) on electricity use: the reduction in consumption relative to the control group was 1%.

In one of the Phase three periods of the Scottish Power trial, a thermometer was also provided but the trial design does not allow an independent assessment of its effect.

Smart meter interventions

Scottish Power smart meter trial groups were first given advice and additional bill data, then a series of financial incentives. In each case, pairs of trial groups differed only in whether they had an RTD. In no case did the reduction in gas or electricity consumption depend on whether an RTD was provided. As noted earlier, it is a matter of definition whether these were truly smart meter groups because of the efforts made to prevent customers from knowing they had a smart meter. In one of the Phase three periods, a thermometer was also provided but the trial design does not allow an independent assessment of its effect.

The other trials each found significant effects of providing an RTD but always with some complication.

The EDF "Wall Panel" group combined an RTD showing both electricity and gas consumption data with energy efficiency advice delivered through the RTD. Savings for gas could not be estimated because of lack of baseline data but electricity savings were significant ($p<0.05$) and estimated at 4.0%. Not all groups that had advice reduced their consumption, so the reduction in this group is most likely attributable primarily to the RTD but advice may well have had a facilitating role.

The effect was greater for electricity-only customers (7.2%) than dual fuel customers (1.0%) and significant ($p<0.01$) only for the former. No definitive explanation can be offered for this, but a number of possibilities may be considered:

- if a customer buys only electricity from EDF, a gas smart meter would not be fitted and the RTD would display only electricity data, hence focusing conservation efforts on electricity rather than gas;

- focusing on electricity might also make the RTD itself simpler to operate;
- the two customer types differed in how easily they could be recruited, and hence in their motivation at the start of the trial.

These are *post hoc* explanations and should be tested further if the distinction between customer types becomes relevant to key decisions.

In the related Usage Reduction Alert (URA) trial, the RTD was set up to sound an alarm if the household exceeded a predefined daily consumption level. There was again a significant ($p<0.05$) reduction in energy use (7.3%) for electricity-only customers (dual fuel customers were not included in this trial).

The EDF trial of an RTD with an integral heating controller is considered in the next section. Similarly, the time of use tariff (TOUT) group is considered separately (Section 8.9); although a basic RTD was provided, this was not the main intervention.

E.ON found significant reductions in electricity consumption in all high use dual fuel (HU_{DF}) groups[48] but this is most likely to be due to regression to the mean rather than the effect of any particular intervention. Other than this, there were significant reductions in electricity consumption only when an RTD was provided (along with the other interventions) and this was seen in two of the three other customer strata. Compared with providing only a smart meter, the reductions were 2.8% for the ‘fuel poor’ (FP) group and 3.3% for Economy 7 (E7). The effect was in the same direction (2.0%) but not significant for not ‘fuel poor’ (NFP). Equally importantly, there were (mainly non-significant) increases in consumption in groups without an RTD. Since there was not a smart meter trial group with RTDs but not advice or monthly bills, it is uncertain whether RTDs would have had the same impact without these supporting interventions.

Reductions in gas consumption were significant in every E.ON intervention involving a smart meter, with no incremental effect of providing an RTD.

SSE found a significant effect of “meter type” (a combination of smart vs non-smart meter and whether or not an RTD was provided) on electricity consumption. This meant that savings relative to the control group were 2.5% with a smart meter, 3.6% with a smart meter and RTD and 2.9% with a prepayment smart meter and RTD. This is based on all comparisons between groups that did and did not have a smart meter or RTD. It therefore represents the enhancement of other interventions by the RTD. The level of enhancement varied with the combination of other interventions but, comparing two groups with smart meters, advice booklet and graphs on bills with a group that also had RTDs, the additional saving was 1.8%.

In contrast, there was no main effect of the RTD on gas consumption in addition to the effect of providing a smart meter, but there was a significant interaction between incentive to reduce and provision of an RTD. This means that the incentive to reduce was effective only when an RTD was provided (3.8% relative savings due to the incentive to reduce, with an RTD also provided, -0.3% without an RTD). There was also a positive effect of the RTD on savings, when the incentive to reduce was in effect (2.6%), which was reversed without the incentive to reduce (-1.5%), but neither difference was significant when tested separately. This interaction has no obvious meaning, especially given that the incentive applied to electricity consumption.

Persistence of effects

A second full year of EDF in-trial data was available for enough households (that also had baseline data) to allow analysis for the Wall Panel group but not URA. The reduction in demand remained significant in the second year, for

---

[48] Four trial groups: smart meter only; smart meter plus monthly bills; smart meter plus monthly bills and advice; and smart meter plus monthly bills, advice and RTD.

electricity-only customers (11.0% saving, $p<0.01$) and marginally significant for both customer types combined (5.8% saving, $p<0.1$). It remained non-significant for dual fuel customers.

The persistence can also be judged by looking at the time-course of differences between the trial and control groups, using quarterly figures. Although this does not take into account the baseline consumption, it does indicate whether any initial effect (if present) remained over time. The observed trends over time confirm the persistence of effects for Wall Panel. For the less certain effect – in URA – any initial effect is eroded over the first few quarters.

The significant reductions in consumption seen by E.ON persisted through to in-trial year 1 plus 3 quarters. In the FP groups, savings if anything increased over time. Reductions in gas consumption were not as persistent and groups with RTDs did not show any greater persistence than the others.

Findings from customer surveys

EDF

- Core triggers to uptake are emerging: *cost* savings, desire for *control* and less *hassle*. Control was particularly strong on all interventions with an RTD. The immediacy of Wall Panel was preferred to retrospective information in the Web, TV and Paper groups.
- When using an RTD to access information, only the first or second screens in the access sequence tend to be used, so the most useful information should be put on those screens.
- In groups with the Wall Panel type RTD, customers varied in their awareness of the device: 49% for URA, 60% for Wall Panel. Splitting the Wall Panel group, awareness was 65% for dual fuel customers but only 54% for electricity-only customers. So awareness was greatest among dual fuel customers but this was not associated with greater reductions in consumption.
- Many of the customers who were aware of the RTD were frequently accessing information from their RTD (39% daily or several times a week) though mainly limited to one householder (the bill payer).
- Based on ratings of being useful, Wall Panel (79%) and URA (82%) were similar.
- The URA intervention did not appear to be sufficiently well understood and any effects in this group are probably due to the RTD, with the alarm tending to *reduce* its effectiveness.
- Records of specific actions taken were difficult to interpret with great confidence but, in two cases of energy-saving action (loft and cavity wall insulation), Wall Panel had higher than average percentage taking the action. Wall Panel was also in the top three for draughtproofing. These are the less expensive actions, with financial support often available, which might therefore be carried out quickly. There is also a non-significant trend for increase in number of appliances to be less likely in trial group Wall Panel.
- The survey question “To what extent do you agree or disagree that your Smart meter technology has enabled you to plan or budget for your energy use in the home?” offers an overall rating of the interventions. On this rating, groups that had an RTD were viewed significantly more positively than those that did not.
- Focus group evaluations of the interventions placed Wall Panel in the top three, URA in the middle two.
- Across all the survey evidence, there is a common theme of the RTD being a key element in the interventions.[49] This is in spite of the fact that many of those who had been given an RTD were not aware of having it. Hence there is also scope for RTDs to have greater impact, simply by ensuring that more households who receive one

[49] The evidence presented in other sections is that this was shown most clearly in the HEC group (RTD combined with a controller for heating and hot water) and least in TOUT where the device was the most basic and not the main feature of the intervention.

retain a greater awareness of its existence and actually make use of it. This would come partly from using well designed devices and partly through the installation/delivery and support processes that are put in place.

E.ON

- Differences in survey response are consistent with the analysis of consumption data, the RTD being more likely to be used, and rated as more useful, in the smart meter groups than in the non-smart meter groups. Since the groups differed in both the type of RTD and the context of other interventions, the reasons for the differences in response cannot be stated with certainty. However, the differences in the RTDs themselves are the more likely explanation.
- The survey also provides evidence relating to the optimum design of RTDs, indicating that cost data and simple graphics are key features. In contrast, CO_2 emissions and the usage alarm were not widely used or highly rated features. One surprise is perhaps the high ratings given to the temperature aspect of the clip-on RTD but this might have been particularly useful when interpreting advice about room temperature.

Scottish Power

- The clip-on RTD offered in Phase 1 was not accepted by 30% of credit customers and 22% of prepayment customers. Of those installed in the initial credit trial group (TG1), only 42% were still operational at Phase 2, when installers fitted smart meters. In the prepayment group, 58% of survey respondents said the RTD was still fitted and working during the Wave 3 interventions (towards the end of the study).
- The mains-powered RTD fitted in Phase 2 for credit customers fared better, 77% being reported as fitted and working at Wave 3. Recall of the RTD being fitted was also higher in Wave 3 for the mains device (86-92%) than for the clip-on (74%). Over the three Phases and three credit trial groups, between 11% and 54% of those who knew they had the RTD rarely or never checked it. Figures were generally better for the mains device (11-26%) than the clip-on (11-54%) and general improved through the trial (47-54% in Wave 1, 11-26% in Wave 3). Nevertheless, overall there was limited potential for the RTD to affect consumption. Part of the problem appears to have been that 32% of users (across all groups that had an RTD) found it difficult or very difficult to change the settings.
- A more positive point is that RTDs did appear to make a difference to customers checking their energy use. Initially, prepayment customers were more likely to be checking their energy use, other than when they received a bill. In the course of the trial, credit customers who received an RTD came up to the same level of checking as prepayment customers. In addition, receiving an RTD was associated with participants noting that something had happened to change their frequency of checking. This effect was seen, even though not all recipients used (or even accepted) the RTD.
- Not all the display features were equally noticed, used or valued. Units of electricity was the feature most likely to be mentioned first, followed by cost of electricity but this is likely to be because of the set-up of the devices. In questions about the use or relevance of the display features, electricity cost was seen as more important. On all measures, CO_2 emissions and (when available) gas consumption and an audible consumption alarm scored poorly. Temperature display was rated more highly than might have been anticipated, perhaps most so at the start, as a one-off response to advice about thermostat settings. The overall “winner” was the traffic lights display on the ecoMeter, providing a simple visual signal about consumption rate.

SSE

- Of those responding to an early survey of customers sent a clip-on RTD, 83% had received the RTD and 46% had received and fitted it (53% of those who received it). Only 31% were still using it. Of those who had installed it, most (86%) thought it had been easy to do but 10% thought the instructions were not clear. Overall, 17% of those who had received the RTD thought it complicated, or thought the instructions were not clear.

- Credit customers with smart meters were asked what had motivated them to join the trial. The most common single reason was saving money (41% and 42% in Projects 3 and 4 respectively) [50], followed by helping the environment (21% and 17%). Customers with RTDs installed were more likely to say the reason was to obtain detailed information about energy consumption (12% vs 5% without an RTD) and to identify excessive energy use (6% vs 3%). These questions were asked more than two years after the decision so they should not necessarily be seen as reliable, but they do give an indication of what issues were uppermost in respondents' minds, which could be useful in presenting changes to consumers in future.
- Householders generally reported more positive changes in attitudes and awareness in the groups that received smart meters, particularly if they also had an RTD. But note the warning from the EDF trial that "Positive attitude and intentions were not always translated into less consumption or lower expenditure", which is a common finding in the wider literature.
- Credit customers with smart meters were asked whether they recalled receiving the advice booklet and a similar percentage (80%) did, whether or not they had also received an RTD. However, customers with an RTD were more likely to still have the booklet (59% vs 79%), still refer to it (8% vs 15%) and say it was quite or very useful (78% vs 86%). Having an RTD appears to have brought about greater interest in the advice booklet.
- RTDs were most often located in the kitchen (29% of clip-on RTDs and a higher percentage of RTDs provided with smart meters – 52% in Project 4 and 47% in the prepayment trial group). Other common locations were the lounge (13% overall), hall (11%) and in a drawer/cupboard (8%). This suggests where householders find it most useful to keep the RTD, which could be incorporated into user guidance.
- The RTD provided with smart meters appears to have been viewed far more frequently than the clip-on device provided with non-smart meters: 37% vs 26% every day, 62% vs 41% at least once a week. This is important in terms of limiting the capacity of the intervention to have any effect on consumption.
- All respondents who had an RTD were asked about its usefulness, regardless of whether they were using it or not. Again the groups with smart meters were more likely to say the RTD was quite or very useful: 75% of credit customers and 70% of prepayment customers with smart meters vs 50% of those without a smart meter.
- Regarding specific aspects of the display, cost information was rated more highly than power/energy information. Where available, the traffic light display was seen as the most useful feature – 32% of credit customers and 42% of prepayment customers rated it the most useful feature and 64% of each rated it useful.
- Turning to the reasons why the RTD was useful, responses were similar in each group. The percentages agreeing with a set of possible reasons were: "It provides reassurance that I'm doing the right things to reduce my energy" (73%), "It gives me more control over my energy usage" (68%), "It helps me to encourage other members of the household" (49%), "It means I have more influence on my bills" (63%), "I would miss it if it was taken away" (62%). This could be incorporated into material used to encourage uptake of RTDs.
- Of those who had an RTD but were no longer using it, 42% of the non-smart meter groups and 32% of credit customers with smart meters had never looked at it whereas 9% and 24% respectively had used in for a few months, 49% and 43% for shorter periods. In the smart meter groups, the dominant reasons for not using the RTD related to the usefulness of the information provided – in general or after a period of use. In non-smart meter groups, the issue is more often the functionality of the device itself (except that the greater complexity of the smart meter RTD perhaps lead to a greater problem with knowing "what the buttons do").

[50] Breakdown of survey findings was based on SSE's original "Project" classification (described in Appendix A4.1) rather than the final Trial Groups. Project 3 consisted of all Trial Groups with smart meters but not RTDs whereas Project 4 had both. Both Projects had a range of other interventions.

8.7.3 Conclusions

The literature is clear that providing an RTD typically brings about a reduction in energy consumption but the percentage savings vary widely and appear to depend on climate (and consequently the main uses of energy – cooling or heating, and whether the heating is electric). From the literature review, tentative estimates of what could be expected in the UK were that a base level effect of RTDs alone could be less than 3% electricity savings whereas supplementary interventions that increase engagement could double or triple the benefit. Far fewer studies have tested effects on gas consumption but they have generally shown a benefit; while it is not feasible to quantify it, savings tend to be of a similar order to those for electricity.

The EDRP findings were generally in keeping with this, and allow greater certainty on the magnitude of effects in the UK, and over a longer period than most of the literature. The effect of clip-on RTDs, without a smart meter, was modest, only one trial showing a small (1%) significant effect. The trials were more positive in showing savings from the mains devices provided with smart meters – generally 2-3% for electricity but higher in the EDF trials (4% overall but 7% for electricity-only customers). The greater impact in the EDF trial may be because the accompanying advice was more effective (see Section 8.2). These effects were persistent. Only the Scottish Power trial showed no positive effect of RTDs with smart meters and this may be related to the fact that the trials were not presented to customers as smart meter trials (see Section 8.6).

Effects on gas consumption were less clear: the EDF data did not allow analysis and, as for electricity, Scottish Power did not find a significant effect. Neither SSE nor E.ON found any incremental effect of RTDs on gas consumption above the effect of a smart meter. SSE did find a significant interaction effect, such that the combination of an RTD with the “incentive to reduce” had a greater effect than either on its own. This effect is hard to explain, given that the incentive to reduce applied only to electricity and it did not have an effect on electricity consumption. In any study of the scale of EDRP, there will be occasional isolated inexplicable findings and they may be random.

The net effect is that the combination of a smart meter and an RTD was associated with significant reductions in energy consumption. In relation to gas, the smart meter itself appears to be the critical element whereas for electricity the RTD is also needed. This is in keeping with theoretical considerations, that real-time feedback is clearly more relevant to electricity consumption than to gas. Applications of gas tend to be subject to occasional adjustments having long-term effects, which are less amenable to influence by real-time feedback.

It is also clear from the survey findings that householders paid less attention to the display of gas information than the display of electricity information (this may be due in part – but only in part – to the order in which button pushes accessed different types of information). The surveys also showed that cost information was used and valued more than unit (kW) information. Displays of CO_2 emissions were generally not widely noticed or used or perceived as useful. This is all consistent with the literature, which additionally shows that portability is a benefit (at least initially) and that appliance-specific feedback can have additional effects on energy savings.

The literature also identifies two key points that were not explored further in EDRP, but which could be put into guidance on using the devices:

- householders may find RTDs more useful in confirming savings after attempts to reduce consumption, rather than using an RTD to initiate savings;
- RTDs can be used to check that everything has been switched off before going to bed or leaving the house.

The stronger effects on electricity consumption in the smart meter trials may be explained by the presence of the smart meter or by the differences in the display devices. The latter is inherently more likely and this explanation is backed up by the survey evidence: the mains RTDs, which displayed both gas and electricity consumption and tended to have more sophisticated functions, were consistently (across trials) more likely to be fitted, retained, used and rated positively. Having an RTD also tended to result in more positive perception of the smart meter. The SSE survey found that reasons for not using the RTD differed between clip-ons (where the functionality of the device itself was the dominant reason) and mains RTDs (where the usefulness of the information provided was dominant).

The implication is that RTDs will have less impact if customers are relied upon to fit them: a significant proportion will simply not be fitted. Furthermore, where they are fitted, guidance needs to cover how to use the information that RTDs provide, not just how to access the information. More generally, the written instructions may sometimes have been unclear or too complex, especially for people who had been sent the device, rather than making an active decision to buy it. Hence support from the installers may be particularly important.

There is little published evidence on the use of audible alarms or visible signals in tandem with RTDs. The limited evidence (generally qualitative) indicates that audible alarms tend not to be used but visible signals with clear meaning may be viewed positively and used as part of an energy-saving strategy. This was backed up by EDRP, the audible alarm in EDF's trial causing no incremental reduction in consumption and attracting only negative response in the EDF survey. The "traffic lights" visual signal of consumption level, in contrast, was often the most positively rated feature.

A thermometer may be a simple and effective feedback device but there is insufficient evidence in the literature to say whether it actually has any effect on energy demand. The Scottish Power trial used a thermometer but not in a way that adds to the body of evidence. However, displays of temperature data via RTDs were generally rated positively and may have been particularly useful in the early stages in responding to advice to reduce thermostat settings.

Evidence on the provision of online real-time feedback is limited and, as yet, inconclusive in the literature. It was not tested in EDRP but surveys suggested that the online provision of feedback would have been more effective if real-time data had been provided.

8.8 Heating controller integrated with RTD

8.8.1 The literature

Principles

One EDRP trial integrated a controller for gas heating and hot water with an RTD. While this may be seen as simply expanding the RTD function, it creates a fundamental difference in the intervention. This was the only EDRP intervention that included technology that could directly alter energy use, as distinct from providing advice, feedback or motivation. By increasing the likelihood of interacting with the RTD, the heating control function potentially also enhances the basic feedback function and provides a very direct link to something the household can do in response to the feedback. Potentially it could also sustain interest in the RTD for a longer period.

Evidence

There are no comparable trials in the literature although participants in one study were positive about the idea of incorporating a thermostat into their RTDs. There is also background evidence that a majority of householders do not use their heating controls as intended, and/or fail to realise the energy-saving potential of programmable controls. A qualitative trial of displays found some unwillingness to use buttons to change screens, and some preferences for a particular model because everything was on one screen. This reflects a widespread caution about using controls, related to concern about malfunction or loss of useful information. There may be a generational difference, with younger householders more at ease with toggling between screens.

Capabilities on project:
Building Engineering

8.8.2 EDRP findings

Non-smart meter interventions

The heating controller intervention was not tested in non-smart meter groups.

Smart meter interventions

Only EDF tested the heating controller (HEC) intervention. There was a non-significant reduction in electricity consumption. No baseline data are available for gas consumption.

Persistence of effects

A second full year of in-trial data was not available for enough households to test for longer term effects in this group.

Findings from customer surveys

EDF

- Core triggers to uptake are emerging: *cost* savings, desire for *control* and less *hassle*. The immediacy of HEC was preferred to retrospective information in the Web, TV and Paper groups.
- Customers expected, and could have benefited from, more engagement and instruction during installation of equipment.
- When using an RTD to access information, only the first or second screens in the access sequence tend to be used, so the most useful information should be put on those screens.
- In groups with the Wall Panel type RTD, customers varied in their awareness of the device, from 49% for URA, through 60% for Wall Panel to 64% for HEC. Splitting the Wall Panel group, awareness was 65% for dual fuel customers but only 54% for electricity-only customers. So awareness was greatest among dual fuel customers (in HEC or Wall Panel) but this was not associated with greater reductions in consumption.
- Many of the customers who were aware of the RTD were frequently accessing information from their RTD (39% daily or several times a week) though mainly limited to one householder (the bill payer).
- Based on ratings of usefulness of RTDs, HEC was the best, with 95% rating it “quite” or “very” useful and TOUT the least useful (63%), with Wall Panel (79%) and URA (82%) intermediate.
- Records of specific actions taken were difficult to interpret with great confidence but HEC was in the top three for draughtproofing but not the other less expensive actions (loft and cavity wall insulation). In the case of double glazing and getting a new boiler – more expensive actions – HEC stood out (along with TV in the case of double glazing).
- The survey question “To what extent do you agree or disagree that your Smart meter technology has enabled you to plan or budget for your energy use in the home?” offers an overall rating of the interventions. On this rating, groups that had an RTD (including HEC) were viewed significantly more positively than those that did not.
- Focus group evaluations of the interventions placed HEC in the top three.
- Across all the survey evidence, there is a common theme of the RTD being a key element in the interventions – most clearly in HEC. This is in spite of the fact that many of those who had given an RTD were not aware of having it. Hence there is also scope for RTDs to have greater impact, simply by ensuring that more households who receive one retain a greater awareness of its existence and actually make use of it. This would come partly from using well designed devices and partly through the installation/delivery and support processes that are put in place.

Capabilities on project:
Building Engineering

8.8.3 Conclusions

There is nothing in the literature on the effect of an RTD with integral heating controller so the EDF trial was the first test of its kind. There was no effect on electricity consumption and the change in gas consumption could not be analysed. Nevertheless, the customer survey revealed a positive response to the intervention, as measured by awareness of the device, perceived usefulness and overall rating. The most likely explanation for the discrepancy is that customers valued the device as a heating controller but did not use it so much as an RTD. This would be in keeping with the literature showing that people are often disinclined to switch between functions available from displays.

8.9 Time of use tariff/incentive

8.9.1 The literature

Principles

Research has examined the effect of various interventions focused on when energy is used, by providing a direct financial incentive to shift energy use away from a period of peak demand. The main two examples are:

- time of use (TOU) tariffs, i.e. varying the unit price with time of day (sometimes also by season);
- critical peak pricing (CPP), i.e. setting much higher unit prices on a limited number of occasions when the energy supplier experiences excessive demand and signals this to the consumer (generally the hottest hours of summer days in regions with a high demand for space cooling).

Only TOU tariffs, not CPP, were used in EDRP.

While the aim is normally to shift consumption away from periods of peak demand, there is also the possibility that total consumption will be reduced, because consumers either become more aware generally of their energy use and savings options, or because they simply take more care to eliminate energy wastage when the price is higher (e.g. turn off appliances that are not being used) rather than compensating by increasing usage at other times. The opposite effect is also possible, if consumers take less care of energy wastage during cheaper off-peak periods; the tariff levels and ratios may be important in this respect but it is difficult to disentangle these from other aspects of tariff scheme and research design.

The intervention is focused on *motive* – mainly financial but potentially environmental if users grasp the relative environmental impact of peak time power generation.

Evidence

Most studies have been carried out in regions (mainly in North America) where electricity used for air conditioning is the main target; this is generally not a major issue in the UK at present. Other studies also have limited relevance because they were aimed at peaks in demand from electric heating in very cold winter climates.

Unless electricity is used for space heating, the opportunities for load shifting are limited. Furthermore, even if electricity is used for heating, this will often have already established load-shifting (using electric storage heating). Response to TOU pricing in the UK depends on ambient temperature, particularly during the daytime – unsurprisingly, colder temperatures prompt people to turn on the heating (with little evidence of shifting consumption into the night). This interaction must be taken into account when interpreting TOU studies.

CPP is associated with larger reductions in peak usage than TOU tariffs but TOU is more relevant to EDRP because CPP was not tested. The impact of either type of tariff is increased by providing some kind of enabling technology (e.g. thermostats that vary the thermostat set point automatically with tariff or allow remote “direct load control”). Again, this is less relevant than simple TOU tariffs in the context of EDRP.

TOU studies of summer energy use in North America have shown peak time savings of between 2.4% and 10.6%, and overall savings from small negative values up to 6% but typically around 3%. One UK study found an overall conservation effect of 3.1%. Another reported a 1% reduction in usage compared with baseline consumption but with no control group comparison.

Overall savings do not vary as much as peak time savings with type of tariff intervention or technology. Indeed, although the large variation in intervention details and samples makes firm conclusions difficult, overall savings may even be higher for TOU than CPP. Mathematically, this would make sense because the higher peak time percentage savings are sustained for a shorter period with CPP. Psychologically there is also a rationale because TOU applies every day whereas CPP applies only on certain days, hence being less likely to prompt development of new habits. Hence, there may be a trade-off between CPP and TOU, with CPP better managing peaks in demand but TOU having a greater likelihood of reducing overall demand in addition to suppressing peaks.

It is clearly difficult to predict the impact of TOU interventions, even for the North American cooling season; much less can the findings be applied directly to UK heating and non-heating use of electricity. While there is a demonstrated potential for reducing peak demand and overall consumption through TOU tariffs, the magnitude of this potential in the UK must be determined in the UK (and possibly within each climatic region).

In general, the overall savings achieved by TOU tariffs are less than for programmes focused on overall energy savings. This is no surprise and the implication is that the aim of an intervention – demand reduction or load shifting – needs to be clear.

8.9.2 EDRP findings

Non-smart meter interventions

Time of use tariffs and incentives were not applied in any non-smart meter groups except in the Scottish Power prepayment groups. The trial design does not allow analysis of any independent effect of the modest (£5) incentive to reduce peak time consumption over a three-month period, which was applied at the same time as a £10 incentive to reduce overall consumption.

Smart meter interventions

E.ON did not use any time of use tariffs or incentives. The Scottish Power trial design does not allow analysis of any independent effect of the incentive to reduce peak time consumption, as noted above. EDF and SSE each tested time of use tariffs for electricity consumption.

Demand reduction

EDF found a small but significant difference in electricity consumption between the time of use tariff (TOUT) trial group and the control group, not accounting for baseline consumption. Comparing with baseline, there was a large reduction in energy use but it was not statistically significant (or meaningful) because the sample size was too small for households where both in-trial and pretrial data were available. The effect on overall consumption is therefore unproven but plausible. No gas data were available for this group – this is of lesser importance because the intervention was aimed specifically at electricity use but gas data would have allowed a test of “spillover” effects from an electricity intervention to gas consumption.

SSE’s “incentive to shift” consumption from the peak period had no overall effect on electricity or gas consumption. For gas consumption, there was a significant interaction between the incentive to shift and provision of information via the web: savings were greater if participants had both the incentive to shift and web information (12.4%)[51] or neither (11.8%), compared with those that had web information without the incentive to shift (9.5%). Savings in groups with the incentive to shift but not web information (11.4%) were intermediate and did not differ from the other groups. This interaction has no obvious meaning, especially given that the incentive applied to electricity consumption.

Load shifting

In the EDF TOUT group, at weekends, there was a significant ($p<0.05$) overall reduction in the percentage of consumption that occurred in the peak tariff period, compared with the Wall Panel group (relative peak time consumption in TOUT was 92% of that in Wall Panel).

Comparing TOUT with the control group, the overall difference in peak period consumption was modified by a significant interaction effect, meaning that peak load was reduced in TOUT only for smaller households (based on the number of people in the age range 16-64) – for larger households, relative peak time consumption actually increased. Modelled relative peak time consumption in TOUT was 86% that in the control group with nobody aged 16-64 in the household (this would be almost exclusively pensioner households), 90% with one person and 94% with two people. The tipping point is three people (98%).

On weekdays, the load-shifting effect is similar but weaker. There was a marginally significant ($p<0.1$) overall reduction in the percentage of consumption that occurred in the peak tariff period, compared with the Wall Panel group (relative peak time consumption in TOUT was 96% of that in Wall Panel). Comparing TOUT with the control group, there is the same significant interaction as was seen for weekends. Modelled relative peak time consumption in TOUT was 89% that in the control group with nobody aged 16-64 in the household, 93% with one person and 97% with two people. The tipping point is three people (101%).

SSE's analysis of smart meter data by tariff period and high/low tariff seasons showed an effect of incentive to shift on consumption. Electricity consumption was higher at night in high season for customers on the incentive. The percentage shift from peak to night electricity usage is estimated as 8.5-10.1%, based on high season consumption at night (overall shift from peak is not estimated). Even though the incentives applied only to electricity, gas consumption was also higher at night in high season for customers on both the incentive to shift and incentive to reduce.

The percentage of electricity consumption that falls in the peak period is reduced by the SSE incentive to shift but by only a small amount – from 19.8% to 19.5% on weekdays and from 19.4% to 18.9% at weekends. The weekday effect of the incentive to shift was greater in the absence of an RTD and in the absence of web information, suggesting some kind of interference effect if too many interventions are in place at the same time. The interaction with RTD was also seen at weekends.

Persistence of effects

A second full year of in-trial data was not available for the EDF TOUT group. Persistence can be judged by looking at the time-course of differences between the trial and control groups, using quarterly figures. Although this does not take into account the baseline consumption, it does indicate whether any initial effect (if present) remained over time. The observed trends over time show that any initial effect on overall consumption is eroded over the first few quarters. In the SSE trial, there were no initial effects and the time-course beyond the first year could not be tested. Similarly, the persistence of load shifting could not be examined.

---

[51] Percentages shown are the average for the trial groups having the interventions indicated, regardless of other interventions that were in effect. They are not adjusted for changes in the control group and therefore appear larger than the percentages cited for other interventions.

Findings from customer surveys

EDF

- Core triggers to uptake are emerging: *cost* savings, desire for *control* and less *hassle*. The cost savings motivation is most directly met by TOUT. The immediacy of feedback in TOUT was preferred to retrospective information in the Web, TV and Paper groups.
- When using an RTD to access information, only the first or second screens in the access sequence tend to be used, so the most useful information should be put on those screens. This would have particularly impacted the TOUT intervention, where the basic RTD required multiple button-pushes to access data.
- In groups with the Wall Panel type RTD, customers varied in their awareness of the device, from 49% to 65%. The percentage aware of the RTD dropped to 38% for TOUT, and the low-tech design (used only for TOUT) was a disappointment to some householders. It appears that the tariff effect could have been better supported by an alternative RTD.
- Based on ratings of usefulness of RTDs, HEC was the best, with 95% rating it “quite” or “very” useful and TOUT the least useful (63%), with Wall Panel (79%) and URA (82%) intermediate.
- Records of specific actions taken were difficult to interpret with great confidence but, in two cases of energy-saving action (loft and cavity wall insulation) the TOUT group had a higher than average percentage taking the action. There is also a non-significant trend for an increase in number of appliances to be less likely in TOUT.
- The survey question “To what extent do you agree or disagree that your Smart meter technology has enabled you to plan or budget for your energy use in the home?” offers an overall rating of the interventions. On this rating, groups that had an RTD (including TOUT) were viewed significantly more positively than those that did not.
- Focus group evaluations of the interventions placed TOUT in the top three.
- Across all the survey evidence, there is a common theme of the RTD being a key element in the interventions – most clearly in HEC (combined with a controller for heating and hot water) and least in TOUT where the device was the most basic and not the main feature of the intervention.

SSE

- With 83% of those in the “Incentive to shift” groups aware of the incentive, this is a high level of awareness relative to other interventions, although only 75% were aware and had some understanding of how the incentive worked. Of those who were aware, 40% thought they had shifted consumption and made savings, 33% that they had shifted but without making savings and 28% that they had not shifted consumption.
- This intervention was generally perceived as complex and this may be worse when it is combined with the incentive to reduce because cost saving is confused with energy saving.
- Customers who had been in neither incentive group were asked “How much cheaper than the peak daytime tariff would the night tariff have to be to encourage you to move some of your consumption?” Group average figures ranged from 19-32% (the mean of the group averages was 25%).

8.9.3 Conclusions

The literature shows clear evidence that time of use tariffs can shift consumption from the peak period, and often also bring about reductions in total energy consumption. However, the evidence is almost exclusively from studies in hot regions (where the dominant energy demand is for air conditioning) and cold regions with electric heating. The limited evidence from the UK suggested that only small reductions in overall electricity demand (3% or less) should be expected.

EDRP tested TOU tariffs only in combination with smart meters and did not provide convincing evidence of an overall reduction in demand. In the case of EDF, the RTD provided was more basic than in the other trials and the customer survey showed that it was less likely that customers would be aware of its existence, and those who were aware rated it as being less useful.

There were, however modest effects on shifting load from the peak period. This comes from the EDF and SSE trials, both of which showed a stronger load-shifting effect at weekends than on weekdays. Estimates of the magnitude of shifting effect vary with trial but were up to 10%. The EDF data showed that the effect is stronger with smaller households (1 or 2 people), thus providing a clear focus for where such interventions should be targeted. The effect was weaker in the SSE trial and this may be because awareness of the intervention was limited and it was seen as overly complex.

8.10 Incentive to reduce consumption

8.10.1 The literature

Principles

Incentives to reduce consumption go a step beyond TOU interventions and give the consumer a direct financial reward (in addition to any financial benefit from buying less energy) for reducing overall consumption. The emphasis here is clearly on the financial *motive* although sometimes the reward is in the form of vouchers to spend on energy-saving products, thus turning motive into *opportunity*. EDRP included several such interventions, always combined with other interventions.

Evidence

It is not unreasonable to suppose that, if the reward is great enough, consumers will reduce energy use – indeed this has been seen from the early small-scale studies of how to achieve energy conservation, which were of relatively short duration. The practical issue is whether it is economically feasible for suppliers to offer a reward that is high enough, and sustained for long enough, to have a lasting effect. It is perhaps because of such doubts that there is little direct research evidence.

Research in other fields commonly shows that, when an extrinsic motivation is removed, people return to their former behaviour. In the case of reducing energy demand, it is possible in principle that a lasting effect might occur if the consumer invests in efficiency measures or more efficient appliances as part of an effort to gain the incentive payment, but of course such incentives would have to be structured for the long term in order to make the investment worthwhile.

The implication, that short-term financial incentives are expected to have only temporary effects, is supported by the limited evidence on the subject but there is also concern that financial incentives can be counterproductive in the longer term by focusing attention on the financial motive, to the detriment of other motives that the consumer might have had prior to the intervention. Not only would this make it difficult to sustain energy savings, it would render more probable the “rebound effect” whereby energy savings are spent on other goods, services or benefits that might themselves be energy-intensive (e.g. a bigger TV or a holiday flight).

There is, however, some evidence that other mechanisms can be used to sustain energy-saving behaviour that has been initiated by financial incentives, specifically community peer support. It may also be safer to target households that are already motivated to save energy but lack the resources to invest in their goal, i.e. to use financial support to provide *opportunity* rather than *motive*.

Rewards for reducing consumption are necessarily linked to a target, even if the target is a minimal change and easy to achieve. Hence the financial incentive effect is confounded with a possible target-setting effect: people can be motivated by succeeding in a challenge without any extrinsic reward (see Section 8.5).

A possibility that has not been explored in research is that smart meters (and associated feedback) could make consumers generally more aware of energy prices and hence more responsive to price changes or offers, making financial incentives (or the negative equivalent – price rises) more effective in reducing demand.

8.10.2 EDRP findings

Non-smart meter interventions

E.ON and SSE did not use incentives to reduce consumption in their non-smart meter groups.

In the EDF “read-reduce-reward” (RRR) group, customers read their own meters and could get a financial reward for reducing consumption. No baseline data were available and other data issues had been identified for this trial group meaning that no clear conclusion can be drawn as to its effectiveness.

In the Scottish Power trial, none of the non-smart meter (prepayment) trial groups made significant gas or electricity savings in response to the various financial incentives (combined with advice and additional bill information, with or without an RTD).

Smart meter interventions

E.ON and EDF did not use incentives to reduce consumption in their smart meter groups.

Scottish Power found significant reductions in electricity consumption in the credit customer trial groups during three of the four Waves of the Green Challenge (GC) financial incentive, each of which lasted three months. This was based on the ratio of consumption to consumption in Phase 1 of the trial. There was no significant effect over the central 12 months of Phase 3 as a whole, based on the difference from pretrial baseline consumption, although the trend was in the same direction. The quarterly findings relate to each individual Wave and therefore give a more fine grain analysis but the control group and baseline non-smart meter data are less reliable over the shorter period. On balance, it is reasonable to concur with the Scottish Power conclusion that there was a temporary effect of the three Waves that focused on electricity consumption (without an additional gas-related incentive).

In the same trial groups, there were significant reductions in gas consumption during one of the four Waves of the GC financial incentive, and during a three-month break in use of incentives. There was also a significant reduction over the central 12 months of Phase 3 as a whole. On balance, it is reasonable to concur with the Scottish Power conclusion that there was a temporary effect of the Wave that had a gas-related incentive, but also the non-incentivised period preceding it.

The major concern with the Scottish Power findings is that they could be due to the Hawthorne effect rather than being a true effect of the intervention, because the intervention periods were so short and it was only at this point in the trial that households were aware that their conservation efforts were being observed. It is also debatable whether these are best viewed as smart meter interventions, given the efforts made to prevent customers from knowing they had a smart meter. They are better seen purely as tests of financial incentives.

In the SSE trial, there was no significant effect of the incentive to reduce on electricity consumption. For gas consumption, there was not a significant main effect but there was a marginally significant interaction between incentive to reduce and provision of an RTD: the incentive to reduce was effective only when an RTD was provided (3.8% relative savings due to the incentive to reduce, with an RTD also provided, -0.3% without an RTD). This interaction has no obvious meaning, especially given that the incentive applied to electricity consumption.

Persistence of effects

Data were not available to test the persistence of the observed effects.

Findings from customer surveys

EDF

- RRR customers were not aware they were part of trial but reading their own meter appears to have raised some awareness about costs and consumption.
- Focus group evaluations of the interventions placed RRR in the middle two.

Scottish Power

- Customer mailings were an important element in the Scottish Power trial and their quality was monitored and improved through the course of the trial. Recall and comprehension during Phase 3 was the highest achieved in the project (and similar among credit and prepayment customers – the latter had recalled the material less well in earlier Waves). Since it was Phase 3 that introduced financial benefits, it may have been this (rather than changes in style) that was responsible. Self-rated comprehension of various aspects of the mailings did not differ between trial groups.

SSE

- While 72% of those in the “Incentive to reduce” groups were aware of the incentive; 37% of them never met the target and a further 9% did not know whether they had ever met the target (46% in total); 14% always or mostly met the target and a total of 45% met it more than once. The target may have been too challenging (or progress too difficult to monitor) to have any effect.

8.10.3 Conclusions

The literature provides little substantive evidence on incentives to reduce consumption except for the general (and obvious) point that sufficient incentive will prompt people to reduce consumption, but only for as long as the incentive is kept in place. There are some concerns that emphasising the financial motive in this way could detract from long-term savings prompted by other motives.

E.ON did not use incentives to reduce consumption. The other suppliers did (only for electricity) but only Scottish Power saw reductions in consumption when the incentives were applied – only in the case of credit customers with smart meters and only for short periods. The Hawthorne effect is a sufficient explanation of the Scottish Power findings.

8.11 Overview of web-based interventions

8.11.1 The literature

The web is not an intervention as such but is a rapidly developing medium for delivery of a range of types of intervention. Therefore, it is worth taking an overview of its potential.

The important differences between paper- and web-based information are probably that the latter is (a) a more active process for consumers than passively receiving printed material, (b) more able to provide information and advice tailored to the individual user, to support response to feedback and (c) more easily and immediately linked to options for action to reduce consumption and the resources for carrying out those actions. Certainly the web offers opportunities to bring together means, motive and opportunity in a way that would be hugely more resource-intensive through other media. The key point is not that the web has greater impact for a given intervention (and there is insufficient evidence on whether or not this is true) but rather that it creates new possibilities.

Creating the optimum web site for achieving energy savings will almost certainly mean combining various types of intervention – feedback, targeted advice, incentives, etc. The combinations have not been studied systematically but rather the existing sites have been designed on the basis of predicting the optimum combination – not purely for energy savings but also for other customer and supplier benefits. The reported impact has been highly variable, partly because the reporting rarely follows a normal scientific pattern but also because the online services vary greatly.

The huge range of outcomes reflects the general variation in outcomes – regardless of the medium used for interventions – in addition to the different web site characteristics, customer populations, climates and ways in which customers interacted with the sites. While it does seem likely that many customers achieve energy savings as a result of web-based interventions, it is impossible to deduce from the literature what the essential "active ingredients" are or what the optimum web-based feedback would look like.

The right question might not be whether web-based interventions work, but rather for whom do they work? The most promising uses of the utility-based websites seem to be with particular subsets of the population and/or specific, focused programmes. Some people would not use web-based service at all, some would make limited use (e.g. to check consumption, but with no intention of reducing it) while others would make effective use of good websites to reduce consumption. Concern about data privacy is likely to be a consideration for some customers; for others, this will seem insignificant in relation to the information that they already exchange online. Alternatively, it may be that most people would use a web-based service but only at certain times (e.g. when moving to a new home or when there is a sharp rise in energy prices) or in certain contexts (e.g. where a more complex energy management demand, such as complex tariffs, benefits from additional information).

In considering targeted use of web sites, it is also important to pitch information at an appropriate level: making the sites simple and attractive for newcomers, offering an overview of basic information but with easy access to detail and tools for people who are actively researching (and seeking to reduce) their usage.

The relationship between RTDs and web-based feedback also needs to be considered. User feedback suggests that the web is not a good substitute for RTDs in relation to their prime function of real-time feedback: using a PC for this takes extra time and trouble and relies on a good internet connection. RTDs are also more portable and more convenient to leave in rooms such as kitchens and utility rooms. This implies that online data should be seen as a complement to RTDs rather than as a substitute, especially if users can download data from their RTD to a personal web page. The RTD is then used for immediate information, while the web-page can be used to supply graphics, a long view of consumption patterns, and the opportunity for detailed exploration of the data.

The range of ways to access web-based energy information and tools is developing rapidly, with applications now available via mobile phones, personal organisers, etc. An increasing proportion of the British population engages routinely with some form of online material. In this context, there is a case for active research into the various web-based services on offer, to achieve a more robust assessment of their impact and how to optimise them for achieving energy savings and supporting the introduction of smart meters.

8.11.2 EDRP findings

Non-smart meter interventions

Only SSE used a web-based intervention with non-smart meter households; this consisted of providing the energy efficiency advice booklet online rather than by post. There was not a significant effect of advice or interaction with medium of delivery.

Smart meter interventions

EDF and SSE tested the effect of providing advice and personalised consumption history online.

In the EDF trial, electricity consumption was non-significantly reduced. No baseline data were available for gas consumption.

In the SSE trial, there was no effect of the web intervention on electricity or gas consumption except for a significant interaction (for gas consumption) between the incentive to shift and provision of information via the web. The interaction means that savings were greater if participants had both the incentive to shift and web information (12.4%)[52] or neither (11.8%), compared with those that had web information without the incentive to shift (9.5%). Savings in groups with the incentive to shift but not web information (11.4%) were intermediate and did not differ from the other groups. This interaction has no obvious meaning, especially given that the incentive applied to electricity consumption.

Persistence of effects

The persistence of effects can be judged only by looking at the time-course of differences between the trial and control groups, using quarterly figures tabulated by EDF. Although this does not take into account the baseline consumption, it does indicate whether any initial effect (if present) remained over time. The observed trends show little variation in electricity consumption but the general trend is upwards. For gas, there is some indication of a downward trend but it is unreliable because of large unexplained peaks in the data.

Findings from customer surveys

EDF

- Core triggers to uptake are emerging: *cost* savings, desire for *control* and less *hassle*. The immediacy of feedback from RTDs was preferred to retrospective information in the Web group.
- Focus group evaluations of the interventions placed Web in the bottom three.
- The limited use that customers made of the Web means that their effects on consumption cannot be fairly judged until more effective implementation is achieved, making the intervention more easily accessible, more focused on the key information that users want to see, and with better linkage between consumption data (real-time and historic), advice and access to external financial and technical support.

SSE

- Overall, only 50% of respondents in smart meter groups were aware of the website, only 9% used it (and only 20% of website users used it once a month or, in the case of one user, more often). This clearly limits the capacity of the intervention to have any impact.

8.11.3 Conclusions

The literature shows the potential benefits of online services to reduce energy demand, but also shows that the potential is rarely realised. EDRP confirmed this, with neither of the suppliers who used web-based interventions (EDF and SSE) seeing any energy savings as a result. The trials also showed that a major reason for failure is likely to be lack of engagement with the web sites, not necessarily a lack of effect among those who use the sites.

---

[52] Percentages shown are the average for the trial groups having the interventions indicated, regardless of other interventions that were in effect. They are not adjusted for changes in the control group and therefore appear larger than the percentages cited for other interventions.

Capabilities on project:
Building Engineering

8.12 Population segment effects

Across all the trials, there was limited evidence of how different population segments were affected by the interventions.

EDF found clearly defined effects, with smaller households being more likely to save energy overall and to shift consumption from the evening peak period. There were also differences between electricity-only and dual fuel customers. These differences were not due to the electricity-only customers not having gas heating but they are otherwise difficult to explain with any certainty.

The E.ON findings clearly varied between customer strata. The customers selected because of their initial high consumption reduced their consumption whatever the intervention but this merely shows what is expected: high consumers have most potential to reduce and this can be explained by “regression to the mean”, rather than the impact of any particular intervention. Behavioural changes were more weakly evidenced in this group, again supporting the conclusion that changes in consumption were artefactual.

More interesting is the difference between E.ON’s ‘fuel poor’ (FP) and not ‘fuel poor’ (NFP) groups. Although there are problems with the definitions of these groups, the FP group is likely to be generally less affluent, with more pressure on energy costs. The generally more positive response to interventions in FP most likely signifies their greater motivation to save money.

SSE found significant effects of postcode (for electricity only) and Mosaic demographic group (for gas and electricity) but not Grid distribution point (an industry geographic code). Including geographic region in the analysis did not affect the outcome and there was not a significant effect of this additional variable. This emphasises the need to control for location and demographic variables but the type of analysis used did not permit identification of specific locations or Mosaic groups that had a higher propensity to reduce consumption. Since postcode was significant but not the more geographically systematic variables, the details of any location effect are in any case difficult to define. In no case did one of these variables have a significant interaction with a trial intervention, i.e. the effects of interventions did not differ between locations or demographic categories.

Most of the EDRP findings (and most of the literature) relate to credit customers – there is relatively little about prepayment customers or prepayment meters. EDF and E.ON did not address this sector. Only SSE included prepayment smart meters (for electricity only) in its trial, with one trial group being given smart meters and an RTD. This group made savings similar to those in the credit customer groups.

Scottish Power made more extensive investigations of prepayment customers but not with smart meters. These customers did not make savings relative to the control group under any interventions, even the financial incentives that brought about temporary reductions in demand in the credit customer groups. Nevertheless, some prepayment customers did meet their savings targets (arguably more consistently than credit customers), so it may be that prepayment customers were making savings anyway and could not save more in response to the interventions.

Prepayment meters tend to be installed where customers are having difficulty paying energy bills, so they generally have lower consumption anyway (and this was certainly the case in the Scottish Power trial). Therefore they have less scope to reduce. However, they may be more practiced at monitoring consumption, more aware of costs and have greater (financial) motivation to save energy. With these counterbalancing factors, savings appear to be possible (as shown by SSE) but not always achieved (as shown by Scottish Power). Logically, efforts to help prepayment customers reduce consumption might focus on means and opportunity, the motive already being present, but this was not explored in the trials.

9 Practical and Technical Issues

This section of the report details the significant practical and technical issues highlighted in EDRP. It concludes by summarising key learning points identified by the energy suppliers.

In reading this section, it is important to note that the primary purpose of EDRP was to investigate consumer behaviour (and consequent changes in energy use) and it was not set up as a technology or roll-out trial. The equipment used was therefore what was either readily available or could be developed for use in the time available and it was not based on a common set of requirements. Nevertheless the experiences of the energy and equipment suppliers in the installation and use of the equipment from EDRP provides valuable practical lessons for the mass roll-out of smart metering. Many of the issues identified are already informing work under the smart metering programme in terms of both the technical design of the smart meter system and the approaches to the roll-out and installation of the equipment. Furthermore, given the advances in technology, a number of the issues would not be encountered if starting EDRP today. It is also worth noting that some of the issues would be encountered in “business as usual” meter replacement activity, whether smart meters or not.

Much of the information presented in this section has previously been summarised in Ofgem’s published March 2010 progress report and the accompanying presentations from the four energy suppliers. However, for completeness, this information is collated in more detail here.

This information is principally taken from the four energy suppliers’ reports and CSE’s six-monthly reports to Ofgem. The reporting of practical and technical issues was inconsistent between suppliers. This makes it difficult to compare and contrast the various challenges encountered by the energy suppliers and identify where a particular supplier has better addressed a particular problem.

The information is presented in three sub-sections:

- customer feedback;
- issues with the development, installation and operation of the smart metering equipment;
- issues with data management.

9.1 Customer feedback

This section focuses on consumer issues that could practically affect the roll-out of smart meters. This section does not address the impact of smart meters on consumer behaviour and energy consumption reduction which is covered in the previous sections.

9.1.1 Customer service logs

It is helpful to understand the number and type of communications received by the customer service teams. This highlights the key issues that the customers faced, to be considered prior to national roll-out. It also provides an insight into the scale of customer service teams required during roll-out and the type of training that may be necessary.

Number of communications

The SSE trial is reviewed first because SSE maintained a detailed log of all communications to and from customers throughout the trial, which was also one of the largest trials. It also presents the most detailed analysis of communications received. In total, just over 20,000 incoming communications were received from customers during the period from March 2008 to September 2010 (mainly by telephone, but also letters and emails). SSE notes that this number may be lower than the actual total, owing to some missing data at the start of the project, calls going to other customer service teams, etc. Given that SSE had approximately 27,000 participants, this suggests that there was an average of approximately 0.75 calls per participant. In practice, some households would have contacted SSE multiple times either relating to the same issue or to multiple issues. SSE notes that the majority of these communications (70%) were not EDRP-related, i.e. business-as-usual calls (billing enquiries, meter readings, payment plans). It is noted that SSE continued with manual reads for billing purposes for all trial participants and so the communications are not representative of a complete shift to smart metering.

Furthermore, around 15,000 communications were sent from the SSE customer service team. Nearly two-thirds of these do not appear to be responding to an individual concern but are instead ‘mass’ communications to those participating in the financial incentive trial groups (incentive to shift or incentive to reduce energy consumption).

E.ON notes that, for customers who experience a meter exchange, an average of 30 additional minutes of contact time is experienced by call centres. E.ON suggests that this occurs for both smart and non-smart meter exchanges and that it believes that the calls arise principally from the process of meter exchange.

EDF notes that there was a peak in calls to customer services for support at the time of installation, after which there is a significant drop-off in volume of calls unless a particular problem has occurred.

Issues raised in communications

SSE provides the most detailed information in its reports on communications to and from its customer service team. Table 9.1 summarises the top reasons specifically related to EDRP that were raised by customers within the smart meter trial groups. This is based on SSE’s grouping of the communications and the rank order would depend how the topics are grouped. The learning and implications are more important than simple ranking but two of the top five issues were from the financial incentive trial groups (incentive to shift and incentive to reduce energy consumption).

By further grouping of the issues in Table 9.1, SSE notes that the top five overall customer communications specific to EDRP are as follows (the other three energy suppliers reported similar issues).

- Problems with smart metering equipment or other aspects of the interventions (shows that getting it right first time will save many calls).
- Communications resulting from EDRP made customers more aware of their consumption (and this greater awareness of the detail of their energy use prompted them to seek further advice or explanations).
- Customers tried to understand the new intervention that they had been given (so simplicity of message and testing understanding is key in a national roll-out).
- Customers queried why they still had estimated bills when they had smart meters installed. This type of call should be addressed once end-to-end billing is successfully deployed with smart meters.
- Customers contacted SSE stating that their billing details were not correct (owing to meter exchange details not being updated on the billing system, incorrect install/removal reads and/or delays in billing).

Table 9.1 Top EDRP-related reasons for customer communications to SSE

Customer comment	*Total*
Customer contacts due to shadow bills[1] not sent due to communications issues, technical difficulties etc.	527
Gas smart metering service removed from customer due to poor communications, meter not bound, meter not installed, etc.	517
Customer contacts SSE regarding concerns over usage.	421
Explanation of incentive to shift/reduce given to customer.	408
Customer queried why reads are estimated on bills when they have a smart meter installed/why manually read when have smart meter installed.	351
Customer contacts SSE regarding RTD faults (or perceived RTD faults).	286
Customer asks “When will the incentives start?” / Customers say they are looking forward to start of incentives.	285
Meter exchange details not updated on billing system, incorrect install/removal reads leading to customer queries and/or delays in billing.	202
Energy Saving Advice/Leaflet supplied. Customer taking (or will take) energy-saving measures or appliances to reduce consumption.	180
Explained aim/function/features/benefit/how to use smart meter to customer.	163
Advised customers how to read smart meter.	131
Explained aim/function/features/benefit/how to use RTD (including site visit request).	88
Customer contacts regarding meter faults (or perceived meter faults).	87
Prepayment customers reported problems topping up meter or adding credit to their meter.	81
Smart meter readings used for billing or to resolve queries.	70

[1] For the purpose of the incentive to shift trial group, time of use tariffs were set to operate as a ‘shadow’ form of billing (shadow bill). This meant that only customers benefiting from the tariff would be charged using its rates and those not benefiting from the tariffs would be charged at standard rates.

The E.ON trial had a similar number of participants to SSE’s but provided more limited detail of customer issues. EDF and Scottish Power were smaller trials and identified some key issues raised. Issues not explicitly covered in the discussion of SSE are listed below. The issues are not stated in any order of importance.

- Customers queried the energy consumption, and associated running cost, of the smart meter equipment installed in the home (EDF, Scottish Power). EDF had one customer in its TV trial groups wanting EDF to contribute to the cost of the customer’s broadband bill each month as EDF was using it as well as the householder (EDF’s TV intervention solution used broadband as a communication device into the home), even if the marginal cost is zero.
- Customers queried why the gas meter displayed consumption in m^3, but the RTD showed kWh (EDF).
- EDF notes that Time of Use participants were very sensitive to issues, with three of the four most serious escalated calls in the EDF trial being from Time of Use customers. On all three occasions, customers thought that an action that EDF had or had not taken was affecting them financially. (It is noted that SSE had a relatively high number of communications from participants in its financial incentive trial groups).
- EDF’s web trial provided the customer with an estimated forecast consumption level. The forecasting did not take into account sudden alterations in weather. During a cold snap in January 2009, the forecast consumption was misaligned with actual consumption, prompting six households to contact the customer service team.
- Some customers queried the low credit alarm on the smart prepayment meter, which they found irritating. The alarm tells customers every half hour by a beep that they have low credit which is more often than a traditional low credit alarm and the beep is loud. The volume of the beep and frequency cannot be controlled other than to switch off the alarm (EDF).

- There were issues arising from turning off the power or gas supply while fitting meters but these should be common to all meter changes. E.ON experienced complaints from customers owing to their boiler pilot light not relighting following meter exchange. In all cases this was due to a poorly maintained boiler or some underlying fault. Scottish Power also reports four instances of the gas boiler pilot light failing to re-ignite. Scottish Power also provides an instance of a customer's security system not resetting after the power had been turned back on. This was due to a fault with the alarm system itself and not specifically the act of replacing the meter. Scottish Power also had an instance where the customer reported that the meter installation had caused a DVD recorder to malfunction.
- A householder reported that the earth clamp on the gas meter had not been reinstated on replacement of the meter; this was rectified. It is thought that the earth clamp had been sheared off prior to the meter fit (Scottish Power).
- A customer reported that the meter was not positioned level following installation. This was not unsafe but the installer revisited the property and levelled the meter (Scottish Power).
- One customer was unhappy with an installation being rescheduled. The reason for rescheduling had been heavy rain on the planned day of installation which prevented the safe exchange of an electricity meter located outside the home (Scottish Power).
- Four customers enquired as to what would happen if the RTD was switched off and whether it should be switched off overnight (Scottish Power).
- EDF notes six prepayment households where the customer was unwilling to allow the supplier to drill holes to install the RTD. It is not known if the other suppliers faced similar issues (it is believed that only EDF fixed the RTD to the wall).
- One prepayment customer lived so close to a pay-point outlet that the customer was back in front of the smart meter display before the credit was received. Once the customer understood the need to wait up to 15 minutes for the credit to download, the customer was happier (EDF).

9.1.2 Access to properties

There will be practical issues in gaining access to properties during the national roll-out both for installation and to rectify any technical fault. Some common issues identified in EDRP are highlighted below.

- *Contacting customers.* All studies highlighted a difficulty in initially contacting customers. For example, EDF undertook its recruitment by telephone and it appears that 25% of its calls were to unobtainable numbers. A further 10% of calls were to numbers that did not accept number-withheld calls, customers who had moved address or customers who had insufficient English.
- *Customers agreeing to a meter exchange.* In all these voluntary trials, the majority of those contacted refused to participate. It is noted that for the national roll-out, there will be a requirement for the meters to be exchanged.
- *Accessing properties.* EDF notes that 18% of those who had agreed to a date for a visit cancelled in advance (31% of these cancelled as 'too inconvenient'; 17% due the terms and conditions – this is discussed further later). Suppliers also had difficulties due to no-one being present when the installer visited the property or the householder being present but deciding not to take part in the trial.

- EDF also notes that even when recruited and installed onto the trial, gaining access to the home (necessary in this instance to implement a technical work-around) was not possible for around 5% of participants, despite significant effort. This finding is consistent with EDF's experience of customer contact generally.

In general, these issues will be expected to be common to any meter exchange. There is insufficient evidence from EDRP on whether there is a greater resistance in general from customers to a smart meter replacement.

- EDF notes that its variable tariff intervention (Time of Use Tariff) was initially marketed to the consumer as a money-saving scheme. A number of customers did not understand why EDF would want them to save money on their bills and were reluctant to participate. EDF subsequently modified the presentation of this intervention to customers.
- EDF required its smart meter customers to sign a set of terms and conditions before installation. This originally caused great issues as it was presented as a legal document, leading to an initial 90% fallout rate between recruitment and installation. The terms and conditions were revised in line with customers' comments. Terms and conditions were subsequently only formally responsible for a fraction of fallout rate. However, EDF expects that many more of its customer cancellations were due to the terms and conditions though customers did not explicitly said they would not sign them. EDF also noted that the terms and conditions were quite manually intensive in terms of managing, answering customer queries and chasing customers to return.

9.1.3 Customer surveys

The previous sections on the energy suppliers' findings (Sections 4 to 7) summarised the results of the customer surveys. However, several practical issues are further highlighted in this section.

- *Time for installation.* The time for EDF's installation was typically 1.5 hours but varied with trial group (the heating controller tended to take longest owing to the boiler interaction)[53]. Occupant feedback was that the installation took longer than anticipated. By contrast, Scottish Power report that the majority of electricity and gas meters were each installed and commissioned in 20-30 minutes (it is noted that EDF's interventions were, in general, more complex). SSE telephone surveys highlighted that the majority of respondents (around 80%) said there was no disruption associated with the installation of smart meters, a further 15% said there was minor disruption and 1-2% said that their intervention had been a major disruption. There is no information on the causes for the differences in responses (in particular, whether they arose from differences in expectation or the actual features of the installation, such as duration, noise, dust or interruption/delay of activities).
- *Amount of equipment installed.* EDF customers reported some concern about the amount of equipment that was involved in the TV intervention. The equipment included the meters and communication hub, a set-top box, and powerline connectors for powerline carrier communications.
- *Displays (location).* EDF customers noted that, owing to signal strength, on occasions the products installed within the home had to be placed in second choice locations. If this relates to the location of the RTD, it may affect engagement with the occupant. It was also noted that in the TV trial groups, most had the intervention installed on secondary TVs and this was part of the reason that the information was less regularly accessed. Secondary TVs were used partly because of technical compatibility issues. E.ON separately notes that 67% of householders liked the portability of their RTD. It suggests that in developing a smart metering solution, the access point for energy consumption data (e.g. RTD or TV) needs to be in a convenient location for the occupant to access.

9.2 Smart metering equipment

9.2.1 Equipment development

At the beginning of EDRP, only electricity smart meters were commercially available as off-the-shelf products that had been previously proven in trials (e.g. in Northern Ireland and Italy). This was not sufficient for the purposes of

---

[53] EDF appears to be the only energy supplier that undertook an initial survey of the property prior to a separate installation visit. This survey took approximately 40 minutes.

EDRP: development was necessary to deliver solutions that provided remotely-read gas and electricity data and customer interfaces (RTDs) to interact with these data. The interfaces developed and tested in EDRP included:

- wireless RTDs;
- internet-based energy management graphs;
- TV interfaces;
- displays integrated in the home heating control system.

There were delays in some of the trials due to the development of the delivered solutions. In particular, the following were noted from the suppliers' reports.

- EDF had several technically complex interventions which resulted in delays to the trials. In particular, a longer than expected time was required for development and testing of the heating controller, TV and web-based interventions.
- SSE initially had a number of technical quality issues related to its meters and displays which were identified only after installation, and delayed the roll-out of some of the trial groups. Following this, SSE increased the level of its quality testing of suppliers' products. Many of these issues appear to be due to the products not being sufficiently tested in advance, both on their own and integrated with each other. SSE employed a greater range of metering types than in the other trials and this may account for the greater number of problems encountered.

9.2.2 Installation of smart metering solutions

Table 9.2 provides details of smart meter installation failures from technical reasons. It shows the percentage of installations that could not be undertaken. In the case of dual fuel homes, this includes cases where there were problems in the replacement of the gas meter and only a smart electricity meter could be installed. Installation failures due to not being able to access the property are not included in this list but, as discussed previously, they were a significant problem. The corresponding information is not available from SSE's reports.

Table 9.2 Reasons for aborting installation

	E.ON	EDF	Scottish Power
Inadequate Wide Area Network signal	1%	17%	4%
Health and safety issues	8%	2%	0%
Inadequate space to install meter	<1%	7%	10%
Inadequate Home Area Network	<1%	5%	6%
Unable to gain access to meter	4%	2%	not known
Premises on an Independent Gas Transporter network	not known	not known	3%
Other	1%	25%	0%

Inadequate Wide Area Network (WAN) signal

The smart meters typically communicated data back to the energy supplier via a GSM modem. The extent to which this WAN signal was a problem varied considerably between the suppliers.

- EDF reports that 17% of aborted installations were due to inadequate signal strength for WAN communications – the highest of the three suppliers. Initially the failure rate was double this value as the electricity meter, where the GSM modem was installed, does not tend to be in a location in the home where the signal strength is strongest. EDF subsequently changed from using a single network provider to using two different network providers and selected whichever was best for the particular home (other suppliers have similarly noted the need for the option of multiple network providers). This resulted in the failure rate falling to around 10% of installations. EDF also notes that a rework of the smart metering firmware was necessary to introduce a more robust mechanism for coping with the network terminating inactive communications in areas with poor signal strength or a heavily burdened base station. A six-hourly reset process was introduced to refresh the connection automatically.

- Scottish Power notes that, despite the instruction to have the network signal pre-checked before swapping the existing meter, some 18 sites were installed where insufficient network coverage was available at the meter. It turned out that the meter operatives on a few occasions checked the mobile coverage outside the property and not at the meter, where the signal strength could well be different.

There is insufficient evidence as to why EDF had a higher level of installation difficulties than the other suppliers. There are many possible factors that could contribute to this, including: (i) a more stringent initial screening process, (ii) lower strength signal transmitted by the communications equipment used, (iii) greater physical barriers to communication relating to the properties selected and the location of the communication equipment within the properties, and/or (iv) properties located in poorer reception areas. For example, E.ON reports that the majority of its installations were in urban locations and the trial did not include high-rise apartment blocks.

Health and safety issues

Both E.ON and EDF report rejecting meter replacements due to the state of the current metering installation. This included potential safety implications from installing a meter replacement. It is noted that such issues are likely to be common to any meter exchange. Where there was a safety concern with the current installation it was separately addressed (often requiring third-party involvement). E.ON notes that in some cases, the current installation was safe but there were safety issues associated with the replacement of the meter (e.g. lead pipes for a gas meter installation) and rectifying this was outside the scope of works planned for (and again may have required third-party involvement).

E.ON had to abort many more meter replacements for gas than electricity, even though only about 75% of the smart metered properties visited were dual fuel and thus required a gas meter replacement. E.ON reports 587 cases of the gas meter installation not being up to standard compared to 101 cases of the electricity meter not being up to standard. Examples of sub-standard electricity meter installations include bad or faulty cut-outs, rotten meter boards or cut-out boards, asbestos, and being unable to pull the main fuse. Examples of sub-standard gas meter installations include bad pipe work or a faulty emergency control valve. Insufficient information is provided to explain why problems are more prevalent for gas meters. It may be that there are more potential safety issues associated with gas meter installations. However, it is also noted that E.ON's gas meter installations were outsourced to a third party supplier and that some meter replacements were aborted on non-safety grounds as the areas of work required fell out of scope of the commercial arrangements.

Scottish Power reports that, where the electricity meter was located outside the property, the installer requires suitable weather conditions. It is deemed unsafe to work on live electricity circuits in conditions of driving rain. In the interests of health and safety, a number of installations had to be rescheduled as per normal practice owing to the inclement weather conditions.

EDF decided not to install smart meters where one of the occupants had a pacemaker, or where other medical equipment was in use at the property. This was part of its approach to recruiting trial participants, aimed at removing operational complications, for example from wireless communication compatibility or the unexpected results of remote software updates. This approach resulted in approximately 2% of installations being cancelled.

Inadequate space

In each trial, there were cases where there was insufficient space to install the new smart meter in the space previously occupied by the non-smart meter. This appears to be more prevalent for gas meter exchanges.

- E.ON reports that insufficient space for the new meter caused 73 aborted installations of a gas meter compared to 30 for the electricity meter.
- Scottish Power reports no problems in replacing electricity meters. However, it reports that the dimensions of the gas smart meter were different to those of common existing gas meters. Gas meters were swapped only where they could be fitted into the same location without much additional effort in repositioning pipe-work or supporting brackets.

E.ON, Scottish Power and SSE note difficulties where the existing gas meter was contained in a semi-concealed meter cabinet. In particular, Scottish Power stated that it was the most prevalent cause of its aborted installations as it required a special bracket to be made and there were concerns over the durability and integrity of the meter as a significant number of semi-concealed cabinets can find themselves holding a sump of water within which the new gas meter may have to endure being submerged for long periods.

SSE reports on glass bricks in front of meters (found in trials in South Wales) which limited the size of meter that could be installed and meant smart meters could not be provided in this situation.

Both E.ON and SSE note that current (non-smart) technology has enabled meters to be produced that are significantly smaller in recent years than before, and some installation points did not have enough space to accommodate larger hardware. In particular, SSE highlights that, owing to the reduced size of electricity meters, there is frequently less room between the distribution service head and the consumer unit, leaving little room for a smart meter, which occupies a larger footprint. SSE reports instances where it had to request the distribution business to re-position the cut-out to enable the smart meter to be installed.

Home Area Network (HAN)

For most smart meter interventions, the Home Area Networks were set up in a similar manner. The electricity meter was either integrated with or hard-wired to the GSM modem to communicate externally. The gas meter and RTD were then wirelessly connected to the electricity meter via low power radio transmission. For some initial installations, EDF hard-wired the meters and the RTD to a central communications hub with a GSM modem.

- EDF had some initial problems with the hard-wired option. It rejected households where the physical layout of the home would have resulted in excessive cable lengths for correct operation (the gas and electricity meters needed to be less than five metres apart). This was addressed mainly by moving to wireless technology for later installations.
- EDF had fewer problems with its wireless installation. EDF does report that the number of sites that failed for HAN issues would have been larger but WAN communications were tested first and some issues are common to both HAN and WAN communication. The main HAN issues were: (i) flats where the electricity and gas meter were located a long way from the living area with the RTD, (ii) large homes where the electricity and gas meters were located too far apart, and (iii) where the electricity meter was located in a metal meter box. EDF notes that issues related to blocks of flats or metal meter boxes would also fail on WAN communications. E.ON also notes that current metering is being installed in metal boxes ('mantel units') which stop most wireless signals.
- SSE reports that the wireless range achieved for gas data to the electricity meter in the trial was too limited in a significant proportion of properties visited. The majority of these installation issues were attributable to the physical location of gas meters which, it was suggested, are more commonly being sited in semi-concealed meter cabinets.
- Scottish Power reports that, even with pre-checks on the HAN signal strength, there were 17 instances where a gas meter would not bind with the rest of the network. In these cases, the new gas meter was left *in situ* to perform as a non-smart meter.

Access to meter

E.ON highlights that a significant cause of aborted installations was the need for prior remedial work that the customer was legally responsible for. This was mainly due to meters having been made inaccessible, sometimes by moveable objects, and sometimes by fixed items such as kitchen cupboards (also noted by SSE) or alterations to walls, etc. having been made to the property. E.ON notes that some of these incidences would simply require that customers are given enough forewarning, and that they are made aware of the access requirements so that they have time to comply. However, customers' resistance may be higher where they have inadvertently blocked access by changing the fabric of the property, as remedial work could be both destructive and costly.

EDF noted another problem in gaining access to the meter if it is in a communal cupboard that is locked and the customer does not have a key.

Premises on an Independent Gas Transporter Network

Scottish Power reports that to deploy gas smart meters on an Independent Gas Network would have required special arrangements with each network. This was considered out of scope for EDRP. Scottish Power notes that approximately 12% of gas meter installations across the UK are on an Independent Gas Network and appropriate arrangements will need to be agreed with these Networks.

Other

There are a significant number of 'other' reasons for installations to be aborted. Some of these have already been reported as part of the customer feedback and are not repeated here.

- E.ON separately quantifies that 0.3% of failures are due to not having access to turn off the electricity or gas prior to meter exchange. E.ON also provides a list of other, non-quantified, issues which contribute towards the total installation aborts: technical problems with the meters that occurred before installation was complete; problems with pipe work; electrical cable too close to gas meter; customer refused installation on visiting property; unable to remove the existing meter; asbestos; floor alteration kit required and the customer refused; could not guarantee boiler relight.
- The number of installation failures that are unexplained for EDF is large (25%). It is unclear what exactly these failures are but it is understood that they are a mixture of technical and administrative issues.

In addition, EDF noted particular difficulty during the installation for the following trial groups which required more effort.

- The TV intervention was the hardest to install as the installation interacts with many items in the home including the metering, the customer's broadband access point, replacing any existing set top box and reconnecting the TV in the same way as the customer had set it up previously.
- The heating controller was also particularly difficult to install as it had to be connected to the boiler. Furthermore, care had to be taken in setting up the HAN as it needed to include the boiler which is not traditionally located near the utility meters and the communications hub needed to be located such that all items were reliably on the network. EDF also notes that during the initial survey stage, the surveyor had to be competent to inspect the wiring configuration of any existing heating controller and have an understanding of boilers. Rather than training metering installers about boilers, it was found easier to train electricians who work with boilers about the basics on metering to do the survey. It is unclear from the reports who actually undertook the installations.

EDF undertook a site survey prior to a separate visit to install equipment. EDF notes that it is very unlikely to install smart meter systems that need surveys first, for reasons including cost, best use of time for skilled resources and also the impact on customer motivation with the delay it causes between recruitment and installation. This may affect the type of smart metering solutions that are suitable for wider roll-out.

Installation for the SSE Community Trials

SSE undertook three community trials (see Appendix D), each with a different approach to communications than for the rest of the trials.

SSE trialled a Radio Frequency (RF) mesh network in community North Leigh. Overall, the trial appears to have been technically successful (98-99% successful communication). However, two implementation issues, particular to a mesh network, are noted.

- During the installation period, until a sufficient number of meters were installed to provide adequate mesh network coverage, there were some communication difficulties. This is expected within a mesh network.
- Within the village, a relatively small number of houses took part. This meant that even the fully implemented mesh network did not provide sufficient coverage for a few outlying households.

SSE trialled a Powerline Carrier (PLC) solution within communities St Athan and Alyth. The meters provide half-hourly consumption data, via PLC to 13 Local Data Concentrators installed at the low voltage substations within the two communities. The Local Data Concentrators upload the consumption data daily via GSM to central storage.

Overall, the PLC communication success rates at the end of the trial were 96-97%. It is noted that communications failed where there was not enough meter concentration to provide adequate repeating. This is likely to be due to the relatively small number of homes that took part in the trial.

9.2.3 Operation of equipment during trials

Many of the operational issues are already highlighted as part of the customer feedback in Section 9.1. These are not repeated here.

General equipment problems

There were some general reliability problems associated with the smart meters and RTDs. These appear to be due to insufficient testing of the equipment, both separately and integrated, in advance of the trials. Furthermore, even though in some cases checks were done at the time of installation, there were both HAN and WAN data communication drop-out issues which, where possible, were rectified.

In some cases the communication issues were due to fluctuating signal strength so the data were stored by the meter and subsequently re-acquired. It is noted that, during periods of poor HAN communication, Scottish Power received customer complaints due to differences in readings between the gas meter and the RTD. EDF had an instance of a customer placing tins of food near the meter and accidentally blocking the signal.

WAN communication reliability

The energy suppliers' reports suggest that the overall reliability of data communication was close to 100% (EDF: 98-99%; E.ON: >99%; SSE: >99%). This is based on the performance of the systems towards the end of the trial when many of the initial communications problems had been resolved or the households removed from the trials. During a national roll-out, it will be necessary to install smart meters in all homes as well as minimising the number of initial 'teething' issues, given the very large numbers of customers involved.

EDF TV trials (including implications for any intervention based on broadband connection to the home)

EDF highlighted a number of issues with the implementation of this intervention. It is noted for the discussion below, in a similar manner to the other trials the smart meter data were transmitted wirelessly to the energy supplier via GSM. However, the energy consumption data were fed back to the consumer via broadband to their home and then, from the broadband point, EDF used PLC technology to transmit data to the set-top box attached to the television.

- Whilst PLC was generally good, it had difficulties with items like fish tank pumps, anti-surge power sockets or other PLC equipment (e.g. BT home hub set-up) interfering with the internal data communications.
- EDF suspects from the trials that, during the peak times of internet usage, the internet noticeably slows down. This could be a much greater problem if the national smart metering solution is based around the internet.
- If broadband was also used to transmit the metered data back to the energy supplier, there could be several problems. If customers do not pay their telephone bill, their broadband service is currently disconnected. Furthermore, customers may wish to switch off their router, e.g. to save energy at night. Finally, not all homes have broadband – EDF rejected 960 customers for the TV trial on the grounds that they had no broadband connection and a further 112 because their connection was USB rather than Ethernet.
- One customer, who is an IT expert, on having the TV intervention installed, worked out how the intervention worked technically. He worked his way back through the IT communications system. He was not able to access data for other customers as the data system was secure. However, it demonstrates potential data security issues for an internet-based communications system for national roll-out.

Firmware downloads

There was one report (by EDF) of a successful firmware download and update of the meter. This is something that would need to be tested further in advance of roll-out.

9.3 Data management

A multi-utility smart meter system is much more than just the smart meters and an RTD. Significant IT development was also necessary to deal with two-way data communication between the meters and the energy supplier (potentially via one or more third parties) and store half-hourly data records from both gas and electricity meters across multiple households. The challenge of integration of these local metering systems into the wider communications and data storage systems was met by the suppliers developing their own bespoke data management systems. The details, accuracy and reliability of these systems have not been fully described in the suppliers' reports.

It is noted that DECC is proposing a new entity, DataCommsCo (DCC), which will deliver central data and communications activities in the domestic smart meter roll-out. The functions will include providing a two-way communications channel between smart meters and a central communications hub to which smart meter data users (suppliers, network companies and other authorised third parties) will have access for specified purposes. Hence the utility suppliers will not, in practice, need to deliver all of the data and communication activities that they provided in EDRP. However, the issues identified in this study will still be relevant for the overall roll-out programme.

Data entry

There were a number of instances of problems caused by incorrect manual entry of data into the databases during the installation process. In particular, SSE reports that manual data collection of fixed asset details (e.g. meter serial numbers), and manual input into back office systems, creates a major risk of data errors. Furthermore, SSE notes that a valid meter reading cannot be delivered in the billing system until all systems involved in the end-to-end process have all data relevant to a particular customer synchronised. SSE recommends implementing automated systems and asset tracking (e.g. bar coding and extranets), automated flows to third parties and automated data validation checks.

Data processing

All four suppliers faced major issues in managing and processing their data during the trials. The suppliers suggest that the issues relate to the 'pilot' nature of the project and the necessary learning curve. Some of the issues relate to the development of new data management systems and business processes. Others relate to the immature nature of the smart meter technology supplying the data. Some key issues are identified below.

- *General faults.* There were a number of one-off 'teething' issues as the bespoke data management systems were implemented.
- *Synchronisation on site.* If communication failure occurs, it is simpler and more economical to resolve the issue whilst the engineer is on site. SSE recommends that as much synchronisation as possible of the full end-to-end data management process is conducted during the installation process.
- *Volume of data.* Some trials had initial problems handling the large amount of data from the smart meters. There were also subsequent problems where the volume of data was higher than normal, owing to the need to re-load metered data from previous days because of data communication problems.
- *Implementation of data validation processes.* Smart meter data have inconsistencies; data validation processes were therefore developed to check the data, and edit as necessary, for correct operation. There is likely to have been a greater need for data validation during the trials owing to the relative immaturity of the smart meter technology, but it is still very relevant to wider roll-out.
- *Automated processes for fault resolution.* When an error occurs, data volumes are such that manual processes cannot cope with the number of exceptions created by a large number of smart meters. Automated processes for fault resolution were recommended.
- *Meter capacity.* If there are problems with the communication networks (HAN or WAN), the data stored within the meter can be accessed subsequently once the fault is rectified. One supplier noted that the data retention within the meters needs to be at least 7 days.

SSE recommends that similar to the regulated half hour (HH) market (large commercial energy users that are required to have 'advanced' energy metering which automatically transmits energy data to the energy supplier), rules will have to be devised for gap filling for missing data for the smart-meter domestic market especially if the data are to be used for settlement processes.

9.4 Energy supplier recommendations

The energy suppliers highlight a number of learning points but this is a fast changing environment and some of the learning points highlighted below are already being addressed.

1) Sufficient time is required for product testing, system development and integration, staff training and development of customer services.

2) Suitably sized equipment needs to be developed for current installation scenarios. For example, energy suppliers have noted difficulties in installing meters both in newer homes with smaller existing meters and in semi-concealed gas meter enclosures.

3) To deliver a successful end-to-end process there will need to be interoperability standards. This is to ensure that products that will work together can be mixed and matched from different suppliers throughout the process. This includes common communication protocols and data formats and should allow for potential future applications, e.g. to help facilitate the smart grid.

4) The smart metering system needs to be relatively straightforward to install. This will minimise cost, disruption to the consumer and potential need to rectify problems subsequently.

5) Installation needs to be carefully programmed.
   a) EDRP highlighted significant problems in gaining access to some properties.
   b) Action should be taken to identify problem sites in advance where possible, to avoid delays to the installation programme.
   c) The installers need to be given adequate training and tools to carry out reliable installations on site and quality management processes put in place.
   d) As much checking as possible should be undertaken during the installation as to the correct operation of the individual products installed, as well as the HAN and WAN communication networks. This is to prevent the need for post-installation access, which is costly and may be problematic to achieve.
   e) Proactive work is needed to identify and engage with other parties (e.g. the Distribution Network Operator or the Gas Transporter) that may also need to be involved in the installation.

6) As part of the installation process, there needs to be a strategy to deal with properties that are difficult to access and/or where there is remedial work that the customer is legally responsible for and is necessary prior to installing the meters. Suggestions include the following.
   a) Clear guidance (e.g. through a marketing campaign) such that customers understand their responsibilities.
   b) Guidelines as to sufficient time to allow customers to provide access or take remedial action where necessary.
   c) Access to an (independent) body to advise customers on legal responsibilities, help vulnerable groups and possibly mediate in disputed cases.
   d) Processes to address cases of outright customer refusal to provide access or undertake remedial work.

7) The overall smart metering system needs to deliver reliable communication.
   a) Problems were identified with HAN and WAN communications. The majority of trials used low power radio transmission for HAN and mobile (SMS) communications for WAN. Whilst, overall, it appeared that these technologies were successful there were a significant number of problems during installation and operation and greater communication reliability is required if these solutions are used during roll-out. Even a small percentage of problems could result in a large number of homes being affected.
   b) Alternative solutions may need to be developed for more difficult installations where the resultant signal strength is reduced because of, for example, meters located in enclosures, large distances between HAN communication devices, intervening building structures or poor mobile reception.
   c) Whilst EDRP particularly focussed on the transmission of energy consumption data to the occupants and the energy suppliers, testing is also required of the ability to remotely update the metering technology, e.g. uploading tariff changes to the meter or updating the installed software.

8) A number of data management issues need to be addressed.
   a) Manual data collection of fixed asset details (e.g. meter serial numbers) and manual input into back office systems creates a major risk of data errors. Automated systems and asset tracking (e.g. bar coding and

extranets), automated flows to third parties and automated data validation checks would support national roll-out. This is noted as being a significant issue as suppliers will have to develop a wide range of new interacting systems to support the roll-out activities. This will include appointment booking systems, mobile platforms (to issue instructions and information to operatives and to enable capture of site-specific data), meter data management and meter asset management systems and new data transmission links with the proposed DCC and other partners.

b) A number of parties are involved in the transfer of data from the meters to the energy suppliers. This creates opportunities for different rules to be applied. There is a need for the development of underlying processes common to all.

c) Smart meter data will have inconsistencies and, therefore, data validation processes need to be agreed to check the data, and edit it as necessary, for correct operation. This is particularly the case where data cleaning needs to happen across several different market participants. Furthermore, rules will have to be devised for gap-filling for the half-hourly data if used for billing purposes.

d) Automated processes for fault resolution need to be developed. When an error occurs, data volumes are such that manual processes cannot cope with the number of exceptions created by a large number of smart meters. Smart metering systems need to have sufficient memory capacity to allow retrieval of meter data once any communication faults are rectified.

e) Testing is also required of data management from the energy supplier to the meter, e.g. uploading tariff changes to the meter.

10 Conclusions

This section presents the main conclusions from EDRP. In particular, it draws together two sets of findings.

1) The impact of the interventions on consumer behaviour and energy consumption. This covers the results from both the energy suppliers' findings and AECOM's findings from its analysis (taking into account the quality assessment) and broader literature review.
2) Key practical and technical issues identified from EDRP. These particularly focus on smart meter systems, given the Government's policy of the national roll-out of smart meters.

10.1 Impact of interventions on energy consumption

10.1.1 Energy efficiency advice

The EDRP findings are consistent with the literature insofar as an effect of energy efficiency advice was not always seen and, when it was seen, the reduction in annual consumption was up to 5%. This was important to confirm in the British context, given the previous dependence on evidence from elsewhere.

The particular combination of advice and historic feedback on consumption that EDF deployed (along with smart meters) reduced electricity consumption by 2.3% overall in the first in-trial year but separately statistically significant only for dual fuel customers (4.6% saving) and not for electricity-only customers (0.9% saving).[54] Furthermore, the effect was persistent into the second in-trial year for both customer types combined (4.0% saving) and dual fuel customers (2.4% saving), and became statistically significant for electricity-only customers (5.0%). The E.ON trial found some weaker evidence of savings due to advice, in combination with smart meters and historic feedback on consumption, but it is difficult to quantify because of the particular schedule of interventions.

These effects were seen in spite of survey findings that there was low customer engagement with the EDF energy efficiency and consumption history material.[55] This perhaps gives a clue as to why the EDF trial was effective: information was provided in simple, short statements, over a period of time – minimal but well presented and easy to absorb a little each month. The SSE advice booklet was more comprehensive but required more effort from householders. An interesting twist was that there was greater engagement with the SSE booklet if an RTD was provided, so some of the effect attributed to the RTD may have been due to the RTD prompting interest in the advice and, conversely, the advice helping householders to use the feedback from the RTD.

The message is that advice should be provided but the details of delivery (e.g. clarity, quantity of information, timing) and combination with other interventions, are critical. The credibility of advice can also be undermined if it carries risks (e.g. blanket advice to draughtproof homes, regardless of potential for increasing risk of condensation or build-up of toxic combustion gases) or its effectiveness is uncertain. EDRP has taken a step forward in understanding these details but there is more work to do and the optimum approach is likely to vary between customer segments and to change over time as more energy knowledge becomes commonplace.

10.1.2 Historic feedback

The EDRP findings are consistent with the literature insofar as an effect of historic feedback was not always found and, when it was found, the impact on consumption was up to 5%. This was important to confirm in the British context, given the previous dependence on evidence from elsewhere.

---

[54] Gas consumption could not be assessed because of the lack of baseline data.

[55] The delivery of the Paper intervention was less reliable in the second in-trial year and this may have unduly influenced the survey responses, since the survey was at the end of the trial. It is possible that engagement had been greater in the first in-trial year.

As noted in Section 10.1.1, the particular combination of advice and historic feedback on consumption that EDF deployed (along with smart meters) reduced electricity consumption and the effect was persistent into the second in-trial year. The other trials found no evidence of savings due to historic feedback.

The effects in the EDF trial were seen in spite of survey findings that there was low customer engagement with the EDF material. The billing material was more obvious to customers than that provided by E.ON (large, in colour and on separate sheets) and the advice was more accessible than that provided by SSE. The SSE survey data did show that customers found the information on bills more useful if they also had advice, which helps to explain the effect on consumption found by EDF. This survey also showed that many customers found the bill data and RTDs to be complementary, with a value in providing both.

Historic feedback can be useful but the details of delivery, and combination with other interventions, are critical. EDRP has taken a step forward in understanding these details but there is more work to do and the optimum approach is likely to change over time as people become more familiar with their consumption levels, and to vary between customers, depending on their current understanding of energy use and interest in reducing consumption.

10.1.3 Benchmarking (comparative or normative feedback)

Only the SSE trial used benchmarking (without smart meters or RTDs). There was a small but significant effect on electricity consumption, comparing the trial group with advice, historic feedback and benchmarking information with the most similar group (which had the advice and historic feedback but not benchmarking information). The additional savings attributable to benchmarking were 1%.

Although a small effect, this is one of the clearest pieces of evidence for an effect of benchmarking, in the context of a literature that is generally positive but allows no quantification. There is concern in the literature that those who consume less than the benchmark amounts may start to consume more; therefore, careful consideration needs to be given to how benchmarking might usefully be deployed in the population in general.

10.1.4 Customer engagement using commitment to reduce consumption

Commitment to reduce consumption did not have any detectable effect on consumption, other than an effect (savings increased) due to making customers aware of the trial. This is a new finding, with no comparable evidence in the literature.[56]

10.1.5 Smart meters

The literature does not provide direct evidence of the impact of installing a smart meter without any other scheduled interventions. Two of the EDRP trials (E.ON and SSE) provide the first evidence on this, showing that some aspect of the experience of getting a smart meter can itself prompt a reduction in energy consumption, particularly gas consumption (savings of around 3%). The clearer effect for gas consumption makes sense in the context that simple one-off changes (e.g. reducing a thermostat setting) can have big effects on gas demand.

Reductions in gas consumption in the E.ON trial persisted (in fact, slightly increased) to the first quarter of the second in-trial year (i.e. for 15 months) and for one or two further quarters in some groups. This suggests that the experience of getting the smart meter prompted some initial action (e.g. turning down a thermostat) but the effect may require support over time from other interventions (e.g. advice or billing information) to be sustained for longer periods.

[56] In short-term trials, there is related tentative evidence of a benefit of realistic but stretching targets, combined with frequent feedback, but there is insufficient evidence to quantify such an effect.

The relevant feature of the customers' experience of the smart meter installation cannot be stated with certainty but there was a clear difference between E.ON and SSE, which set out to explain to customers what they were receiving, and Scottish Power, which ran the trial as a blind trial (every effort was made to prevent customers knowing they had a smart meter, by presenting it as a routine meter replacement). While smart- and non-smart conditions could not be compared in the EDF trial, the survey data showed that (a) customers expected, and could have benefited from, more engagement and instruction during installation of equipment and (b) some respondents were happy to benefit from no more meter readers or estimated bills and nothing else.

10.1.6 Real-time display (RTD) devices

The literature is clear that providing an RTD typically brings about a reduction in energy consumption but the percentage savings vary widely and appear to depend on climate (and consequently the main uses of energy – cooling or heating, and whether the heating is electric). From the literature review, tentative estimates of what could be expected in the UK were that a base level effect of RTDs alone could be less than 3% electricity savings whereas supplementary interventions that increase engagement could double or triple the benefit. Far fewer studies have tested effects on gas consumption but they have generally shown a benefit; while it is not feasible to quantify it, savings tend to be of a similar order to those for electricity.

The EDRP findings were generally in keeping with this, and allow greater certainty on the magnitude of effects in the UK, and over a longer period than most of the literature. The effect of clip-on RTDs was modest, only one trial showing a small (1%) significant effect. The trials were more positive in showing savings from the mains devices provided with smart meters – generally 2-3% for electricity but higher in the EDF trials (4% overall but 7% for electricity-only customers). The greater impact in the EDF trial may be because the accompanying advice was more effective (see Section 8.2). These effects were persistent. Only the Scottish Power trial showed no positive effect of RTDs and this may be related to the fact that the trials were not presented to customers as smart meter trials (see Section 8.6).

Effects on gas consumption were less clear: the EDF data did not allow analysis and, as for electricity, Scottish Power did not find a significant effect. Neither SSE nor E.ON found any incremental effect of RTDs on gas consumption above the effect of a smart meter. SSE did find a significant interaction effect, such that the combination of an RTD with the "incentive to reduce" had a greater effect than either on its own. This effect is hard to explain, given that the incentive to reduce applied only to electricity and it did not have an effect on electricity consumption. In any study of the scale of EDRP, there will be occasional isolated inexplicable findings and they may be random.

The findings support the expectation from theoretical considerations, that real-time feedback is clearly more relevant to electricity consumption than to gas. Applications of gas tend to be subject to occasional adjustments having long-term effects, which are less amenable to influence by real-time feedback.

It is also clear from the survey findings that householders paid less attention to the display of gas information than the display of electricity information (this may be due in part – but only in part – to the order in which button pushes accessed different types of information). The surveys also showed that cost information was used and valued more than unit (kW) information. Displays of CO_2 emissions were generally not widely noticed or used or perceived as useful. This is all consistent with the literature, which additionally shows that portability is a benefit (at least initially) and that appliance-specific feedback can have additional effects on energy savings.

The literature also identifies two key points that were not explored further in EDRP, but which could be put into guidance on using the devices:

- householders may find RTDs more useful in confirming savings after attempts to reduce consumption, rather than using an RTD to initiate savings;

- RTDs can be used to check that everything has been switched off before going to bed or leaving the house.

The stronger effects on electricity consumption in the smart meter trials may be explained by the presence of the smart meter or by the differences in the display devices. The latter is inherently more likely and this explanation is backed up by the survey evidence: the mains RTDs, which displayed both gas and electricity consumption and tended to have more sophisticated functions, were consistently (across trials) more likely to be fitted, retained, used and rated positively. Having an RTD also tended to result in more positive perception of the smart meter. The SSE survey found that reasons for not using the RTD differed between clip-ons (where the functionality of the device itself was the dominant reason) and mains RTDs (where the usefulness of the information provided was dominant).

The implication is that RTDs will have less impact if customers are relied upon to fit them: a significant proportion will simply not be fitted. Furthermore, where they are fitted, guidance needs to cover how to use the information that RTDs provide, not just how to access the information. More generally, the written instructions may sometimes have been unclear or too complex, especially for people who had been sent the device, rather than making an active decision to buy it. Hence support from the installers may be particularly important.

There is little published evidence on the use of audible alarms or visible signals in tandem with RTDs. The limited evidence (generally qualitative) indicates that audible alarms tend not to be used but visible signals with clear meaning may be viewed positively and used as part of an energy-saving strategy. This was backed up by EDRP, the audible alarm in EDF’s trial causing no incremental reduction in consumption and attracting only negative response in the EDF survey. The “traffic lights” visual signal of consumption level, in contrast, was often the most positively rated feature.

A thermometer may be a simple and effective feedback device but there is insufficient evidence in the literature to say whether it actually has any effect on energy demand. The Scottish Power trial used a thermometer but not in a way that adds to the body of evidence. However, displays of temperature data via RTDs were generally rated positively and may have been particularly useful in the early stages in responding to advice to reduce thermostat settings.

Evidence on the provision of online real-time feedback is limited and, as yet, inconclusive in the literature. It was not tested in EDRP but surveys suggested that the online provision of feedback would have been more effective if real-time data had been provided.

10.1.7 Heating controller integrated with RTD

There is nothing in the literature on the effect of an RTD with integral heating controller so the EDF trial was the first test of its kind. There was no effect on electricity consumption and the change in gas consumption could not be analysed. Nevertheless, the customer survey revealed a positive response to the intervention, as measured by awareness of the device, perceived usefulness and overall rating. The most likely explanation for the discrepancy is that customers valued the device as a heating controller but did not use it so much as an RTD. This would be in keeping with the literature showing that people are often disinclined to switch between functions available from displays.

10.1.8 Time of use tariff/incentive

The literature shows clear evidence that time of use tariffs can shift consumption from the peak period, and often also bring about reductions in total energy consumption. However, the evidence is almost exclusively from studies in hot regions (where the dominant energy demand is for air conditioning) and cold regions with electric heating. The limited evidence from the UK suggested small reductions (3% or less) in overall electricity demand.

EDRP did not provide convincing evidence of an overall reduction in demand. In the case of EDF, the RTD provided was more basic than in the other trials and the customer survey showed that it was less likely that customers would be aware of its existence, and those who were aware rated it as being less useful.

There were, however modest effects on shifting load from the peak period. This comes from the EDF and SSE trials, both of which showed a stronger load-shifting effect at weekends than on weekdays. Estimates of the magnitude of shifting effect vary with trial but were up to 10%. The EDF data showed that the effect is stronger with smaller households (1 or 2 people), thus providing a clear focus for where such interventions should be targeted. The effect was weaker in the SSE trial and this may be because awareness of the intervention was limited and it was seen as overly complex.

10.1.9 Incentive to reduce consumption

The literature provides little substantive evidence on incentives to reduce consumption except for the general (and obvious) point that sufficient incentive will prompt people to reduce consumption, but only for as long as the incentive is kept in place. There are some concerns that emphasising the financial motive in this way could detract from long-term savings prompted by other motives.

E.ON did not use incentives to reduce consumption. The other suppliers did (only for electricity) but only Scottish Power saw reductions in consumption when the incentives were applied – only in the case of credit customers and only for short periods. The Hawthorne effect is a sufficient explanation of the Scottish Power findings.

10.1.10 Overview of web-based interventions

The literature shows the potential benefits of online services to reduce energy demand, but also shows that the potential is rarely realised. EDRP confirmed this, with neither of the suppliers who used web-based interventions (EDF and SSE) seeing any energy savings as a result. The trials also showed that a major reason for failure is likely to be lack of engagement with the web sites, not necessarily a lack of effect among those who use the sites.

10.2 Population segment effects

Across all the trials, there was limited evidence of how different population segments were affected by the interventions.

EDF found clearly defined effects, with smaller households being more likely to save energy overall and to shift consumption from the evening peak period. There were also differences between electricity-only and dual fuel customers. These differences were not due to the electricity-only customers not having gas heating but they are otherwise difficult to explain with any certainty.

The E.ON findings clearly varied between customer strata. The customers selected because of their initial high consumption reduced their consumption whatever the intervention but this merely shows what is expected: high consumers have most potential to reduce and this can be explained by “regression to the mean”, rather than the impact of any particular intervention. Behavioural changes were more weakly evidenced in this group, again supporting the conclusion that changes in consumption were artefactual.

More interesting is the difference between E.ON’s ‘fuel poor’ (FP) and not ‘fuel poor’ (NFP) groups. Although there are problems with the definitions of these groups, the FP group is likely to be generally less affluent, with more pressure on energy costs. The generally more positive response to interventions in FP most likely signifies their greater motivation to save money.

SSE found significant effects of postcode (for electricity only) and Mosaic group (for gas and electricity) but not Grid distribution point (an industry geographic code). Including geographic region in the analysis did not affect the outcome and there was not a significant effect of this additional variable. This emphasises the need to control for location and demographic variables but the type of analysis used did not permit identification of specific locations or Mosaic groups that had a higher propensity to reduce consumption. Since postcode was significant but not the more geographically systematic variables, the details of any location effect are difficult to define. In no case did one of these variables have a significant interaction with a trial intervention, i.e. the intervention effects did not vary with location or demographic category.

Most of the EDRP findings (and most of the literature) relate to credit customers – there is relatively little about prepayment customers or prepayment meters. EDF and E.ON did not address this sector. Only SSE included prepayment smart meters (for electricity only) in its trial, with one trial group being given smart meters and an RTD. This group made savings similar to those in the credit customer groups.

Scottish Power made more extensive investigations of prepayment customers but not with smart meters. These customers did not make savings under any interventions, even the financial incentives that brought about temporary reductions in demand in the credit customer groups. Nevertheless, some prepayment customers did meet their savings targets (arguably more consistently than credit customers), so it may be that prepayment customers were making savings anyway and could not save more in response to the interventions.

Prepayment meters tend to be installed where customers are having difficulty paying energy bills, so they generally have lower consumption anyway (and this was certainly the case in the Scottish Power trial). Therefore they have less scope to reduce. However, they may be more practiced at monitoring consumption, more aware of costs and have greater (financial) motivation to save energy. With these counterbalancing factors, savings appear to be possible (as shown by SSE) but not always achieved (as shown by Scottish Power). Logically, efforts to help prepayment customers reduce consumption might focus on means and opportunity, the motive already being present, but this was not explored in the trials.

10.3 Practical and technical issues

The deployment of a smart metering system has proved complex. Much of this is because, at the beginning of EDRP, limited solutions existed. There needed to be significant development of smart meters, customer interfaces and data management systems to handle the remotely accessed energy consumption data. Given the advance in technology, many of the issues would not be encountered if starting EDRP today. The remaining issues identified are already informing work under the smart metering programme.

The study has shown that one size does not fit all with regard to smart metering technology. Different geographical locations experience different signal strengths which can affect the ability of the meter to send or receive information. The location of a smart meter within the property is also important, e.g. communications can be affected if the meter is in a metal box, in a basement, or is too far from the RTD. Installing smart metering in basements, blocks of flats, or communal housing tends to be more difficult with regard to gaining access, external communications and utilising the Home Area Network (HAN). The location or set-up of an existing meter installation can mean that remedial work is required to replace it, e.g. if the smart meter is larger than the original meter and does not fit into the meter box. Practical issues affecting installation are generally easy to check for in a property. However, in some cases, technical problems may not be recognised until after the metering system has been installed. This can result in additional effort to rectify problems, including potentially revisiting the property, which should be avoided if possible.

Significant data management issues were encountered in the project. It is important to ensure accurate and timely processing of information relevant to the installation (e.g. the meter serial numbers installed in the property) for use by the supplier (e.g. for billing purposes). Given the large volume of energy data, it is easy to become swamped by faults in the data when they occur and, for example, the development of automated processes for fault resolution proved successful (e.g. automatically requesting the meter to resend missing data). Furthermore, EDRP highlighted the importance of data validation processes to check and edit the data as necessary, for smooth operation.

Delivering the trials has demanded a range of different skills from the suppliers. Installers need the right technical skills to complete jobs and need additional training to be able to install both gas and electricity meters and to deal with any unforeseen safety or legacy issues. Softer skills are essential too, so that installers can clearly explain the new technology to customers. Customer advice services will need to build up their knowledge of the new technology to be able to answer consumer enquiries.

Gaining access to properties proved challenging. EDRP has highlighted difficulties in initially contacting customers, obtaining the customer's agreement to the meter exchange and customers keeping appointments. Many of these issues, and others identified in EDRP, would be encountered in "business as usual" meter replacement activity, whether smart meters or not, although over a longer period and therefore more easily managed.

A significant issue going forward will be the customer interaction with the smart metering equipment (including RTDs). All trials reported problems of customers having difficulties in understanding the new equipment provided. This is not simply about the design of equipment, although that has been improving: care is needed in how RTDs are 'sold' and explained to the wide cross-section of population such that they know how to access and use the data that the display provides, and take an interest in doing so.

11 Glossary

ACORN	A demographic classification system. For further details see: http://www.caci.co.uk/acorn-classification.aspx
Analysis of Variance (ANOVA)	A statistical test that can be used with normally distributed data to test whether two or more samples come from the same population or, alternatively, whether the samples differ from each other on the dependent variable being examined.
Benchmarking	Comparison of customer's consumption with that of similar households.
CAMEO	A demographic classification system. For further details see: http://www.callcredit.co.uk/products-and-services/consumer-marketing-data-and-segmentation/cameo-classifications/cameo-uk
Chi-squared test	A non-parametric test of whether an observation is made with the same frequency in two or more samples.
Credit customer	Pays for the use of energy after it is consumed.
DF	Dual fuel (i.e. the customer buys both electricity and gas from the same supplier). In EDRP smart meter trials, both electricity and gas smart meters are fitted.
E7	Economy 7 customers.
EO	The supplier managing the trial supplies the customer with electricity only. Gas may also be supplied to the property but by a different supplier. In EDRP smart meter trials, only an electricity smart meter is fitted.
Hawthorne effect	People may change their behaviour merely because they know they are being observed or tested, regardless of any specific attempts to change their behaviour.
HEC	Heating controller (an EDF trial group). See Section 3.1.
HU	E.ON customer group with high baseline energy use. See Section 3.2.
FP	'Fuel Poor'. Used to classify E.ON trial groups. See Section 3.2.
Interventions	Measures to reduce energy consumption.
Mann Whitney U test	A non-parametric statistical test to determine whether two samples come from the same population or, alternatively, whether the samples differ from each other on the dependent variable being examined.
Mosaic	A demographic classification system. For further details see: http://www.experian.co.uk/business-strategies/demographic-classifications.html
Multivariate statistical analysis	Analysis to investigate the effects of more than one variable simultaneously.
National roll-out of smart meters	The Government is committed to rolling out smart meters to both domestic customers and smaller businesses. For further details see: http://www.decc.gov.uk/en/content/cms/tackling/smart_meters/smart_meters.aspx
NFP	Not 'fuel poor'. Used to classify E.ON trial groups. See Section 3.2.
Non-parametric statistical methods	Techniques that do not rely on data belonging to any particular (e.g. normal) distribution or intervals in the data all having the same value.
Ocean	A demographic classification system. For further details see: http://www.caci.co.uk/ocean.aspx
Peak period	The period of the day in which there is the highest energy demand.
Prepayment customer	Pays for the use of energy in advance.
Real-time feedback	A means of providing current (i.e. live) information of the energy used in the home.
Regression to the mean	The statistical phenomenon whereby values above (or below) the mean are likely to decrease (increase) if measurements of the same group are made a second time.
RTD	Real-time display. Also often referred to as "in-home display" (IHD).
Smart meters	In the context of EDRP, these were electricity and gas utility meters that allowed the remote collection of half-hourly energy consumption data.

Time-of-use (TOU) tariffs	Variable tariffs that change with time of day and/or between days. For example, the tariff structure can be set to be highest at periods of peak demand to encourage customers to reduce consumption during that period.
TOUT	Time of use tariff (an EDF trial group). See Section 3.1.
Tukey test	A test applied when ANOVA of more than two groups shows an overall significant effect, to identify which groups differ from which other groups.
URA	Usage Reduction Alert (an EDF trial group). See Section 3.1.