Energy Demand Research Project: Final Analysis

Appendix A: Findings Reported by the Energy Suppliers

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

Appendix A: Findings Reported by the Energy Suppliers

This Appendix summarises the main EDRP findings reported by the four energy suppliers to Ofgem. The suppliers reported to Ofgem every six months throughout the project. These reports contain commercially sensitive information and so have not been published.

- Key findings from customer surveys and focus groups are stated first because they provide context for the consumption findings.
- This is followed by the suppliers' findings from analysing their energy consumption data. Findings are presented first for interventions that did not involve a smart meter and then for interventions that did involve a smart meter (or, in the case of Scottish Power, a prepayment meter reconfiguration).
- There is also a brief evaluation by AECOM of the quality issues that should be borne in mind when interpreting or applying suppliers' findings. Appendix B provides a more thorough final quality assessment of the findings following AECOM's analysis.

The trial interventions, trial groups and other aspects of the trial method are described in Section 3 and only a brief reminder of the trial groups is provided in this section.

AECOM's own analysis of the data and findings (including a summary incorporating the suppliers' findings) are described in Sections 4 to 7.

A1 Findings by EDF

EDF tested only one intervention that did not include a smart meter, the Read-Reduce-Reward (RRR) scheme.

EDF tested seven combinations of interventions in groups that also received a smart meter and accurate bills, without visits by meter readers. The trial group names are abbreviated as follows.

- Paper (Additional bill data + Energy efficiency advice).
- Wall Panel (RTD + Energy efficiency advice, sent via the RTD).
- TV (Consumption information + Energy efficiency advice, sent via the participant's TV).
- URA (Wall Panel type RTD + Usage reduction alert + Energy efficiency advice, sent via the RTD).
- TOUT (Basic RTD + Time of use tariff + Energy efficiency advice).
- HEC (Wall Panel type RTD + Heating controller integrated with RTD + Energy efficiency advice, sent via the RTD).
- Web (Consumption information + Energy efficiency advice, online).

A1.1 Survey findings

Introduction

EDF carried out two main customer surveys (both by telephone): during recruitment (therefore covering the whole sample) and at the end of the trial, seeking to collect data from all households and achieving 1,046 out of 1,947 (54%). In addition, 41 depth interviews (mainly face-to-face) and customer focus groups were conducted to gain a richer understanding of participants' response to the trial. Where the technology employed in the trial provided data on participants' use of the technology, this was also analysed (referred to as "click-through surveys"), providing data for the Wall Panel (n=18), TV (n=127) and Web (n=88) groups. The sample size was too small to draw useful conclusions for Wall Panel.

As context for findings related to specific interventions, 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 is particularly strong for all interventions with an RTD.
    - The real-time energy consumption information of Wall Panel, HEC and TOUT was preferred to retrospective information in the Web, TV and Paper groups.

Specific interventions

Regarding each trial group or intervention, the following points are reported. In addition, the interviews suggest that customers expected, and could have benefited from, more engagement and instruction during installation.[1]

RRR

Customers were not aware they were part of trial – they had signed up primarily for practical reasons (no more meter readers or estimated bills). Reading one's own meter appears to have raised some awareness about costs and consumption and 41% agreed it helped plan or budget for energy use. Focus groups placed this intervention in the middle two of the eight interventions: neutral for *Usability* and *Engagement*, negative for *Impact on behaviour*.

Paper

Customer research indicated very little engagement with the additional bill data and advice provided. Customers "at best glanced at" the energy efficiency advice (although this may have been enough, given the short, focused tips on each occasion). Focus groups placed these interventions in the bottom three: neutral for *Usability*, negative for *Engagement* and *Impact on behaviour*.

Wall Panel

In groups with the Wall Panel type RTD, customers were frequently accessing information (39% daily or several times a week) though mainly limited to the first and second screen and only by one householder (the bill payer). Reported impacts on use of household appliances are said to have "tended to occur early on" but it is not clear whether they then persisted. Focus groups placed this intervention in the top three: positive for *Usability* and *Impact on behaviour* but neutral for *Engagement* because it needs a way to maintain impact in the long term. Overall, 79% found it very or quite useful.

There is a weak positive correlation between how useful respondents found their RTD[2] and how often they looked at it (Kendall's tau = 0.294, Spearman's rho = 0.337, both $p<0.001$).[3] There is no significant correlation between how useful respondents find their RTD and two other variables available from the survey (the number of electrical appliances or how financially well off the household describes itself).

TV

The click-through survey showed limited use of this intervention. While 85% of households in the TV group used it at least once, having an average of 15.3 sessions, 36% used it fewer than three times (58% ≤ 7 times; 78% ≤ 20 times, 22% > 20 times). It was used a little more for reviewing electricity consumption than for gas, for energy consumption (kWh) than cost, and for historic consumption rather than monthly, weekly or forecast. There was little accessing of the carbon emissions screens, usage history or energy efficiency messages. It was used more with screens requiring fewer clicks to access (which may explain why kWh was apparently preferred to cost). This implies that any wider roll-out should put the most impactful screens earlier in the sequence (e.g. cost before welcome screen). Usage of the intervention increased initially (although this may just be more customers entering

[1] Installers' training did not focus on customer interaction; they gave out spare user guides and answered questions, otherwise customers were referred to customer services team. This was intended to minimise "risk of bias" but it may have reduced the effectiveness of RTD-based interventions in particular, by limiting both the ability to interact with the intervention and the motivation to do so. Also, in the absence of training, installers might have invented their own advice, which could increase variance in householder response, or even give customers wrong messages.

[2] This analysis covered all households with an RTD, not just the Wall Panel group.

[3] The p-value represents the confidence in a finding (a value of less than 0.05 is generally accepted as an indication that we should be confident in a finding).

the trial) then declined over the trial period to less than a third of peak. The median screen view lasted 21 seconds (mean = 24 seconds) and the median time between uses was 9 days (mean = 15 days).

Customers thought the intervention seemed very attractive at the outset but actual engagement was low and declining. There was disappointment at the lack of real-time and appliance-specific data (energy consumption information was provided to the TV on the following day). Focus groups placed it in the bottom three: neutral for *Engagement*, negative for *Usability* and *Impact on behaviour*. An issue noted by EDF is that more interesting information is always available through this medium.

URA

This intervention was not sufficiently well understood and appears to have had very little effect; 3 of the 5 interviewed in depth in this group were not even aware of the alarm function and only one was engaging positively with it. Impact was limited by lack of understanding about the threshold and how it is calculated, coupled with potential irritation. Focus groups placed this intervention in the middle two: neutral for *Usability* and *Engagement*, negative for *Impact on behaviour* (the alarm function does not seem to be a positive addition to get long term engagement).

TOUT

The RTD used in the TOUT group was (purposefully) low-tech, which was a disappointment to some householders. This may affect the householders' ability to interact with the intervention and motivation to do so but this would need to be tested further because people might like the idea of a high-tech device but not use the functions in practice. There was a perceived positive impact on behaviour and consumption but some customers were concerned that TOUT was having a negative financial effect on them (even though this was not possible within the terms of the tariff). Focus groups placed this intervention in the top three: positive for *Engagement*, and *Impact on behaviour*, neutral for *Usability*.

HEC

This had the highest level of use of all interventions because customers had to interact with it to control their heating. Feedback was mostly positive – simple to use and empowering, but this related mainly limited to controller function rather than consumption feedback. Focus groups placed this intervention in the top three: positive for *Usability*, *Engagement*, and *Impact on behaviour*.

Web

Only 17 customers ever used the web-site. Only 2 of the 5 interviewed in depth in this group ever accessed their account, even though all were fairly frequent web users. There were problems with logging in and those who did access the site expressed confusion over interpreting the graphs of energy consumption. Potential for impact in this study is therefore very limited. Focus groups placed it in the bottom three: negative for *Usability*, *Engagement*, and *Impact on behaviour*. An issue noted by EDF is that more interesting information is always available through this medium.

Specific actions

The final survey asked whether specific actions had been taken during the past two years (i.e. the trial period). EDF reported some variation between trial groups for these actions but statistical significance was not reported and no account was taken of whether there was scope for the action (e.g. whether the existing boiler was new at the start of the trial, or the walls were already insulated). Nevertheless, the variations may help to explain variations in savings within the trial sample – see Table A1.

In two cases (loft and cavity wall insulation), the three groups with a 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). This is a hint – and not more – that some more long-term benefit may come from these interventions, which was not observable during the trial period.

The percentages of post-trial survey respondents who agreed or strongly agreed that the "smart meter technology" had helped to plan or budget for energy use were: HEC 75%, Wall Panel 65%, URA 59%, TOUT 61%, others <42%. This again supports the view that HEC may bring a more long-term benefit.

Table A1 Percentage of each trial group that had carried out the specified actions during the trial period

Trial	*Loft insulation*		*Cavity wall insulation*		*Double glazing*		*Draughtproofing*		*New boiler*	
	N	*%*	*N*	*%*	*N*	*%*	*N*	*%*	*N*	*%*
Paper	102	26.0	72	25.0	121	7.4	111	17.1	120	21.0
Wall panel	176	29.0	123	19.5	195	12.8	182	20.3	194	22.0
TV	89	15.0	50	6.0	102	18.6	100	20.0	100	22.0
URA	101	20.0	53	13.2	121	9.9	115	11.3	116	16.0
TOUT	101	24.0	60	21.7	110	14.5	107	14.0	108	20.0
HEC	94	18.0	60	15.0	98	17.3	97	21.6	97	25.0
Web	55	13.0	34	8.8	57	14.0	52	11.5	57	21.0
Control	173	17.0	108	13.9	199	15.6	188	15.4	197	22.0
Average		20.3		15.4		13.8		16.4		21.1

Shaded cells are more than 3% above the average.

A1.2 Non-smart meter interventions – effects on energy consumption

Findings

The analysis compared trial group with control group consumption with no reference to baseline (pretrial) consumption. It did, however, control for outdoor temperature, indirect demographic group, house type and age, number and language of occupants, and insulation measures. Electricity consumption was significantly higher (by 5.7%) in the RRR group compared with the control group whereas gas consumption was significantly lower (by 2.1%).

Plots of gas consumption show the difference from the control group to arise mainly from winter consumption (quarters 1 and 4 of each year).

Quality assessment

There is a complex mixture of concerns with the non-smart meter findings findings, the most important being a lack of pretrial baseline data. The findings from other trials show that, even with strenuous efforts to match trial and control groups, large differences in pretrial consumption can remain. Data quality issues, a small sample and the lack of a non-smart meter control group have a substantial further negative impact on confidence in the findings that is unlikely to be overcome by including baseline data in the analysis.

A1.3 Smart meter interventions – effects on energy consumption

Findings

Whole trial period

The analysis compared trial group with control group consumption with no reference to baseline (pretrial) consumption but controlled for outdoor temperature, indirect demographic group, house type and age, number and language of occupants, and insulation measures. Data were corrected for weather and season and log transformed to reduce the impact of outliers.

Corrected electricity consumption was lower in some groups, higher in others (relative to the control group):

- 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.

Corrected 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.

Change in consumption during trial period

To assess the persistence of effects, the above results of regression analysis need to be interpreted in the context of quarterly consumption, as shown in Tables A2 to A4 for electricity and Tables A5 to A7 for gas. The tables are based on all households for which there are data in a given quarter; comments given below take account of the fact that these households will not be the same for all quarters. Because of the staggered start of each trial, the numbers of households build up over the first two quarters of each trial so the early trends may be distorted by this. This is indicated in the tables below by showing the first two quarters (and rolling annual figures based on them) in grey. There will also be changes later in the trial if households drop out but these changes are less and more gradual.

The regression findings, reported above, control for other variables so they are probably the best guide to the overall difference but the quarterly figures are not always consistent with the regression findings. For electricity, the figures support the regression analysis for TOUT but not Paper, Wall Panel or HEC. In the case of TOUT and URA, the percentage figures suggest an initial benefit that was not fully sustained.

Table A.2 Electricity consumption (average corrected daily consumption per quarter, kWh)

Trial	*Q1 2008*	*Q2 2008*	*Q3 2008*	*Q4 2008*	*Q1 2009*	*Q2 2009*	*Q3 2009*	*Q4 2009*	*Q1 2010*	*Q2 2010*	*Q3 2010*	*Mean*
Control	15.0	13.2	11.8	13.9	14.5	10.9	10.8	13.7	14.0	11.2	10.8	12.7
Paper	16.3	13.3	12.6	15.2	15.5	11.3	11.1	14.4	15.0	11.5	10.8	13.3
Wall panel		11.9	11.9	14.5	14.6	11.2	11.1	14.0	13.6	11.2	10.3	12.4
TV						12.8	12.6	15.3	15.3	12.2	11.8	13.3
URA		6.3	10.2	12.5	13.6	10.5	10.5	13.4	13.6	10.9	10.3	11.2
TOUT		7.8	8.1	9.5	10.4	8.5	8.3	10.9	10.9	8.6	8.3	9.1
HEC				8.3	11.8	10.9	10.8	14.0	14.3	11.2	10.3	11.4
Web			11.7	15.0	16.3	12.4	11.9	15.1	16.3	12.7	12.2	13.7

Table A.3 Electricity consumption as a percentage of control group consumption

Trial	*Q1 2008*	*Q2 2008*	*Q3 2008*	*Q4 2008*	*Q1 2009*	*Q2 2009*	*Q3 2009*	*Q4 2009*	*Q1 2010*	*Q2 2010*	*Q3 2010*	*Mean*
Paper	108	101	107	109	107	104	103	105	107	102	100	105
Wall panel		90	101	105	101	103	102	103	97	100	96	98
TV						117	116	112	110	109	109	105
URA		47	87	90	94	96	97	98	97	97	95	88
TOUT		59	69	68	72	78	77	80	78	76	77	72
HEC				60	82	100	99	102	103	99	95	90
Web			99	108	113	114	110	111	117	113	113	108

Table A4 Rolling annual electricity consumption as percentage of control group[4]

	Year ending							
Trial	*Q4 2008*	*Q1 2009*	*Q2 2009*	*Q3 2009*	*Q4 2009*	*Q1 2010*	*Q2 2010*	*Q3 2010*
Paper	106	106	107	106	105	105	105	104
Wall panel		99	102	103	102	101	101	99
TV						113	112	110
URA		80	92	94	96	97	97	97
TOUT		67	71	73	77	78	78	78
HEC				84	95	101	101	100
Web			109	111	112	113	113	113

For gas, the figures give an impression contradictory to the regression analysis for HEC. The percentage figures for the Wall Panel group show a marked peak in the first two quarters of 2009 but perhaps savings towards the end of the trial. In the Web group, the overall higher gas consumption is largely due to a marked peaks in the earlier part of the trial.

Table A5 Gas consumption (average corrected daily consumption per quarter, kWh)

Trial	*Q1 2008*	*Q2 2008*	*Q3 2008*	*Q4 2008*	*Q1 2009*	*Q2 2009*	*Q3 2009*	*Q4 2009*	*Q1 2010*	*Q2 2010*	*Q3 2010*	*Mean*
Control	117.3	43.7	21.0	83.9	99.7	29.8	13.9	71.2	106.4	34.5	15.3	57.9
Paper	98.3	34.1	22.2	94.2	107.0	31.3	16.1	77.9	114.9	37.2	17.3	59.1
Wall panel		22.6	21.2	95.8	237.6	83.3	13.9	67.7	98.2	31.5	14.9	68.7
TV						11.0	17.9	80.6	116.6	38.3	17.4	47.0
HEC				49.5	71.9	24.2	13.1	60.6	92.4	30.1	14.8	44.6
Web			22.4	102.9	278.2	101.2	18.2	76.3	118.2	41.8	17.4	86.3

[4] The rolling annual consumption is a full year's consumption, advanced by three months in each column. Each year overlaps the adjacent year by nine months.

Table A6 Gas consumption as a percentage of control group consumption

Trial	*Q1 2008*	*Q2 2008*	*Q3 2008*	*Q4 2008*	*Q1 2009*	*Q2 2009*	*Q3 2009*	*Q4 2009*	*Q1 2010*	*Q2 2010*	*Q3 2010*	*Mean*
Paper	84	78	106	112	107	105	116	109	108	108	113	102
Wall panel		52	101	114	238	279	100	95	92	91	97	119
TV						37	128	113	110	111	114	81
HEC				59	72	81	94	85	87	87	97	77
Web			107	123	279	340	131	107	111	121	114	149

Table A7 Rolling annual gas consumption as percentage of control group

	Year ending							
Trial	*Q4 2008*	*Q1 2009*	*Q2 2009*	*Q3 2009*	*Q4 2009*	*Q1 2010*	*Q2 2010*	*Q3 2010*
Paper	94	104	109	109	108	108	109	109
Wall panel		152	187	189	187	119	93	93
TV						102	112	111
HEC				70	79	86	87	87
Web			215	220	221	142	113	112

Variation in effects between participants

Electricity consumption varied with indirect demographic group (Acorn/Ocean categories), house type and age, number of occupants, number of lodgers, number of electrical appliances and how financially well off the household describes itself. Gas consumption varied with indirect house type and age, number of floors, number of bedrooms, number of occupants, number of children, the presence of cavity walls or draughtproofing on external doors, and having a new gas boiler in the last two years.

Of all these variables, only the number of occupants was found to have an interaction effect with trial group. EDF applied separate regressions where household size is 1 or 2 and when it is 3 or more. For both gas and electricity, the trial groups overall consumed less relative to the control group where there were only 1 or 2 occupants ($p<0.001$ for each regression). The results are summarised in Table A8: in 7 out of 12 cases the intervention-attributable difference in energy consumption is more negative for smaller households.[5] In one case there is no difference to the first decimal place (but in the same direction as the majority if greater precision in the figures is accepted).

[5] The regression does not necessarily imply that the difference is statistically significant in each case.

Table A8 Average percentage difference in consumption from control group over one-year in-trial period

Trial	*Electricity or Gas*	*Number of occupants*	
		≤2	*≥3*
Paper	Electricity	0.1% less	0.2% less
	Gas	0.6% more	0.1% more
Wall panel	Electricity	0.5% less	0.5% less
	Gas	0.3% less	0.1% less
TV	Electricity	5.4% more	6.4% more
	Gas	9.2% more	11.1% more
URA	Electricity	0.1% less	0.2% less
TOUT	Electricity	0.3% less	0.9% less
HEC	Electricity	4.9% more	6.2% more
	Gas	7.2% more	7.3% more
Web	Electricity	4.8% more	5.9% more
	Gas	6.1% more	8.1% more

Shaded cell indicates the group with the greater reduction relative to the control group.

Quality assessment

There is a complex mixture of concerns with the EDF findings for overall reduction in consumption, the most important being a lack of pretrial baseline data. The findings from other suppliers' trials show that, even with strenuous efforts to match trial and control groups during the recruitment process, differences in pretrial consumption can remain.

This creates a difficult backdrop for consideration of the findings for each specific intervention; in order to have confidence in a particular finding, there would need to be strong mitigating factors relating to the specific intervention and/or a very strong effect that could not obviously be accounted for by the identified biases. Neither of these applies: while the influence of other factors on confidence in the findings varies with intervention, none is sufficient to overcome the lack of baseline data.

The Web and TV consumption findings should not be seen as representing the potential of TV-based interventions. Any further consideration of using web- or TV-based feedback would need to develop more effective ways of targeting the households most likely to benefit and delivering a more effective system, including real-time feedback. The study tested only one approach in each case, with no targeting of those customers most likely to use the technology provided. In the case of TV, it should also be noted that that the set-top box itself is estimated to demand 200 kWh per year.

For the other trial groups, while there are some concerns over aspects of the trial design, there is a good case for further analysis if baseline data can be obtained. In relation to wider application of similar interventions, design improvements identified through the EDF trial could achieve greater impact.

Observation of the course of change during the trial is not affected in the same way by lack of baseline data. In the case of Paper and Wall Panel, if there is a genuine effect of the intervention, that effect appears to be sustained (or even increase) over time. In the case of TOUT, any effect appears to be greater but subject to some reduction in magnitude over the first few quarters.

The effect of number of occupants is the only clear finding so far of an effect of a household variable that appears to interact with intervention (and fuel type).

A2 Findings by E.ON

E.ON tested four interventions that did not include a smart meter, with trial groups as follows (a customer engagement intervention, TG3, was withdrawn before trial completion and results are not reported).

- TG1: Additional bill data (graphs on quarterly bills showing historic energy consumption information).
- TG2: Energy efficiency advice (monthly tips sent by post) + Additional bill data.
- TG4: Clip-on RTD + Additional bill data.
- TG6: Energy efficiency advice + Clip-on RTD + Additional bill data.

Customers were stratified at postcode level using CSE's fuel poverty indicator and subsequently treated as separate 'fuel poor' (FP) and not 'fuel poor' (NFP) groups. They are referred to in this report as customer strata.

E.ON tested four combinations of interventions in groups that also received a smart meter, as follows.

- TG5: Smart meter + Accurate billing and no meter reading visits.
- TG7: Smart meter + Accurate billing and no meter reading visits + Monthly bills + Additional bill data (graphs on bills showing historic energy consumption information).
- TG8: Smart meter + Accurate billing and no meter reading visits + Monthly bills + Additional bill data + Energy efficiency advice (monthly tips sent by post).
- TG9: Smart meter + Accurate billing and no meter reading visits + Monthly bills + Additional bill data + Energy efficiency advice + Mains RTD (electricity and gas).

Customers with smart meters were recruited in four strata: (a) FP (b) NFP (c) high user, dual fuel (HU_{DF}) and (d) Economy 7 with no gas supply from E.ON (E7). These customer strata were subsequently treated as separate groups in the analysis. The FP group was further divided into electricity-only (FP_{EO}) and dual fuel (FP_{DF}) customers. Only TG5 and TG9 included NFP and E7 customers.

A2.1 Survey findings

Introduction

E.ON conducted telephone surveys after interventions commenced (Jun-Jul 2008 for non-smart meter groups, including control groups, Jan 2009 for smart meter groups). The overall sample size was 1,203. At the end of the trial (Jun-Jul 2010), a further survey telephone survey was conducted with a sample of 2,739 households (60-101 per trial group, 21-100 per control group).

The final survey is of interest here because it offers insight into how participants perceived the interventions and what, if anything, they did in response to them. E.ON's own analysis of the survey data concentrated on providing detailed descriptive statistics, broken down by trial group and customer strata. Findings are tabulated separately for trial groups with and without smart meters.

Differences between trial groups and customer strata were subject to statistical tests but these had low power because they were applied to separate tables of groups and strata, resulting in low household numbers in each comparison. This is probably why only a few differences were declared significant out of a large number of tests (11 significant out of around 460 tests). With about 2% of tests being declared significant at the 5% level, no individual test can be regarded as statistically meaningful (one would expect 1 in 20 (5%) of tests to give a significant result at the 5% level even if there were no differences between any of the groups). A few findings of interest are reported here but please refer to the new analysis in Section 5.2 for a more extensive examination of the survey findings.

Awareness of the smart meter

In only one case is there a cluster of significant effects for one variable. Responding to the question “Do you have a smart meter?”, TG9 participants (who had an RTD) were consistently more likely to answer yes than the other smart meter groups (who did not have an RTD), by a margin of around 20%. The first concern here is that the question is not ideal because some participants might have known they had a new, different meter but not called it a smart meter. The significant effect may therefore be interpreted in one of two ways, either (a) the RTD made householders more aware that they had a smart meter and/or that it was called a smart meter, or (b) householders thought the question was referring to the RTD itself as being the smart meter.

Awareness of the RTD

Of the participants who had the clip-on RTD, 22% were aware of it and used the device every day or several times per week. Awareness of the RTD used with smart meters differed between customer strata: 58% in FP, 52% in NFP, 54% in HU_{DF} and 40% in E7. The figures were almost reversed for the percentage checking daily, if aware: 38% in FP, 49% in NFP, 58% in HU_{DF} and 69% in E7. As a result, the percentage aware *and* checking daily did not vary significantly between groups, but the figures were generally higher than for the clip-on RTD: 21% in FP, 25% in NFP, 32% in HU_{DF} and 28% in E7.

Sociodemographics of HU_{DF} groups

HU_{DF} households tended to have higher incomes, more appliances and be likely to buy energy-efficient appliances. This is some indication that they may be typified as high users because of wealth, rather than because of poverty (i.e. having old, inefficient homes and appliances, and possibly spending many hours at home because of poverty and/or retirement).

A2.2 Non-smart meter interventions – effects on energy consumption

Findings

The analysis was based on comparing individual trial groups with their respective control group (with no factorial analysis, comparison across trial groups or accounting for possible confounding variables). The principal dependent variables were electricity and gas consumption, aggregated over the trial period, compared with a business-as-usual control group appropriate to each trial group, relative to each group’s pretrial baseline.

None of the interventions had a significant effect in either of the customer strata, on either gas or electricity consumption. In addition, none of the interventions had a significant effect in either customer group on either gas or electricity consumption when data were analysed quarter by quarter, except in a few isolated quarters in some conditions (some increases, some decreases).

Quality assessment

The capacity of the trial to find significant effects is reduced by noise introduced by the design and method as a whole, although this would be partly balanced by most of the identified biases tending to increase the likelihood of a significant finding. The customer survey found that none of the trial groups differed significantly from the control group in the number of behaviour changes; confidence in this result itself is low but it tends to increase confidence in the non-significant changes in energy consumption.

The net effect is a modest overall reduction in confidence in the null results, especially for the aggregated annual data. The quarter-by-quarter analysis by E.ON eliminates some of the concerns about the aggregated data (i.e.

where they arise from pretrial or initial biases) but introduces new issues around smoothing of seasonal effects by interpolation of non-smart meter data (as described below in Section A2.3) data.

Confidence is further reduced by factors specific to implementation of the interventions, in the sense that better implementation of the same type of intervention might have reduced consumption. The null finding is, nevertheless, more robust for the RTD intervention than for the others.

Looking across the two sets of analyses (annual and quarterly data), the conclusion is consistent that none of the interventions reduced consumption. Furthermore, each intervention was tested (individually or in combination) in at least two trial groups and two customer strata.

On balance, there is reasonable confidence in the null findings. Nevertheless, the possibility cannot be excluded that different implementation of the same type of intervention would be more successful.

A2.3 Smart meter interventions – effects on energy consumption

Findings

The analysis was based mainly on comparing individual trial groups with their respective control group (with limited factorial analysis), comparison across trial groups or accounting for possible confounding variables. The principal dependent variables were electricity and gas consumption, aggregated over the trial period, compared with a business-as-usual control group appropriate to each trial group, relative to each group's pretrial baseline.

The smart meter trial groups show greater seasonal extremes than the non-smart control group 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 drift between meter reading and trial period. If so, then analysis at sub-annual level would be valid only for 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 the seasonal differences found but this needs to be better understood for any future trials.

Electricity

There were statistically significant reductions in electricity consumption in some trial groups and some customer strata with smart meters, as shown in Table A9:

- FP (only the dual fuel customers with full package of interventions);
- HU_{DF} (similar percentage reduction in all four intervention groups).

Table A9 Changes in electricity consumption relative to the control group

Customer stratum	TG5 (Smart meter only)	TG7 (Smart meter + monthly bills)	TG8 (Smart meter + monthly bills + advice)	TG9 (Smart meter + monthly bills + advice + RTD)
FP_{EO}	0.14	0.06	-0.02	
FP_{DF}	0.04	-0.06	-0.07	-0.44 (-4.1%)
NFP_{DF}	-0.03			-0.17
HU_{DF}	-0.76 (-4.3%)	-0.89 (-4.8%)	-0.94 (-5.1%)	-0.80 (-4.7%)
E7	-0.01			-0.32

Shaded cells indicate significant difference between trial group and control (% change also shown).
Negative values mean a reduction in consumption (mean daily energy use, kWh).

Gas

There were statistically significant reductions in gas consumption in some trial groups and some customer strata with smart meters, as shown in Table A10:

- FP (reduction not increasing with number of interventions);
- HU_{DF} (similar reduction in all groups with some intervention in addition to the smart meter).

Table A10 Changes in gas consumption relative to the control group

	TG5 (Smart meter only)	TG7 (Smart meter + monthly bills)	TG8 (Smart meter + monthly bills + advice)	TG9 (Smart meter + monthly bills + advice + RTD)
FP_{DF}	-2.48 (-5.3%)	-1.78 (-3.8%)	-3.97 (-8.4%)	-2.89 (-6.0%)
NFP_{DF}	-0.86			-2.23 (-4.5%)
HU_{DF}	-0.29	-2.23 (-3.9%)	-2.39 (-4.2%)	-2.43 (-4.4%)

Shaded cells indicate significant difference between trial group and control (% change also shown).
Negative values mean a reduction in consumption (mean daily energy use, kWh).

Quality assessment

The study's capacity to find significant effects is subject to a complex mixture of sources of noise and bias.

On balance, there is reasonable confidence in the null findings for gas. There is less confidence in the null findings for electricity, especially in the FP groups, because of different influences in the design (the complications of separating FP_{EO}, FP_{DF} and FP_{DF} without a gas smart meter) and the greater range of conditions producing non-significant findings. The non-significant results require further checking. In any case, there is a reasonable chance that different implementation of the same type of intervention would be more successful.

Overall there is good reason to doubt the significant effects on electricity consumption for HU_{DF} customers. These customers started with high levels of consumption and it is expected that at least some of the reduction results from 'regression to the mean'.[6] Only the effect of the full package of interventions in the FP_{DF} group retains any confidence. Unexplained conflicts in findings for the NFP and E7 groups require further investigation.

Overall there is some doubt over the significant effects on gas consumption, especially in the HU_{DF} groups, but further analysis is merited.

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

[6] 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.

A3 Findings by Scottish Power

Scottish Power's trial was conducted over three phases, with the same households being subject to a series of different interventions.

Phase 1 implemented interventions without smart meters in four trial groups (TG3 had not yet entered the trial).

- TG1 (Credit customers) and TG4 (Prepayment customers): Additional bill data (two mailings of year-on-year consumption history) + Energy efficiency advice (seasonal) + Clip-on RTD.
- TG2 (Credit customers) and TG5 (Prepayment customers): Additional bill data + Energy efficiency advice.

In *Phase 2*, Scottish Power upgraded the meters in all four trial groups – smart meters for credit customers and reconfigured meters for prepayment customers. The reconfiguration of prepayment meters allowed consumption data to be broken down by time of day. For TG1, the RTD was upgraded to a mains powered device displaying gas consumption in addition to electricity. TG1 and TG2 both received more granular consumption history data. TG4 and TG5 both received prepayment consumption transaction history.

- TG3 (Credit customers) entered the trial in the middle of Phase 2 and had the same interventions as those used in TG1 at this stage.

In *Phase 3*, the technology deployed remained the same and the main change at this stage was the introduction of the "Green Challenge" (GC) financial incentives to reduce consumption. The challenge was spread over four "Waves", combining three 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: incentive to pledge to reduce gas consumption (in Wave 3).

Throughout the three phases, there were business-as-usual control groups for credit (CG1) and prepayment (CG2) customers.

A3.1 Survey findings

Introduction

Scottish Power conducted three main surveys, referred to as survey Waves 1, 2 and 3 (note that these 'Waves' were different to the four Green Challenge 'Waves' in Phase 3). Each had a similar content except where questions were changed to reflect the different interventions that were in effect at each stage. All three main surveys were a mix of face-to-face and telephone interviews (the two types are not separated in the reported findings).

- Wave 1: near end of Phase 1 (Dec 07 – Jan 08), n=515.[7]
- Wave 2: Phase 3, during GC Wave 1 (Dec 08 – Jan 09), n=523.
- Wave 3: Phase 3, during GC Wave 4 (Dec 09 – Jan 10), n=435. Supplemented by 13 depth interviews and two focus groups (9 participants in each group).

Additional information is drawn from monitoring of customer acceptance and use of the RTDs provided.

Results are not reported here if a more in-depth analysis of the same questions is reported in Section 6.

[7] In addition, at the end of Phase 1 there was a telephone survey of 49 customers who had been provided with RTDs but reported that the device was not working. The results of this survey are not reported.

Energy-saving measures

At both Wave 2 and Wave 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.

Advice

Although 37% said the mailings were telling them nothing new, 66% would like to receive more of the same information, and qualitative research pointed to the their role in reminding consumers what can be done. The focus groups noted the importance of being kept regularly informed on energy matters and relating to energy use in a personal rather than generic way. Repetition of simple tips and advice "was not necessarily considered an issue" but "viewed if accessible and presented in the right form as being a gentle and appreciated memory jogger".

Clip-on RTD

At installation, 30% of TG1 and 22% of TG4 refused the RTD. The reasons not known but it may be because there was no warning about the installer arriving and whoever was at home at the time did not want to take sole responsibility for it. Incomplete take-up limits the capacity of the intervention to have any impact and might, in effect create two subgroups, thus increasing the within-group variance. These subgroups cannot be identified in the Scottish Power database.

Of the clip-on RTDs installed in TG1, 42% were still operational when replaced in Phase 2. Of these, 36% were kept by the customer, even though they were getting a new RTD. Of the remaining operational units, about 50% had new batteries, indicating that battery life was low but some householders thought it worth replacing them.

Mains RTD

Asked what information provided by the RTD had had most influence on efforts to save energy, the 'traffic lights' indicator of consumption was well ahead on 45%, followed by the cost of electricity used (27%), units of electricity used (16%), cost of gas used (6%), units of gas used (4%) and CO_2 emissions (2%).

Focus group respondents were pleased to have received the RTD and did not report any difficulty in understanding it, except in relation to greenhouse gas emissions. They felt it had affected their energy consumption and helped them to monitor consumption relative to the targets set in Phase 3. The traffic lights were the most spontaneously mentioned feature. The RTD was used in conjunction with the letters on consumption history and the two were seen to work together.

Intervention communications

The percentage of respondents who recalled the mailings 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) – see Table A11. It is uncertain whether this is because the information was recalled because it was used, or it was used because it was recalled (more likely the former, since the letters were offering financial benefits).

Self-rated comprehension of various aspects of the mailings (on a scale of 0 to 5) did not differ between trial groups but was higher in Wave 3 than in earlier Waves (4.4 in Phase 3 vs 4.2[8] for the Green Challenge (GC) letter at the

[8] Adjusted for different aspects rated.

end of Phase 2 and 4.1 for earlier communications). Focus groups also said that the GC mailings were easy to understand, and financial motivation to save energy was the motivation most likely to be mentioned. They used words like simple, realistic, achievable, personal, measurable and enjoyable, in describing their GC experience.

Table A11 Percentage of respondents who recalled Scottish Power mailings

	Wave 1	Wave 2	Wave 3
TG1	59%	51%	84%
TG2	62%	66%	83%
TG3		79%	80%
TG4	45%	57%	83%
TG5	45%	47%	76%

A3.2 Non-smart meter interventions – effects on energy consumption[9]

Findings

The analysis used the consumption ratio (rather than difference) between the in-trial and pretrial period, aggregated over 90 days and compared to a control group. Analysis was then based on three "Test Points" in Phase 1, with separate analysis of each trial group and no factorial or multivariate analysis.

Neither electricity nor gas consumption ratio differed between any of the trial groups and their control group.

Quality assessment

This is the only phase of the trial that is (a) free of interference from experiences in earlier phases and (b) has a pretrial baseline in the analysis (Phases 2 and 3 effectively used Phase 1 as the baseline). Overall the metadata tend to give confidence in the finding. The implication is that the three interventions independently would also have had no effect. A more effective analysis could be performed by combining all data from the trial period, rather than separating the period into three Test Points and omitting three months of data between the Test Points. However, the graphical displays of data do not suggest that additional significant effects would be found. Further analysis should also use difference from baseline, rather than ratio.

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

A3.3 Smart meter interventions – effects on energy consumption

Findings

All findings are based on mean consumption ratio for electricity and gas, aggregated over a month relative to pre-intervention baseline and compared to control group. Analysis was based on two Test Points in Phase 2 and five in Phase 3. The analysis was based on individual trial groups with no factorial or multivariate analysis.

[9] Prepayment meters were non-smart throughout the trial but Scottish Power reconfigured those in the trial groups for Phases 2 and 3, at the same time as credit customers in the trial groups received smart meters. Because of the coordinated timing, findings for reconfigured prepayment meters are reported later, alongside the findings for smart meters.

Phase 2

Neither electricity nor gas consumption ratio differed between any of the trial groups and the control group.

Phase 3 – electricity

There were statistically significant reductions in electricity consumption ratio at some Test Points for credit customers (see Table A12) but not prepayment customers. Test Points 6, 7 and 10 represent Waves 1, 2 and 4 of the Green Challenge respectively. Test Point 9 represents Wave 3 and Test Point 8 represents the period between Waves 2 and 3. The effect was similar with and without an RTD and for both volunteers (TG3) and those who joined the trial unaware (TG1 & TG2). The intervention started in winter and a sharp reduction in demand coincided with this winter start, diminished over summer and autumn and re-emerged in first 2 months of following year. The reduction at peak was approximately 10% (1 kWh/day per household).

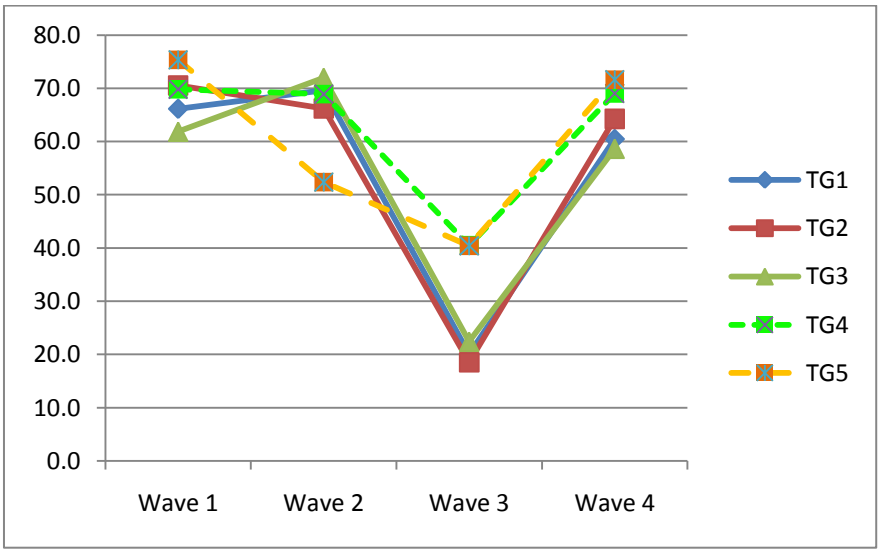
Table A12 Ratio of electricity consumption in Phase 3 to baseline consumption

	CG1	TG1	TG2	TG3
Test Point 6: 1 Feb 09	1.026	0.941	0.911	0.936
Test point 7: 1 Apr 09	1.027	0.944	0.939	0.945
Test point 8: 1 Aug 09	0.984	0.949	0.970	0.952
Test point 9: 1 Nov 09	0.972	0.958	0.980	0.970
Test Point 10: 1 Feb 10	1.008	0.934	0.917	0.942

Shaded cells indicate significant difference between trial group and control.

Although there was no significant reduction among prepayment customers, the percentages of customers who met the challenge (Figure A1) suggests that the drop in success rate in Wave 3 was less for prepayment than for credit groups and the average success rate across the four waves was higher for prepayment customers. This suggests that prepayment customers do have scope for reducing consumption and perhaps they were pacing themselves differently because it was more difficult for them to make savings.

Figure A1 Percentage of customers who met the Green Challenge in each Wave



Phase 3 – gas

There were statistically significant reductions in gas consumption ratio at two Test Points for credit customers (see Table A13) but not prepayment customers. These test points represent Wave 3 (Test Point 9) and the period between Waves 2 and 3 (Test Point 8), when no intervention was active. The effect was similar with and without an RTD and for both volunteer participants (TG3) and those who joined the trial unaware (TG1/TG2). The demand reductions emerge gradually, in contrast with the sharp onset for electricity. Differentiation between trial groups and control diminished in magnitude during the autumn and winter months.

Table A13 Ratio of gas consumption in Phase 3 to baseline consumption

	CG1	TG1	TG2	TG3
Test point 8: 1 Aug 09	0.979	0.933	0.940	0.933
Test point 9: 1 Nov 09	0.961	0.902	0.912	0.899

Shaded cells indicate significant difference between trial group and control.

Quality assessment

Phase 2

Overall the metadata tend to reduce confidence in the non-significant effects. As with Phase 1, Phase 2 combines interventions in a way that means the separate effects are not identified although this could be done for the RTD effects. Since the combination did not have a significant effect, it is inherently unlikely that either would do so separately. The time lag in smart meter installation (5 months in a 7-month trial period) and lack of a full year of data, undermine the finding.

Further analysis of data from the last quarter of the phase, based on those households that have a smart meter throughout the quarter, could compare groups with and without an RTD (without comparison to the control group, which did not have smart meters).

It remains possible that reduction in demand could be brought about by more effective delivery and content of advice and consumption data, particularly for the prepayment groups, delivery of accurate bills and/or take-up of the RTD. In this context, if enough customers who used the RTD (at all or throughout) can be identified, further analysis of this group might be worthwhile (using difference from baseline, rather than ratio).

Phase 3 – electricity

Apart from a range of sources of bias, the possibility of Hawthorne effects cannot be discounted: people can change their behaviour because they are being observed rather than because of the intervention itself. This effect tends to be more prominent at the start of a study and, as each Wave of intervention was only 3 months, it is difficult to know whether the impact arises from the intervention itself or the Hawthorne effect. While prepayment customers might have been subject to the same influences, they would have had less scope to respond and might reasonably have felt less inclined to do so.

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

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

Overall the metadata leave little confidence in the significant reductions in Phase 3. 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.

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

Phase 3 – gas

The issues here are similar to those for electricity, the major difference being that the effect on gas consumption emerged gradually over the summer, rather than the sharp change seen at the start for electricity consumption. This is less likely to be due to the Hawthorne effect but is more likely to be a seasonal smoothing effect. It could alternatively be that the main gas intervention (the gas pledge) was introduced in August and its impact on heating would initially be minimal but increasing through the autumn.

The variation over time might be seasonal as suggested by Scottish Power but this does not fit with the different seasonality compared to electricity consumption, with customers concentrating on gas consumption outside the heating season. It may be questioned why an effect on gas consumption should be seen more in the warmer months than in winter. It is possible that, in warmer months, it is easier to cut hot water use (e.g. by using electrically heated water) and heating demand (by using electric heating or by accepting lower but still comfortable temperatures with the heating off). This is pure conjecture but using electricity instead of gas would be counter-productive in cost and CO_2 terms.

The lack of significant reductions in the prepayment groups is in stark contrast – the prepayment customers differed obviously in the means of payment (therefore meter type) but also in having lower initial consumption and a range of biases in different directions. On balance, the non-significant finding is credible in comparison with the credit groups.

A4 Findings by SSE

The findings presented here relate to interventions targeting individual households. SSE's community trials are covered separately, in Appendix D.

SSE tested six combinations of interventions in the absence of a smart meter.

- Energy efficiency advice (an advice booklet).
- Clip-on RTD.
- Energy efficiency advice + Additional bill data (graphs on quarterly bills showing historic energy consumption).
- Energy efficiency advice + Additional bill data + Benchmarking (customer's consumption compared with that of other households of comparable demographic group in the region).
- Energy efficiency advice + Additional bill data + RTD.
- Energy efficiency advice + RTD.

SSE tested the effects of seven interventions applied in homes provided with a smart meter (but with continued meter reading visits).

- Monthly bills.
- Additional bill data: graphs on monthly bills showing historic and half-hourly energy consumption.
- Energy efficiency advice: an advice booklet sent by post and/or provided online.
- Mains RTD.
- Time of use tariff: incentive to shift electricity use from peak period consumption.
- Incentive to reduce consumption: reward for year-on-year reduction in electricity consumption.
- Web information: personalised consumption history available online.

The smart meter trials had a partly factorial design and the actual group-level additional interventions are set out in Section 3.4 (Table 3.8).

Customers were stratified according to whether they were *Aware* of the trial, *Unaware* or aware and *Committed* to reducing energy. In some trial conditions, only the Aware and Committed groups were used.

A4.1 Survey findings

Introduction

SSE carried out two surveys at the start of the trial (autumn 2007). In order to preserve the "Unaware" group, the survey covered only "Aware" and Committed" households. The larger survey used self-completion questionnaires, posted to all Aware and Committed households when they joined the trial. Initially, completion of the questionnaire was required but this was abandoned because it was a barrier to recruitment. The completed, or part-completed, questionnaire was returned by "around 4,000" households. In addition, the smaller survey comprised 420 face-to-face interviews. Both surveys served only to provide general descriptive data on the trial sample and the SSE reports do not break down the data by trial group.

Early in the trial, SSE also carried out two surveys focusing on specific interventions.

- Following delivery of clip-on RTDs (December 2007), 105 customers were interviewed by telephone to check the success of delivery and initial reactions. This was followed up by a larger survey (1,320 interviewed out of 2,000 approached) but results have not been provided.
- During the initial months of providing additional data through graphs on bills (November 2007 onwards), telephone surveys were carried out (172 customers out of 554 approached but data reported for only 150). This survey checked recall of receiving the bill and new information and any initial impact of the information.

During the latter stages of the trial (Feb-May 2010), SSE commissioned 8 group discussions, 24 depth interviews and 1186 telephone interviews to gather data on how households had responded to the interventions. Only the telephone interview findings are reported. AECOM did not obtain the data from the final survey and therefore more attention is given to SSE's survey findings than was given to findings by the other suppliers.

Survey of clip-on RTD recipients

Of those responding to the survey, 83% had received the RTD and 46% had received and fitted it (53% of those who received it); most of the remainder were too busy (31%) or found it too complicated (11%). At the time of the survey, only 31% of the sample were still using the RTD. The low percentage installed means it is perhaps more likely that, if fitted, the householder would understand and engage with the device (of those who had used it at all, over 79% reported they were finding it useful).

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.

The survey also gathered data on actions taken since receiving the RTD but without it being clear whether the RTD or some other trial intervention was responsible for the action, or the action would have been taken anyway. The number of customers taking each action is reported, but not the number of customers who had taken some kind of action. Switching off lights and appliances are the most common actions reported – 31% and 28% respectively of those who had received the RTD – followed by installing low energy light bulbs (9%), asking others to switch things off (8%) and using economy programmes on appliances such as washing machines (7%).

Survey of recipients of additional bill data

Out of 150 respondents, 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.

The survey also gathered data on actions taken since receiving the new bills but without it being clear whether the bill was responsible for the action or some other trial intervention or the action would have been taken anyway (in reality, the percentages for some actions are so high that a direct link to the graphs on bills defies credibility). Switching off lights and appliances are the most common actions reported – 53% and 57% respectively of those who had received the additional bill data – followed by low energy light bulbs (41%), using economy programmes on appliances such as washing machines (25%), loft insulation (17%) and asking others to switch things off (11%).

Final survey

Survey findings are broken down by "Project", which was the original SSE structure of the trial, rather than the trial groups and interventions that formed the basis of the analysis of energy consumption data. Hence, cross-comparison is problematic.

The Projects were defined as follows.

Project 1: Clip-on RTD.

Project 2: Energy advice (energy efficiency booklet) and no RTD.

- 2a: With opportunity to access online advice too.
- 2b: With graphs on quarterly bills.
- 2c: With graphs on quarterly bills and benchmark comparisons.

Project 3: Smart meter without RTD.

Project 4: Smart meter with RTD.

Project 5: Prepayment smart meter with RTD.

Project 6: Control group (having no additional information or improved metering).

Project 8: Rejecters (i.e. those who had declined to participate in the trial).

Another limitation is that tests of statistical significance are not reported, so only trends can be reported.

Reasons for joining the trial

Householders in Projects 3 and 4 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 (21% and 17%). Respondents in 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.

Reported change in attitudes and awareness

Householders in Projects 1-4 were asked "Thinking about the trial overall, from a personal point of view, how successful do you feel it has been in raising your awareness of how much energy you use and how you can reduce it?" The percentage responding in the top two categories of a five-point scale (Somewhat or Very Successful) was used as the principal outcome measure. This percentage was highest for Projects 3 (71%) and 4 (77%) and much lower for other projects (57% for Project 1 and 58%, 45% and 52% for Projects 2a, 2b and 2c respectively).

A series of questions about changes in attitudes was asked to householders in all Projects, with the general form "We would like to understand more about how your attitudes towards energy consumption have changed over the last 2 years – please tell me how much you agree or disagree with the following statements …" Table A14 shows the percentage responding in the top two categories of a five-point scale (Agree Somewhat or Completely) for questions about general attitudes. Table A15 shows the same information for three more specific questions about behaviour. In each case, the three highest percentages are shaded and figures below (or equal to) the column mean are shown in grey. The mean values shown are unweighted row or column means.

The trend is for groups with a smart meter (Projects 3, 4 and Prepay) to give the highest general ratings. For the specific practical point of looking at energy ratings, Projects 1 and 2a are highest, closely followed by credit customers with smart meters. Project 4 is in the lead for generally trying to reduce energy use, although there is relatively little variation between projects for this question. On the question of actually saving money, Projects 4 and Prepay were again in the top three, along with Project 2c, which had benchmarking information (which may have given them a different perspective on savings, not necessarily more savings). The provision of an RTD appears to have had a positive influence.

Table A14 Changes in attitudes and awareness (general attitudes)

	I am now more aware about my own energy consumption	*I am more aware of my own impact upon the environment*	*My attitudes to reducing energy consumption have changed*	*Reducing my energy consumption makes me feel good about making a positive contribution to helping the environment*	*Mean*
Project 1	77%	64%	61%	59%	65%
Project 2a	68%	66%	54%	52%	60%
Project 2b	64%	67%	54%	52%	59%
Project 2c	72%	57%	67%	52%	62%
Project 3	81%	67%	63%	60%	68%
Project 4	81%	65%	69%	61%	69%
Prepay	83%	77%	72%	64%	74%
Control	65%	59%	52%	54%	58%
Rejecters	70%	65%	55%	58%	62%
Mean	73%	65%	61%	57%	64%

Table A15 Changes in attitudes and awareness (specific points)

	When buying new appliances I now look at the energy rating information	*I now do more to actively reduce the amount of energy that I use in my home*	*The changes I have made to my energy consumption have helped me to save money on my bills*
Project 1	85%	75%	57%
Project 2a	80%	72%	52%
Project 2b	70%	76%	52%
Project 2c	72%	77%	63%
Project 3	78%	74%	57%
Project 4	78%	78%	61%
Prepay	72%	72%	59%
Control	72%	67%	47%
Rejecters	74%	69%	52%
Mean	76%	73%	56%

Specific actions taken

The survey recorded actions recalled as having been 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

Householders in Projects 3 and 4 were asked whether they recalled receiving the advice booklet and a similar percentage did in each Project (79% and 81% respectively). However, Projects 3 and 4 differed in the percentage who still had the booklet (59% vs 79%) and who still referred to it (8% vs 15%). Having an RTD appears to have brought about greater interest in the advice booklet. Also, Project 4 respondents were more likely to say the booklet was quite or very useful (78% vs 86%).

For Projects 3 and 4, a question was asked about what type of information was found most useful. The question sequence means that the question relates to any energy information received in the past 12 months, not only information from SSE. Nevertheless, it does indicate what kind of information householders find most useful. Percentages reported are the percentage finding the particular aspect of information useful. Several types of response with figures above 3% relate to general characteristics of the booklet: “Good reference point / something to keep” (30%), “Easy to understand / well written” (14%), “Unsure, thought it was useful at the time” (14%). 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; 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

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 (Projects 2b and 2c) than in Project 2a, which had advice only (60% vs 34%). The figures reported do not allow this comparison for smart meter groups. Oddly, the figure 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 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.[10] Projects did not differ greatly 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 reinforced 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 Projects 3 and 4 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 Project 3, 81% in Project 4 and 83% in Prepay. Only 2% of each group recalled anything worse than “minor disruption”. Differences in recall of the smart meter information booklet were more distinct, 41% recalling it in Project 3, 73% in Project 4 and 85% in Prepay. Overall the impact appears to have been more positive when an RTD is provided with the smart meter.

[10] In common with a number of questions reported as being about RTDs, data are presented for Project 3, where RTDs were not provided. It must be assumed that the devices referred to in this case were present independently of the trial. In most other cases, we have not reported findings for Project 3 because of this uncertainty as to what the question relates to.

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. Just over half (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

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 Prepay). Other common locations were the lounge (13% overall), hall (11%) and in a drawer/cupboard (8%). In the latter case, it is not clear whether the device was permanently in storage or was taken out and used when needed.

Frequency of looking at the RTD is shown in Table A16 for households that had an RTD fitted. The RTD provided with smart meters appears to have been viewed far more frequently than the clip-on device provided with non-smart meters. This is important in terms of limiting the capacity of the intervention to have any effect on consumption. It also creates unidentified subgroups in the analysis of consumption, thus increasing within-group variance and making it more difficult to detect any effects that do occur. This, in turn, decreases confidence in non-significant effects.

Table A16 Frequency of looking at the RTD

Frequency	*Project 1*	*Project 4*	*Prepay*
Every day	26%	37%	37%
Several times a week / Once a week	15%	26%	23%
Less often	14%	19%	16%
Never / Don't Know / Not Stated	44%	19%	21%

All respondents who had an RTD fitted 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 device was quite or very useful: 75% in Project 4, 70% in Prepay vs 50% in Project 1. 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 Project 4 and 42% of Prepay 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%)[11], "It means I have more influence on my bills" (63%), "I would miss it if it was taken away" (62%).

Of those who had an RTD fitted but were no longer using it, 42% of Project 1 and 32% of Project 4 had never looked at it whereas 9% and 24% respectively had used in for a few months, 49% and 43% for shorter periods. There were too few respondents to give percentages for Prepay.

Reasons for not using the RTD provided (Table A17) are instructive. In the smart meter groups, the dominant reasons 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").

[11] This figure differed between Projects: 43% in Project 1, 46% in Project 4 and 74% in Prepay.

Table A17 Reasons given for not using the RTD

	Project 1	*Project 4*
Usefulness		
Not telling me anything useful	15%	22%
Can't be bothered / not bothered	4%	5%
After a while it doesn't tell you anything new	4%	8%
Looked at it at start as something new	2%	3%
Done all I can to reduce energy consumption / can't do any more	11%	19%
Functionality		
Doesn't work	20%	5%
Not set up / it is unplugged	13%	3%
Don't really like gadgets / new technology	9%	3%
Hard to understand the information	15%	5%
Don't know what the buttons do / can't work it	4%	11%
Other	29%	27%

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. Differences between projects 3 and 4 were small.

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%).

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

A4.2 Non-smart meter interventions – effects on energy consumption

Findings

The analysis was based on a single multivariate analysis that combined both smart and non-smart meter interventions. The analysis controlled for pretrial consumption and determined a control group for each trial condition. It also controlled for demographic group and location. Consumption was corrected for weather and season and aggregated over the trial period.

None of the interventions had a significant effect except that provision of the RTD was associated with a small but significant effect on electricity consumption (Table A18).

Table A18 Percentage reduction in energy use

	Electricity	Gas
Standard meter + RTD	5.0	4.9
Standard meter	4.0	6.7

Shaded cell indicates significantly different from Standard meter.

Quality assessment

While the quality of analysis gives reasonable confidence, the capacity of the study to find significant effects is reduced by noise introduced by the design and method as a whole, although this would be balanced by most of the identified biases tending to increase the likelihood of a significant effect of interventions. The net effect is to increase confidence in non-significant findings but reduce confidence in the one small significant effect.

Confidence is further reduced by factors specific to implementation of the interventions, in the sense that there is substantial scope for better implementation of the same type of intervention to reduce consumption. This applies particularly to the RTD intervention and this might be demonstrated by separate analysis of those household that ever used the device, if that is possible.

On balance, there is reasonable confidence in the null findings. Nevertheless, the possibility cannot be excluded that different implementation of the same type of intervention would be successful.

A4.3 Smart meter interventions – effects on energy consumption

Findings

Demand reduction

The analysis was based on a single multivariate analysis that combined smart and non-smart meter interventions. The analysis controlled for pretrial consumption and determined a control group for each trial condition. It also controlled for demographic group and location as potential confounding variables. Consumption was corrected for weather and season and aggregated over the trial period.

The only intervention to have a significant effect on either gas or electricity consumption, aggregated over the trial period, was “meter type” (a combination of smart vs non-smart “standard” meter and whether or not an RTD was provided) – see Table A19.

Table A19 Percentage reduction in energy use

% reduction	Electricity	Gas
Smart meter	6.5	9.7
Smart meter + RTD	7.6	9.9
Smart meter (prepay)[12]	6.9	7.0
Standard meter	4.0	6.7

Shaded cells indicate significantly different from Standard meter.

[12] The smart meter was for electricity only.

Awareness of trial also had a significant effect, the electricity demand reduction (relative to baseline) being higher for the Aware and Committed groups than the Unaware group (reductions were 6.4%, 6.1% and 4.0% respectively).

Postcode and (for electricity only) Mosaic group also had significant effects but Grid supply point (an industry geographic code) did not. This emphasises the need to control for location and demographic variables. There were no significant interaction effects[13] and no tests of whether trial groups differed from each other in addition to the control. The SSE report suggests that a significant effect is not seen for the tariff incentives because of the relatively small numbers of households on the tariff incentives (i.e. a significant effect may have been seen if more households had been participating in trial groups with tariff incentives).

Load shifting

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 peak 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% (no comparable figure is given for gas).

An initial analysis for electricity consumption included smart prepayment meters, which were later excluded because of concerns that the small numbers in this group could be distorting the analysis. In this analysis the Incentive and interaction effects were also seen, in addition to effects of Meter type and Awareness (and possibly the medium of presentation of advice – on paper or on the web).

Quality assessment

Demand reduction

The quality of SSE’s analysis gives reasonable confidence but some further analysis is needed to ensure that the optimum group of confounders is included in the analysis and issues around selection of control groups and control periods are addressed. This would allow greater certainty in both significant and non-significant effects.

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

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

On balance, there is reasonable confidence in SSE’s null and significant findings but further analysis should improve confidence on both sides and might produce additional significant effects.

Load shifting

Because smart meter data are needed for this analysis, no comparison could be made with a pretrial baseline. The independent analysis of seven different periods further undermines the analysis and transparency of findings; a unified analysis might clarify the outcomes.

---

[13] E.g. the main effect of Mosaic group was significant but not the interaction with any intervention, meaning that energy demand declined more in some Mosaic groups than others, regardless of the intervention. This means that the finding may not be useful for informing the roll-out.