SAVE SDRC 2.2: SAVE Updated Customer Model

Tom Rushby (t.w.rushby@soton.ac.uk)
Ben Anderson (b.anderson@soton.ac.uk, @dataknut)
AbuBakr Bahaj (a.s.bahaj@soton.ac.uk)

Document History

Date	Author	Note
December 22, 2017	Tom Rushby	Version 2.3
December 20, 2017	Tom Rushby	Version 2.2
December 18, 2017	Tom Rushby	Version 2.1
December 14, 2017	Tom Rushby	Version 2.0 for review
December 1, 2017	Tom Rushby	Version 1.1 for review
November 25, 2017	Tom Rushby	Draft v1.0 for internal comments

Terms of Use

Form of attribution/citation:

Rushby, T. and Anderson, B. (2017) SDRC 2.2: *SAVE Updated Customer Model*, Scottish and Southern Electricity Networks Low Carbon Networks Fund SAVE project SDRC, Southampton: University of Southampton

Contents

Figures

Tables

1 Introduction

1.1 Abstract

This report provides documentation of the work carried out in the development of the Solent Achieving Value from Efficiency (SAVE) Customer Modelling Framework (SCMF). This report builds on the requirements set out in SDRC 2.1 (Anderson, 2014) and describes the first iteration of work intended to achieve those objectives using data from the SAVE projects first trial period. In doing so, it provides updates to previous reports: SDRC 5 (Sample Recruitment) and SDRC 4 (Create Commercial Energy Efficiency Measures).

In this first section, the SAVE project is introduced within the broader research context of demand response and energy efficiency as an option for mitigating low-voltage distribution network constraints. With reference to academic literature, SAVE is shown to align with best-practice, meaning that the project has been designed to produce results that are generalisable to the wider customer base. An analysis of the SAVE sample population is presented providing an evaluation of whether the sample is representative of the population within the study.

Second, a summary of the SCMF is provided, detailing the elements that are developed within the work reported, together with details of the implementation of the modelling framework (Section 2).

Third, the sources of data used in the SCMF are identified, describing both project-generated data streams and third-party sources of data (Section 3).

The fourth section reports on analysis of the impact of the interventions tested during the projects first trial period. The analysis focusses on detecting treatment effects (i.e. changes in consumption within the treatment groups) and which household characteristics are associated with higher or lower response to treatment. The results give network planners insight into how customers might be segmented and represented in a number of business as usual network modelling scenarios (Section 4).

The fifth section details the development and preliminary results of the spatial microsimulation element of the SCMF (Section 5). It describes in detail the use of the individual household consumption profiles collected by SAVE to estimate area-based aggregate profiles. In this section, the modelling is applied to small geographical areas roughly equivalent to the feeder scale and used to simulate a number of scenarios including the roll-out of the interventions across the wider customer base.

A summary of the modelling and results to date is provided, describing how the outputs align with the SCMF requirements and interface with the network model. Finally, the limitations of the existing customer model are identified, along with an overview of further work to be carried out (Section 6).

1.2 Research context

1.2.1 Background

The large-scale rollout of electric vehicles and domestic heat pumps plays a key role in the UK's decarbonisation strategy. Growth in such technologies, together with the increased penetration of small-scale, distributed renewable electricity generation, pose operational problems to distribution network operators (Thomson and Infield, 2007; Watson et al., 2016). The load growth and changes to domestic load profiles associated with these technologies is difficult to predict, however, domestic demand response can help mitigate the impact on the low-voltage (LV) distribution network (Pudjianto et al., 2013; Teng et al., 2016).

Distribution network operators (DNO's) are therefore looking to domestic demand management and response as a way to defer costly LV network reinforcement (Bradley et al., 2013). In addition to the technical challenges, electricity network planners will also need to understand the potential costs and benefits of demand response in the domestic sector and its feasibility in mitigating generation and distribution issues in future networks.

The purpose of the SAVE project is therefore to:

1. Test the effectiveness of a range of demand response interventions in reducing and/or shifting electricity demand consumption during 16:00 – 20:00 on winter weekdays in the Solent Area (county of Hampshire and the unitary authorities of Southampton, Portsmouth and the Isle of Wight);
2. Enable any statistically robust results to be generalised to the wider SSEN customer population;
3. Use these results to develop a generalisable customer model to assess a range of intervention scenarios;
4. Use this model to develop local area demand profiles under baseline and intervention conditions as input to a network model.

1.2.2 SAVE approach design

Although many aspects of demand side response (DSR), such as enabling technologies and consumer acceptability are being investigated (e.g. Balta-Ozkan et al., 2013), relatively little is known about the mechanisms involved in demand response behaviour (Carroll et al., 2014), or which customers are more (or less) likely to respond (Powells et al., 2014; Walker, 2014). Many studies do not collect enough detail on the socio-economic characteristics of households, or the physical properties of dwellings, that are required to predict the impact of DSR interventions. Robust empirical evidence is therefore scarce (Torriti, 2017), indeed, recent reviews have indicated that current best-practice methods of evaluation used in many sectors are yet to be applied to energy efficiency interventions (Delmas et al., 2013). In response to these shortcomings the SAVE project adopts a framework for best-practice as recommended by Frederiks et al. (2016). Whilst generally intended to be applicable to energy related behavioural studies, the framework is clearly relevant for demand side response trials and is adopted in the SAVE project (see Appendix A.1 for more detail). The key elements of the framework are outlined below:

Hypothesis and statistical power: Statistical power calculations enable the investigator to calculate (or estimate) the sample size that is needed to robustly detect a specific difference between two samples within given confidence levels. Calculation of the required sample size for the control and intervention group(s) requires the estimation of the likely effect size and so derives directly from the study hypotheses.[1] A number of studies have suggested that sample sizes in previous energy efficiency studies may be too low to provide adequate power and so statistically robust conclusions cannot be drawn (see Delmas et al., 2013; Frederiks et al., 2016).

Program design: In order to generate findings that can be generalised to the wider population, randomised[2] and representative[3] samples must be drawn from the target population avoiding the introduction of self-selection or other biases. [4] Once the sample has been established, participants should then be randomly allocated to intervention (treatment) and non-intervention (control) groups. Given a sufficiently large sample, any differences in energy consumption between the control and trial groups can be ascribed to an intervention effect. Random allocation is crucial to this process, as it will ensure that, given a large enough sample, there should be no meaningful differences between the control and trial groups before interventions begin.[5]

Methodology: Having established a sample, it is recommended that the achieved sample be tested prior to interventions. Any bias that is likely to be derived from the sample recruitment should be stated and tested against high quality external data. This will provide external validity of the sample and thus deliver confidence in the future generalisability of the results. In the case of household samples, doing so will require the collection of information on key features of the dwelling as well as socio-economic and demographic attributes of the occupants that are known or suspected to influence the behaviour under investigation. The second step in the

[1] Along with a hypothesised effect size, calculations require the mean and standard deviation for electricity consumption for a household sample.

[2] Non-random sample selection can violate assumptions of the independence of observations (households) and in principle negate the use of most inference methods without additional corrective action. Where no such corrections are made, there is potential to conclude that an intervention had an affect when this may not have been the case.

[3] Many samples were found to be drawn from utility companies' own regional, or national customer bases and designed to be representative of these, not the general population (see AECOM, 2011; Schofield et al., 2014).

[4] Previous trials have often set inclusion criteria for potential participants, resulting in samples that exclude or under-represent certain groups within the target population. For example: customers on a dual-rate 'economy 7' tariff (AECOM, 2011) or 'night-saver' tariff (CER, 2011), and households with microgeneration technologies (Schofield et al., 2014).

[5] Few of the major studies reviewed reported these crucial steps and where it was undertaken, it is clear that in many cases, control and intervention groups have not been randomly allocated (see e.g. AECOM, 2011; Schofield et al., 2015). Therefore, in most of the reviewed studies the extent of bias and therefore the ability to generalise to the general customer population is unknown.

methodology is to use this data to ensure that control and intervention groups are equivalent on these key dimensions, providing internal validity of the trial design before interventions begin.

1.3 Establishing the SAVE sample

The SAVE project has implemented the best-practice framework described above using the following elements.

1.3.1 Overall design approach

Based on the best practice framework described above, a randomised control trial (RCT) design was adopted encompassing the recruitment of a large and representative random household sample of four equal groups from the Solent Area. One group was to be a control and the other three to undergo a series of winter trial interventions. Intervention impacts were to be objectively measured using (i) whole house electricity consumption (kWh) and power (kW) monitors installed at each dwelling, (ii) time-use diaries and (iii) household surveys.

1.3.2 Hypothesis and sample size

Prior to the commencement of fieldwork for the SAVE project, the size of the required sample was established through statistical power analysis. The analysis used pre-trial data from the Irish smart meter trials (CER, 2011). Details of the calculations are presented in the Appendix (A.1.3). The results indicated that the project required 1,000-1,200 households per sample group and thus a total of 4,000-4,800 households to ensure that an intervention effect size of 6% or larger would be robustly detectable. This also ensured that an expected attrition level of 5% per year would still leave a large enough sample size at the end of the study.

1.3.3 Sample recruitment

A stratified random sampling approach based on addresses was used to establish a representative household sample of the required size. The overall population was to be domestic households in the geographical areas mentioned above. No household or respondents were excluded from the random sampling process with the exception of:

- Known student or multi-occupancy (shared) housing which were excluded on the basis of transience (high turnover) and associated difficulty in obtaining appropriate informed consent over time;
- Blocks of flats with primary (whole building) and secondary (specific dwelling) doors due to difficulties of access (flats).

Households at the randomly selected addresses were then contacted by an appointed fieldwork agency with a view to recruitment to the study. No additional publicity or appeal for volunteers was conducted to ensure that the random sampling approach was not contaminated by self-selected volunteers (see Appendix A.1.1 for more detail).

For the households that agreed to participate, Navetas 'Loop' electricity monitoring equipment[6] was installed and an initial recruitment survey was conducted. As Table 1 shows, this covered a wide range of attributes and characteristics to enable both the future analysis of intervention efficacy for different social groups but also to support the assessment of sample and trial group allocation bias.

Table 1 SAVE recruitment survey summary

Attributes	Details	Justification
Dwelling attributes	Tenure, electricity supply, bill payment, presence of PV, EV, smart meters/in-home displays, source of heating, typical times occupied, when built, insulation, water heating, appliances owned	Enable assessment of sample bias; Enable analysis of effects of differing dwelling forms
Household attributes	Number in household, relationship to respondent, age, sex, working status and, where relevant, mode of transport to work/place of education, length of occupancy, accommodation type, number of rooms, availability of cars, household income	Enable assessment of sample bias; Enable analysis of effects of differing household types and social factors; enable spatial microsimulation using matched Census variables
Household behaviours	Frequency of use of appliances between the hours of 4pm and 8pm, and ease of avoiding use at these times	Enable assessment of sample bias; Enable analysis of effects of differing household behaviours and lifestyles;
Household response person attributes	Attitudes towards the environment and extent of sustainable behaviours, preferred methods of contact, highest qualification, ethnicity, religion, disability status	Enable assessment of sample bias; Enable sample management; Enable analysis of effects of differing social factors.

1.4 SAVE outcomes and response rates

This section provides an evaluation of the composition of the SAVE household sample using the household response and recruitment survey data and is an update to that provided in SDRC 5.

Figure 1 shows the location of recruited households across the study region and demonstrates the appropriate rural and urban spread of the sample.

[6] Navetas Loop Energy Saver is an off-the-shelf, internet connected home energy monitoring system (see https://www.loopenergysaver.com).

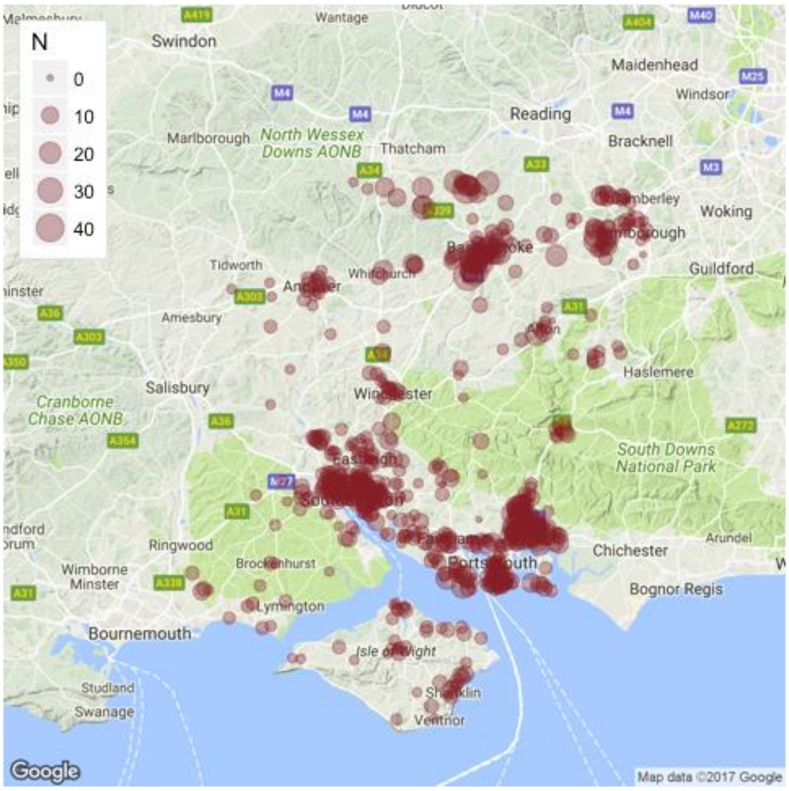

Figure 1: Location of recruited households

Table 2 summarises the outcomes of the 42,470 addresses issued for recruitment purposes. Of these, contact was achieved with 9,041 (22%) and of these 5,091 (56%) agreed to installation. Of the refusals, the greatest proportion by far (86%) stated they were not interested with only just over 5% not prepared to help named project partners (SSEN, University of Southampton) and just over 1.2% citing privacy concerns.

Table 2 Recruitment contact outcome

	Number of addresses	% of addresses
Installation complete	5,091	12%
Refusal	2,530	6%
Unable to fit clamp	8	0.02%
No contact	1306	3%
Unused addresses	33,535	79%
TOTAL	**42,470**	**100%**

Further analysis of non-response and refusal using a logistic regression model including area level predictors in the absence of any information on non-responding or refusing households can be found in Appendix A.1.4.

1.5 Assessing sample bias

Using the collected survey data, the recruited SAVE sample was compared to data from the high quality Understanding Society survey (Institute for Social and Economic Research, 2015)[7] to assess the extent to which the sample is representative of the population of the areas from which it was drawn. The Understanding Society survey contains equivalent demographic information to that collected for the SAVE sample and also carried 'environmental attitudes' items that were also replicated to enable the assessment of the extent to which households agreeing to participate in SAVE are more (or less) environmentally minded than the general population.

1.5.1 Demographic profile

The analysis in this section compares the unweighted SAVE and weighted[8] Understanding Society (USOC) samples. It is assumed that the weighted USOC data represents the 'true' distributions of the population and comparison with the unweighted SAVE sample enables assessment of inferred bias in the SAVE sample. The figures include 95% confidence intervals for the estimates to indicate uncertainty. In general, the figures show a reasonable match between the SAVE and Understanding Society samples although the results point to a number of specific but expected differences (See Appendix A.1.5, Table 31). As an example, Figure 2 shows that the deliberate exclusion of a number of types of household (see Appendix A.1.1) has caused SAVE to under-represent households with older (75+) household response persons (HRPs) but to slightly over represent HRPs aged 45-74. Similarly SAVE also slightly over-represents home owners but under-represents private renters (Figure 3) and under-represents single person households (Figure 4).

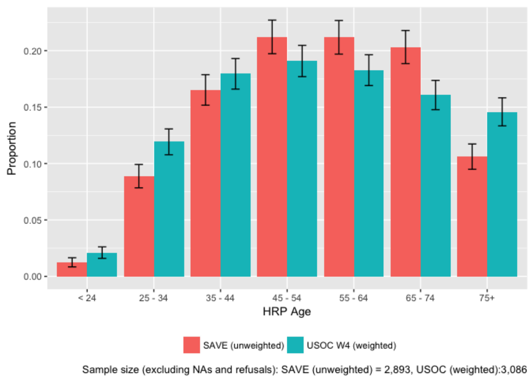

Figure 2 SAVE vs Understanding Society Wave 4 - HRP age

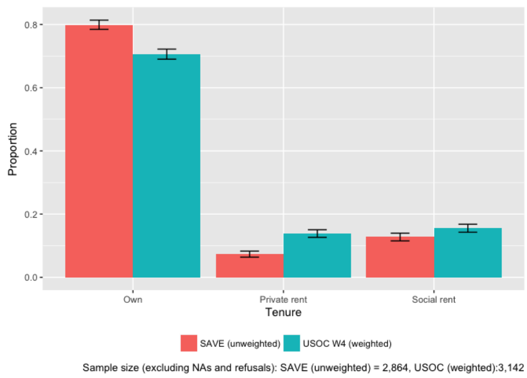

Figure 3 SAVE vs Understanding Society Wave 4 – Tenure

---

[7] To date, only a sub-sample of SAVE households have completed the full recruitment survey. Due to the sampling methodology, the sample distributions are not expected to change substantially as additional surveys are completed.

[8] Weighted for survey and design-based non-response.

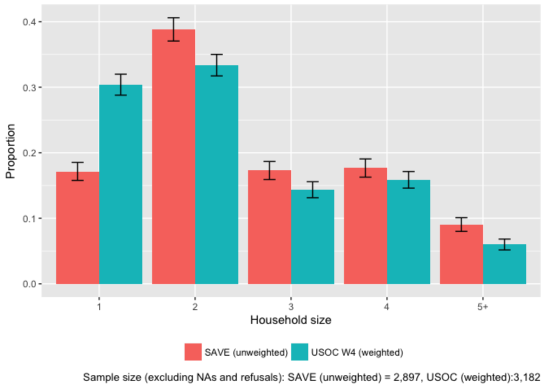

Figure 4 SAVE vs Understanding Society Wave 4 – Household size

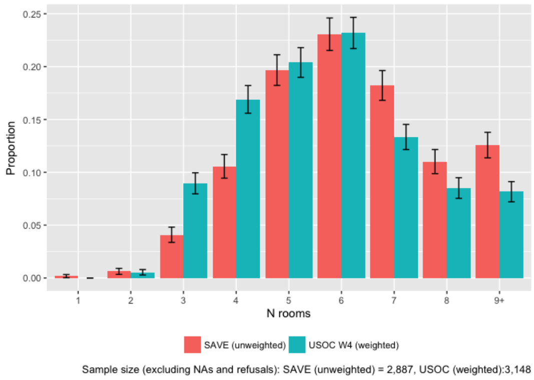

Figure 5 SAVE vs Understanding Society Wave 4 – number of rooms

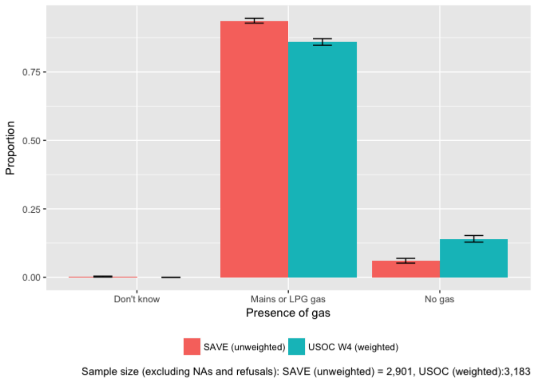

Figure 6 SAVE vs Understanding Society Wave 4 – Mains gas

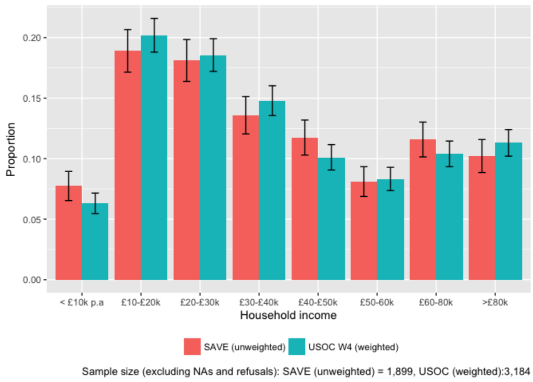

Figure 7 SAVE vs Understanding Society Wave 4 – Gross household income

There are fewer substantive differences for dwelling size (Figure 5) but SAVE appears to under-represent households without gas (Figure 6). Again, this may be caused by the tendency to filter out flats during the recruitment process but it may also be caused by differences in the geographical coverage of the two samples as USOC covers the whole of the South-East region of England. SAVE shows a reasonable fit to the USOC income distribution (Figure 7) with slight over-representation at the lower ends of the scale and under-representation at the higher ends of the scale. As before, this may be caused by the wider geographical coverage of the USOC sample but tests suggest that the differences are not statistically significant.

Finally, Figure 8 shows the mean 'environmental behaviour' scores across a battery of six questions for both SAVE and Understanding Society Wave 4 respondents. As can be seen, the median for SAVE is slightly higher as might be expected given that the randomly selected households have been asked to take part in an energy-related study.

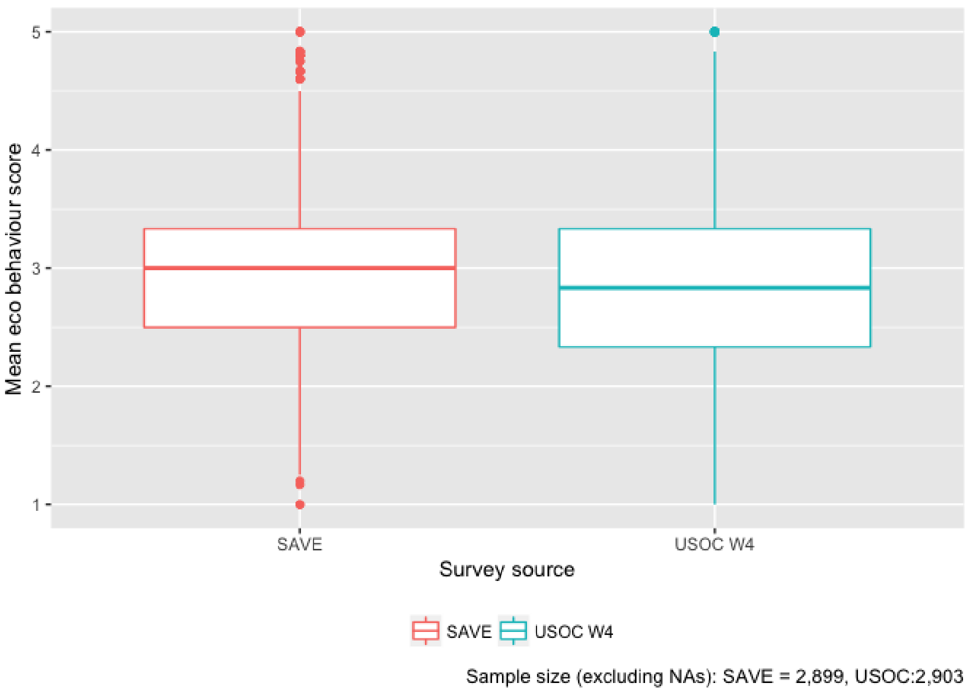

Figure 8 SAVE vs Understanding Society Wave 4 - Mean HRP 'eco score' (unweighted data)

However, a more informative regression model using the pooled samples (see Appendix A.1.4, Table 30) confirms that that Understanding Society respondents have slightly lower scores but other co-variates have a much stronger effect. Thus, the apparent difference suggested by Figure 8 may be caused by the small differences in sample composition identified above.

1.6 Summary

Overall, the analysis reported in this section suggests that the SAVE sample is representative of the wider population from which it is drawn with some minor caveats. As a result of the recruitment process, the sample under-represents the 75+ age group, slightly over-represents home owners, larger homes and slightly under-represents those without mains or liquefied petroleum gas (LPG). The sample also has slightly higher pro-environmental attitudes than the general population which is to be expected given the nature of the study.

Furthermore, similar analysis of the SAVE trial groups (Appendix A.1.5, Table 33) reveals no statistically significant differences between groups with the exception of the presence of mains gas. Comparison along additional variables, including primary heat source, environmental attitudes and overall electricity consumption shows that there is no apparent systematic bias within groups. This means that subsequent analysis is able to assume that any difference between the consumption levels in different trial groups is due to the interventions and not to any systematic differences between the groups.

2 SAVE Customer Model Framework

The requirements of the SAVE Customer Model Framework are set out in *SDRC 2.1 SAVE Customer Model Framework Specification* (Anderson, 2014). The key requirements of the framework are reproduced here in summary and commentary is provided where these have changed as the project has progressed.

2.1 Key requirements

The requirements for the SAVE customer model are as follows:

1. The ability to produce 'baseline' half-hourly electricity consumption profiles at the individual household level for any day of the year (or aggregation of days) as input to the Network Model;
2. To produce similar profiles for trial intervention groups as input to the Network Model, taking account of intervention and community trial effects *where feasible*;
3. To produce similar profiles for designated Census areas in the Solent region under a range of demand response scenarios including those trialled by the SAVE project;
4. The estimation of electricity consumption increase/decrease at specific times of day that can be attributed to the SAVE intervention trials for overall effect reporting;
5. The analysis of the household economic, demographic and behavioural factors that mediate these changes to provide insights relevant to future DNO interventions;
6. The ability to estimate changes in temporal (half hourly) demand that might ensue from other (non-trialled) behavioural changes;

The framework proposed to meet these requirements included the following features:

- *Microsimulation* – households modelled as units (micro) with specific observed consumption values rather than modelled as groups or types of households with 'average' consumption values;
- *Time-of-day* – household level consumption represented at the half hour level over 48/7/365 to enable analysis of scenarios which affect consumption at specific times of day and/or season;
- *Spatial* – aggregate demand profiles for local areas produced by combining micro-level data with UK Census data to produce a *spatial microsimulation* model.

The remainder of the report describes the development of the SAVE customer model to date to meet these requirements.

2.2 Technical implementation

RStudio[9] is used by the University of Southampton (UoS) to implement the SAVE customer modelling framework. The use of this software allows the input and linking of the various project generated data streams, along with other data from third-parties used in the modelling and analysis. The use of RStudio also allows output from the SAVE customer model in a format compatible with Microsoft Excel allowing interface with the Network Model. The data used in the modelling framework is described in the following section.

To ensure that the results are transparent and reproducible, as far as possible all analysis is produced from original data without any pre-processing of the data files. Any instances where processing of the data has been performed externally to the RStudio package are noted within the relevant sections of this report.

3 Data and methods

3.1 SAVE project generated data

As set out in *SDRC 2.1 SAVE Customer Model Framework Specification* (Anderson, 2014), the proposed data inputs into the Customer Model were as follows:

1. Household socio-economic and demographic data from recruitment and repeated waves of household surveys;
2. Household response person time-use activities as recorded at 10 or 15 minute intervals in a 1-day time-use diary to be implemented as part of the repeated waves of fieldwork;
3. Dwelling level electricity consumption data (in kWh) provided at the half hour level by default and at the 10 or 15-minute level during the period of time time-use diary;
4. For a sub-sample of the households, selected appliance level electricity consumption data provided via smart plugs.

The delivery of these data inputs has been achieved during the project mobilisation and fieldwork carried out to date, with the exception of the sub-sample of appliance level electricity consumption data. This is due to the replacement of the original electricity consumption monitoring instruments with another system with different functionality and specification (Change Request 1, Scottish and Southern Energy Power Distribution, 2015). The specification of the achieved data sources feeding into SAVE customer model has altered as a result and is now as follows.

[9] RStudio is an open-source integrated development environment based upon the R programming language for statistical analysis.

1. Household socio-economic and demographic data from initial recruitment survey and repeated 'update' surveys implemented during the repeated waves of fieldwork;
2. Household response person time-use activities recorded at 15 minute intervals during the period of the time-use diary;
3. Dwelling level electricity consumption data at 15 minute intervals (in watt-hours, Wh).

3.2 Household data

In this report, the latest update file from the fieldwork contractor (BMG Research) is used. This data file contains the socio-economic and demographic data for the households participating in the fieldwork. Update surveys were conducted in late 2016/early 2017 with those households who had been originally recruited in late 2015 to ensure that basic household attributes such as number of occupants were updated.

3.3 Time-use diary data

Time-use diary data was collected during SAVE's first trial period with two specific objectives:

- analysis of differences in reported activities between trial groups, especially during the event day intervention,
- match reported activities to power demand (or consumption) data to understand the relationship between practices and actual electricity demand.

For the purposes of this report, analysis is directed to differences between trial groups. To compare the groups, activity metrics are defined and these include:

- the number of electricity-using activities reported in a given time period - as a coarse-grained indicator,
- the number of specific electricity-using activities (e.g. tumble drying etc) reported in a given time period - as a fine-grained indicator,
- the number of activities reported at home/out of the home in a given time period - as an indicator of 'not being home' as a response to the trial intervention messaging and incentives.

Analysis of the time-use diary data for Trial Period 1 is presented as part of the evaluation of the impact of the event day intervention (Section 4.5). More information about the data, along with limitations and data checking is provided in Appendix A.2.1.

3.4 Monitored Data

The analysis contained in this report makes use of two datasets collected by the projects equipment supplier, Navetas. First, data collected via the 'portal': the system through which customers access their data and receive email (or text message) notifications. The second, consumption data collected via the internet-connected 'Loop' electricity monitoring kit (hitherto referred to as 'clamps').

3.4.1 Portal data (event day email opening)

Data from the Navetas portal is used to add information on whether households opened the event day notification email (Table 3), and allows any associated effect to be modelled in the analysis (see section 4.5). The data contains only a crude "yes" or "no" response for the email notifications (or not applicable for trial groups who did not receive event day notification emails).

Table 3 Portal data: number of households opening event day notification email

Trial group	No	Not applicable	Yes
BMG Group 1: Control	0	10610	0
BMG Group 2: LED trials	10146	0	476
BMG Group 3: Data informed engagement and price signals	10243	0	376
BMG Group 4: Data informed engagement	0	10619	0

3.4.2 'Loop' electricity consumption data

The 'Loop' data used in the analysis consists of watt-hour (Wh) readings observed at 15-minute intervals for each participating household. This data provides the measure of electricity consumed by 3934 individual households during the trial period. More detail on the Loop data used in the analysis can be found in Appendix A.2.2.

3.5 Third party data used in analysis and modelling

Third-party data is also used in the analysis reported here. The data and sources are summarised below.

3.5.1 Weather data

Met Office weather data is used in the 'baseline' descriptive analysis. The hourly data used was collected at Middle Wallop between the dates 2016-09-30 and 2017-03-31, and was downloaded from the Met Office Weather Observations Website (WOW). The hourly weather data is pre-processed prior to use to create daily and weekly means, for details see Anderson and Rushby (2017).

3.5.2 Understanding society data

The analysis of the recruited SAVE households uses data from the Understanding Society 'south east England' regional sample (Institute for Social and Economic Research, 2015) which was collected in 2012-2014.

3.5.3 Census 2011 output area data

Census data is used in the spatial microsimulation element of the SAVE customer model. This is in the form of output area (OA) level tables containing aggregate household counts for a range of household and household response person attributes. The data has been downloaded from Nomisweb (https://www.nomisweb.co.uk/census/2011/data_finder).

3.6 Methods

Given the RCT design of the SAVE trials, intervention effects have been analysed by comparing the difference between control and intervention groups. Any difference in consumption between the control and intervention groups is assumed to be a result of the intervention alone (Frederiks et al., 2016). It is assumed that all households in the study experienced the same environmental conditions during the trial weeks and therefore there is no need to correct for any differences in environmental conditions.

Note that due to the randomised control trial design of the study, it is not necessary to control for other potential confounding characteristics of the households in each trial group. However, the analysis in section 4 below includes a small selection of household attributes to understand if certain characteristics are associated with a stronger reduction effect.

4 Trial period 1 (TP1) analysis

The following section provides analysis of the impact of the first period of trial interventions. This analysis will feed into the customer model by informing the key dimensions (characteristics) by which customers are to be segmented in order to capture differences in how they respond to interventions to shift or reduce peak demand such as those trialled during the SAVE project. It is therefore oriented toward providing an understanding of if, and how, treatment effects vary across intervention types (i.e. mode of messaging, event type, targeting etc.) and across household types. The analysis of trial impacts will therefore inform several microsimulation scenarios that allow the modelling of the rollout of the interventions across areas within the SAVE study region (see Section 5). Finally, the analysis will show why, using data from TP1, there is currently little visible difference between the spatial microsimulation models for each trial group.

4.1 Trial groups and schedule

Following the best practice approach outlined above, the SAVE household sample was randomly allocated to one of four trial groups at recruitment. These are:

- Group 1: Control
- Group 2: LED trials
- Group 3: Data informed engagement and price signals
- Group 4: Data informed engagement

Group 2 (LED) received a voucher towards the cost of purchasing a small number of LED light bulbs. This was sent during the week commencing 6th February (trial week 6).

Groups 3 and 4 received identical messaging to:

- encourage general evening peak reduction
- encourage reduction on a specific 'event day'

Finally, Group 3 also received a financial incentive:

- £10 if they reduce their evening peak (16:00 - 20:00) electricity consumption by 7%

Table 5 shows the timing and nature of the planned interventions. The interventions consist of leaflets and postcards as well as emails sent to groups 3 and 4 on a weekly basis. In addition, all respondents are able to access and review their usage via the Navetas portal.

4.2 Trial 1 Overall Schedule

The time-use surveys were originally planned to coincide with these intervention weeks but these were paused in late February to allow an accrual of sufficient contacts for the final event day time-use survey.

Table 4 also shows the content for each of the weekly contacts. Reference '5' in the final row in the table refers to the 'event day' on Wednesday 15th March 2017.

Table 4 Trial communication schedule

		Trial week	Booklet	Leaflet	Postcard	Email/Loop notification	Text	Loop portal
Jan	2	1						
	9	2			1 to 6	1 to 6		
	16	3			4E	4E		
	23	4			4A	4A		
	30	5			4B	4B		
Feb	6	6			4C	4C		
	13	7			4D	4D		
	20	8	Half term					
	27	9		1 to 4				
Mar	6	10				5	5	

As described in detail in SDRC 4 (Create Commercial Energy Efficiency Measures), the eventual as-implemented trial groups were as follows:

- Group 1: Control group
- Group 2: LED trial through postal communications and consumer engagement through online communications including event day messages (but without price signalling)[10]

[10] The impact of the LED engagement was minimal with only five participants purchasing discounted LED bulbs out of 1,137 households receiving promotion mail. This translates to 0.4% take up of the offer. This group is therefore treated in the analysis as online engagement only.

- Group 3: Consumer engagement through both online and postal communications and price signalling including event day messages (with price signalling)
- Group 4: Consumer engagement through postal communications only

Therefore the analysis seeks to establish whether there is a significant difference in evening peak period electricity consumption between the trial group (1) and each of the intervention groups as appropriate.

4.3 Descriptive analysis

Summary statistics for the 15-minute consumption data during the TP1 period (January to March 2017) are provided in Appendix A.3.1, which also details the removal of zero Wh observations prior to analysis.

4.3.1 Weekly consumption trends

The following figures show the overall mean weekly Wh consumption by trial group and trial week. Figure 9, shows the overall mean of each household's mean weekly 15 minute Wh consumption disaggregated by trial group.

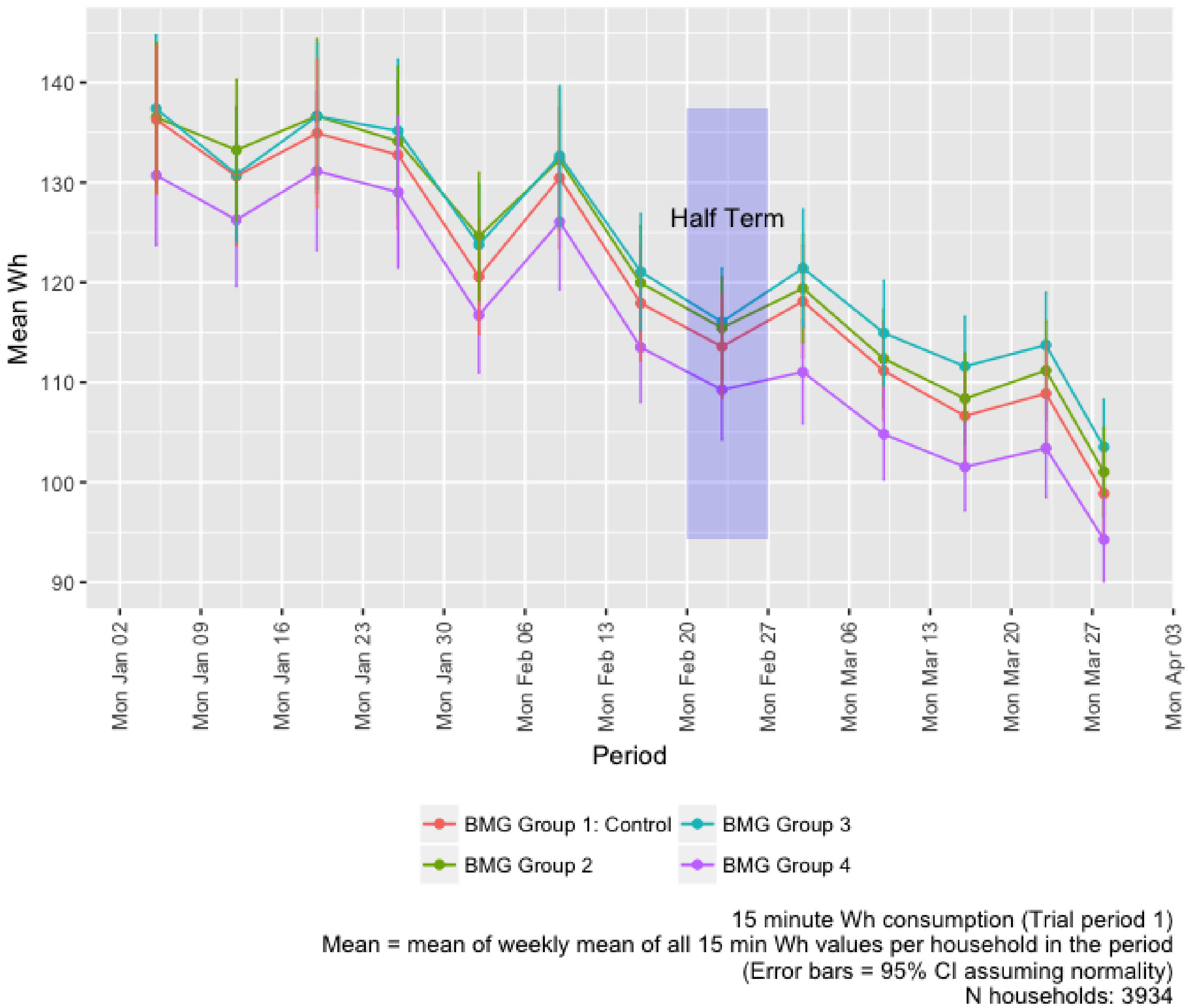

Figure 9 Mean weekly Wh by trial group

There is an overall trend towards a reduction in mean electricity consumption over the period from January to March as day length and temperatures increase. Consumption is also lower in half-term week perhaps reflecting lower occupancy due to holidays or less intense consumption for households with school age children. Whilst trial group 4 appears to have consistently lower mean consumption, the 95% confidence interval (error bars) suggest that these patterns hold true across all trial groups as they substantially overlap.

The week to week anomalies are clearer when plotted as the change in mean consumption from week to week with the change in mean temperature (Figure 10).

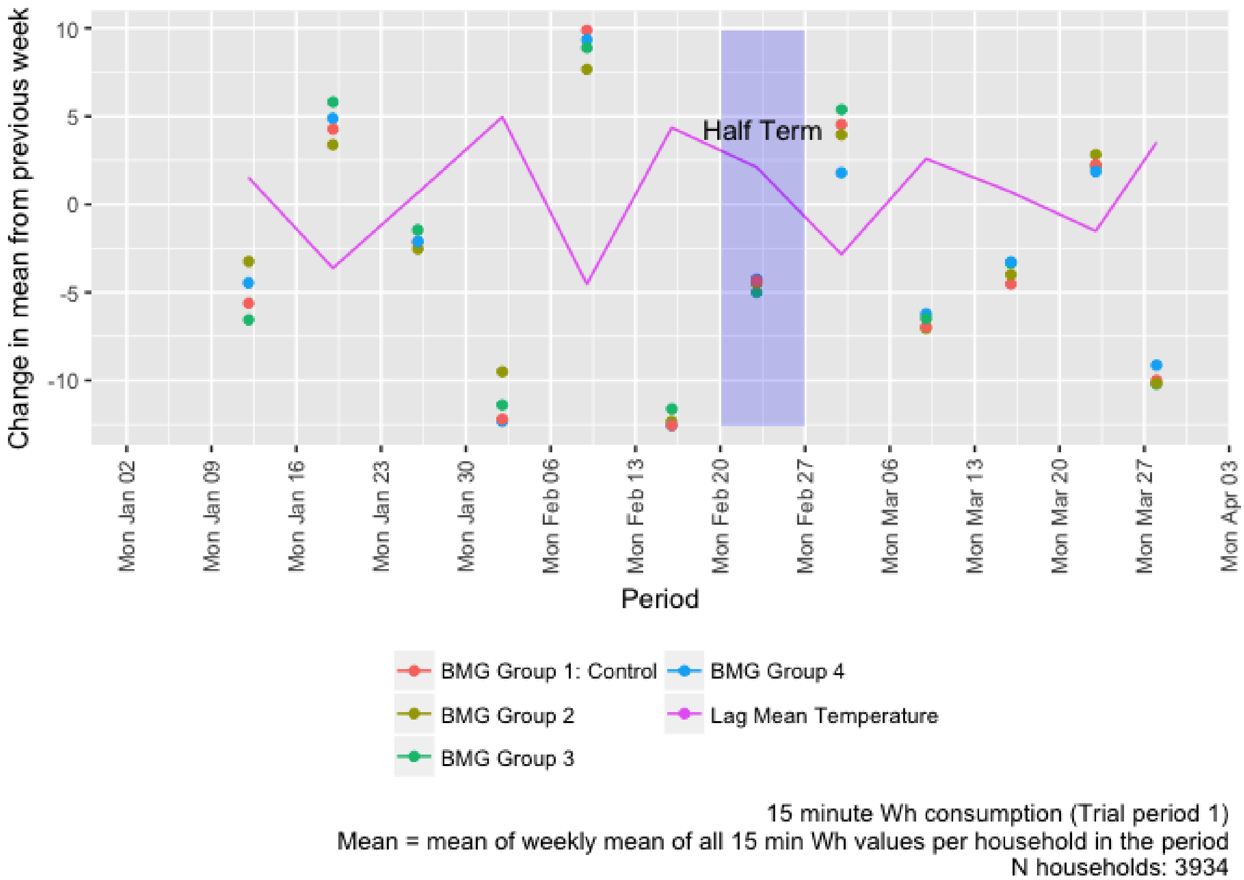

Figure 10 Trend in weekly change in mean consumption by trial group and mean temperature change

Figure 10 suggests that the weeks starting January 30th and February 13th were substantially warmer than the previous week(s) whilst the week starting January 16th, February 6th (and 27th) may have been substantially colder. This is further examined by analysing the trend in consumption for households who used different kinds of heat sources (Table 5). NA indicates households for whom there is currently no survey data available.[11]

Table 5 Number of households by mean heat source in TP1 period clamp data

Main Heat Source	N households
NA	982
Electricity (storage heaters)	124
Gas boiler	2647
Oil/wood/solid fuels/biomass/other	163
Other electricity (e.g. heat pump)	18

Figure 11 shows scatter graphs of temperature change against electricity consumption change by main heat source. Households heated by 'Other electricity' (n = 18) respond to the temperature change quite substantially compared to all other heat sources, with 'storage heaters' the next most responsive. Note that the response of 'Oil' etc. is also relatively strong, suggesting that electricity is being used as an additional heat source in these homes.

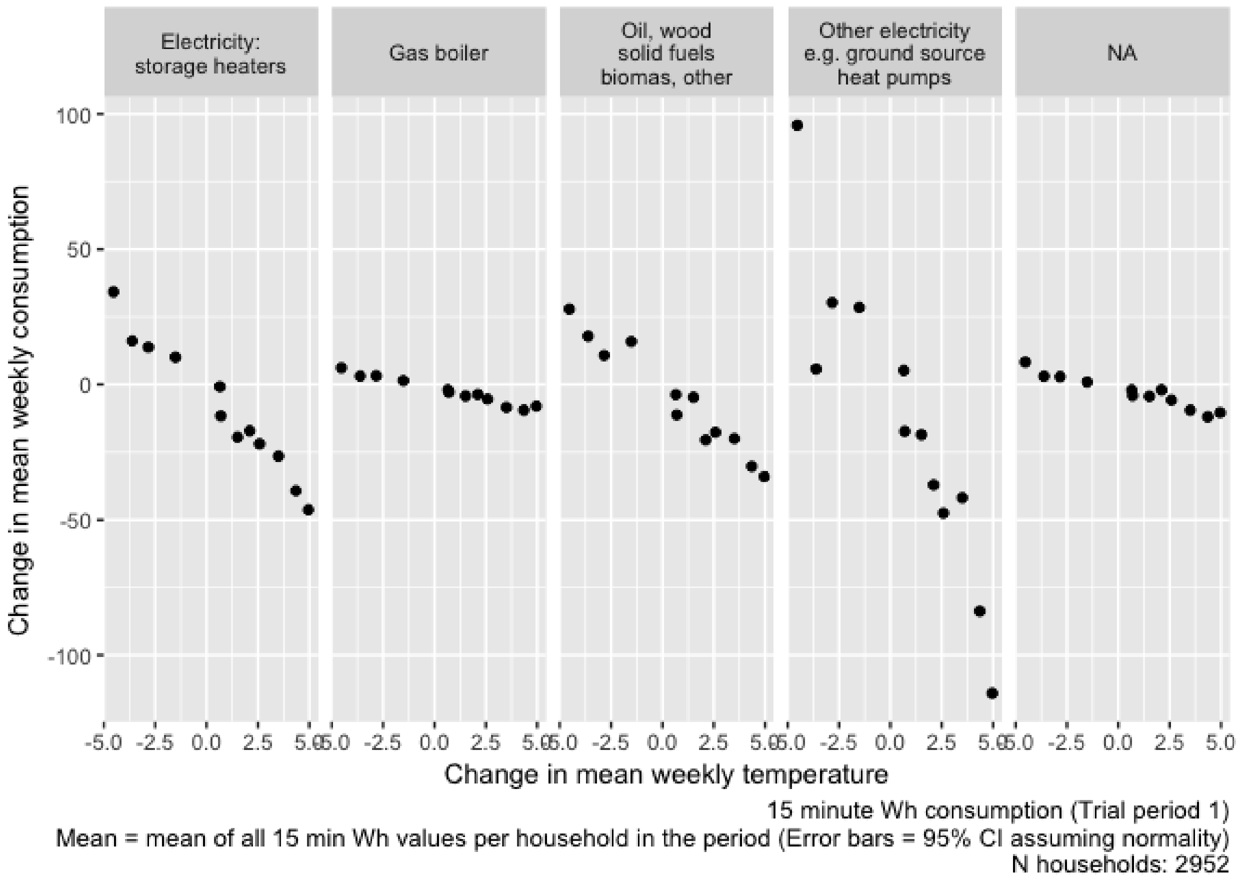

Figure 11 Scatter graph of weekly change in mean consumption by heat source and mean temperature change.

[11] At the time of writing, not all households had responded to a full (and thus an update) survey. As a result, the analysis of TP1 consumption data alone can include all 3,934 households, however, any analysis requiring survey data will use a maximum of 3,010 households.

4.3.2 Daily consumption trends

The analysis in the following section is based upon consumption data for the control group only, which received no interventions. Due to the stratified random sampling approach used to recruit households and the subsequent random allocation to trial groups, this group can be assumed to represent 'normal' customer behaviour for the study region.

4.3.2.1 Daily consumption profiles

Figure 12 shows the mean consumption (Wh) per 15-minute period for all days in January, February and March 2017 (separately) for the control group. It should be noted that the consumption in March would have been affected by the clock change on Sunday 26th.

The figure explains to some extent why the overall daily mean Wh consumed on Sundays (and Saturdays) tends to be higher than that for weekdays (see Appendix A.3.2, Figure 42). Mean consumption is higher during the days at the weekends, and especially so on Sundays with a slightly earlier peak and with no diminishing of the evening peak demand level compared to weekdays. The only time of day when weekend consumption is lower than on weekdays is between 05:00 and 09:00 reflecting the lack of habits and routines associated with the working week.

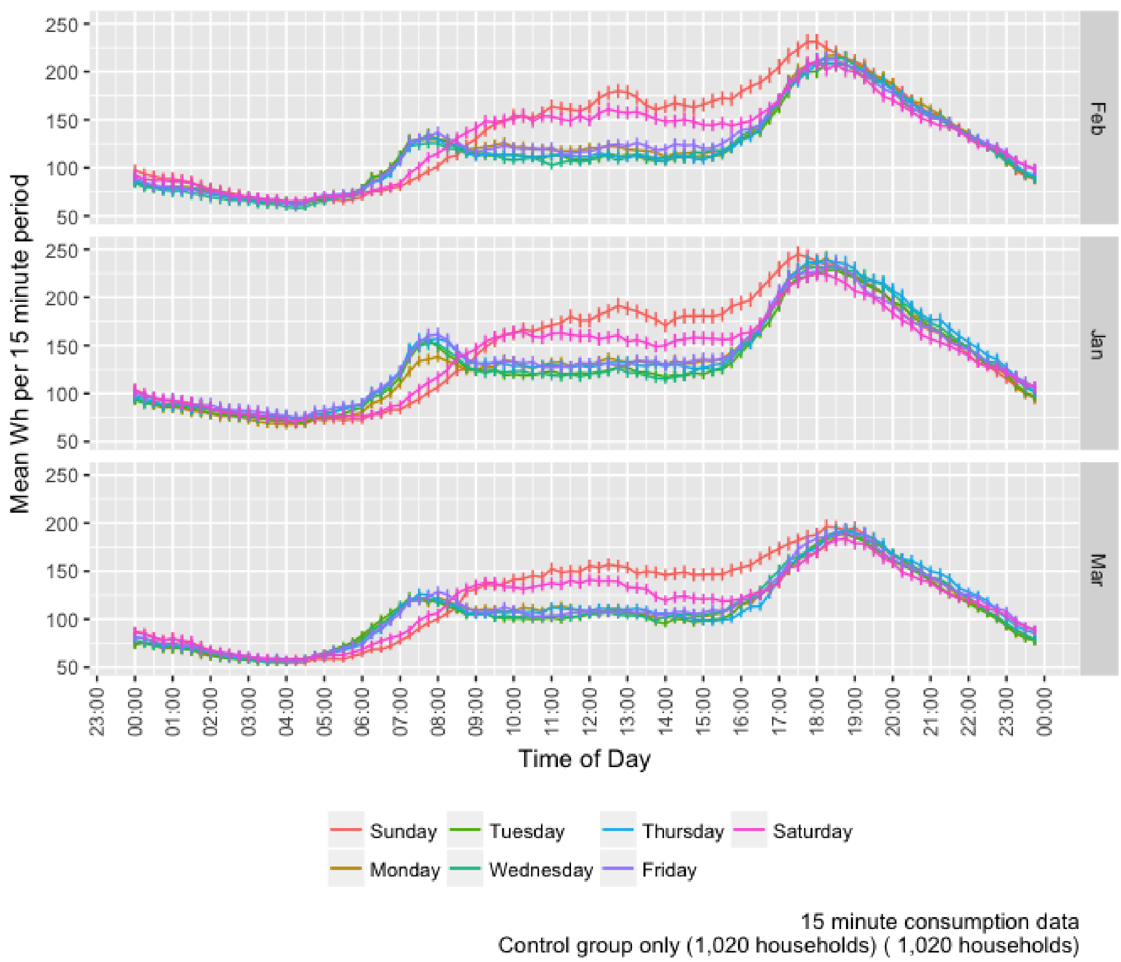

Figure 12 Mean 15 minute Wh observations by day of the week and month

On this basis, differences in temporal consumption profiles are expected depending on whether the household is in work. This is confirmed by examining mean household consumption by work status (see Appendix A.3.2, Figure 44).

4.3.2.2 Daily profiles: heat source

Alongside an interest in how different habits of different kinds of people influence daily demand profiles, it seems likely that dwelling characteristics such as main heat source can also play a role. In this section, an analysis of overall electricity consumption by households with different main sources of heat is presented. To keep the number of households in each sub-group as high as possible (Table 6), all groups are used and the analysis is restricted to January 2017 when trial period 1 had only just started. During this period intervention effects might be assumed to be minimal.

Table 6 Main heat source (NA = no survey data to date)

Heat source	N households
NA	982
Gas boiler	2647
Other electricity (e.g. heat pump)	18
Oil/wood/solid fuels/biomass/other	163
Electricity (storage heaters)	124

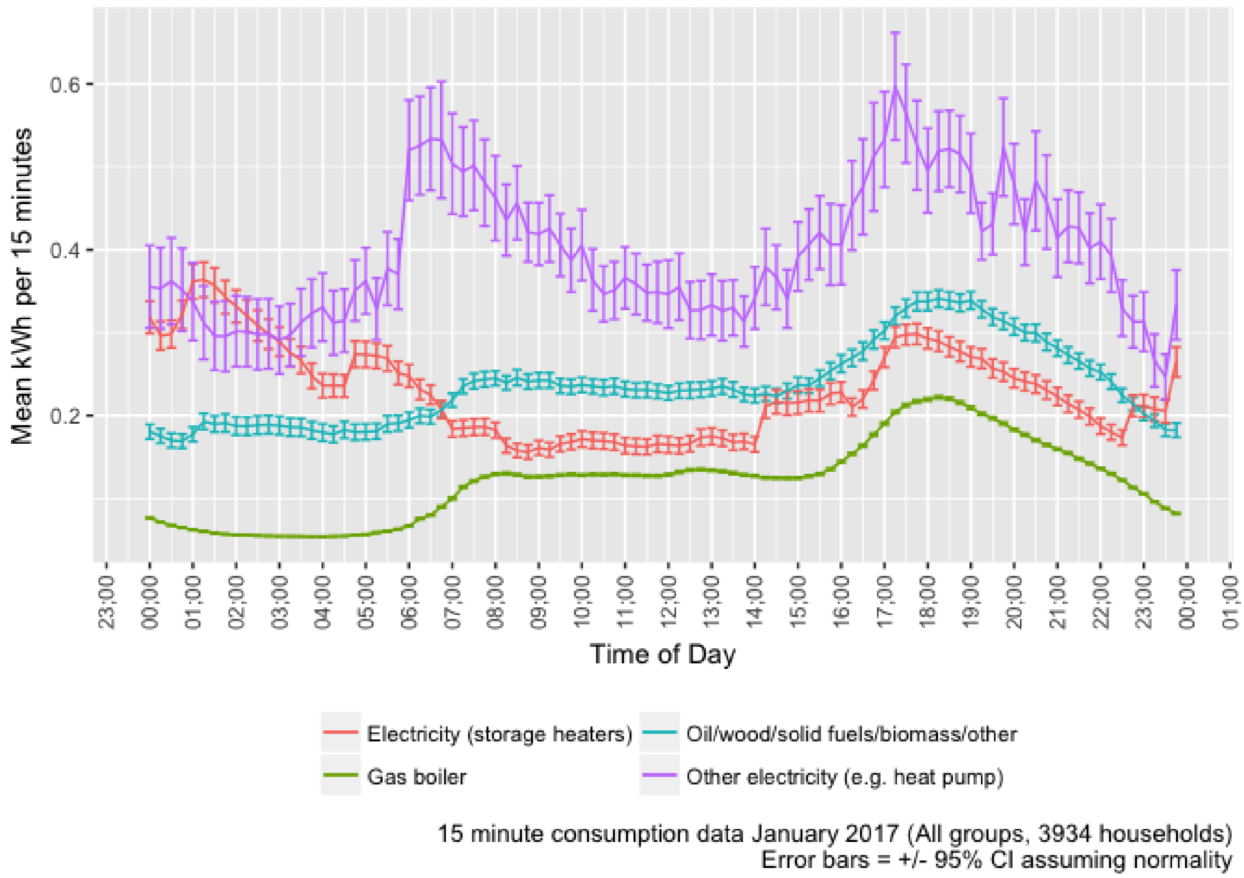

Figure 13 Mean kWh consumption per 15 minutes by heat source

There are some notable differences between consumption profiles for households with different main heat sources although the width of the 95%

CIs indicates both the relative rarity of electric heating in this sample and also the variation between households that use it (Figure 13). Overall, those with gas boilers are least likely to be using electricity as a source of either heat or hot water whilst both electricity heat source groups may be using electricity for hot water and also space heating. Those with oil/wood and other fuels may use electricity to heat domestic hot water and possibly also as a back-up heat source.

4.4 Impact of weekly 4-8 messaging

This section presents analysis of the week-by-week 'impact' of the different messaging forms on:

- evening peak consumption – are there differences between trial groups on a week by week basis as different messages were received (in different forms),
- reported evening peak activities (time use data) – are there differences between trial groups on a week by week basis.

4.4.1 Introduction

The following analysis focuses on the mixed-mode messaging received by the three intervention groups over the entirety of Trial Period 1 from January to March 2017. These messages were intended to encourage trial respondents to reduce and/or shift their electricity consumption during the key evening peak periods of 16:00 - 20:00 or '4-8'.

As noted above, this intervention was comprised of a series messages with a weekly focus:

- w/c 2nd January: background information
- w/c 9th January: 'understanding'
- w/c 16th January: charging mobiles, tablets and laptops
- w/c 23rd January: laundry
- w/c 30th January: dishwasher
- w/c 20th February: tumble drier
- w/c 6th March: watch TV

The different trial groups received either online-only messaging, post-only or both online and postal messaging. To aid interpretation the intervention groups have been recoded as shown in Table 7.

Table 7 Clarification of 4-8 messaging intervention types by trial group

BMG Group	TP1 4-8 messaging intervention
BMG Group 1	Control (no intervention)
BMG Group 2	Online only messaging
BMG Group 3	Online and postal messaging
BMG Group 4	postal only messaging

4.4.2 Descriptive analysis

Figure 14 shows the overall mean weekly Wh consumption for the 16:00 - 20:00 periods by intervention group and trial week. The figure includes coloured boxes displayed at the bottom showing the focus of the weekly 4-8 messaging in the relevant time periods.[12]

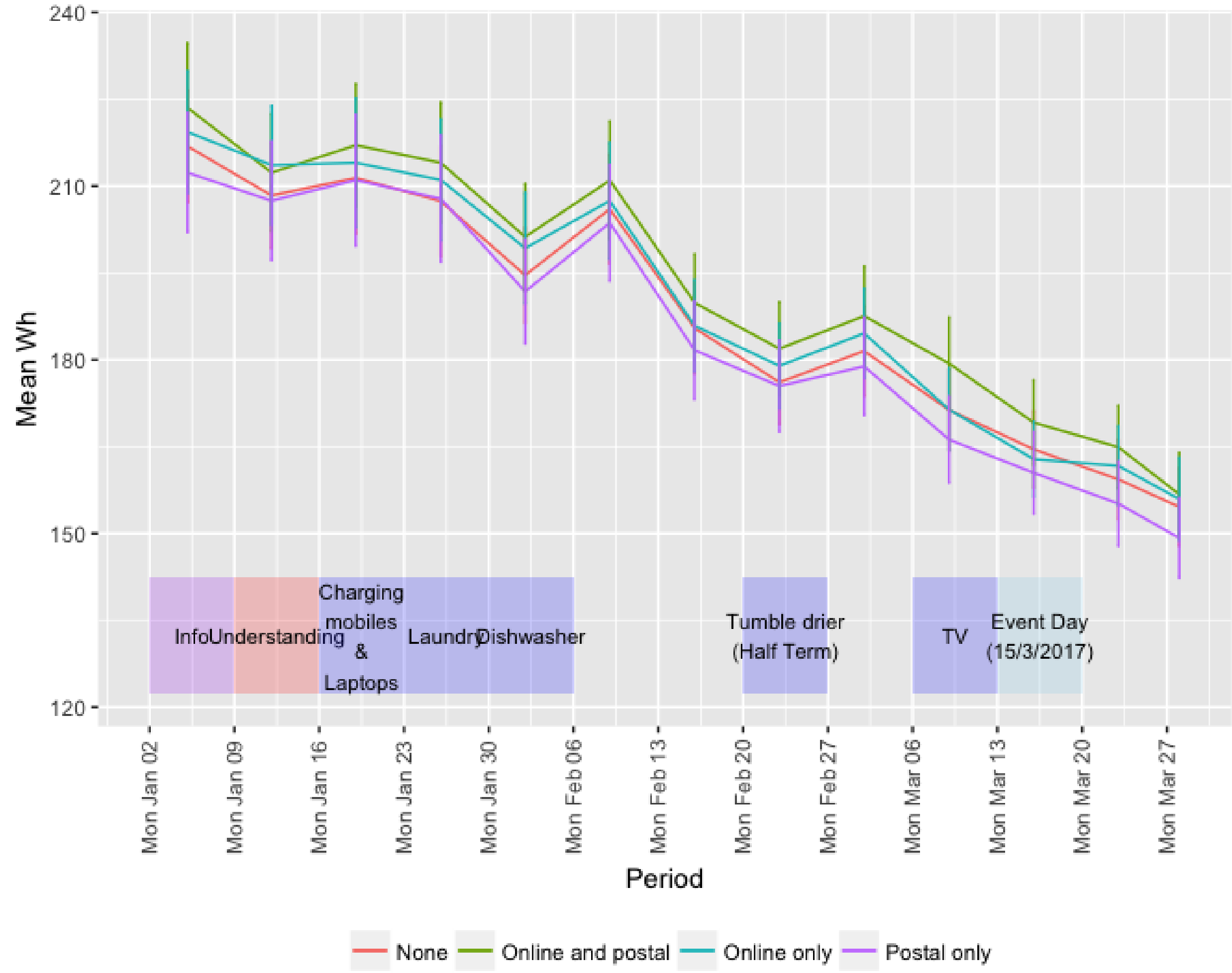

Figure 14 Mean weekly Wh by intervention type

As noted earlier, the peaks and troughs are evidence of a temperature effect in some of the trial weeks. As before, there are few apparent differences between trial groups on a week-by-week basis for the 16:00 - 20:00 period specifically although it is noticeable that the 'Online and postal' group appears to have consistently higher consumption (see Appendix A.3.3, Table 43). The 'Postal only' group still shows a slightly lower overall consumption level with the differences reduced when considering only this evening peak period. However, the extent to which the 95% confidence intervals overlap suggest that few of the trial interventions in the coded weeks will prove to be statistically significant.

[12] Cases are removed where the mean weekly 16:00 - 20:00 Wh consumption is zero and also where the weekly 16:00 - 20:00 standard deviation is zero (indicating a constant consumption level or single observations).

Figure 15 shows the differences between mean Wh for the 16:00 - 20:00 period by week and intervention group relative to the control group. Some intervention groups show a consistently higher level of consumption than others, and the 'postal only' group is the only one to show more or less consistently lower consumption than the control group, especially towards the end of the period.

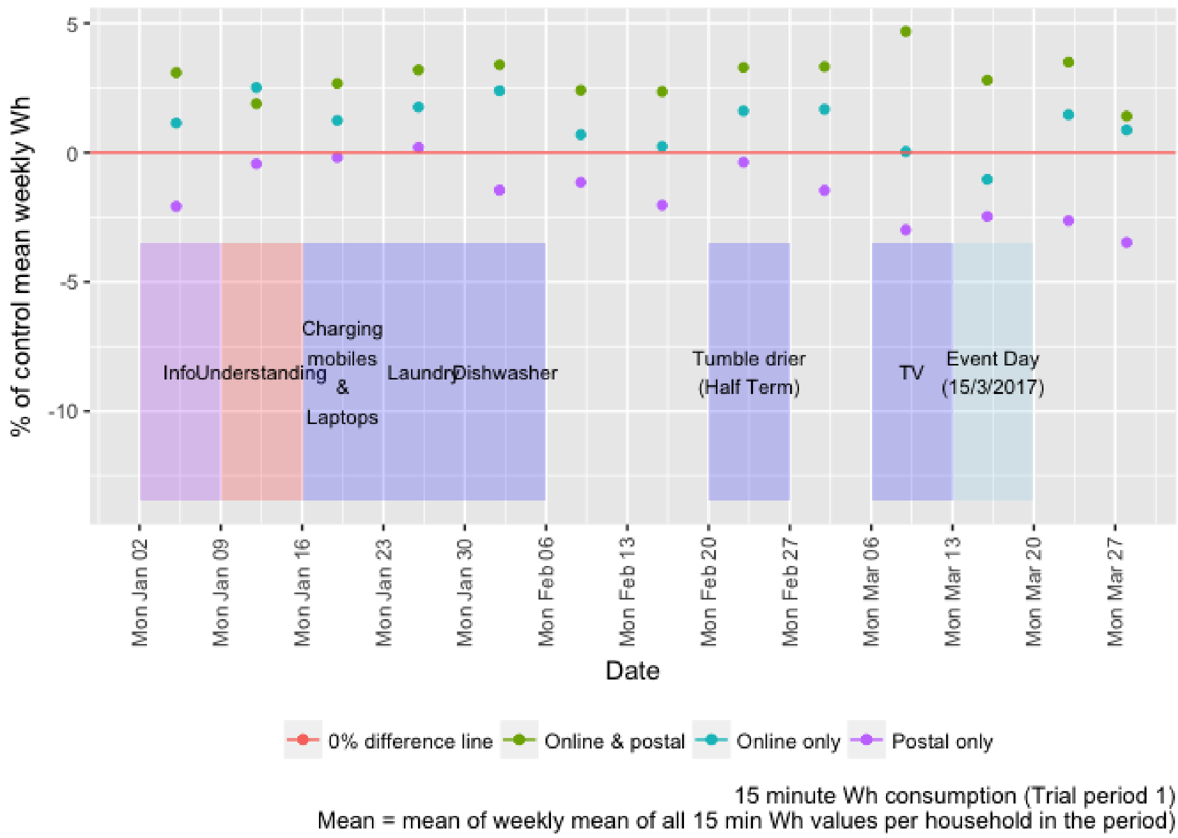

Figure 15 Difference in mean weekly consumption relative to control group (16:00 - 20:00)

A description of the consumption data has been provided, and the patterns exhibited in different groups has been visualised. The analysis will now move to formally model and perform statistical tests for relationships between intervention groups, household characteristics and consumption.

4.4.3 Intervention tests (regression approach)

In this section, tests were conducted for intervention effects in the 4-8 messaging weeks using seven regression models, one for each week:

1. w/c 2nd Jan: background information =Trial Week 01
2. w/c 9th Jan: 'understanding' = Trial Week 02
3. w/c 16th Jan: charging mobiles, tablets and laptops = Trial Week 03
4. w/c 23rd Jan: laundry = Trial Week 04
5. w/c 30th Jan: dishwasher = Trial Week 05
6. w/c 20th Feb: tumble drier = Trial Week 08
7. w/c 6th March: watch TV = Trial Week 10

The analysis uses log (mean Wh) as before. As this produced some log mean Wh values that are less than 0, these are removed in the subsequent analysis.

In the following sections, models are run for each trial week. In each case, the analysis follows the same approach:

- an initial simple model using just trial intervention (treatment group) as a predictor of the 'main' effects of the interventions,
- a secondary, more complex model that includes a number a household attributes which are modelled as interaction terms with treatment group membership.

The second model includes attributes from the household surveys, and therefore has a smaller number of observations.

4.4.4 Model summaries

The results for the weekly models are combined and reported in the following sections with the results of the simple regression model first. The full table of results can be found in Appendix A.3.3.1.

Due to the size of the tabulated results for the complex model, the results are reported in sections, focusing on a selection of different predictor variables from the table for examination in detail.

4.4.4.1 Simple model

The results from the initial (simple) model are combined and visualised in Figure 16. Each trial group is shown with a set of bars representing each trial week as shown in the figure legend. The wider, shaded bars indicate the effect coefficients reported by the model and the black lines represent the 95% confidence interval (CI) for the coefficients. The results show that membership of one of the treatment groups (when compared to the control group) does not predict any significant difference in consumption, measured as log mean Wh per week (indicated on the figure by the 95% CI bars crossing zero).

Whilst the results are not statistically significant, the 'online and postal' group has consistently higher consumption than the 'none' group (control) and that the difference increases over the trial period. In contrast, the 'postal only' group has consistently lower consumption than the control group (c.f. Figure 15).

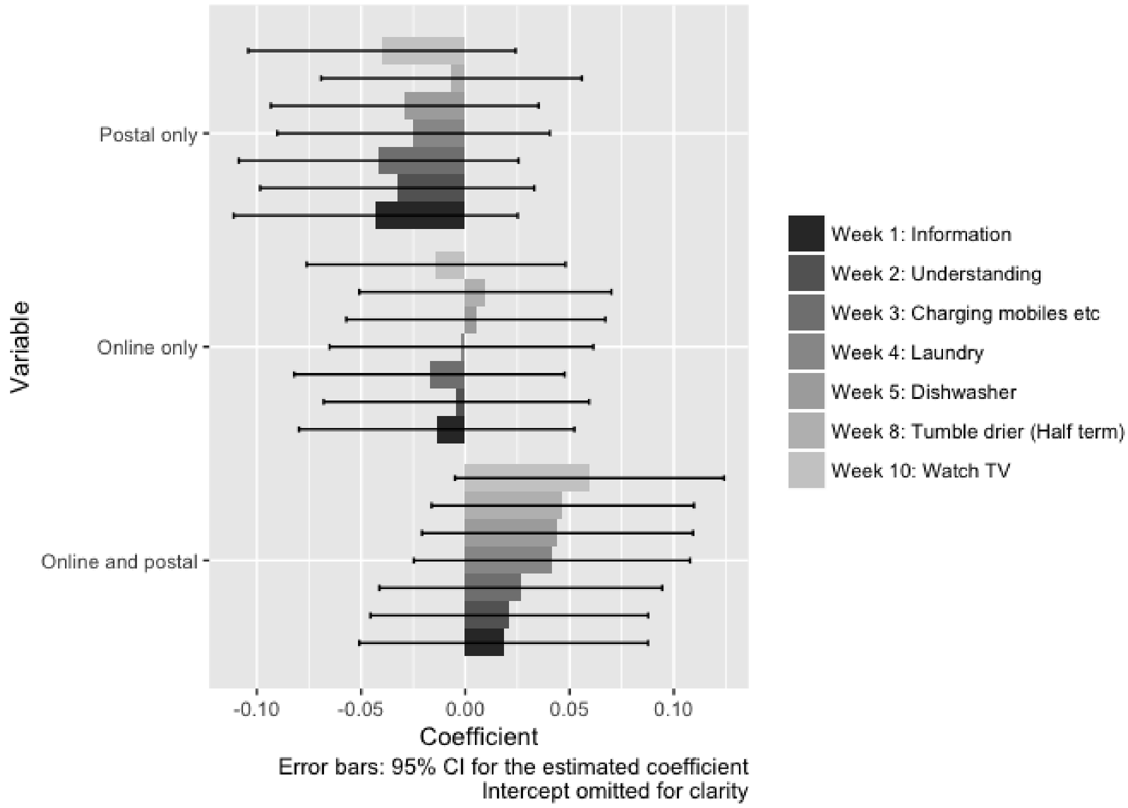

Figure 16 Model 1 results for the 4-8 messaging trial groups.

4.4.4.2 Household attribute model

This section describes the results of the second, more complex model that includes a number a household attributes modelled as interaction terms with treatment group membership to understand if any of these attributes had an additional reducing effect.

The full results are difficult to decipher in a single visualisation and due to its size, it is included in tabular format which can be found in Appendix A.3.3.2 (Table 45). A number of predictor variables have been selected from the tables for further examination. It should be noted that all the results described here are net effects when all other co-variates are controlled - they are not bi-variate results. Figures of these results are included in Appendix A.3.3.3.

Household size: the first example of these results reveals the role of household size (see Appendix A.3.3.3, Figure 45), showing that larger households tend to use more electricity (statistically significant). There are generally no statistically significant interaction effects with the intervention groups, the exception being week 8 (half-term) in four-person households receiving postal only treatment.

*Dwelling Type: t*he second example examines the role of dwelling type showing that most dwelling types use less than detached homes (the reference category) and this is especially true of flats/other dwellings (see Appendix A.3.3.3, Figure 46). In each case, there are no statistically significant interaction effects except for those in *flats/other* who were in the *postal only* group who appeared to consistently consume more.

Primary heat source: As shown in Section 4.3.1, the rates of week-by-week change in electricity consumption is associated with heat source. The effects of different heat sources are therefore examined in the household attribute regression model. In the case of households with *oil, wood, solid fuels*, *biomass etc* as a main heat source, these households tend to use more electricity in general, perhaps because they use electricity as an additional heat source (Appendix A.3.3.3, Figure 47). There are no statistically significant effects for the interaction terms although it is noticeable that there are (non-significant) but consistent reduction effects for Week 8 for the *online only* and *online and postal* interventions.

Figure 17 shows the results for those households with electric storage heaters. The large error bars are indicative of smaller number of households in this group. In this case, the main effect also shows that they tend to use more electricity. However, comparing the values of the coefficients, the effect is not as large as for those households who use oil. Interestingly, those households who received the *Online and postal* interventions appeared to use more in all weeks with the error bars indicating that the effect is statistically significant except in weeks 8 and 10. There are no negative effects at all.

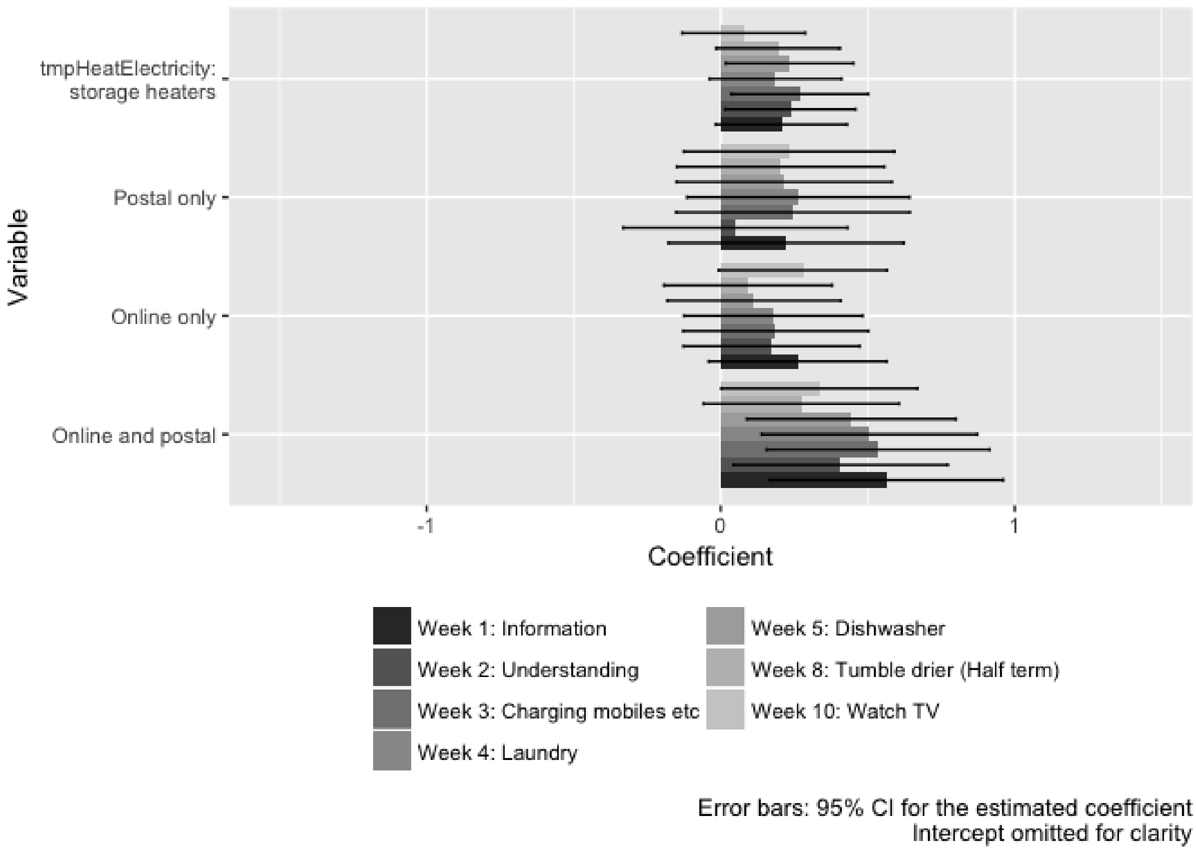

Figure 17 Regression model results for intervention: electric storage heating interaction.

The results for *Other electricity* again shows that these households tend to use more overall and also that there are no statistically significant interaction effects (Appendix A.3.3.3, Figure 48). In this case, the even larger confidence intervals are also strong indicators that the small number of households in this group is causing greater uncertainty. However, it is noticeable that those who received *only online* messages appeared to use less (although not significant) in Weeks 1-4 whilst those who received *Online and postal* messaging used less in weeks 5, 8 and 10 (although again n/s). Consistent (but non-significant) higher consumption is observed in households who received the *postal only* messaging.

Comparing these results for *Other electricity* and *Oil* categories suggest a greater degree of variation than was the case for storage heaters. This could mean that these households had more ability to reduce electricity consumption than did those with storage heaters, perhaps because they have more direct control over the use of electricity to heat their dwellings.

Eco-attitudes: households with higher environmental behaviour scores have lower consumption, with the effect being statistically significant. No significant interaction effects are observed although consumption is generally lower in the *postal only* group (Appendix A.3.3.3, Figure 49).

Income: the results for households of income > £80k variable show non-significant effects but consistently higher consumption. Interestingly, from the interaction terms, the effect of being in this income category appears to be slightly greater for households in the *online and postal* treatment group (Appendix A.3.3.3, Figure 50).

Ethnicity: two ethnicity sub-groups also appear to show a response in the regression model. Consistently lower consumption is observed in the *Asian/Asian British* group (Figure 51), but the effect is generally non-significant (except in week 2). The interaction effects are also not significant but show consistently lower consumption in the *online and postal* and *online only* treatment groups and higher in the *postal only* group. In contrast, the *Ethnicity Other* group has consistently slightly higher consumption, although again non-significant (Figure 52). The wider confidence interval also indicates a smaller group and/or higher variability in consumption for these households. Adding the interaction terms, consistently lower consumption is observed, particularly in *postal only and online and postal* treatment groups. Although these effects are non-significant they may point to a small response in households in this ethnicity group.

Presence of children: finally, the results show that the presence of children within households is associated with lower consumption when all other variables are held equal, although only statistically significant in week 8. There are no significant effects in the interaction terms (Figure 53).

4.5 Impact of event day interventions

This section provides an updated analysis of the impact of the 'event day' using the most recent household survey data. The analysis previously presented in SDRC 4 has been updated[13] to include:

- Additional household attributes in models to understand if different kinds of households responded more strongly
- Comparison of time-use diary responses for different trials to assess extent of any detectable changes in behaviour

As earlier, all zero Wh observations have been removed before analysis (see Appendix A.3.4 for sample size). The treatment received by each trial group is shown in the Table 8.

Table 8 Clarification of event day intervention types by trial group

BMG Group	TP1 event day intervention
BMG Group 1	Control
BMG Group 2	Online messaging + event day notification
BMG Group 3	Online & postal messaging + event day notification + £ incentive
BMG Group 4	postal messaging, no event day notification

4.5.1 Descriptive Analysis

This section reports descriptive analysis of the difference in consumption between the trial groups for the event day. Overall, on the event day:

- BMG Group 2 mean Wh for the 16:00 – 20:00 period was 96.37% of the BMG Group 1 (Control) mean – a difference of 3.63 %
- BMG Group 3 mean Wh for the 16:00 – 20:00 period was 96.58% of the BMG Group 1 (Control) mean – a difference of 3.42 %

Figure 18 and Figure 19 show the mean consumption in each 15 minute period for the event day and the day before/after to provide a visual depiction of any shifting of consumption to periods outside the event day peak period and/or to the previous or subsequent day. Adding confidence intervals to the figures indicates that any visible differences between the means are unlikely to be statistically significant (see Appendix A.3.4, Figure 54 and Figure 55).

[13] Results reported here benefit from using the most recent household survey data and therefore a slightly larger sample size.

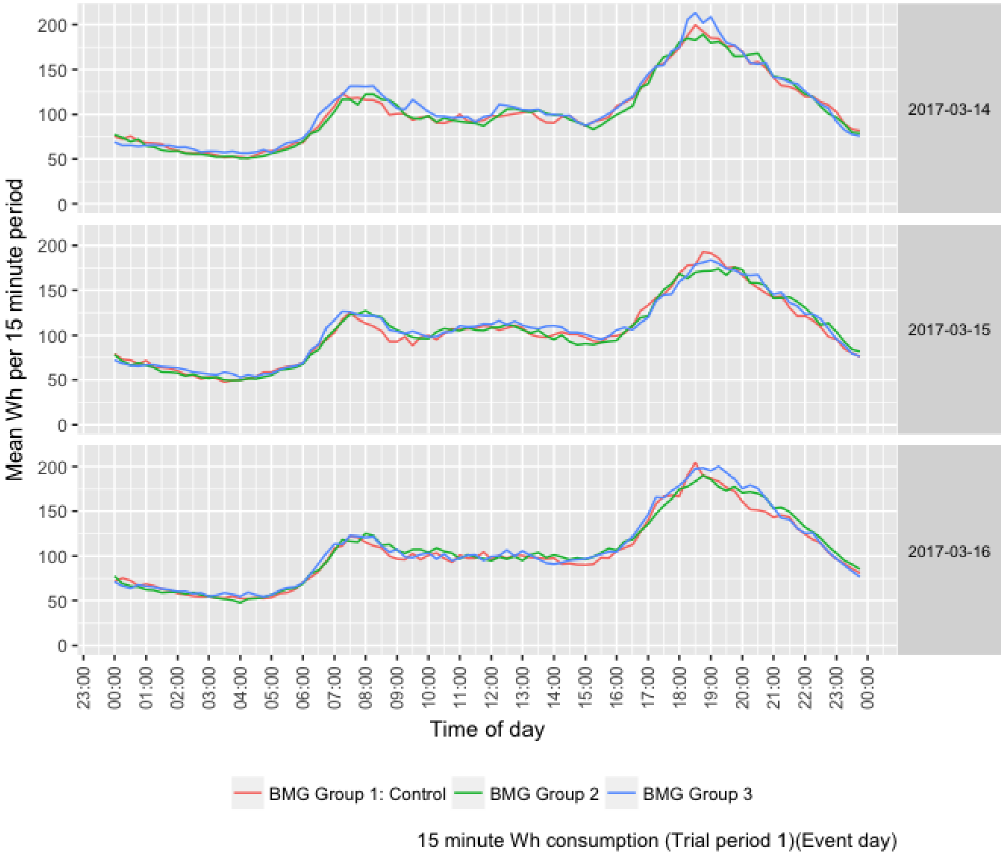

Figure 18 Mean 15 minute Wh consumption profile by trial group – event day (+/- 1 day)

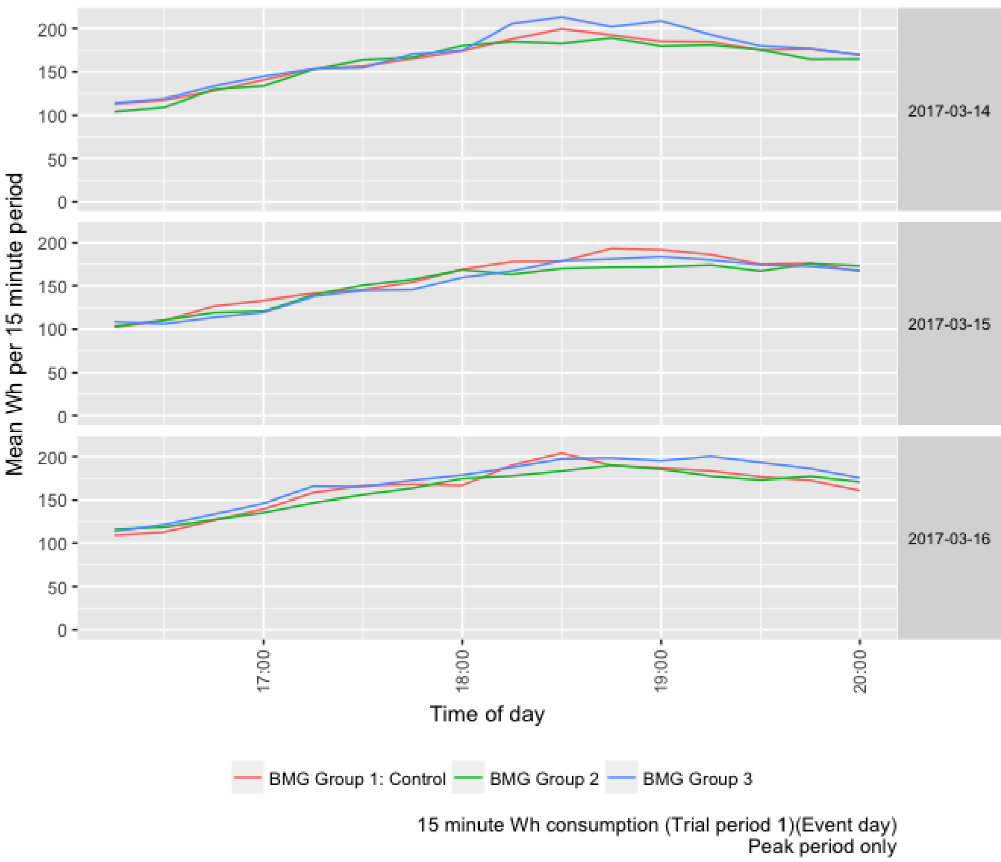

Figure 19 Mean 15 minute Wh peak hours profile by trial group – event day (+/- 1 day)

Figure 20 shows the overall mean for the 16:00 – 20:00 periods of each day compared to the 4 hours before/after.

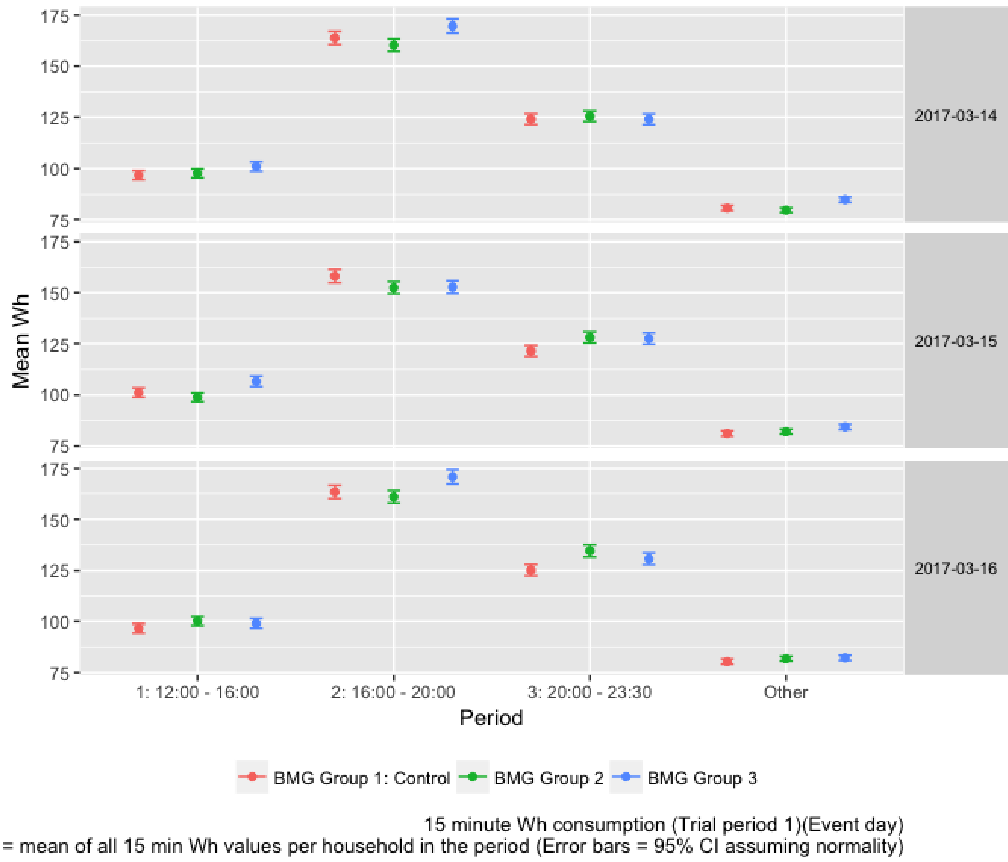

Figure 20 Mean Wh per 15 minute period during event day (where other refers to the time period 00:00 to 12:00).

In summary, the results in these figures provide the following observations:

- On the day preceding the event day: Group 3 appeared to use more than the other groups during the evening peak period which would be the case if consumption had been shifted to this day from the event day;
- On the day of the event: Groups 2 and 3 used slightly less than the Control group during the targeted peak period but only Group 3 used more in the period just prior to the peak period. Both Groups 2 and 3 appeared to use slightly more than the Control in the period just after the peak;
- On the day after the event: Group 3 again used slightly more than the other two groups during the peak period which would be the case if consumption has been shifted to this period from the day before.

However, the extent to which the 95% confidence intervals overlap suggests that not all of these effects will prove to be statistically significant. The next section formally tests for differences using a standard regression modelling approach.

4.5.2 Regression Analysis

This section reports the results of a regression based analysis which assesses the factors associated with mean consumption for three 4-hour time periods during the event day: *pre-peak* (12:00 – 16:00hrs), *peak* (16:00 – 20:00hrs) and *post-peak* (20:00 – 00:00hrs). The simplest such model uses membership of the trial groups as the sole predictor and tests the extent to which experiencing an intervention is associated with lower or higher consumption compared to the control group. The second model makes use of additional data from the Navetas portal to determine the extent to which opening the event day notification email is associated with lower or higher consumption. Clearly, this would be an indicator of more active reception of the event day message and is the only system-based data collected for the receipt of the event day notification.

It should be noted that the regression approach used requires that the dependent variable (mean Wh) be normally distributed. The models use log(mean Wh) to transform the distribution prior to model estimation, this means that coefficients reported in subsequent tables and figures in the next sections represent the effect of co-variates (such as group membership) on log(Wh) and not Wh.

The models are separated into an assessment of differences between the intervention and control groups as follows:

- Differences in the period 12:00 – 16:00 on the event day. The hypothesis being that the intervention groups would use more electricity in this period to avoid the peak period;
- Differences in the period 16:00 – 20:00 on the event day. The hypothesis being that the intervention groups would use less electricity in this period;
- Differences in the period 20:00 – 00:00 on the event day. The hypothesis being that the intervention groups would use more electricity in this period to avoid the peak period;

In addition, for each time period 4 models were constructed to estimate the following:

- Model 1: A basic regression model testing only the effect of membership of the trial groups.
- Model 2: A model which also tests the effect of the household having opened the event day email (the only communication where interaction can be confirmed);
- Model 3: A model which also tests for the effects of pro-environmental attitudes as reported by the household response person.
- Model 4: A model which also tests for the effects of the presence of children (as a predicted constraint on evening peak consumption reduction).

The results for all of the above models are given in Appendix (A.3.5). Note that models 3 and 4 currently have substantially lower numbers of households compared to the others, as there are still significant numbers of non-responses within the household survey data. This is likely to lead to larger confidence intervals.

4.5.2.1 Event day pre-peak period results

The results for the regression models for the hours 12:00 to 16:00, or the '*pre-peak*' period are provided in full in Appendix A.3.5.1 (Table 47).

The results summarised below refer to the models described above for the *pre-peak* period. Comparisons are made in contrast to the control group.

In models 1 and 2 (that do not take account of pro-environmental attitude), membership of trial group 3 is associated with, on average, 6% higher mean consumption in the *pre-peak* hours (95% CI, -3% to +15%) suggesting shifting of consumption to the hours before the peak. Within this group, those who opened the Loop portal email had an additional 2% higher pre-peak consumption. None of these results are statistically significant.

In model 3, pro-environmental score is observed as a statistically significant predictor of lower consumption. From the interaction coefficients, group 3 households have slightly higher consumption while those in group 2 have slightly lower consumption. These interaction effects are not statistically significant.

In model 4, the presence of children in households is associated with higher consumption (c.10%) during pre-peak hours but the negative interaction terms show that households with children in the treatment groups consumed slightly less than households with children in the control group (not statistically significant).

4.5.2.2 Event day peak period results I (16:00 – 20:00)

The results of identical regression models for the event day *peak* period, show that, in contrast with the control group, membership of the treatment groups (trial groups 2 and 3) is associated with lower mean consumption in the peak hours. The results also show that, as with the pre-peak models, the pro-environmental (eco_mean) score is a statistically significant predictor of lower consumption. From the interaction terms for both treatment groups, it can be seen that households with higher eco-mean scores have, on average, slightly higher consumption than those with lower eco-mean scores (shown as positive in Appendix A.3.5.2, Table 48), although these interaction coefficients are not statistically significant.

The presence of children in households is associated with higher consumption and is statistically significant. Households with children within the treatment groups however, have lower consumption compared to the other households in these groups, again the interaction is not significant.

4.5.2.3 Event day peak period results II (18:00 – 20:00)

Further examination of the 24-hour profile shown in Figure 19, suggests that the largest difference was found in the final two hours of the four-hour peak period. Figure 21 shows the mean 15 minute Wh for this period by treatment group.

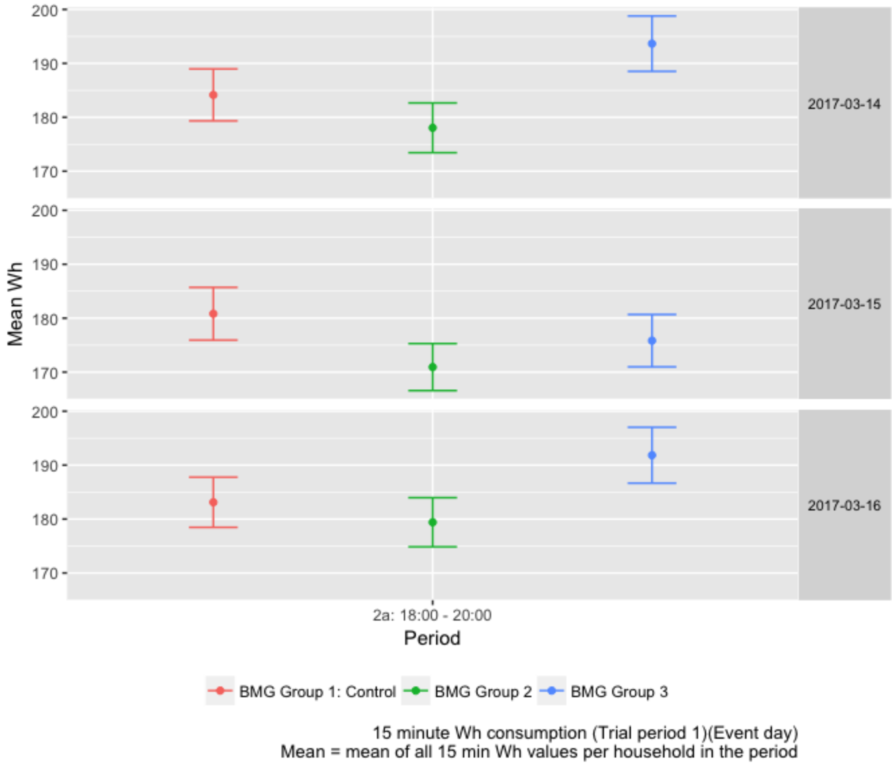

Figure 21 Mean Wh per 15 minute period during event day: 18:00 – 20:00

The results of replica regression models (suffixed with 'a' for clarity) show similar effects to those in the 16:00 – 20:00 peak period models, although there is some variation in the strength of the effects between the treatment groups (see Appendix A.3.5.3, Table 49). In model 1a, including only trial group membership, the coefficients are slightly larger for both Group 2 and Group 3 indicating greater difference between both treatment groups compared to the control group in the 18:00 – 20:00 time period. In model 2a, the effect coefficient for Group 3 was the same as in the longer time period (with a slightly wider confidence interval) but the coefficient for Group 2 was larger suggesting a larger difference in consumption compared to the control group in the final two hours of the period. The interaction coefficient is also larger suggesting that those households who opened the email notification may have responded more in the final two hours of the peak period than in the first two hours. As previously, these results are not statistically significant.

4.5.2.4 Event day post-peak period results

The regression model results for the *post-peak* models reveal again that pro-environmental households have lower consumption (statistically significant) in this period but this factor has no significant interaction with membership of the treatment groups (full results are provided in Appendix A.3.5.4, Table 50).

4.5.3 Time-use data

During the trial period, only the Control Group and Trial Groups 3 & 4 received the time-use diary. Table 9 shows the numbers of cases in different sub-groups for the event day intervention.

Table 9 Total diary cases by trial group (note no Group 2)

Intervention	Group	N Diaries
Event day	BMG Group 1: Control	319
Event day	BMG Group 3	273
Event day	BMG Group 4	303

4.5.3.1 Aggregated electricity demanding activities

Note that the simple indicator used here is a simple count of all electricity-using acts reported (see Appendix A.3.6, Table 51).[14] It does *not* take account of duration, therefore 1 act of 'cooking' which occupied the whole of the 16:00 – 20:00 period looks like 'less' than several acts of different kinds. This can only be fixed by doing a more sophisticated analysis which includes duration of acts.

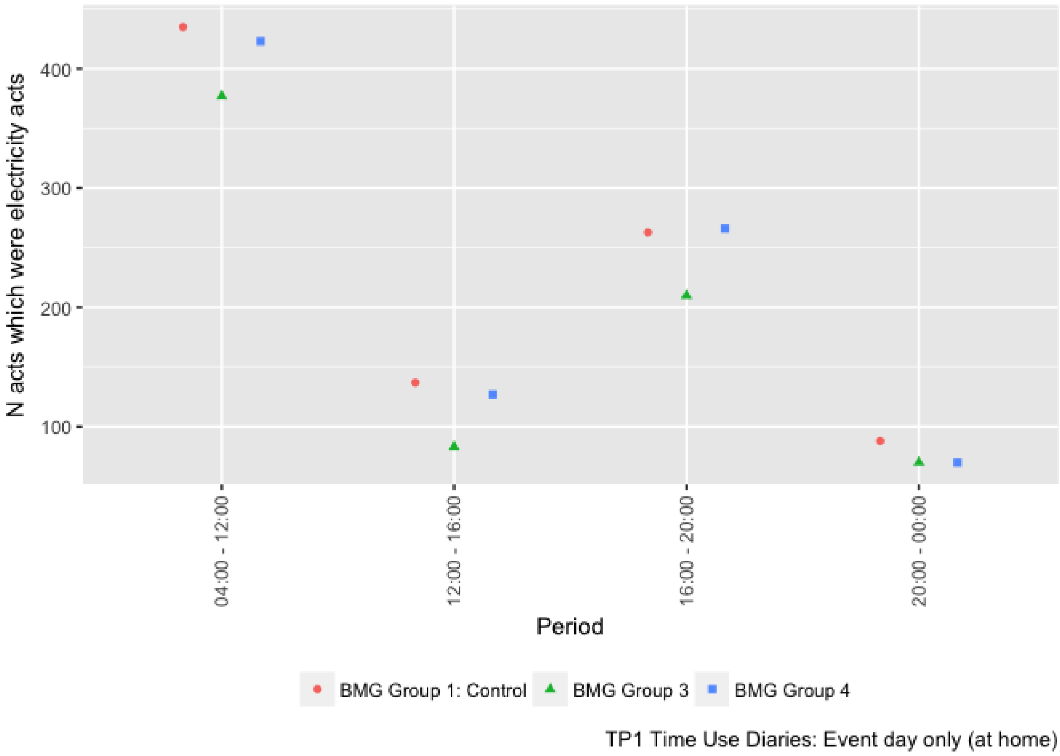

Figure 22 Number of electricity-using acts at home by trial group

[14] Any explicitly electricity-using activity has been re-coded into one code. Although 'Shower' is included, it should be noted that this may not reflect the actual time of energy demand in households with immersion heaters and hot water storage and may not be relevant for households with gas-heated hot water.

Figure 22 shows that Group 3 appeared to report fewer electricity using acts at home during the event period. Note that the sum chart will emphasise differences as it is the sum of the number of acts. The differences are formally tested using a similar regression model to that used for testing the differences between kWh consumption above but using a Poisson model appropriate for the Pareto curve distribution (see Appendix A.3.6).

No statistically significant effects were found in these models (see Appendix A.3.7, Table 52 for full regression results). However, group 3 has a lower likelihood of performing an electricity act during the *peak* hours, particularly those households opening the Loop notification email. Those households are also shown to be more likely to perform an electricity using act in the *post-peak* periods (and to a lesser extent in the *pre-peak* period).

4.5.3.2 Event day: Specific activities

In this section, results are presented for a selection of 'electricity demanding' acts to assess the extent to which any particular effects relating to specific activities can be detected. Results of all of the regression models can be found in the Appendix (A.3.7).

First, acts involving cooking are examined (hob, oven, toaster, microwave, kettle). A similar pattern is observed in cooking acts as in the sum of all acts above, with no statistically significant effects (see Appendix A.3.7.2). The model coefficients imply a smaller difference in cooking related acts recorded between group 3 and the control group in *peak* hours and larger difference in the *post*-peak period than in the previous chart. Group 4 show increased likelihood of performing electricity-using cooking acts during peaks hours, but again, the effect is not significant.

Results for activities involving a dishwasher also reveal no statistically significant effects, however, in contrast to the control group, households in the two treatment groups are more likely to report an act of using the dishwasher in the *pre-peak* period, and less likely in the *peak* and *post-peak* periods, consistent with shifting use of this appliance away from the peak (see Appendix A.3.7.3). Those households in group 3 who opened the notification email were less likely than others in the group to use a dishwasher in the *pre-peak* and *peak* periods and more likely to in the *post-peak* period, although examination of the table reveals very large confidence intervals around the interaction term (most likely due to the very few acts reported by households in this group in the *pre-peak* and *peak-hours* time periods).

For laundry, electricity-using acts are divided into two related activities: washing with a machine; and tumble drying. As with dishwasher related activity, the low number of reported acts results in large confidence intervals around the effects reported in the models (see Appendix A.3.7.4 and A 3.7.5). Whilst there are no statistically significant effects reported in the models, results are consistent with washing machine use being avoided in the *peak* period for both groups. In contrast to dishwasher use, acts relating to washing machine use are also less likely to be performed by groups 3 and

4 (compared to the control group) in the *pre-peak* period, and more likely in the *post-peak* period (except for those in group 3 who opened the email). For the activity of tumble drying, no significant effects are reported compared to the control group. The large confidence intervals and low incidence of reported acts should be noted for this activity (Appendix A.3.7.5, Table 60).

4.5.3.3 Location

There is a suggestion in the literature that people may avoid consumption at home by being out of the home. This is tested using the 'location' code in the diary data: if people do respond by going/staying out then fewer acts should be recorded at home. The model results (see Appendix A.3.7.6) show a statistically significant effect for Group 3 in the *peak* period (model 1) and for those who opened the loop email in particular (model 2), confirming that these households were less likely (c. 11% and 16% respectively) to perform energy acts at home than those in the control group.

4.5.3.4 Summary

Comparing reported energy-using acts of two trial groups receiving treatment (3 and 4) with the control group suggests that although group 3 reported fewer acts during the *pre-peak* and *peak* hours on the event day. The models show no statistically significant effects relating to the total number of energy-using acts, or acts related to a number of specific activities and appliances. The model results do, however, provide evidence to support the claim that households in trial group 3 avoid household energy consumption by being away from home.

The results presented in this section show differences between the trial groups that are consistent with the hypothesis that treatment groups would report more energy-using acts than the control group in the *pre-peak* and *post-peak* periods, and fewer in the *peak* period. Generally, this is apparent for group 3, particularly those households who opened the event day email notification. In group 4, who did not receive event day notifications, the models show mixed results that are not consistent with this reducing/shifting electricity-use hypothesis.

The modelling of the data reveals that there is difficulty in detecting changes with any confidence particularly in specific activities (e.g. tumble drying) where there are small numbers of households performing these activities on any particular day (and in any of the four-hour time periods). The relative rarity of these acts increases the uncertainty in the models and suggests that to be useful, time use diaries need to be concentrated on specific event days with sufficient sample sizes.

5 Small area estimation of future demand profiles

5.1 Overview

This section uses a spatial microsimulation approach to combine the SAVE household level survey data, the observed kWh data and Census 2011 small area tables to produce estimated small area level demand profiles for each output area (OA) in the SAVE study region (6,136 geographical areas). This work therefore, comprises a major update to SDRC2.1 which used the same approach but with time-use diary data to infer electricity consumption at different times of day.

This iteration of the SAVE customer model framework also implements a change in modelling scale, from the larger 'lower super output area' (LSOA) previously proposed, to the smaller 'output area' (OA). This change reflects the preferred scale at which the LV network model is to be implemented and constitutes a major change from the originally planned LSOA level modelling. In particular, this means that the area level profiles will be based on a smaller number of households (c 100-200) rather than the larger LSOAs with households in the range 500-1000. Reducing the number of households may result in unusual patterns that may come to dominate the OA level average demand profiles.

In contrast to SDRC2.1's reliance on imputed half-hourly kWh derived from time-use data, SDRC2.2 uses actual kWh consumption data to provide outcomes under the following four scenarios:

- 'Baseline' –using the full sample of households (all treatment groups) to create weighting for households to simulate consumption profiles using data collected on the Sunday 8th January 2017 (assuming no intervention effects);
- Model 1 –> trial group 1 'control' –using the consumption profiles for only the control group households in the 3 days over the TP1 event day (14th to 16th March 2017) to create a baseline scenario
- Model 2 -> trial group 2 event day roll out –using only households from trial group 2 over the same event period to simulate the roll out of the treatment applied to this group (Online msg + loop);
- Model 3 -> trial group 3 event day roll out – as Group 2 above but simulating a roll out of the different event day treatment applied to this group (Online & postal msg + loop + £).

The resulting profiles will comprise estimated 1/2 hourly mean kWh aggregated to the OA level which can provide estimates of potential demand reduction at the local area level. In addition, the profiles can then be used as input to the SAVE Network Model tool to enable SSEN to assess the cost/benefit of targeted vs whole-customer population interventions in the context of network re-enforcement costs.

5.2 Modelling approach

As in SDRC2.1, the small area estimation of demand profiles uses an iterative proportional fitting (IPF) based spatial microsimulation approach (Anderson, 2013; Birkin and Clarke, 2011; Birkin and Wu, 2012). IPF is a widely-used, deterministic approach to allocating individual micro-data to geographical zones (Lovelace and Dumont, 2017).

Our implementation of small area estimation within the SAVE customer modelling framework consists of the following work stages:

1. Constraint variable selection: the identification of variables common to both the SAVE survey and the Census which are reasonable predictors of evening peak consumption;
2. Small area estimation: Iterative proportional fitting of survey data using Census area level tables and identified constraint variables to produce a synthetic 'energy census' household level dataset;
3. Linking of demand profiles from the SAVE data to the synthetic households to enable the production of area (OA) level average demand profiles.

The modelling work conducted to date is presented in the following sections.

5.3 Data used

This section uses the household data as set out in the data methods section 3.2 above together with:

- Navetas Loop 15 minute Wh data for January 2017 to test constraint variables and model output (baseline scenario
- Navetas Loop 15 minute Wh data for 14th to 16th March 2017 to test model output (event day scenarios)
- UK OA level Census data as set out in the data methods section above.

5.4 Constraint selection

The IPF process requires the use of variables (characteristics) at the individual level within the microdata (household surveys) which are also available at the aggregate level within the spatial data (Census). To enable this, the household surveys were designed to reproduce many of the characteristics captured by the Census.

To select the constraints, Lovelace and Dumont (2017) note that it is preferable to:

- Use constraint variables (explanatory or predictor variables) that are relevant to the target outcome variable (i.e. electricity consumption)
- Strike a balance between using enough constraints to adequately model the differentiation between geographical areas while avoiding problems with 'over-fitting' (see limitations below).

First, the variables that exist in both the micro data and constraints are examined for their value in predicting the outcome of interest – evening peak electricity consumption.

5.4.1 Establish optimum constraint variables

In this section, the Census variables contained in the SAVE household survey are tested to select the best 'constraints' to use in the weighting process. The outcome variable is defined as:

- evening peak (16:00 – 20:00) kWh per household operationalised as mean(kWh) for January 2017 over all non-zero kWh values

Candidate variables must be in both Census and SAVE survey (see Table 10 and Table 11) and SAVE data must be code-able in the same form as Census small area counts.

Table 10 Candidate SAVE variables (HRP)

Variable	Available at OA	Available at LSOA
HRP employment status	Y	Y
HRP qualifications – not yet tested	Not for HRPs	?
HRP NS-SEC	Y	Y
HRP ethnicity	Y	Y
HRP religion – not yet tested	Y	Y
HRP age	Y	Y

Table 11 Candidate SAVE variables (household)

Variable	Available at OA	Available at LSOA
N people	Y	Y
Presence of children	Y (composition)	Y
Main heat source (gas vs elec vs other)	No	Y
Dwelling type	Y (accommodation)	Y
N rooms	Y	Y
Tenure	Y	Y
N cars	Y	Y
Composition (married/single/civil partnership etc) – untested	Y	Y
Presence of gas (DECC LSOA data)	No	Y

The ability of these candidate variables to predict mean half-hourly kWh for the period 16:00 – 20:00 is tested to reduce the list of potential constraint variables. Normal practice is to run a regression model to isolate the best predictor variables.

5.4.2 Estimate regression models

Having established the need to use transformed consumption values (see Appendix A.4.1)[15], two regression models were tested, including all of the potential variables in the first model, with the second model only including those variables that exist at the OA level from Census data (see above). The project is working on validating the coding of the NS-SEC[16] categories based

[15] Transformed log(mean) Wh consumption values are used to avoid the problem of a non-normally distributed dependent variable.

[16] NS-SEC variable is derived from occupation and employment status information see http://www.ons.gov.uk/ons/guide-method/classifications/current-standard-classifications/soc2010/soc2010-volume-3-ns-sec--rebased-on-soc2010--user-manual/section-14--deriving-the-ns-sec--self-coded-method.pdf.

on the survey responses. The NS-SEC constraint is to be integrated into future iterations of the model.

The results for the models show that one or more of the categories in the following variables are statistically significant predictors (full results can be found in Appendix A.4.1, Table 62):

- Household attributes: household size, presence of children, heat source, access to mains gas, dwelling type, tenure, number of cars, income
- Household response person attributes: HRP ethnicity, HRP employment type, HRP age

However, as Table 10 and Table 11 show, the *income*, *heat source* and *access to mains* gas variables cannot be used as they are not available at the OA level from the UK 2011 Census.

5.4.3 Select final constraints

In order to identify the most parsimonious set of constraints for use in the IPF process (Lovelace and Dumont, 2017), both models were run using a stepwise regression[17] function to select the final model form which will then have the constraint variables that provide the most powerful model. While there is criticism of selection models of this type when used for inference, in this case the technique is simply used to look for the strongest correlations – causation is not of particular interest. Both models are tested to understand the extent to which not having some of the variables in model 1 at OA level might cause problems. A summary of the selected variables and their rank by model fit is presented in Table 12, the full stepwise regression results can be found in Appendices A.4.1.2 and A.4.1.3. The ordering of variables is important as the constraint fitted last in the IPF process will always fit perfectly, therefore the constraints are ordered in relation to their contribution to model fit (largest contribution last).

[17] i.e. using a stepwise entry of candidate variables to a regression model and then deducing optimum combinations and order, using contributions to overall model fit changes or AIC-assessment to select the best fitting combinations of variables.

Table 12 Stepwise regression model results (summary)

Variable	Model 1 rank	Model 2 rank
Number of people	1	1
Number of rooms	2	2
Main heat source	3	N/A
Number of cars	4	3
Access to mains gas	5	N/A
HRP ethnicity	6	4
Income	7	N/A
Presence of Children	8	6
Tenure	9	x
HRP age	10	5
HRP employment status	x	7
Dwelling type	x	8

In Table 12, 'x.' indicates a variable excluded from the final model by the stepwise function. 'N/A' indicates variables not available at OA level and excluded from model 2. Comparing the results from the models in Table 12 reveals the variables that are able to improve prediction of consumption but are not available at the OA level, these include:

- main heat sources (gas vs electricity vs other)
- access to mains gas
- are in different income brackets

Heat source and access to mains gas are both available at the larger LSOA level (from Census 2011 and DECC/BEIS sub-national energy statistics respectively) but household income is not.

5.4.4 Summary

The original models show the contributions each variable makes to the prediction of evening peak electricity consumption. Those of note are:

- Household attributes which are strong (& statistically significant) predictors: No gas, 1+ child, other electric/gas heat, n people; n cars & tenure are marginal
- HRP attributes are relatively non-predictive except: retired, mixed ethnicity, Asian, NS-SEC coding

The final model based on variables available at OA level produced using stepwise regression suggests the use of the following variables as constraints:

- Household attributes: number of people, presence children, number of rooms, dwelling type, number of cars
- HRP attributes: ethnicity, employment status, age

These constraints are used in the initial IPF spatial microsimulation process.

5.5 Creating household weights using IPF

The IPF weighting procedure is used to create for each OA a weight per household such that the aggregate tables of the counts/proportions of the constraint variables used match the original Census tables. The procedure uses only those households for whom Navetas Loop data is available in each of the test time periods. Households who have elected to opt out of the trials are also filtered out.

The IPF weighting procedure is repeated for the four test scenarios described in Section 5.1. The first scenario (all households) will create a weight for each household in the complete sample whilst the three subsequent scenarios will calculate a different weight as they only use specific sub-samples (i.e. each trial group) separately.

5.5.1 Set up data for IPF

5.5.1.1 Constraint data (Census OA)

First, the constraint categories are created from the aggregate level spatial (Census OA) data and merged into a single data table.[18] Checks were carried out on each of the constraints prior to running the IPF weighting, these are detailed in Appendix A.4.2.

5.5.1.2 Microdata (SAVE households)

Through the participating households, the SAVE project sample generated the following microdata to be used in the spatial microsimulation model:

- 4,163 surveyed households, of which
- 2,925 surveyed households have kWh data for January 2017, of which
- 2,914 surveyed households have HRP age recorded
- 2,775 surveyed households have HRP ethnicity recorded
- 2,775 surveyed households have HRP employment recorded

A large number of survey households are lost due to non-response to the HRP ethnicity item. After the removals of NA and refusals the data therefore contains 2,743 surveyed households with complete constraint data and with kWh data for January 2017.

Note that some categories available for the Census (for example, the HRP age variable) do not match the microdata. Where necessary the SAVE survey categories are re-coded to ensure the correct functioning of the IPF procedure.

5.5.1.3 IPF input data

The data described above is used to set up the two input data files required for the IPF weighting process:

- constraints – area level counts of categories
- survey data – with the constraints in the same form

[18] The relevant OA level tables for each constraint variable were downloaded from nomisweb (https://www.nomisweb.co.uk/census/2011/data_finder) and merged into one data table.

The IPF process is run, using the prepared input data described above, for each scenario to create a table of weights for all households for all OAs. The constraint data remains the same for each run while the households within the survey data 'pool' changes according to the scenario being run.

5.5.2 Baseline Model (all households)

In the baseline model, the IPF procedure is run using the SAVE household survey (microdata) from all households (all trial groups).

5.5.2.1 Run IPF

The baseline model is run using 2743 households for 6136 areas, the first few rows/columns of the output are reproduced in Table 13 as an example. The table shows how a set of households within a particular OA have been given a specific weight. For one of the households the weight is zero indicating that there are no households with their combination of constraint characteristics in this OA.

Table 13 First few rows/columns of IPF weights table

Household identifier (bmg_id)	OA Code	ipfWeight
956600012	E00085919	0.0000000
956600053	E00085919	0.0931889
956600056	E00085919	0.1006993
956600058	E00085919	0.0499062
956600090	E00085919	0.0787645
956600093	E00085919	0.0215044

The IPF weighting produces a long form data file. The unique household identifiers (bmg_ids) are repeated and the weights are held in a single column (the 'ipfWeight' variable shown in Table 13).

As noted above, the IPF procedure produces some weights that are equal to zero. Note that the number of rows does not represent the 'true' number of households within the areas as many of the survey households have weights less than 1. The sum of all of the weights should give an estimate of the 'true' number of households in all of the OAs. The results are checked below.

The final stage prior to analysis is the linking of electricity consumption data associated with each SAVE household. Before this however, a series of tests are run to check the IPF procedure has produced valid output.

5.5.2.2 Test baseline household weighting for single output area

Within the IPF output table, each household has weights that reflect the extent to which it is required to ensure that the sum of the constraint variables within each 'target' OA matches the Census tables for that OA. An initial check is conducted to ensure that the IPF weighting results in the correct aggregate household counts in each category of two of the constraint variables: household size and ethnicity.

First, an OA is selected to test and the original survey data is linked to each household in the weights file. Table 14 and Table 15 show the number of households in each sub-category of the two constraints (household size and ethnicity) for the OA. Table 14 shows the numbers of households in each for the original SAVE survey data. Table 15, shows the unweighted IPF results; i.e. the count of households in each sub-category in the SAVE survey ‘pool’ for this specific OA. The results in the two tables should be identical unless households have been removed due to their weights equalling zero. If this is the case, they are likely to have been rare categories.

Table 14 Household counts unweighted survey (household size x ethnicity): all households

	1	2	3	4	5	6	7	8+	NA
White British/Irish	451	1115	436	438	143	38	13	5	0
Mixed	3	5	5	4	0	0	0	0	0
Asian/Asian British	2	14	10	13	10	5	1	0	0
Black/Black British	1	2	2	7	3	2	0	1	0
Other	4	3	3	2	2	0	0	0	0
Refused	0	0	0	0	0	0	0	0	0
NA	0	0	0	0	0	0	0	0	0

Table 15 Household counts unweighted IPF results (household size x ethnicity): all households, E00085919

	1	2	3	4	5	6	7	NA
White British/Irish	451	1115	436	438	143	38	13	0
Mixed	3	5	5	4	0	0	0	0
Asian/Asian British	2	14	10	13	10	5	1	0
Black/Black British	0	0	0	0	0	0	0	0
Other	0	0	0	0	0	0	0	0
Refused	0	0	0	0	0	0	0	0
NA	0	0	0	0	0	0	0	0

Note that Black/Black British households are not present in Table 15 indicating that there are no Black/Black British households resident in this OA and this is also the case for households with 8+ people.

To check that the IPF weights give the correct aggregate household counts a weighted table is calculated for the OA for both constraints. The results should be equal to the corresponding counts for the OA in the constraints table. If more than one constraint variable is used (as here) small errors would be expected in the weighted tables. First, the number of people per household in the constraints (Census) data (Table 16) is checked:

Table 16: OA table of household size (people) for E00085919: Census data

OACode	people 1	people 2	people 3	people 4	people 5	people 6	people 7	people 8+
E00085919	32	48	28	12	2	1	1	0

And the weighted IPF results for the SAVE sample households (Table 17):

Table 17: OA table of household size for E00085919: weighted SAVE sample

OACode	1	2	3	4	5	6	7
E00085919	32	48	28	12	2	1	1

The check is repeated for ethnicity in Table 18 and Table 19:

Table 18: OA table of HRP ethnicity for E00085919: Census data

OACode	HRP White	HRP Mixed	HRP Asian/ Asian British	HRP Black/ Black British	HRP Other Ethnicity
E00085919	121	1	2	0	0

Table 19: OA table of HRP ethnicity for E00085919: weighted SAVE sample

	White British/Irish	Mixed	Asian/Asian British	Black/Black British	Other	Refused
E00085919	121	0.999999	2.000001	0	0	0

Very small disparities are observed between the Census and weighted SAVE sample tables. This is due to the constraint not being fitted perfectly by the IPF process. The Census data contained in Table 16 and Table 18 also show that the 'Black/Black British' HRPs and the households with 8 or more members are not present in this OA and therefore not required – hence their zero weighting and absence from Table 17 and Table 19.

5.5.2.3 Internal baseline validation

As indicated above, there may be small disparities between the 'true' and 'simulated' household counts. The simulated population for the *baseline* model, calculated using the weighted ethnic distribution by household size for the entire SAVE region (all OAs), is 790,066 households. An exact match to the Census count. Further systematic validation tests are performed, with details given in Appendix A.4.3.1. The pattern of error between simulated and true household counts will be a combination of the order of the constraints together with a lack of rare combinations of constraints in the relatively small survey data used. As expected, the errors for categories of the household size constraint variable ('npeople') are extremely small as this constraint is fitted last. Overall mean error rates across all OAs are extremely small for all constraints although it should be noted that the min/max statistics suggest some OAs where caution should be used due to relatively large errors – e.g. for HRPs aged 16-24.

5.5.3 Trial groups

The same method is followed to create three further models, each time using sub-groups of households from the SAVE sample to create pools of survey households for the three trial groups. Summaries of the outputs from the IPF procedure for each of the additional models can be found in Appendix A.4.3. The simulated household populations for the trial groups are between 782,590 and 788,457 (compared to the overall Census count of 790,066 households), an under-estimate of less than 1%. The systematic error checks show that the mean errors are larger than for the initial all households model but still within +/-1 household per OA. In some cases the min/max errors indicate larger disparities caused by the use of a smaller household sample.

5.6 Results

The initial results from the four small area estimation models are presented in this section, with key features of the output from each identified. The implications of these results on the customer modelling framework are identified in section 5.7, including the contribution to the network model and identification of further work.

To illustrate the output from the small area estimation, two highly contrasting OAs are selected as the 'target' areas:

- the OA with highest % of single person households: E00167003
- the OA with the lowest % of single person households: E00115898

The OAs have been selected in this way to provide test cases that tease out any limitations in the modelling technique. The household counts for these OAs are shown in Table 20 and the resulting weighted household counts are expected to match these.

Table 20 Census counts and % single-person households for selected OAs

OA Code	Total household count	Number of single-person households	% single-person households
E00115898	85	0	0
E00167003	200	182	91

The OA with the lowest percentage of single-person households (0 households, 0%) has 85 households in total, whilst the OA with the highest percentage (182 households, 91%) has rather more at 200.

As each of the four illustrative models described in Section 5.1 above will draw upon the consumption data from a different pool of SAVE sample households, the weighting file generated by the IPF procedure for each separate model is applied to each of the two OAs in turn. The following sections describe briefly the results gained from each model. The results for each model include tables to illustrate that each of the different treatment groups produce different 'pools' of SAVE households, and that the weights resulting from the IPF process change according to their different characteristics.

5.6.1 Baseline model (all households)

Having established that two quite different OAs have been selected, kWh profile data for the first (non-holiday) Sunday in January 2017 (8/1/2017) is attached as a 'baseline' test. Half-hourly (sum) kWh consumption data is merged to the households that were pushed through the IPF process.[19]

First, the weighted counts for each household size type (single, two person etc) are checked. Table 21 contains the number of households in the SAVE sample 'pool' (N unweighted column) for each household size in both test OAs, along with the mean, minimum and maximum weights that the IPF procedure has assigned to them. The final column contains the weighted household counts (N weighted) i.e. the sums of the weights of each household size. The sum of these counts for each OA should equal the 'target' total household counts in the OA level Census data in Table 20.

[19] Only households with survey data *and* clamp (consumption) data are used so the resulting linked bmg_id <-> kWh data may be a reduced household set.

Table 21 Summary household counts and weights by household size, baseline (to 4 dp)

OA Code	label	HH Size	N un-weighted	Mean Weight	Min Weight	Max Weight	N weighted
E00115898	Lowest % single people	1	461	0.0000	0	0.0000	0
E00115898	Lowest % single people	2	1139	0.0114	0	0.9290	13
E00115898	Lowest % single people	3	456	0.0504	0	2.5322	23
E00115898	Lowest % single people	4	464	0.0754	0	3.6649	35
E00115898	Lowest % single people	5	158	0.0696	0	3.3618	11
E00115898	Lowest % single people	6	45	0.0444	0	1.3844	2
E00115898	Lowest % single people	7	14	0.0714	0	0.4685	1
E00115898	Lowest % single people	8+	6	0.0000	0	0.0000	0
E00167003	Highest % single people	1	461	0.3948	0	101.7724	182
E00167003	Highest % single people	2	1139	0.0088	0	9.7089	10
E00167003	Highest % single people	3	456	0.0044	0	1.2425	2
E00167003	Highest % single people	4	464	0.0043	0	1.9884	2
E00167003	Highest % single people	5	158	0.0127	0	2.0000	2
E00167003	Highest % single people	6	45	0.0444	0	1.7990	2
E00167003	Highest % single people	7	14	0.0000	0	0.0000	0
E00167003	Highest % single people	8+	6	0.0000	0	0.0000	0

The weighted household counts show that the weighting process has simulated the correct number of households (by household size). Note, the correct count is expected for this constraint variable only, as household size is the final constraint to be fitted (because the IPF weighting method fits the final constraint perfectly). Errors are to be expected when checking household counts using the other constraint variables (see systematic validation checks).

The unweighted household counts in Table 21 above, reveals that the OA level results for single person households are based on the 461 unweighted single person households in the pool of SAVE households. The table also shows that a number of households of all sizes in both OAs have been assigned zero weights as expected.

The mean of the weights represents the ratio of the number of target households to the number of SAVE households contributing consumption profiles and therefore indicates the extent to which households are, on average, weighted up or down to fit the OA. It should be noted from the maximum value of weights that one or more single-person households have received large weights for the OA with the highest percentage of single people. The profiles of these households will therefore dominate the subsequent results for this OA and if they have unusual consumption patterns then unusually shaped aggregate profiles would also be expected.

Having attached the 15 minute Wh data from January 2017 to the households, the consumption profiles shown in Figure 23 are the total (sum) half hourly kWh across all households for each OA for Sunday the 8th January 2017. This gives an estimate of the total consumption expected at the OA level (~ LV feeder) for these OAs on this day.

The results show that the OA with the higher number of households (Highest % single people) generally exhibits a lower total consumption profile illustrating the effect of larger households on consumption. Clear differences are observed in the profile in the late afternoon and evening periods potentially indicating different energy-using habits and routines.

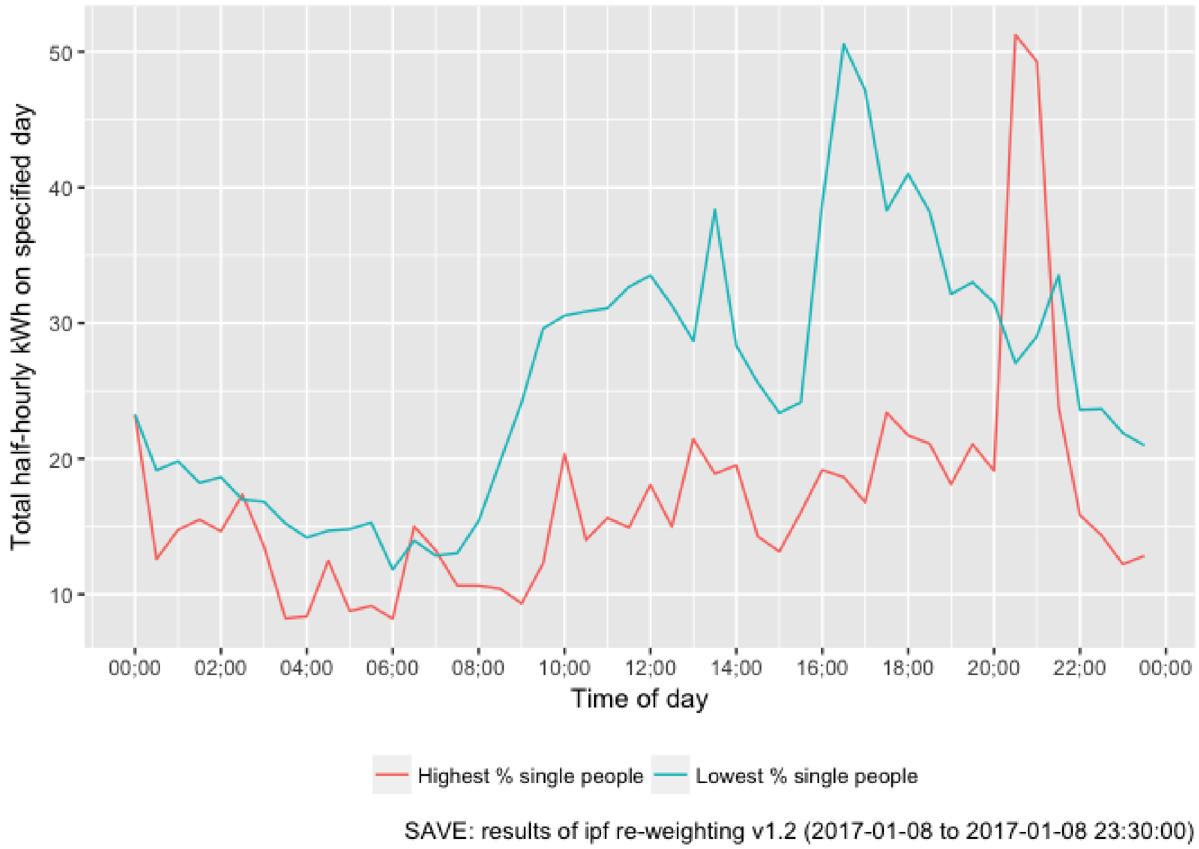

Figure 23 Simulated OA consumption profiles, baseline data, all trial groups

As noted above, the profile for the OA with the highest % of single person households is dominated by the single person households in the SAVE sample and at least some of them recorded high spikes in consumption between 20:00 and 22:00 on this day (Figure 23). In contrast, the OA with a more mixed household composition shows a smoother consumption profile for this day reflecting the daily profiles reported for the sample in Section 4.3.2.

The analysis is repeated by household size type and illustrated using a stacked area chart shown in Figure 24. The results show the contribution of each household type (indicated by coloured areas) to total consumption in a given OA. The results indicate that the consumption profile for the OA with the highest percentage of single-person households is, as expected, composed predominantly of consumption by the single-person households. In contrast, the profile for the OA with the lowest percentage of single-person households is dominated by the consumption of three, four and five-person households.

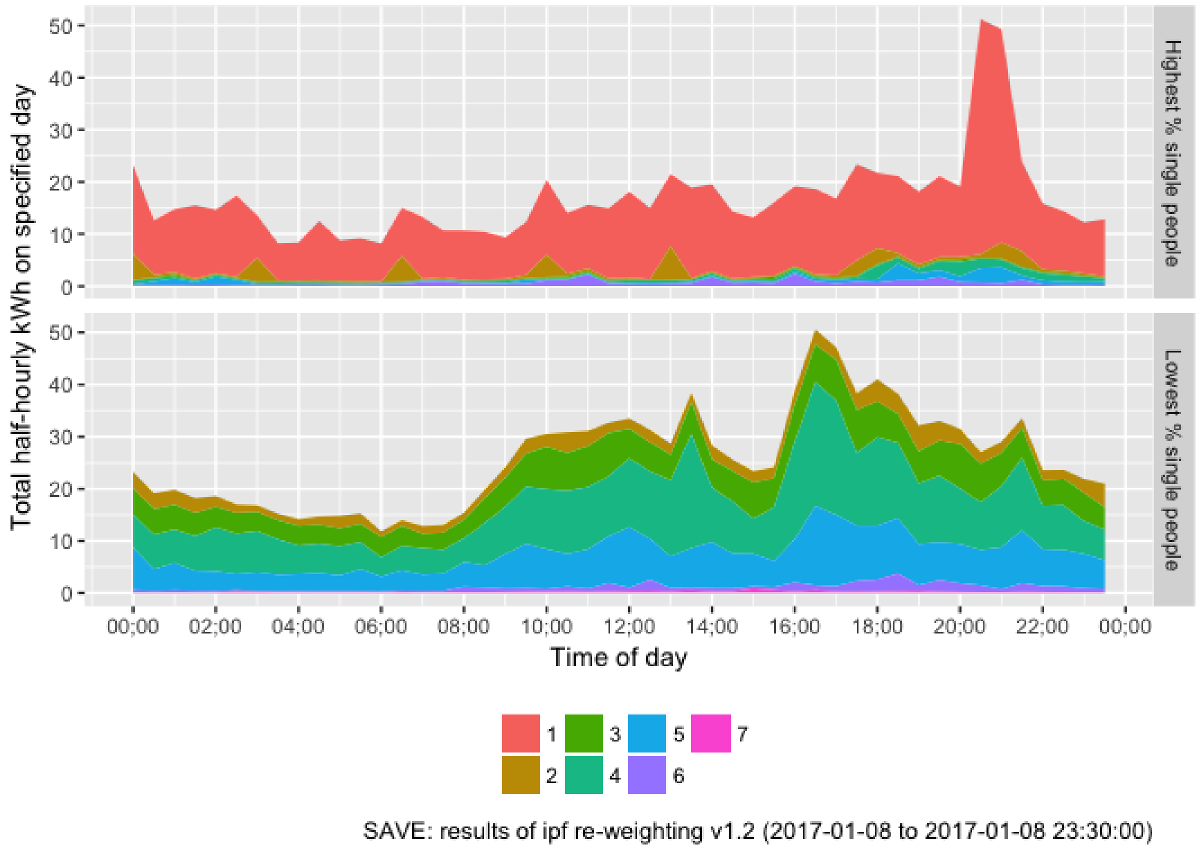

Figure 24 Simulated OA consumption profiles by household size (colours indicate number of people in household), baseline data, all groups

The analysis is repeated for the mean kWh for households by size (Figure 25). The colour of each line represents household size. The results show that single person households have lower average consumption compared to larger households. The mean profiles also show that they do not exhibit greater irregularity (i.e. large variation over short time periods) than the other household sizes, although the results do not reveal whether the mean is concealing any unusually high consuming households. The large peak in the profile for the OA with the highest % of single people is therefore, probably driven by the high weighting of a number of the SAVE single person households on this particular day (perhaps in combination with high consumption).

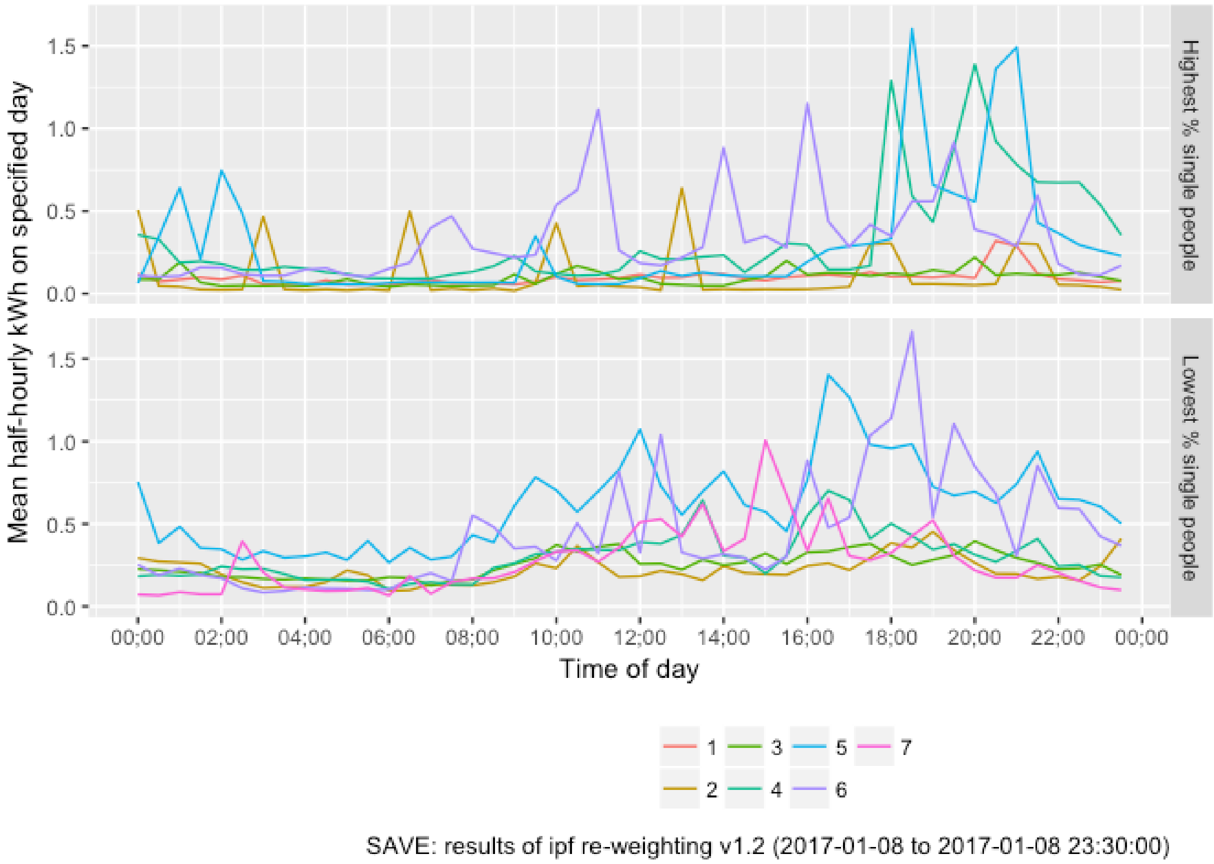

Figure 25 Simulated mean consumption profiles by household size (colours indicate number of people in household), baseline data, all groups

5.6.2 Model 1 (Trial Group 1: control)

The analysis as set out in the 'baseline' case above, is repeated here but using only those households in trial group 1, the control group, attaching kWh profile data to the SAVE sample households as before. However, in this modelling, consumption data for the March 15th event day and the two days either side is used. In contrast to the single-day results in the 'baseline' case discussed above, this therefore demonstrates the capability of the model to generate simulated profiles over a number of days.

It should be noted from the unweighted household counts (Table 22) that the selection of only those households in the control group results in a much smaller pool contributing consumption data to the simulation.

Table 22 Test unweighted household counts of n people (unweighted household counts – should match)

N people	Highest % single people	Lowest % single people
1	122	122
2	295	295
3	135	135
4	113	113
5	46	46
6	12	12
7	5	5
8+	2	2

Inspection of the maximum weights column (Table 23) reveals that the IPF process has generated some large weights but that the weighted household counts create the correct totals for each OA (Table 20).

Table 23 Summary household counts and weights by household size, model 1 (rounded to 4 dp)

OACode	label	HH Size	N un-weighted	Mean Weight	Min Weight	Max Weight	N weighted
E00115898	Lowest % single people	1	121	0.0000	0	0.0000	0
E00115898	Lowest % single people	2	288	0.0453	0	4.6013	13
E00115898	Lowest % single people	3	130	0.1778	0	14.5490	23
E00115898	Lowest % single people	4	108	0.3289	0	16.1930	35
E00115898	Lowest % single people	5	44	0.2539	0	10.4287	11
E00115898	Lowest % single people	6	11	0.1818	0	0.8193	2
E00115898	Lowest % single people	7	5	0.1964	0	0.7812	1
E00115898	Lowest % single people	8+	1	0.0000	0	0.0000	0
E00167003	Highest % single people	1	121	1.5369	0	100.3618	182
E00167003	Highest % single people	2	288	0.0349	0	3.8589	10
E00167003	Highest % single people	3	130	0.0155	0	1.1121	2
E00167003	Highest % single people	4	108	0.0189	0	1.9993	2
E00167003	Highest % single people	5	44	0.0000	0	0.0000	2
E00167003	Highest % single people	6	11	0.1818	0	1.2533	2
E00167003	Highest % single people	7	5	0.0000	0	0.0000	0
E00167003	Highest % single people	8+	1	0.0000	0	0.0000	0

As indicated before, there are 182 single person households in the OA with the highest % of single people. In this model, the consumption profile for these households will be based on the 121 unweighted single person households (see Table 23). It should be noted that at least one of those households received a very high weight in this OA (refer to 'Max Weight' column), so the profiles of this household type will therefore be dominated by these few households in the subsequent results for this OA.

The resulting profiles are presented in Figure 26. The simulated consumption profiles by day are presented by OA in separate panels for clarity.

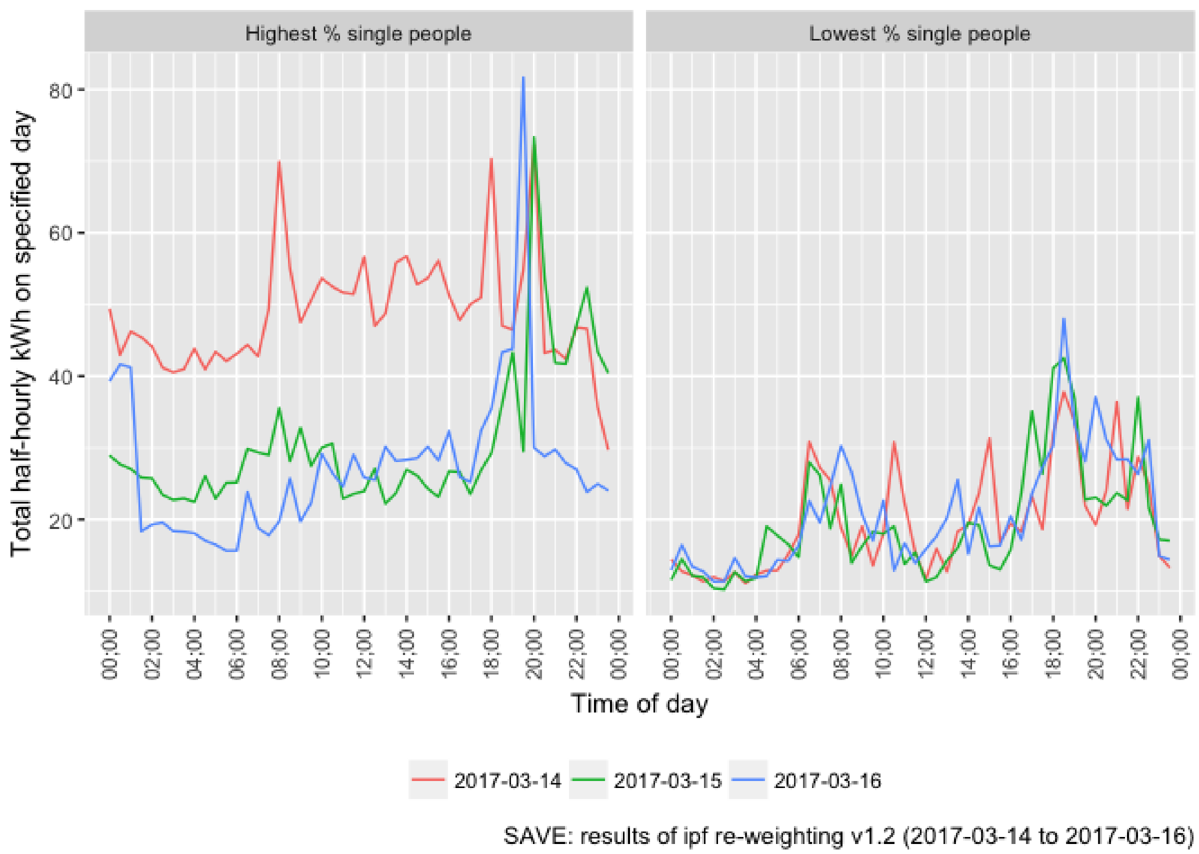

Figure 26 Simulated OA consumption profiles, TP1 event data, trial group 1 (control)

Clear differences are observed between days in the consumption profile for the OA with the highest percentage of single-person households (Figure 26, left-hand panel). In contrast, there is less variation by day in the OA with the lowest percentage of single-person households (Figure 26, right). The consumption profiles for both OAs also appear to show more variability at the shorter time-scale resulting in 'peakier' profiles.

The analysis is repeated by household size showing the results using a stacked area chart (Figure 27) to show the contribution of each household size and then show a mean kWh profile by household (Figure 28). It can be seen that the variability in the aggregate OA consumption profile does not appear to be caused by the variability in the mean profiles by household size. The variability would again seem to be caused by households with highly variable consumption in combination with high weighting. With the large weight (>100) given to a single household in the OA with the highest percentage of single-person households, it is likely that the day-to-day variation in the consumption profile of this household alone could cause the variation observed in the simulated OA level profile.

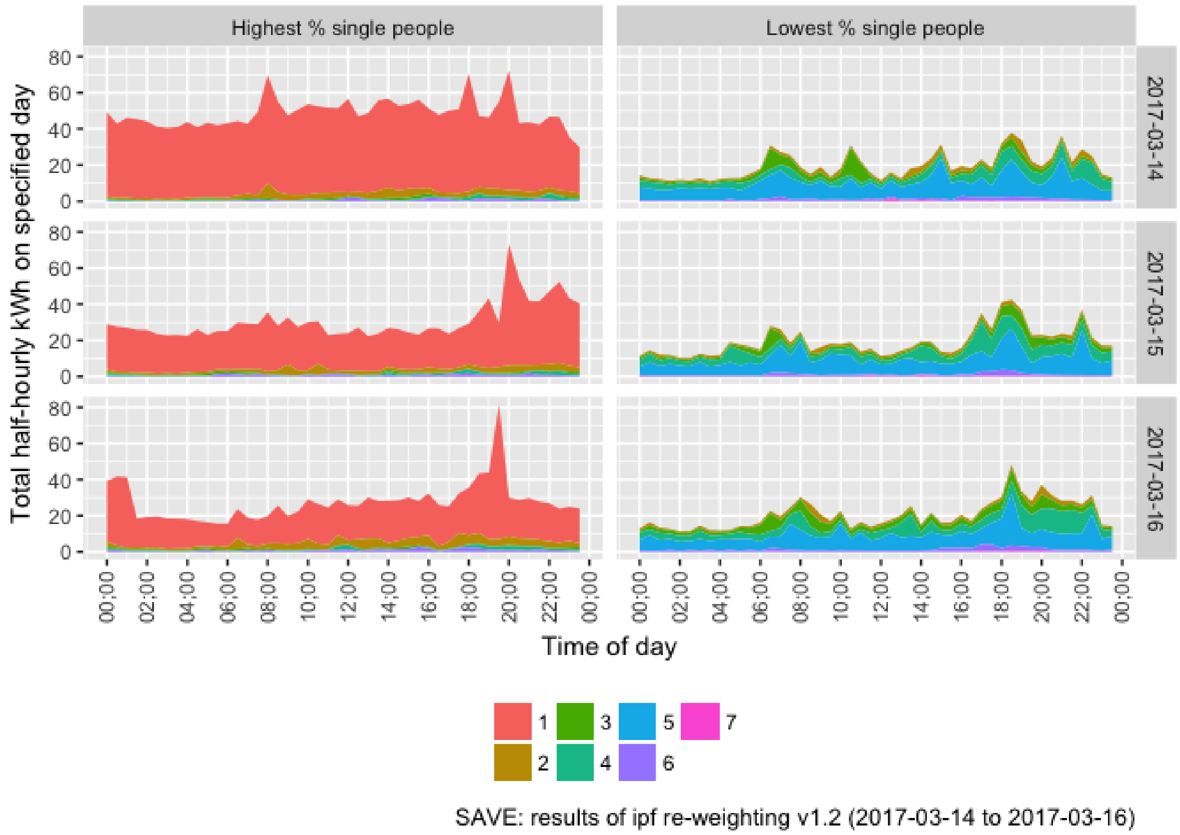

Figure 27 Simulated OA consumption profiles by household size (colours indicate number of people in household), TP1 event data, group 1

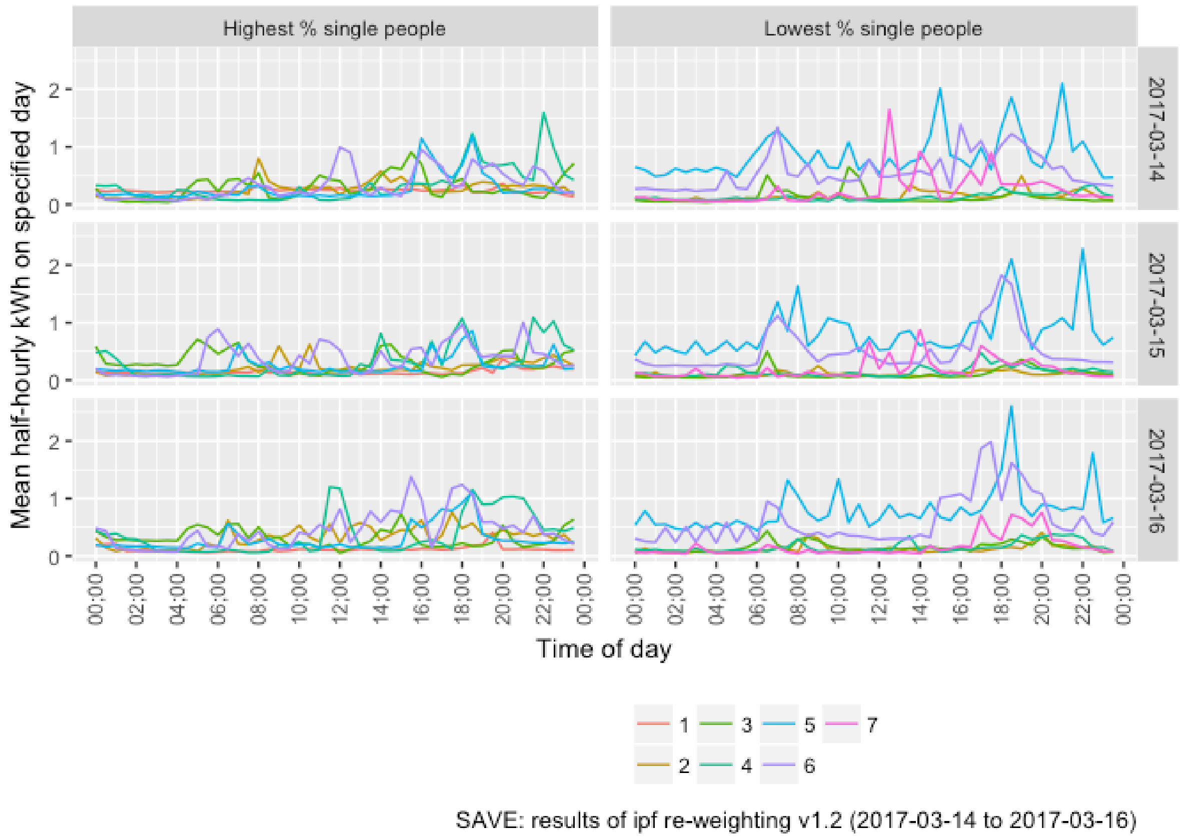

Figure 28 Simulated mean consumption profiles by household size (colours indicate number of people in household), TP1 event data, group 1

5.6.3 Model 2 (Trial Group 2)

In this section, the process is repeated using only households in trial group 2, those who received the online messaging and Loop event day notifications. The counts of contributing SAVE households by size are shown in Table 24.

Table 24 Test unweighted household counts of n people (unweighted household counts)

N people	Highest % single people	Lowest % single people
1	122	122
2	331	331
3	117	117
4	130	130
5	41	41
6	14	14
7	2	2
8+	2	2

Table 25 shows the unweighted and weighted household counts and IPF weights. The 'target' counts for each OA remain the same (Table 20).

Table 25 Summary household counts and weights by household size, model 2 (to 4 dp)

oaCode	label	HH Size	N un-weighted	Mean Weight	Min Weight	Max Weight	N weighted
E00115898	Lowest % single people	1	118	0.0000	0	0.0000	0
E00115898	Lowest % single people	2	324	0.0402	0	2.7259	13
E00115898	Lowest % single people	3	115	0.1936	0	11.0454	23
E00115898	Lowest % single people	4	127	0.2762	0	12.7502	35
E00115898	Lowest % single people	5	41	0.2713	0	10.6932	11
E00115898	Lowest % single people	6	13	0.1538	0	0.8510	2
E00115898	Lowest % single people	7	2	0.5000	0	1.0000	1
E00115898	Lowest % single people	8+	2	0.0000	0	0.0000	0
E00167003	Highest % single people	1	118	1.5439	0	144.3120	182
E00167003	Highest % single people	2	324	0.0310	0	7.8974	10
E00167003	Highest % single people	3	115	0.0175	0	2.0000	2
E00167003	Highest % single people	4	127	0.0158	0	2.0000	2
E00167003	Highest % single people	5	41	0.0494	0	1.5757	2
E00167003	Highest % single people	6	13	0.1538	0	1.4605	2
E00167003	Highest % single people	7	2	0.0000	0	0.0000	0
E00167003	Highest % single people	8+	2	0.0000	0	0.0000	0

In this treatment group, single-person households received an even higher maximum weighting in the OA with the highest percentage of single people.

The resulting consumption profiles illustrated below, show a very high peak in consumption in the OA with the highest percentage of single-person households occurring on the 16th March. This peak dominates the OA level consumption profile (Figure 29). However, lower consumption is observed between 18:00 and 20:00 on the event day itself (15th March) for this OA suggesting some success in modelling demand response on this day. This suggests that single person households may have responded more strongly to the incentive although this is not indicated in the results reported in Section 4.5.2 as household size was not included in the models.

Figure 30 clearly shows that, as expected for this OA, the single-person households contribute the main share of consumption, while Figure 31 clearly shows the variability that would be captured by the network model if mean consumption by household size were used as OA level inputs.

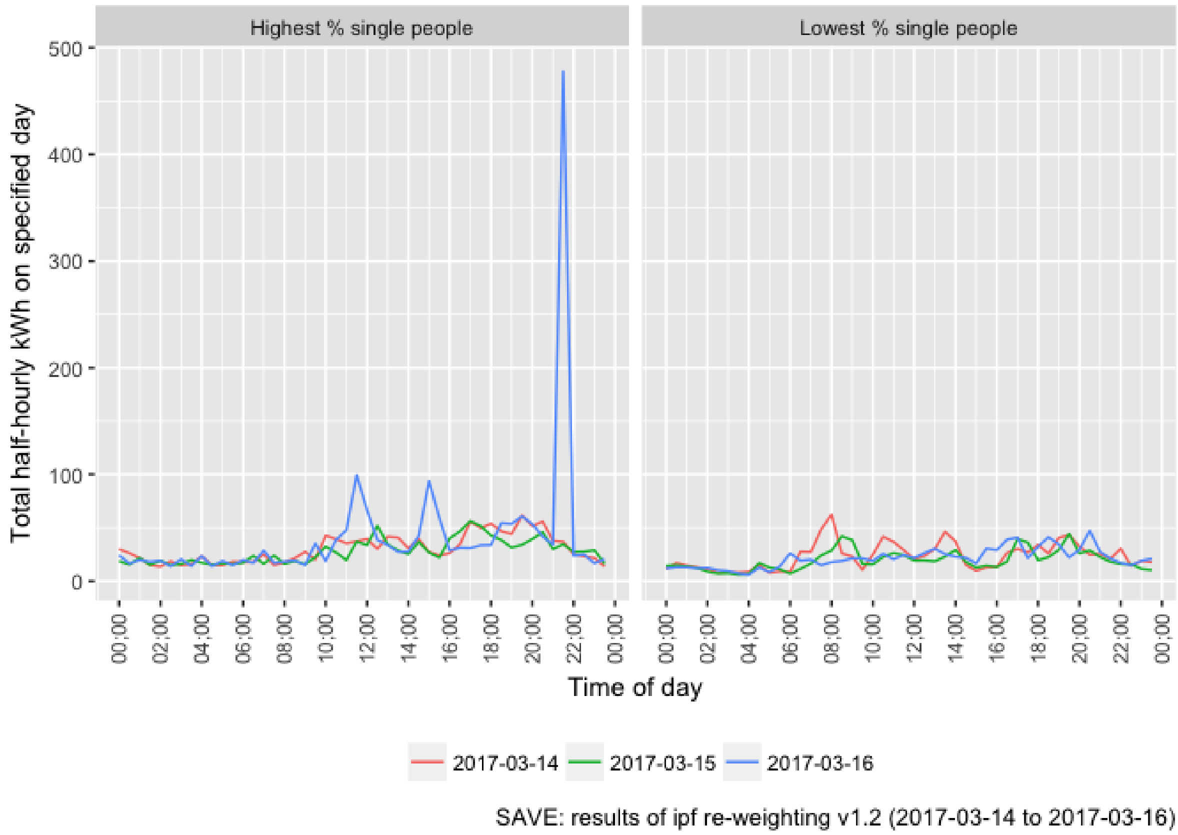

Figure 29 Simulated OA consumption profiles, TP1 event data, trial group 2

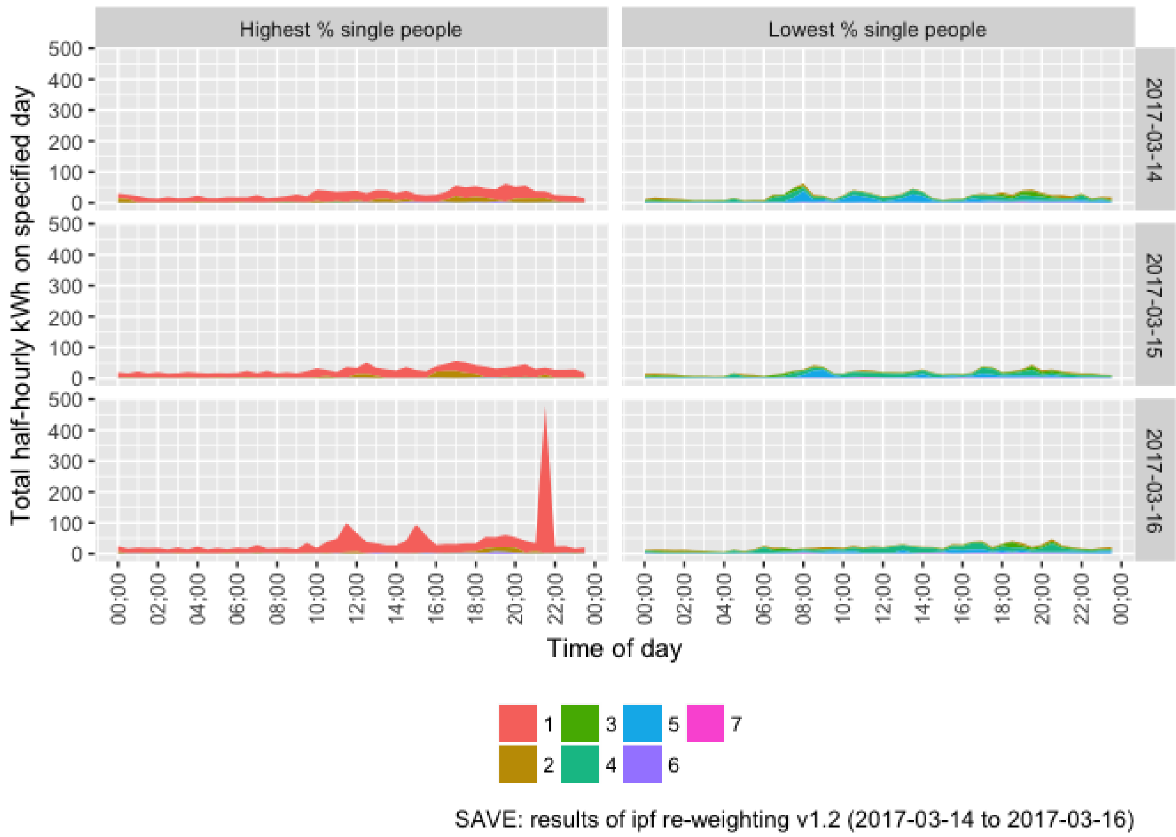

Figure 30 Simulated OA consumption profiles by household size (colours indicate number of people in household), TP1 event data, group 2

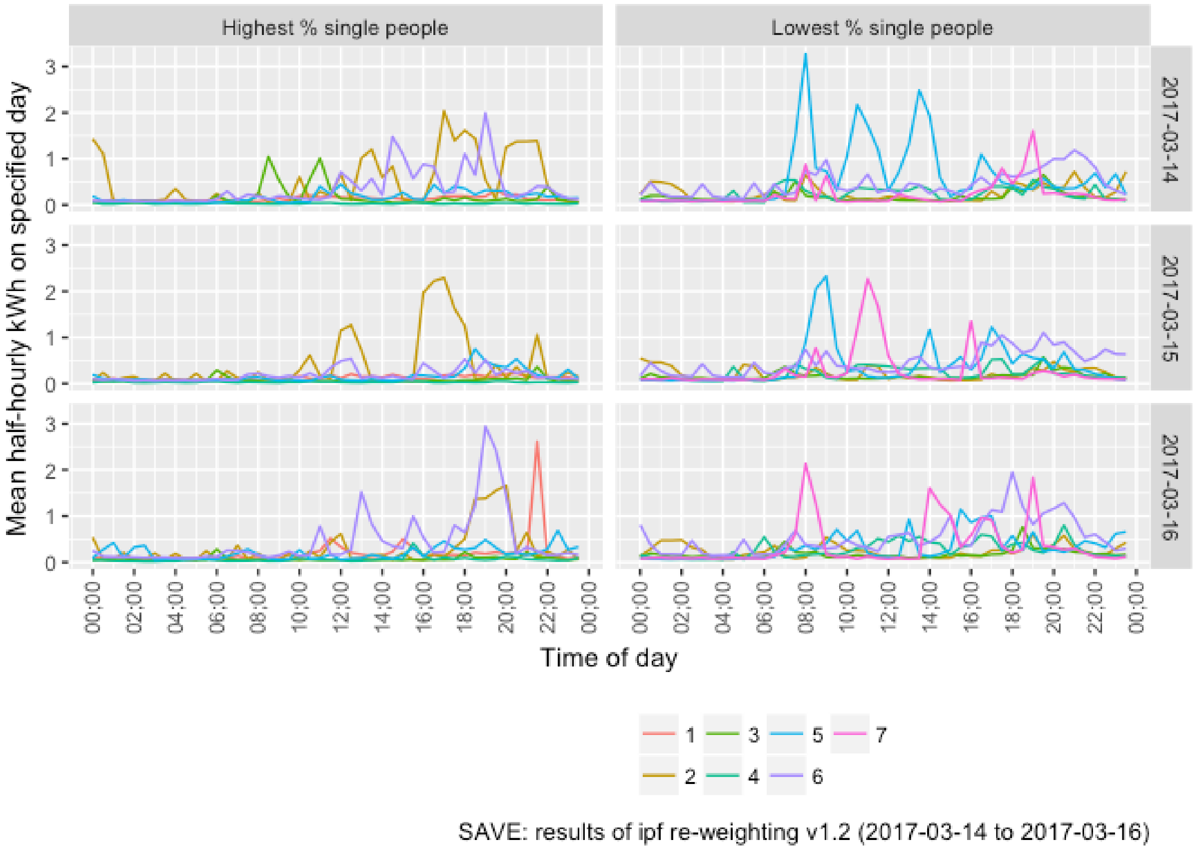

Figure 31 Simulated mean consumption profiles by household size (colours indicate number of people in household), TP1 event data, group 2

5.6.4 Model 3 (Trial Group 3)

In this section, the process is repeated once more using only households in trial group 3, those who received both online and postal messaging, the Loop notifications and a financial incentive. The contributing household counts are shown in Table 26, followed by a summary of weighting results in Table 27. Note that the maximum weight for single-person households in the OA with the highest percentage of single-people is again above 100.

Table 26 Test unweighted household counts of n people (unweighted household counts)

N people	Highest % single people	Lowest % single people
1	100	100
2	261	261
3	108	108
4	107	107
5	34	34
6	9	9
7	3	3
8+	1	1

Table 27 Summary household counts and weights by household size, model 3 (to 4 dp)

oaCode	label	HH Size	N un-weighted	Mean Weight	Min Weight	Max Weight	N weighted
E00115898	Lowest % single people	1	96	0.0000	0.0000	0.0000	0
E00115898	Lowest % single people	2	259	0.0496	0.0000	4.1159	13
E00115898	Lowest % single people	3	107	0.2154	0.0000	11.8231	23
E00115898	Lowest % single people	4	106	0.3302	0.0000	12.2560	35
E00115898	Lowest % single people	5	34	0.3235	0.0000	9.2198	11
E00115898	Lowest % single people	6	9	0.2222	0.0000	1.9631	2
E00115898	Lowest % single people	7	3	0.3333	0.0087	0.9738	1
E00115898	Lowest % single people	8+	1	0.0000	0.0000	0.0000	0
E00167003	Highest % single people	1	96	1.8960	0.0000	116.6331	182
E00167003	Highest % single people	2	259	0.0391	0.0000	3.7572	10
E00167003	Highest % single people	3	107	0.0187	0.0000	2.0000	2
E00167003	Highest % single people	4	106	0.0189	0.0000	1.2447	2
E00167003	Highest % single people	5	34	0.0588	0.0000	2.0000	2
E00167003	Highest % single people	6	9	0.2222	0.0000	1.9981	2
E00167003	Highest % single people	7	3	0.0000	0.0000	0.0000	0
E00167003	Highest % single people	8+	1	0.0000	0.0000	0.0000	0

Figure 32, Figure 33 and Figure 34 illustrate the results for this treatment group in the same fashion as earlier analysis. More pronounced lunchtime and evening peaks are observed in the OA with highest percentage of single-person households (left panels, all figures) and again, some indication of lower consumption during the 17:00 – 20:00 period on the event day (15th March) itself. Unlike the previous model, this is not only to be seen in the OA with the highest % of single person households but is also detectable, albeit at lower levels, in the more heterogeneous OA in the right-hand panel (all figures).

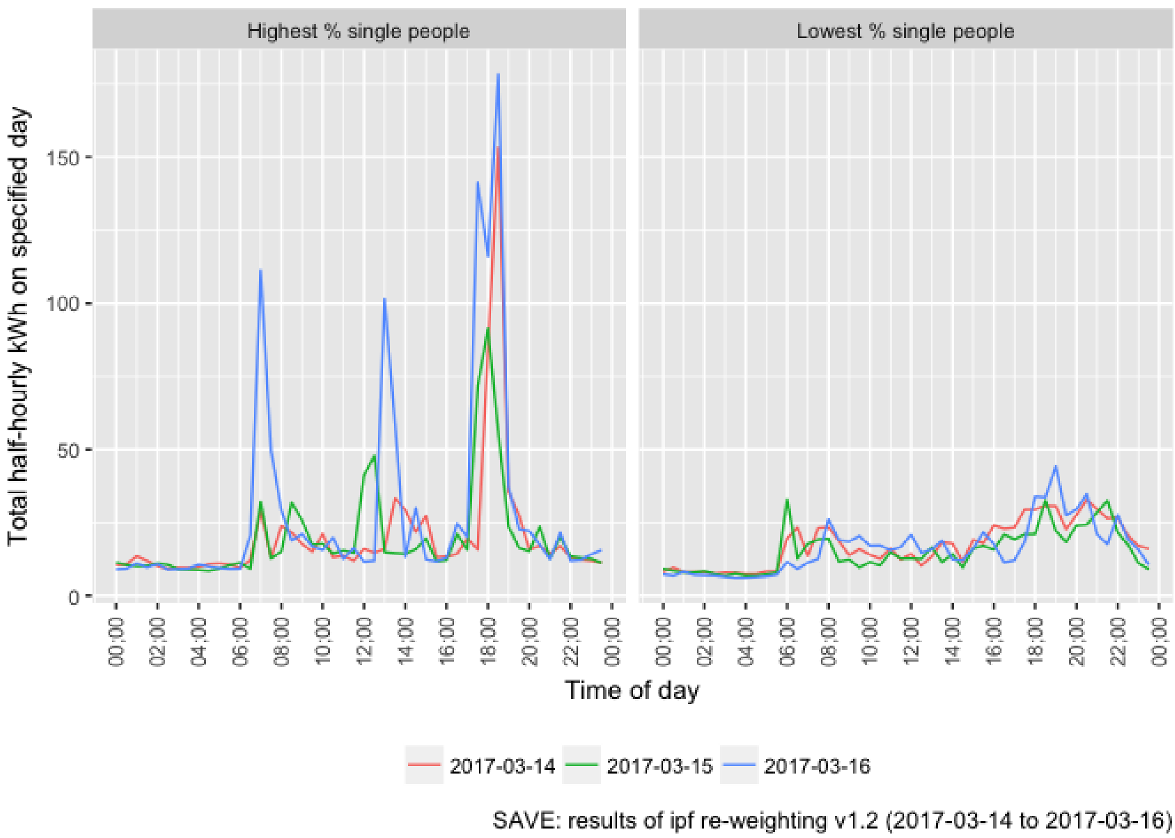

Figure 32 Simulated OA consumption profiles, TP1 event data , trial group 3

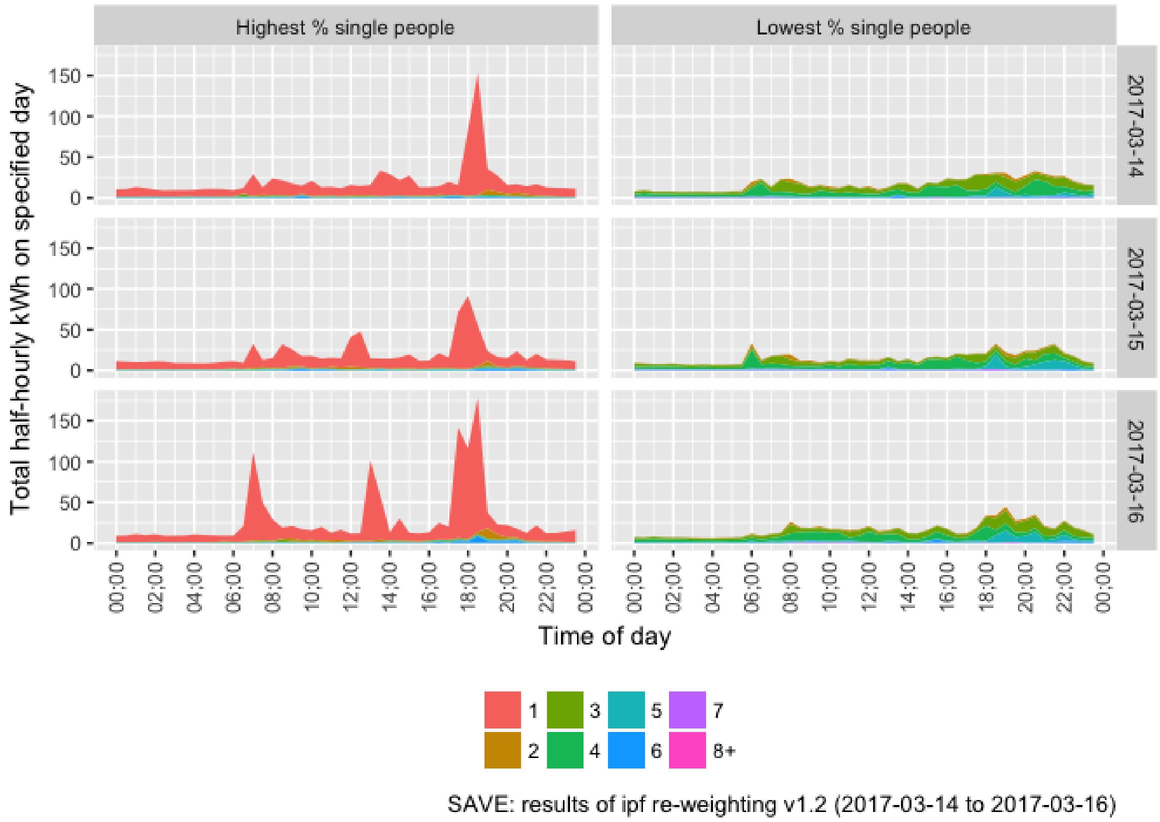

Figure 33 Simulated OA consumption profiles by household size (colours indicate number of people in household), TP1 event data, group 3

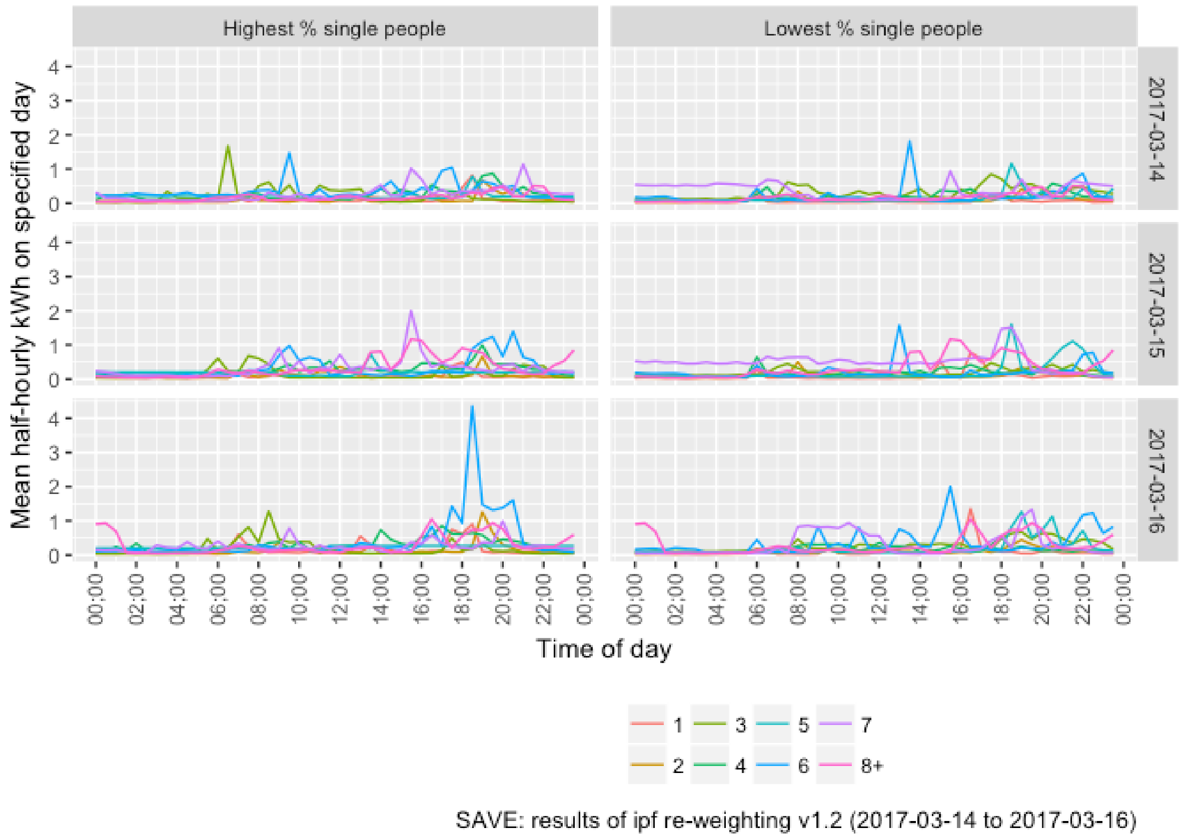

Figure 34 Simulated mean consumption profiles by household size (colours indicate number of people in household), TP1 event data, group 3

5.6.5 Comparison of models 1 to 3 results

In this final results section, the results of models 1 to 3 are combined to provide a comparison of the control and interventions groups in the selected OAs. Figure 35 and

Figure 36 compare the total half-hourly kWh for the three models (Control, and trial groups 2 and 3) for the days surrounding the event day in trial period 1.

The first chart (Figure 35) presents the data in the same way as the previous charts with each date as a coloured line within each OA/group panel and does not enforce constant vertical scales on each row panel (group) to attempt to allow differences to be revealed. This highlights differences between consumption before, during and after the event day for the same groups. Group 1 (control) shows little variation from one day to the next in the OA with the lowest percent of single people but an apparently high consumption level on the 14th March in the OA with the highest percent of single people households. Group 2 may have reduced consumption during 18:00 – 20:00 on the event day (2017-03-15) in the OA with the highest % of single people (left hand panel) although this is difficult to discern due to the spike at 22:00. Group 3 shows the clearer reduction on the event day in the same OA with some evidence of a similar reduction in the other OA (right hand panel).

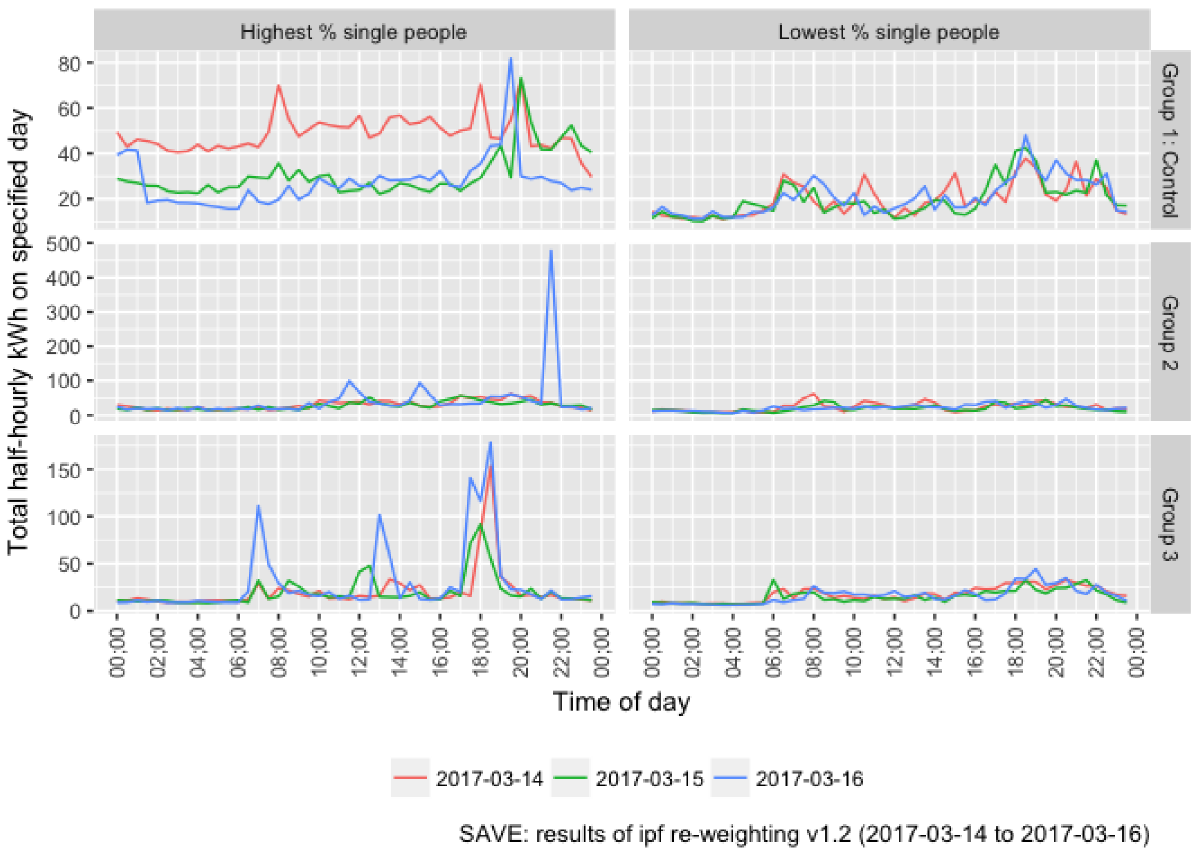

Figure 35 Simulated OA consumption profiles, TP1 event data, by day (indicated by line colour) within trial group

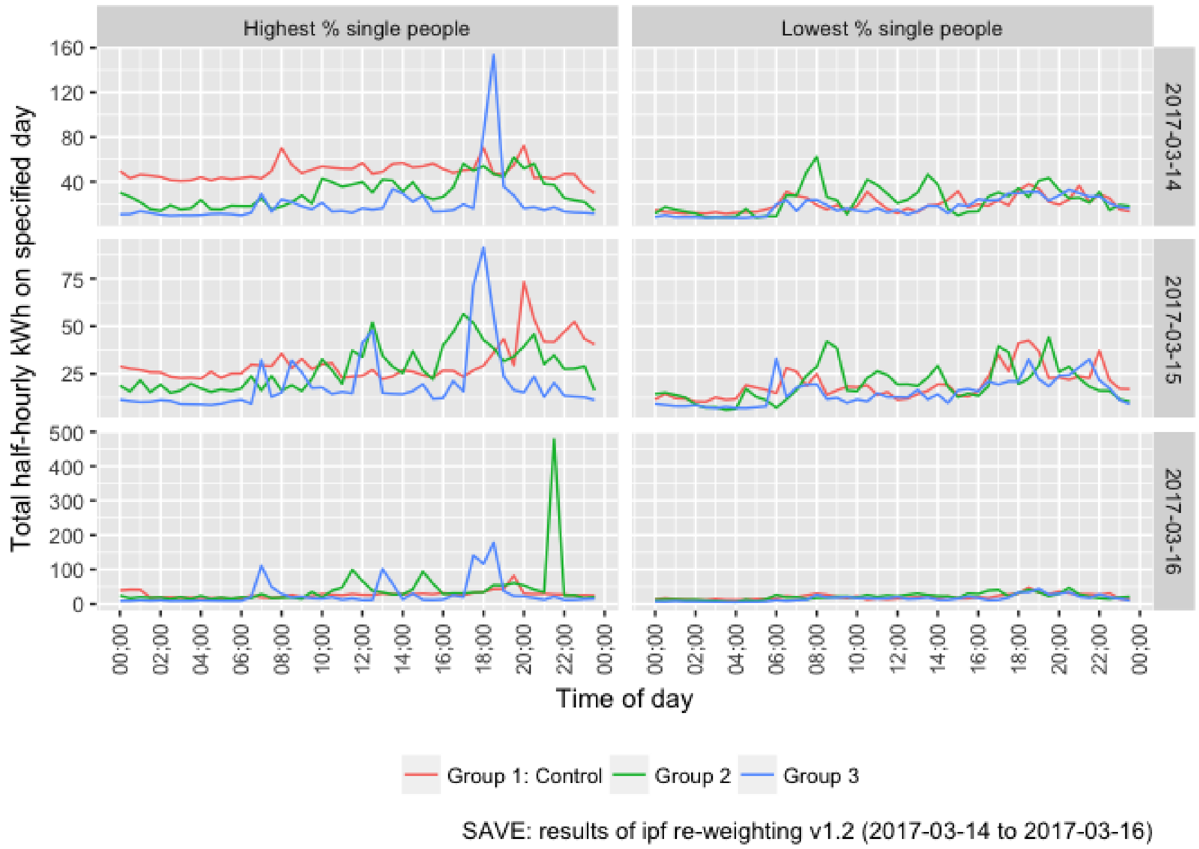

Figure 36 Simulated OA consumption profiles, TP1 event data, by trial group (indicated by line colour) within day

The second chart (

Figure 36) switches the presentation so that each group is a coloured line within each OA/group panel and also does not enforce constant vertical scales on each row panel (day) to attempt to allow differences to be revealed. This highlights differences in consumption between groups for the same day. In this case some evidence of higher consumption is seen just after the 16:00 intervention start on the event day (middle panel) followed by a steady decline for Group 2 in the highest % single person OA (left panel). Group 3 on the other hand shows a notable spike at 18:00 which may reflect the habits of a very few households in this group as it is essentially replicated on all three days. There is however some evidence of lower consumption from 19:00 onwards on the event day for both intervention groups in this OA and this effect appears larger for Group 3. In contrast, there is limited evidence of lower consumption in the intervention groups in the other OA during the 16:00 -20:00 intervention period.

However, comparison of the days or groups is clearly hindered by the large peaks observed in treatment groups 2 and 3. As already mentioned in the results sections for each group above, the high values of weights (possibly combined with particularly ‘peaky’ consumption in an individual household) has caused the magnification of these peaks. Comparison of the simulated OA profiles by treatment is therefore difficult prior to the resolution of this limitation of the small area estimation modelling.

5.7 Summary

Overall the baseline results indicate that estimated OA level mean consumption profiles may be useful for network planners as they can provide indications of probable consumption profiles in different small areas. This may allow planners to understand why certain areas show consumption peaks at different times to others due to different social composition and consequential habits and routines. They may also be able to indicate areas where certain technical innovations may be more beneficial – for example, areas with higher estimated mid-day consumption may be preferred areas for PV installations and could be suitably incentivised. Conversely, areas with lower than average mid-day consumption may require careful management to ensure local load balancing where PV installations are wide-spread. Similarly, the profiles could also indicate areas which are most at risk to evening peak EV charging impacts (in the absence of management) as they already have high evening load due to the commuting patterns of the residents.

The results for each intervention model extend this to enable network planners to assess the likely response to the interventions tested in different small areas. Thus future work could identify those areas where the model estimates greatest demand response as a guide to the potential targeting of future 'business as usual' interventions in the face of anticipated network events. Conversely the results could also be used to identify areas where no response would be expected, helping to avoid the costs of attempting to incentivise households who are unlikely to respond.

5.7.1 Input to network model: current status

The spatial microsimulation model development reported here underpins the delivery of the SAVE Network Model by providing aggregated 30-minute consumption data from the 2 trial and 1 control groups for each OA for an agreed time frame.

As described, this has currently been implemented for the pre/post and event day data for 14/15/16, March 2017 and comprises mean kWh per half hour for each trial group for each OA for these dates. As a result, if fed directly into the Network Model as this stage, the three models represent:

1. Baseline (Model 1): the estimated mean half hourly consumption profile for these dates with no event day intervention. This therefore comprises our estimate of 'normal' consumption on these dates for each of the 6,136 OAs in the study area;
2. Model 2: the estimated mean half hourly consumption profile for each OA for these dates had the intervention received by trial group 2 been implemented across all households in all OAs in the study region;
3. Model 3: the estimated mean half hourly consumption profile for each OA for these dates had the intervention received by trial group 3 been implemented across all households in all OAs in the study region.

5.7.2 IPF-based weighting method: review

The IPF method produces accurate aggregate household counts for the test OAs and produces consistent weighting across the trial groups. It should be noted that the single-person households in the OA with the highest percentage of single people receive a maximum weight of greater than 100 and also that error levels in the aggregate household counts (tested in the systematic error checks for each run of the IPF weighting) are higher where the process is run for each treatment group compared to using all households.

5.7.3 Simulated consumption profiles: issues

As expected, the models based on the trial groups are strongly affected by:

- a few households with 'odd' habits;
- High weights assigned to households, in combination with;
- Rare categories in certain OAs (in our case by household size)

In combination these issues can create a 'multiplier-effect' such that in extreme cases the model can assign a very high weighting to a single household, effectively duplicating the consumption profile of a single-household by 100 times (or more). While this single household may (or may not) represent the mean consumption profile of one instance of that household type well, the exact synchronicity in the timing of the simulated consumption of 100 different profiles does not reflect the diversity in the timing of demand found in real-world situations. The result, as illustrated in the test scenarios presented here, is unrealistically high peaks in some of the simulated OA level consumption profiles.

Strategies to address this issue are under consideration as part of future work but might include:

- excluding outlier households – although this risks reducing the realistic inherent heterogeneity;
- reducing the number of constraints used in the IPF – although this has potential to reduce the extent to which the (fewer) constraints can effectively model the socio-economic distribution of peak demand profiles;
- increasing the 'pool' size by substantially increasing the proportion of households who have completed recruitment surveys to as close to 100% as possible.

6 Conclusions

6.1 Trial period 1 evaluation

During trial period 1, a clear trend is observed showing a reduction in mean electricity consumption consistent with changes in usage due to changing environmental conditions (warmer ambient temperatures, longer daylight hours). Disaggregating the consumption data by main heat source and modelling the week-to-week change in consumption with Met Office temperature data revealed that the rate of change in electricity consumption in response to changing ambient temperature is associated with main heating source. It has been illustrated that household and household representative person attributes such as employment type and main heat source are associated with differing daily consumption profiles. Disaggregated by day-of-the-week, a significantly earlier and higher peak in electricity consumption is observed on Sundays during January and February.

In terms of the impact of the interventions trialled, the analysis was conducted in two stages: (1) analysis of mean weekly electricity consumption to assess the impact of the 4-8 messaging treatments applied to each trial group; and (2) analysis of mean 15-minute electricity consumption during the peak period (and pre- and post-peak) which is geared to assess the response of household in each treatment group to the 'event day' intervention. Both stages of the analysis use regression models to assess factors associated with consumption and the net treatment effects attributable to the interventions.

Overall, the analysis of the 4-8 messaging shows that membership of one of the treatment groups does not predict significant differences in consumption over the trial period. The regression modelling carried out indicates that over the trial period, some household attributes have consistent and significant interaction effects with the treatment groups, and have effects with coefficients opposite to the hypothesis of reduced consumption in treatment groups (i.e. higher consumption). Of note are the following:

- households living in flats interaction with *postal only,*
- electric storage heating interaction with *online and postal.*

The results also reveal household attributes that, as expected, are associated with significantly higher and lower levels of consumption, these are:

- larger households (number of people) -> higher consumption
- Heat source 'other electric' -> higher consumption
- Heat source 'oil/wood/solid fuel' -> higher consumption
- Dwelling type 'flat/other' -> lower consumption
- Higher eco-score -> lower consumption

In analysing the impact of the event day, regression models were run for three periods, *pre-peak*, *peak* and *post peak* time periods with two of the intervention groups:

- Group 2 – online + loop notification
- Group 3 – online and postal + loop notification + £ incentive

Generally, no statistically significant effects were found in any of the three periods related to membership of the treatment groups. The regression model results confirm statistically significant effects on overall consumption from household attributes including eco-mean score and presence of children, but neither of these has significant interactions with treatment groups.

However, the results do show some numerical differences between treatment groups consistent with the hypothesis of expected treatment effects. In the pre-peak period, Group 2 have lower consumption compared to the control group in the pre-peak period. Group 3 have higher consumption, consistent with the hypothesis of increased consumption in the *pre-peak* period. Households in both groups that opened the event notification email had higher consumption than those that didn't.

Both treatment groups show lower consumption than the control group during *peak* hours on the event day, with a larger difference in the 18:00 to 20:00 period, suggesting that if any treatment effect is present, it is largest in this period.

In the *post-peak* period, both groups have higher consumption than the control group. However, as stated above none of these results are statistically significant.

The time-use survey data contributes to the event day analysis. Overall, no statistically significant effects were found relating to the total number of energy-using acts, or acts related to a number of specific activities and appliances. The results are consistent however with the hypothesis that treatment groups would report more energy-using acts than the control group in the *pre-peak* and *post-peak* periods, and fewer in the *peak* period. This is apparent for group 3, particularly for households who opened the event day email notification. In addition, statistically significant results were found in trial group 3 to support the claim that households avoid household energy consumption by being away from home and thus reducing evening peak electricity consumption across the board.

6.1.1 Contribution to Network Model

The spatial microsimulation model development reported here underpins the delivery of the SAVE Network Model by providing aggregated 30-minute consumption data from the 2 trial and 1 control groups for each OA for an agreed time frame.

Currently this has been implemented for the pre/post and event day data for 14/15/16 March 2017 and comprises mean kWh per half hour for each trial group for each OA for these dates.

As a result, if fed directly into the Network Model as this stage, the three models represent:

1. Baseline (Model 1): the estimated mean half hourly consumption profile for these dates with no event day intervention. This therefore comprises our estimate of 'normal' consumption on these dates for each of the 6,136 OAs in the study area;
2. Model 2: the estimated mean half hourly consumption profile for each OA for these dates had the intervention received by trial group 2 been implemented across all households in all OAs in the study region;
3. Model 3: the estimated mean half hourly consumption profile for each OA for these dates had the intervention received by trial group 3 been implemented across all households in all OAs in the study region.

It should be noted that at present these profiles are mean values across all households within an OA. SSEN have expressed an interest in calculating such profiles as means for a small number of household categories ('customer types') within each OA which can then be allocated to specific network nodes using data to which a network modeller may have 'business as usual' access. These categories should use those variables which provide the strongest differentiators between different levels of evening peak consumption in order to provide non-random heterogeneity. Based on the analysis of the best constraint variables to use (see Section 5.4, Table 12), this implies using categories based on the number of occupants, the number of rooms, the number of cars and HRP ethnicity. However cross-classifying this many variables will lead to extremely small and possibly zero-count cells in many OAs. As a result, it is recommended that, at most, the number of occupants and the number of rooms be used, both of which may be relatively easy to determine from other data that a network modeller may have to hand.

6.2 Small area estimation

It has been demonstrated that the inherent variability of household electricity demand over short time-scales (hours/days), and between households that have similar characteristics, raises challenges for the small area estimation method. This is particularly the case where it is applied to areas where there is a lower diversity of households such that a few households' observed profiles will dominate the model. This has been illustrated with the use of the OAs with the highest and lowest percentage of single-person households. In these areas, it is observed that the inherent temporal variability of consumption is compounded by the high weighting values applied by the IPF weighting process.

The loss of the 'load diversity' effect – where asynchronous variability in individual household consumption results in smoothing of the aggregate profile – is magnified through the small area estimation method. This magnification may be exacerbated by the use of small areas (OAs vs LSOAs) which further reduces the number of households in the SAVE 'pool' contributing consumption profile data.

Nevertheless, the results suggest that the modelling approach appears to be able to capture intervention effects with small but observable differences between OA level consumption profiles for the two extreme OAs tested on the pre/post and event days. This is particularly visible in the OA with the highest % of single person households and thus with the highest household homogeneity and also, potentially, with the greatest proportion of households whose 'demand response' actions would disrupt the routine of only 1 person.

While the variability in the consumption data itself represents the real-world case, the 'multiplier-effect' of the small area estimation observed in some cases can produce artificially high peaks in simulated OA level profiles.

However, these limitations are expected to pose a problem in a relatively few cases, and further work will be required to fully evaluate the extent and effect of the issues encountered in this sample. A more detailed examination of this issue will be undertaken for the next iteration of the SAVE customer model (SDRC 2.3).

The reduction in size of output area from LSOA level to OA level to better represent the LV network may be a contributing factor to the problem highlighted above due to a potential (untested) associated reduction in diversity (in terms of household characteristics). This may have resulted in a reduction of the sample (pool) for and/or an increase in weights of some household types in certain areas and will require further testing.

6.3 Further work

Future work to be reported in SDRC 2.3 will examine, and where possible, address the limitations identified in the small area estimation modelling. Development of the customer model will be carried out with ongoing attention to its compatibility with the network model.

SDRC 2.3 will also include analysis of TP2 & TP3 in similar format to SDRC 2.2 and update small area estimates of demand reduction under different interventions.

7 References

AECOM, 2011. Energy Demand Research Project: Final Analysis. AECOM, St Albans.

Anderson, B., 2014. SDRC 2.1: SAVE Customer Model Framework Specification, Scottish and Southern Energy Power Distribution's Low Carbon Networks Fund SAVE project SDRC. University of Southampton, Southampton.

Anderson, B., 2013. Estimating Small Area Income Deprivation: An Iterative Proportional Fitting Approach, in: Tanton, R., Edwards, K. (Eds.), Spatial Microsimulation: A Reference Guide for Users. Springer, London.

Anderson, B., Lin, S., Newing, A., Bahaj, A., James, P., 2016. Electricity consumption and household characteristics: Implications for census-taking in a smart metered future. Comput. Environ. Urban Syst. https://doi.org/10.1016/j.compenvurbsys.2016.06.003

Anderson, B., Rushby, T., 2017. Process Metoffice WOW data for the SAVE study region (Analysis report), Scottish and Southern Energy Power Distribution's Low Carbon Networks Fund SAVE project SDRC. University of Southampton, Southampton.

Balta-Ozkan, N., Davidson, R., Bicket, M., Whitmarsh, L., 2013. Social barriers to the adoption of smart homes. Energy Policy 63, 363–374. https://doi.org/10.1016/j.enpol.2013.08.043

Bibby, P., Brindley, P., 2013. Urban and Rural Area Definitions for Policy Purposes in England and Wales: Methodology.

Birkin, M., Clarke, M., 2011. Spatial Microsimulation Models: A Review and a Glimpse into the Future, in: Stillwell, J., Clarke, M. (Eds.), Population Dynamics and Projection Methods. Springer, London.

Birkin, M., Wu, B., 2012. A review of microsimulation and hybrid agent-based approaches. Agent-Based Models Geogr. Syst.

Bradley, P., Leach, M., Torriti, J., 2013. A review of the costs and benefits of demand response for electricity in the UK. Energy Policy 52, 312–327. https://doi.org/10.1016/j.enpol.2012.09.039

Carroll, J., Lyons, S., Denny, E., 2014. Reducing household electricity demand through smart metering: The role of improved information about energy saving. Energy Econ. 45, 234–243. https://doi.org/10.1016/j.eneco.2014.07.007

CER, 2012. Smart Meter Electricity Consumer Behaviour Trial data. Irish Social Science Data Archive, Dublin.

CER, 2011. Electricity Smart Metering Customer Behaviour Trials (CBT) Findings Report (Information Paper No. CER11080a). Dublin.

Delmas, M.A., Fischlein, M., Asensio, O.I., 2013. Information strategies and energy conservation behavior: A meta-analysis of experimental studies from 1975 to 2012. Energy Policy 61, 729–739. https://doi.org/10.1016/j.enpol.2013.05.109

Di Cosmo, V., O'Hora, D., 2017. Nudging electricity consumption using TOU pricing and feedback: evidence from Irish households. J. Econ. Psychol. 61, 1–14. https://doi.org/10.1016/j.joep.2017.03.005

Frederiks, E.R., Stenner, K., Hobman, E.V., Fischle, M., 2016. Evaluating energy behavior change programs using randomized controlled trials: Best practice guidelines for policymakers. Energy Res. Soc. Sci. 22, 147–164. https://doi.org/10.1016/j.erss.2016.08.020

Herter, K., McAuliffe, P., Rosenfeld, A., 2007. An exploratory analysis of California residential customer response to critical peak pricing of electricity. Energy 32, 25–34. https://doi.org/10.1016/j.energy.2006.01.014

Institute for Social anbd Economic Research, 2015. Understanding Society: Waves 1-5, 2009-2014 (SN: 6614 No. 7th Edition). University of Essex.

Lovelace, R., Dumont, M., 2017. Spatial Microsimulation with R. CRC Press.

Powells, G., Bulkeley, H., Bell, S., Judson, E., 2014. Peak electricity demand and the flexibility of everyday life. Geoforum 55, 43–52. https://doi.org/10.1016/j.geoforum.2014.04.014

Pudjianto, D., Djapic, P., Aunedi, M., Gan, C.K., Strbac, G., Huang, S., Infield, D., 2013. Smart control for minimizing distribution network reinforcement cost due to electrification. Energy Policy, Special Section: Transition Pathways to a Low Carbon Economy 52, 76–84. https://doi.org/10.1016/j.enpol.2012.05.021

Schofield, J., Carmichael, R., Tindemans, S., Woolf, M., Bilton, M., Strbac, G., 2015. Experimental validation of residential consumer responsiveness to dynamic time-of-use pricing, in: 23 International Conference on Electricity Distribution.

Schofield, J., Carmichael, R., Tindemans, S., Woolf, M., Bilton, M., Strbac, G., 2014. Residential consumer responsiveness to time-varying pricing. Rep. A3 "Low Carbon London" LCNF Proj. Imp. Coll. Lond.

Scottish and Southern Energy Power Distribution, 2015. Change Request 001. Optimising Project Learning Through Trial Iteration Adjustment and Household Monitor Equipment Change.

Smith, T., Noble, M., Noble, S., Wright, G., McLennan, D., Plunkett, E., 2015. The English Indices of Deprivation 2015 (Technical Report). Department for Communities and Local Government, London.

Teng, F., Aunedi, M., Strbac, G., 2016. Benefits of flexibility from smart electrified transportation and heating in the future UK electricity system. Appl. Energy 167, 420–431. https://doi.org/10.1016/j.apenergy.2015.10.028

Thomson, M., Infield, D.G., 2007. Impact of widespread photovoltaics generation on distribution systems. IET Renew. Power Gener. 1, 33. https://doi.org/10.1049/iet-rpg:20060009

Torriti, J., 2017. Understanding the timing of energy demand through time use data: Time of the day dependence of social practices. Energy Res. Soc. Sci. 25, 37–47. https://doi.org/10.1016/j.erss.2016.12.004

Walker, G., 2014. The dynamics of energy demand: Change, rhythm and synchronicity. Energy Res. Soc. Sci. 1, 49–55. https://doi.org/10.1016/j.erss.2014.03.012

Watson, J.D., Watson, N.R., Santos-Martin, D., Wood, A.R., Lemon, S., Miller, A.J., 2016. Impact of solar photovoltaics on the low-voltage distribution network in New Zealand. IET Gener. Transm. Distrib. 10, 1–9.

8 Appendices

The appendices provided here are to be read in conjunction with the main body of the report. Each sub-section of the appendix is referenced within the main text and should not be read independently. Sub-sections should not be read sequentially.

A.1 Designing the SAVE approach

Drawing on the broad literature, Frederiks et al. recommend the use of a framework setting out key dimensions to be considered in the design of behavioural interventions.

Table 28: Assessing behavioural interventions: best practice framework (after Frederiks et al., 2016)

Feature	Recommendation
Formulate Hypotheses	Clearly specify the expected effect of the interventions on behaviour including their magnitude, direction and nature.
Program Design	Plan a sample size sufficient to give the statistical power required to test the hypotheses (to enable robust conclusions); Draw a random & representative sample (to enable generalisation) of the population of interest without self-selection (to avoid bias); Use a randomized control trial design wherever possible in order to be able to compare intervention with non-intervention groups; Randomly allocate participants to control or trial groups without self-selection (to avoid bias);
Methodology	Define and assess sample 'representativeness'; Collect baseline data on key socio-economic and demographic attributes to assess sample 'representativeness'; Establish that control and intervention groups are equivalent in key respects prior to interventions.

A.1.1 Sampling and recruitment

The stratified random address-based approach used Census Output Areas (COAs) as the basic building block and the Postcode Address File (PAF) as a source of addresses within COAs. Following stratification by Index of Multiple Deprivation 2015 (Smith et al., 2015) and Rural Urban Classification 2011 (Bibby and Brindley, 2013), a total of 1,165 COAs were then randomly selected from each stratum proportionate to the number of addresses accounted for by each.

In each of the selected COAs a random sample of up to 50 addresses[20] was then randomly selected to give an initial sample of 58,233 addresses. This

[20] Three COAs had fewer than 50 addresses

sample was then screened against commercially available databases (e.g. Acorn, via UKChanges) to remove multi-occupancy households, flats/tower blocks and student households wherever possible. Each of the resulting 42,470 addresses was then randomly allocated to one of the four sample groups, and allocated a unique anonymised household ID, which links all data collected from the household.

It should be noted that, while attempts were made to remove flats from the selected sample, where a flat was sampled, it was included within the sample where it was possible to undertake an installation such as flats in converted houses or some smaller blocks.

The details of households that agreed were collected via Computer Aided Personal Interview (CAPI) and the monitoring equipment was then installed. This consisted of:

- A meter clamp which was fitted around the live electricity cable coming from the meter box, and attached to a battery-powered electricity monitor;
- A mains-powered gateway which transmits the usage information captured via the monitor via Ethernet cable fitted to the household's broadband router (where possible) or via SIM card enabled monitor (where not possible to use the household's broadband).

As the installation process was expected to take up to 30 minutes, householders were then re-contacted either by email or telephone and invited to complete an initial recruitment survey via Computer Aided Web Interview (CAWI) or Computer Aided Telephone Interview (CATI). The average length of the CATI survey was 33 minutes.

A.1.2 Statistical power and hypothesis

Statistical power calculations enable the investigator to calculate (or estimate) the sample size that is needed to robustly detect a specific difference between two samples within given confidence levels and risk of a false positive (Type I error) or false negative (Type II error) result. Without such prior calculations a study may recruit too few participants to be able to robustly detect the hypothesised intervention effect (Delmas et al., 2013).

Calculation of the required sample size for the control and intervention group(s) requires the estimation of the likely effect size and so derives directly from the study hypotheses. In the case of DSR interventions for example, the project needs to detect a given % reduction in energy demand or consumption in a given time period and reasonable estimates of the reduction that can be derived from previous studies or analysis of readily available datasets.

The sample size calculation also requires an agreement on the significance level of the statistical tests to be used. By convention this is normally set to 0.05 (5%) and represents the probability of concluding that there is a real effect when in fact there was not (Type I error). In commercial or policy terms, this represents the risk of taking a potentially expensive action that will have no effect at all.

Finally, the calculation also requires agreement on the statistical power to be used. By convention this is normally set to 0.8 (80%) and represents the risk of concluding that there was no trial effect when in fact there was (Type II error). From a commercial or policy point of view, this represents the risk of taking no action when an action would probably have had the observed affect.

Given these values the required sample size can be calculated if using an estimate of the mean and standard deviation of the outcome to be measured. In most cases this will need to be estimated from previous studies or, more rarely, can be estimated from initial data collected on a pilot sample.

A.1.3 Establishing the required sample size

The CER Irish Smart Meter Trial residential household pre-intervention data (CER, 2012) was used to estimate the mean and standard deviation of total residential electricity consumption on weekday evenings (16–00 - 20:00) in December 2009 before the CER Trial interventions started.

Overall this data comprises 6,088,225 observations from 4,225 households. Non-zero readings were selected for the period 16–00 - 20:00 and any households with less than 30 days in the data were removed, as they were likely to have a lower total consumption in the month due to the lower number of observations. Having summed the total evening consumption per household per month, further analysis (not shown) suggested that to avoid a strong skewing effect, those values that lie between the 1% (21.667 kWh) and 99% (685.768 kWh) percentiles of the distribution should be selected. This produced a sample of 4,115 households with mean monthly evening consumption of 254.55 kWh and standard deviation of 131.73 kWh.

The sample size required to detect a range of effect sizes, given the usually accepted p value of 0.05 and power of 0.8, was calculated. Note that a one-sided test (a test for an increase or a decrease) rather than a two-sided test (a test just for difference without specifying the 'direction') was chosen, since the project seeks specifically to detect a decrease in consumption.

Table 29: Relationship between sample size and detectable effect size [power = 0.8, Irish CER trial Residential Sample, December 2009]

Group size	Detectable % effect (p = 0.05)
200	12.89
400	9.11
600	7.43
800	6.44
1000	5.76
1200	5.25
1400	4.86
1500	4.70

The results (Table 29) suggest that 1400 households are required per trial group to detect an effect size of 4.9% and 1000 to detect an effect size of 5.76%. These outcomes derived from the CER data suggests that a SAVE trial group sample size of 1200 with p = 0.05 (5%) and power = 0.8, would be able to robustly detect a reduction in consumption of 5.25% or higher. Based on the review of the empirical evidence and the intention to recruit a representative rather than self-selecting and thus more engaged and responsive sample, the effect size of 6% formed the basis for the sample size targets.

A.1.4 Sample outcomes and response rates

Table 30 contains the full results from the logistic regression models for non-response and refusal to take part in the SAVE study.

Table 30 Logistic models estimating probability of non-response and refusal

	Non-response	Refusal
% in LSOA with electric central heating[21]	2.002*** (1.342, 2.661)	-0.949* (-1.716, -0.181)
Number of gas meters in LSOA[22]	0.0001 (-0.0004, 0.001)	-0.0002 (-0.001, 0.0002)
Mean annual gas (kWh in LSOA)	-0.0001*** (-0.0001, -0.0001)	-0.0001*** (-0.0001, -0.00004)
Mean annual electricity (kWh) in LSOA	0.0001 (-0.00000, 0.0003)	0.0002** (0.0001, 0.0004)
% households in fuel poverty in LSOA[23]	0.021 (-0.001, 0.042)	0.033*** (0.015, 0.051)
IMD 2 (IMD 1 – least deprived)[24]	0.346** (0.115, 0.578)	-0.121 (-0.321, 0.079)
IMD 3	0.366** (0.106, 0.626)	-0.524*** (-0.751, -0.297)
IMD 4	0.360** (0.104, 0.617)	-0.198 (-0.418, 0.022)
IMD 5 (most deprived)	0.487** (0.188, 0.785)	-0.384** (-0.630, -0.139)
Town and fringe (rural hamlets and isolated dwellings)	-0.424 (-1.016, 0.168)	-0.071 (-0.616, 0.475)
Rural village	-0.247 (-0.845, 0.351)	-0.266 (-0.837, 0.305)
Urban city and town	-0.241 (-0.807, 0.325)	0.615* (0.094, 1.136)
Isle of Wight (Hampshire)	-1.191*** (-1.695, -0.688)	-1.872*** (-2.311, -1.433)
Portsmouth	-1.621*** (-2.058, -1.184)	-2.385*** (-2.761, -2.008)
Southampton	-1.678*** (-1.985, -1.372)	-2.751*** (-3.051, -2.450)
Constant	-1.412** (-2.469, -0.354)	-0.768 (-1.689, 0.152)

Estimate * (95% CI) p<0.05; p<0.01; p<0.001 Robust SE

In general households were less likely to respond in areas of higher mean gas consumption, in the Cities of Portsmouth and Southampton as well as the Isle of Wight compared to Hampshire. They were more likely to respond in areas with a higher percentage of electric central heating and those in all index of multiple deprivation (IMD)[25] quintiles compared to 1 (least deprived). Similarly, households were less likely to refuse (once contacted) in areas with a higher percentage of electric central heating, lower mean gas consumption, IMD quintiles above 1, and if they were in the Cities of Portsmouth and Southampton as well as the Isle of Wight. Households were more likely to refuse in areas of higher fuel poverty rates and in urban

[21] Census 2011, LSOA level

[22] DECC/BEIS, 2015, LSOA

[23] DECC/BEIS, 2017, LSOA

[24] IMD 2015, LSOA

areas. In general, these results may indicate effects such as the tendency to be out of the home at work during the likely contact hours and to consider themselves 'too busy' when contacted. However, those who live in areas with more electric central heating were more likely to respond and less likely to refuse whilst the IMD estimates suggest that the non-response and refusal effects may counteract each other. Further analysis (below) shows the extent to which these processes may have affected the bias of the sample on key dimensions.

A.1.5 Assessing sample bias

To date, a total of 2,901 full surveys have been completed, which equates to 57% of households where an installation has been completed. The sample is compared with Wave 4 of Understanding Society 'south east England' regional sample (Institute for Social and Economic Research, 2015) which was collected in 2012-2014 compared to the SAVE recruitment period of 2015-2016. The results are shown in Table 31.

Table 31 Analysis of SAVE vs Understanding Society Wave 4 distributions (weighted data)

Variable	*Chi sq statistic*	*Df*	*p.value*
HRP Age	64.856	6	0.000
Tenure	84.625	2	0.000
Household size	152.87	4	0.000
N rooms	164.58	8	0.000
Presence of gas	113.17	2	0.000
Household income	11.748	7	0.109

Table 32 Linear regression model estimating factors predicting pro-environmental scores

	Estimate (95% CI)
USOC W4 (compared to SAVE)	-0.086** (-0.164, -0.007)
Survey year	0.011 (-0.012, 0.034)
Female	0.065*** (0.030, 0.099)
−65 - 74	-0.153** (-0.279, -0.027)
75+	-0.112* (-0.242, 0.018)
Age Refused	-0.560** (-1.034, -0.086)
Constant	-19.348 (-65.787, 27.090)
Observations	5,802
R^2	0.025
Adjusted R^2	0.023
Residual Std. Error	0.661 (df = 5791)
F Statistic	14.724*** (df = 10; 5791)
Note:	*p<0.1; **p<0.05; ***p<0.01

Having established the extent to which the SAVE sample is representative of the population it was drawn from, the next step is to assess the success of the random allocation of recruited households to trial groups.

Analysis of the same variables as above but comparing the trial groups within the SAVE sample reveals no statistically significant differences between groups on these dimensions with the exception of the presence of mains gas (Table 33). In this analysis, main heat source has been added (electric storage vs gas vs oil vs other electric) and the results for this dimension show no inter-group difference. In addition, analysis of mean 'eco index' reveals no significant difference between groups[26] (Figure 37). Finally, differences in overall log mean electricity consumption[27] were tested showing no statistically significant differences between the trial groups[28] (Figure 38).

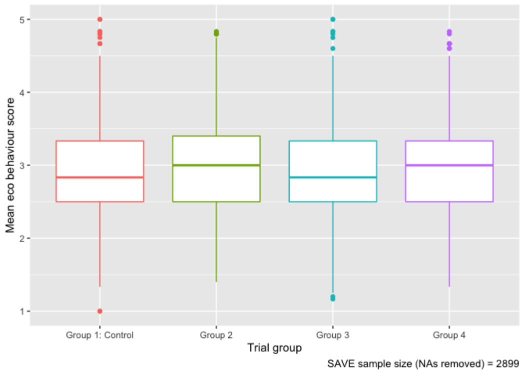

Figure 37 Trial groups – Eco mean

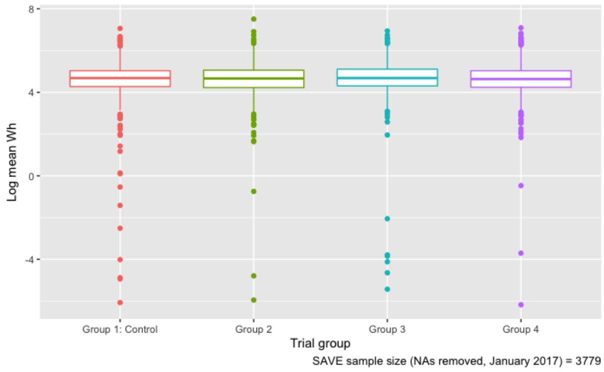

Figure 38 Trial groups – Log mean Wh

Table 33 Within-sample (between trial group) tests of homogeneity of proportions

Variable	*chi sq statistic*	*df*	*p.value*
HRP age	16.717	21	0.728
Household size	13.304	12	0.347
N rooms	24.548	24	0.431
Tenure	8.390	9	0.495
Mains gas	13.126	6	0.041
Income	25.489	27	0.547
Main heat source	8.881	9	0.448

[26] Analysis of variance (ANOVA): F = 1.616, p = 0.183

[27] Using log mean Wh to avoid the effects of skewed Wh consumption (data for January 2017).

[28] ANOVA: F = 0.561, p = 0.641

A.2 Data

A.2.1 Time-use diary data

The time-use diaries were implemented in two phases: 300 collected per trial group distributed across the 4-8 messaging period (mid-January to mid-February 2017); and a further 300 per group to collect data relating to the event day on the 15th March. Diaries were focussed on the event day in an attempt to capture any differences in the daily practices of households in the groups subjected to event day treatment.

A.2.1.1 Limitations of diary data

Anticipated limitations of the time-use diary data include:

- the time use diaries only captured 1 person's activities. No data was captured of the activities of other people in multi-person households. It may be that analysis restricted to solely single person households will be required, and this will substantially reduce the number of diaries available for analysis;[29]
- the activities reported in the time use diary may not have direct consequences for *electricity* demand at that point in time. For example, cooking in many households would use gas rather than electricity and showers or baths will use hot water heated either by gas or potentially at other times of the day. This may mean that analysis should be restricted to non-gas households but this would also substantially reduce the number of diaries available for analysis.

There are several respondents with more than 1 diary as they responded to the diary during the 4-8 messaging period and also the event day. In these cases, care is taken to preserve both diary responses. In total, the data contained 1307 diary respondents.

A.2.1.2 Diary data checks

Table 34 shows the number of diaries collected across the different sub-groups. Note that there were no diaries for Group 2 due to the trial implementation problems.

[29] Implementation of the time-use diaries in trial period 2 will attempt to address this limitation.

Table 34 Total time-use diary cases by trial group (note no Group 2)

Diary timing	Trial Group	n. of diaries
4-8 messaging	BMG Group 1: Control	315
4-8 messaging	BMG Group 3	324
4-8 messaging	BMG Group 4	324
Event day	BMG Group 1: Control	319
Event day	BMG Group 3	273
Event day	BMG Group 4	303

The number of diary cases by date of collection and trial group is shown in Figure 39.

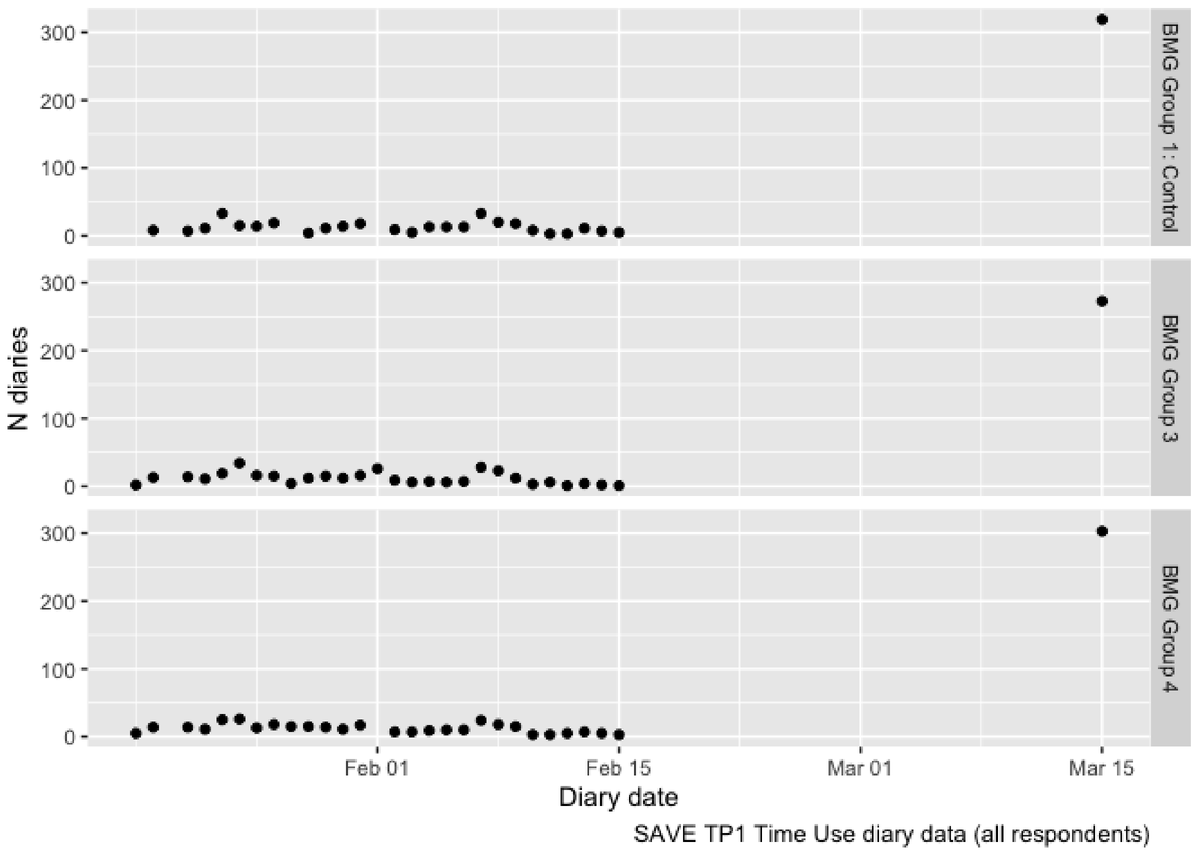

Figure 39 Number of time-use diaries by diary date

Due to the low number of diaries recorded over the weeks of the trial (Figure 35), analysis was focused on the event day, where the majority of time-use diaries were targeted.

The time-use diaries were not all collected on the day following the event day. Figure 40 shows the lag between diary date and the actual diary interview and that the event day diaries in particular were not necessarily all carried out on the day after the event, some took place a few days afterwards.

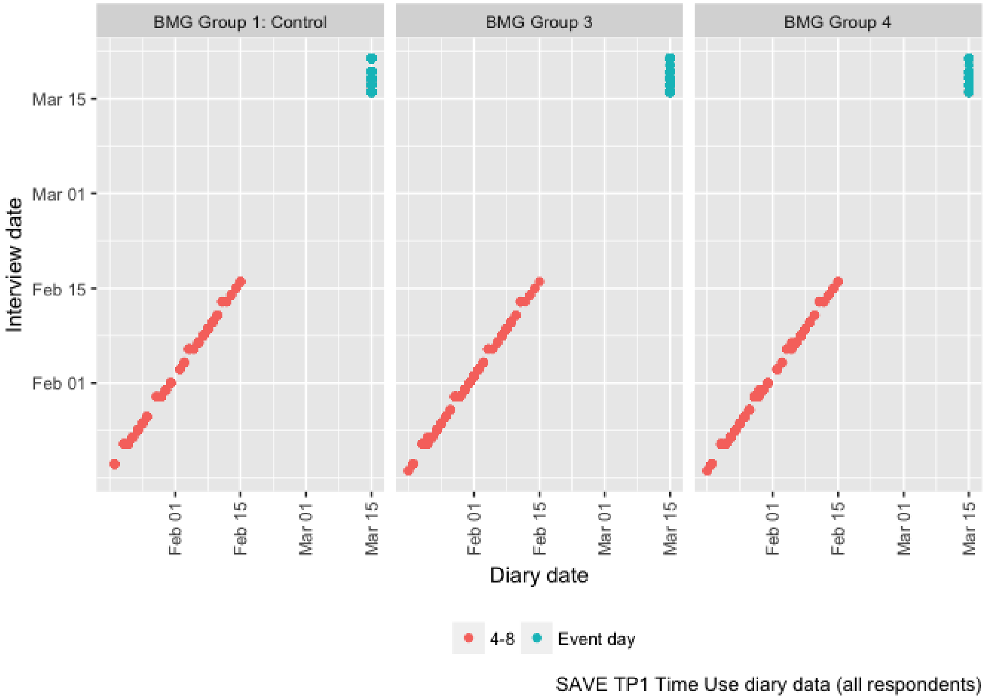

Figure 40 Time-use diaries: interview date vs. diary date

A.2.1.3 Time-use diary non-response bias

To assess diary non-response, the diary respondents are added to the household survey data. To test for non-response bias, an Analysis of Variance (ANOVA) test was conducted on the eco-attitudes/behaviours score; the results show no statistically significant difference between trial groups or diary respondents vs non-diary respondents by mean (see Table 35 and ANOVA test results below). This suggests no particular bias for those who responded to the diary within each trial group.

Table 35 Mean eco score by diary (non)respondents

Trial Group	Time-use diary status	meanEco score
BMG Group 1: Control	Completed diary	2.919930
BMG Group 1: Control	No diary (but have survey)	2.960296
BMG Group 2	No diary (but have survey)	2.949924
BMG Group 3	Completed diary	2.950255
BMG Group 3	No diary (but have survey)	2.912694
BMG Group 4	Completed diary	2.967112
BMG Group 4	No diary (but have survey)	2.983457

Analysis of Variance (ANOVA) test on the eco-attitudes/behaviours score: between trial groups or diary respondents vs non-diary respondents

```
##                   Df Sum Sq Mean Sq F value Pr(>F)
## tp1Diary           1    0.0  0.0302   0.066  0.798
## bmgGroup.respF     3    1.2  0.3885   0.843  0.470
## Residuals       4039 1860.8  0.4607
## 38426 observations deleted due to missingness
```

Further analysis would require us to know exactly who was contacted to do a diary but refused.

A.2.2 Navetas 'Loop' electricity consumption data

Table 36 shows the observations (nObs) for the trial period by trial week, along with columns to check coding of the trial weeks (minDate and maxDate).

Table 36 Loop data: date ranges and observations per trial week

trialWeek	nObs	minDate	maxDate
Week 01	2319814	2017-01-02	2017-01-08
Week 02	2473280	2017-01-09	2017-01-15
Week 03	2540315	2017-01-16	2017-01-22
Week 04	2580673	2017-01-23	2017-01-29
Week 05	2609415	2017-01-30	2017-02-05
Week 06	2605661	2017-02-06	2017-02-12
Week 07	2594862	2017-02-13	2017-02-19
Week 08	2583075	2017-02-20	2017-02-26
Week 09	2566971	2017-02-27	2017-03-05
Week 10	2553041	2017-03-06	2017-03-12
Week 11	2538373	2017-03-13	2017-03-19
Week 12	2524284	2017-03-20	2017-03-26

This analysis was run using the 100% sample of loop data (containing 3934 individual clamps/households) for 2017-01-02 UTC–2017-04-09 UTC. Subsets of this data were used in parts of the analysis and modelling that follow. Where a reduced dataset has been used, it is identified clearly within the 'Data used' sub-sections at the head of each main section of analysis.

As noted above, as little pre-processing is performed on the data prior to analysis as possible. There is no interpolation or imputation is used for the electricity consumption data prior to statistical modelling, however some filtering is performed on the data to remove observations with zero consumption.

The pre-processing of clamp data involves:[30]

- ✓ assembling single data file from daily files,
- ✓ creation of date and time variables for easier manipulation and better modelling,
- ✓ identify interpolated observations,
- ✓ calculation of observation lag time for diagnostics,
- ✓ removal of invalid clamps/households (out of sample),
- ✓ create consistent household identifier for linking clamp and survey data,
- ✓ merging of trial group information.

[30] For detail see the RMarkdown script: 'extractNavetas15m.Rmd'.

A.3 TP1 analysis

A.3.1 Descriptive analysis

The summary statistics shown in Table 37 show a roughly constant number of households with a slight decline during the trial period as expected through slow attrition. Mean Wh values are considerably higher than the median suggesting a distinct skew in the Wh observation values. The standard deviation is roughly double the mean suggesting a great deal of variability in 15-minute electricity consumption. Whilst the maximum values do not seem unusually high, 1-2% of observations in a given month are zero Wh. Further analysis of the data reveals that relatively small number of households report any zero Wh observations (Table 38) and this is concentrated in just a few households.

Table 37 Summary statistics for 15 minute Wh observations for all households in January, February and March 2017

obsMonth	N households	N Wh observations	Mean Wh	S.D. Wh	Median Wh	Min Wh	Max Wh	% 0 Wh
Jan	3,911	10,654,691	131.77	204.06	60	0	5958	1.83
Feb	3,899	10,383,358	120.26	187.32	55	0	5809	1.72
Mar	3,824	10,526,105	109.25	168.05	51	0	4388	1.18

Table 38 Number of clamps recording 0 Wh observations

zeroObsLab	N households
> 1 zero Wh observations	772
No zero Wh observation	3934

This is confirmed by the histogram in Figure 41 which demonstrates that the majority of households report no 0 Wh values but a very few report more than 10% of observations as 0 Wh.

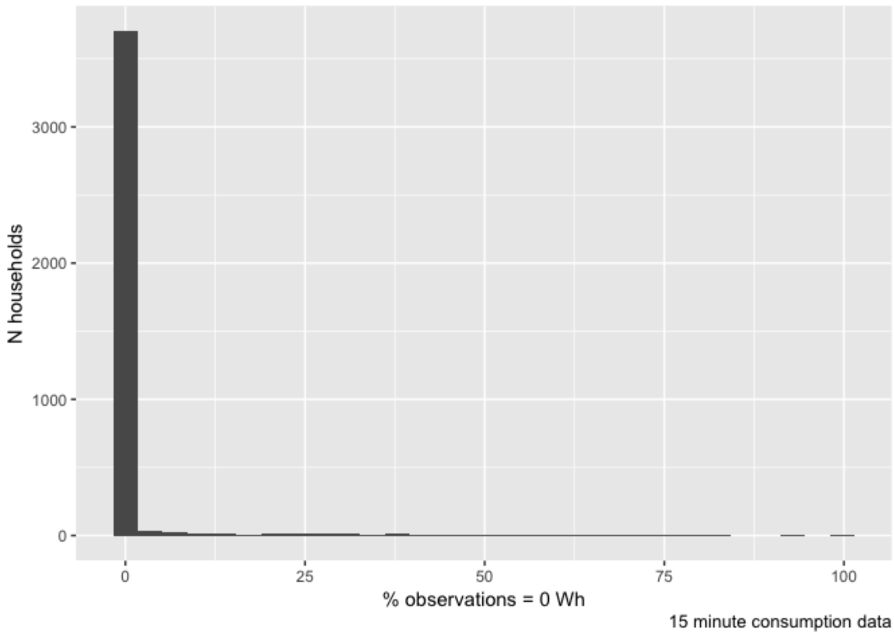

Figure 41 Distribution of households by % observations which were zero Wh

The few zero Wh observations are removed as probably being indicators of incorrectly fitted clamps. Clamps (households) reporting the zero Wh observations are not removed, just the observations themselves. No high values are removed as they do not appear unreasonably large. Summary statistics after removing 0 Wh observations are shown in Table 39.

Table 39 Summary statistics for 15 minute Wh observations for all households in January, February and March 2017 (0 Wh observations removed)

obsMonth	Number of Clamps	N observations	Mean Wh	S.D. Wh	Median Wh	Min Wh	Max Wh
Jan	3,911	10,460,076	134.22	205.15	61	1	5958
Feb	3,898	10,204,595	122.37	188.27	56	1	5809
Mar	3,824	10,401,501	110.56	168.63	52	1	4388

Not all of these households will have responded to a full (and thus an update) survey as Table 40 shows.

Table 40 Number of surveyed households with clean clamp data

	Telephone survey	Web survey	F2F survey	No survey data
BMG Group 1: Control	425	376	3	216
BMG Group 2	433	369	37	232
BMG Group 3	373	301	6	221
BMG Group 4	382	298	7	255

This means that analysis of the TP1 consumption data alone can use all 3,934 households, but any analysis requiring survey data will only be able to use a maximum of 3,010 households.

A.3.2 Daily consumption trends

The following chart shows the mean consumption (Wh) per day for all days in January, February and March for the control group which received no interventions.

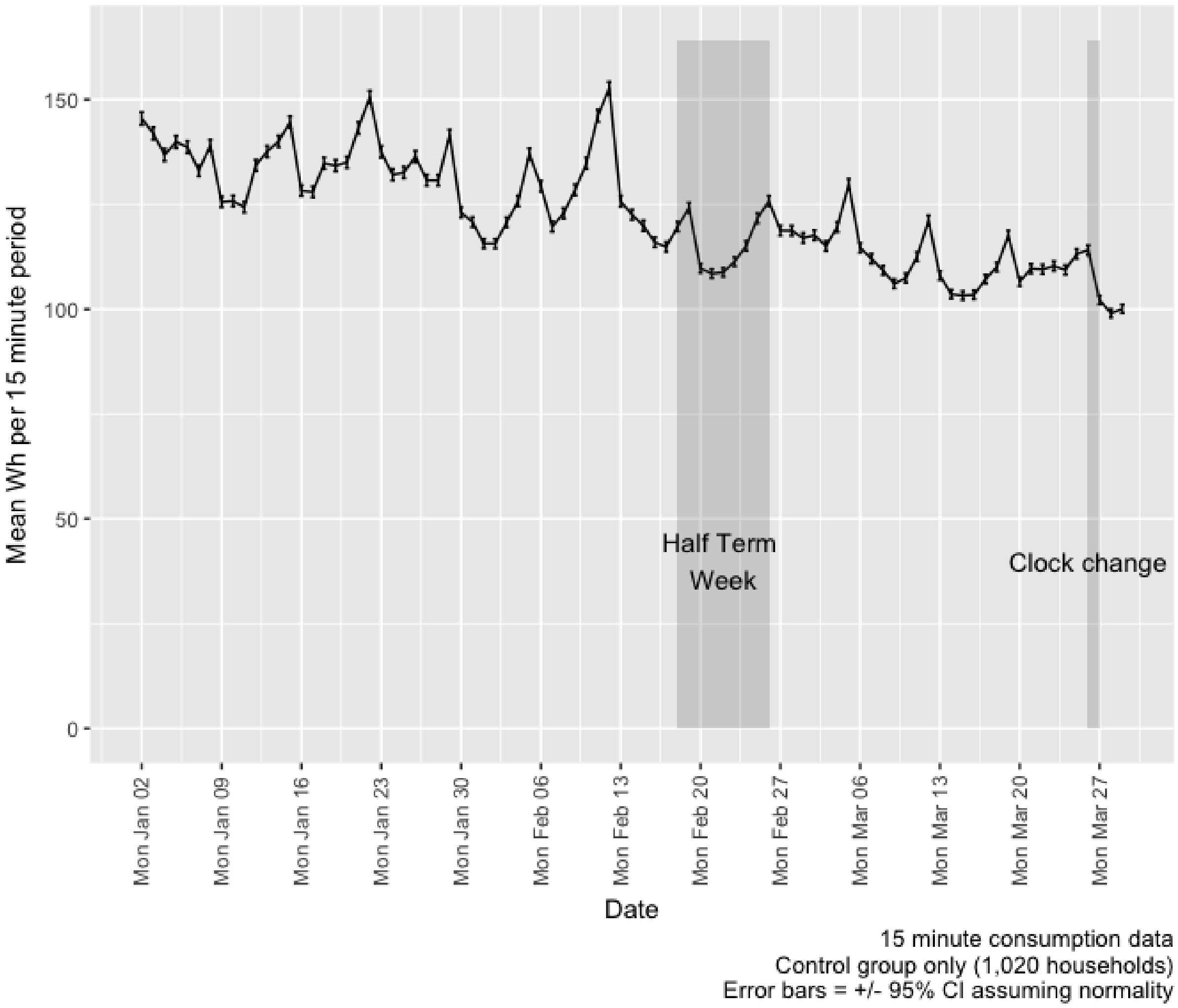

Figure 42 Mean 15m Wh consumption per day

The chart clearly shows spikes in mean Wh consumed on Sundays with lower overall levels of consumption during the weekdays. There are also clearly different consumption levels during half term indicating the potential effects of children and carers being at home during the day and/or absences due to holidays away from home.

Distinguishing between households with and without children. Table 41 shows how many control group households do/do not have children. Households marked 'NA' are those who are yet to respond to the survey.

Table 41 Presence of children within control group

Presence of children	N households
NA	192
0 children	588
1+ child	240

Figure 43 re-creates the previous figure but compares the results for those households where there are children aged under 16 with those where there are not. It can be seen that mean consumption for households with at least one child present is generally higher. Interestingly there does not appear to be a substantial change to this pattern during half term week (2–/2 - 26/2).

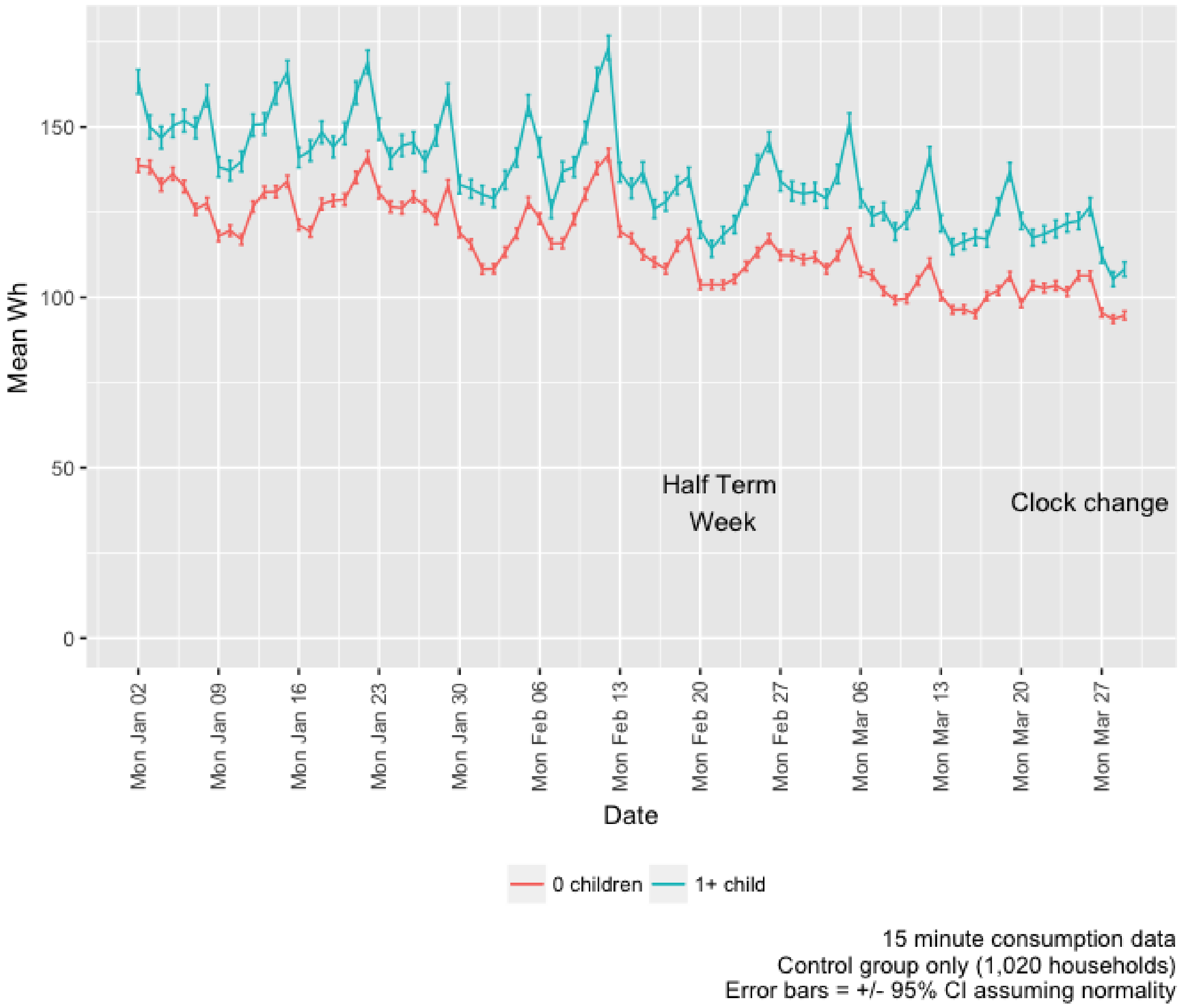

Figure 43 Mean 15m Wh consumption per day by presence of children

Figure 44 uses the working status of the household response person (only) as an indicator (Table 42). Households in full time and part-time employment (first and second panels in Figure 44), exhibit a higher and more distinctive morning peak in consumption, and greater separation of the daytime consumption profiles between weekdays and weekends. Households where the household representative person (HRP) is retired (third panel) appear to have a lower evening peak than other households. Households with unemployed HRP (final panel) display the most variable daily profiles, although the small number of households in this group will contribute to the variability.

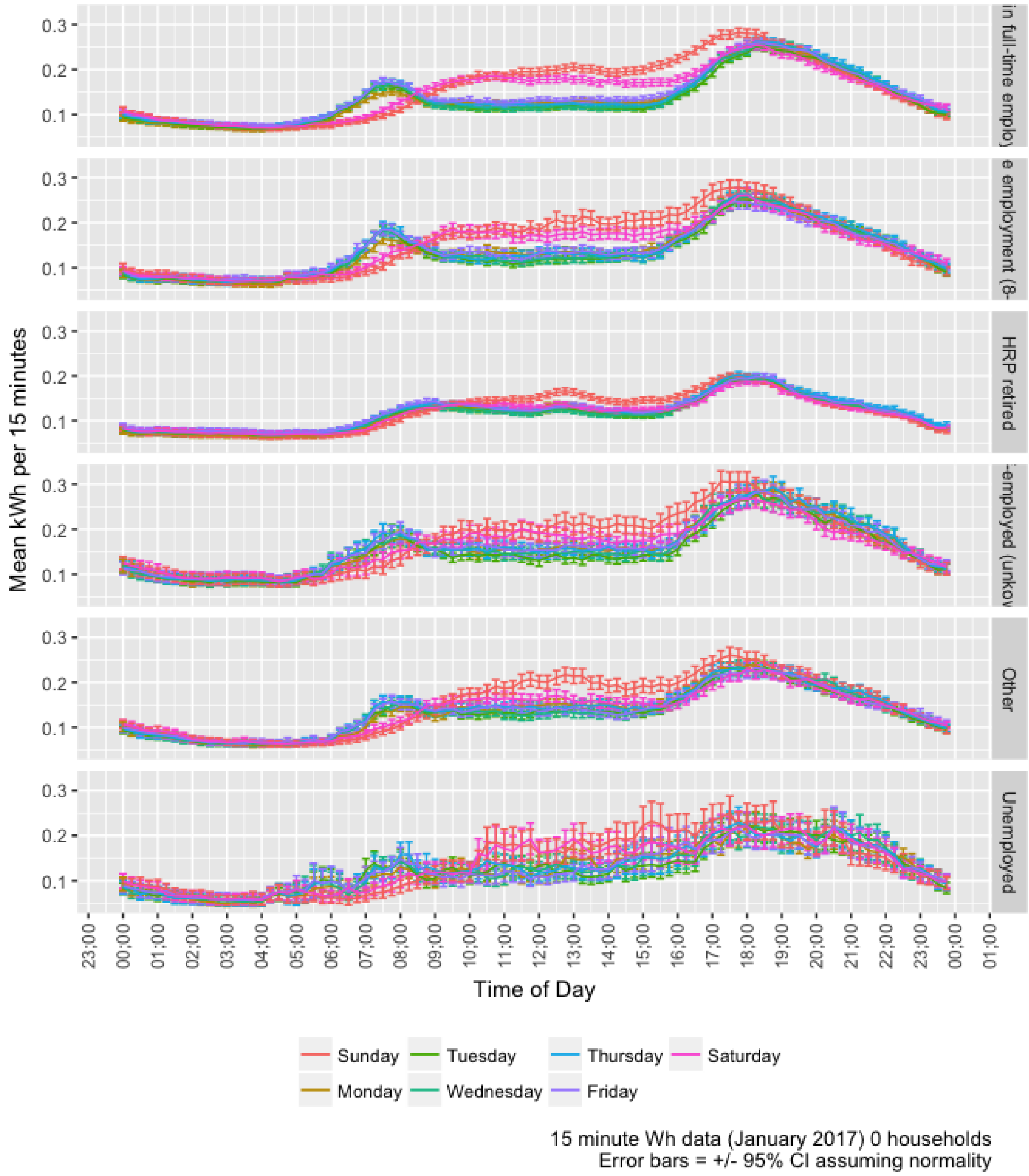

Figure 44 Mean kWh consumption per 15 minutes by household type (weekdays only)

Table 42 Working/retired HRPs (NA = no survey data to date)

HRP work status	N households
NA	861
HRP self-employed (unknown hours)	294
HRP in part-time employment (8-29 hours/week)	404
HRP retired	1065
HRP self-employed (unknown hours)	209
HRP in full-time employment	1038
Other	289
Unemployed	68

A.3.3 Weekly 4-8 messaging impact

Table 43 Mean Wh per week and % differences from control group

Trial Week	Control	Online only	Postal only	Online and postal	Online only	Postal only	Online and postal
	Mean Weekly Wh				% difference from control		
Week 01	216.86	219.34	212.34	223.57	1.14	-2.08	3.09
Week 02	208.39	213.63	207.48	212.33	2.52	-0.43	1.89
Week 03	211.42	214.05	211.03	217.08	1.25	-0.18	2.68
Week 04	207.42	211.08	207.83	214.06	1.76	0.20	3.20
Week 05	194.61	199.27	191.79	201.23	2.39	-1.45	3.40
Week 06	206.03	207.46	203.66	211.00	0.69	-1.15	2.41
Week 07	185.45	185.88	181.68	189.83	0.24	-2.03	2.36
Week 08	176.12	178.96	175.46	181.92	1.61	-0.37	3.29
Week 09	181.52	184.57	178.87	187.56	1.68	-1.46	3.32
Week 10	171.34	171.41	166.22	179.37	0.04	-2.99	4.69
Week 11	164.53	162.82	160.47	169.15	-1.04	-2.47	2.81
Week 12	159.34	161.68	155.15	164.92	1.47	-2.63	3.50
Week 13	154.59	155.94	149.22	156.77	0.88	-3.47	1.41

A.3.3.1 Regression model results: simple model

The results from the initial (simple) model are combined in Table 44. As a guide, any co-efficient that has more than 1 * in the tables is statistically significant at the 95% level or greater. Similarly, if the 95% CI for the coefficient shown in the chart does not include 0, then it can be concluded that the effect is statistically significant at the 95% level. Note that the 95% CIs referred to in the model results charts are not subject to the same potential invalidity as those used in the descriptive analysis above.

Table 44 Model 1 results

	Dependent variable:						
	logMeanWhPerWeek						
	Week 1	Week 2	Week 3	Week 4	Week 5	Week 8	Week 10
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Online and postal	0.018 (-0.051, 0.088)	0.021 (-0.045, 0.088)	0.027 (-0.041, 0.095)	0.042 (-0.025, 0.108)	0.044 (-0.021, 0.109)	0.047 (-0.016, 0.110)	0.060 (-0.005, 0.124)
Online only	-0.014 (-0.080, 0.052)	-0.004 (-0.068, 0.059)	-0.017 (-0.082, 0.048)	-0.002 (-0.065, 0.061)	0.005 (-0.057, 0.067)	0.010 (-0.051, 0.070)	-0.014 (-0.076, 0.048)
Postal only	-0.043 (-0.111, 0.025)	-0.033 (-0.098, 0.033)	-0.042 (-0.109, 0.026)	-0.025 (-0.090, 0.041)	-0.029 (-0.093, 0.035)	-0.007 (-0.069, 0.056)	-0.040 (-0.104, 0.024)
Constant	5.162*** (5.114, 5.209)	5.121*** (5.075, 5.167)	5.124*** (5.078, 5.171)	5.104*** (5.058, 5.149)	5.052*** (5.007, 5.096)	4.958*** (4.915, 5.001)	4.931*** (4.886, 4.975)
Observations	3,554	3,716	3,763	3,833	3,846	3,819	3,789
R^2	0.001	0.001	0.001	0.001	0.001	0.001	0.003
Adjusted R^2	0.00005	-0.0001	0.0003	0.0002	0.0005	0.0001	0.002
Residual Std. Error	0.732 (df = 3550)	0.721 (df = 3712)	0.740 (df = 3759)	0.728 (df = 3829)	0.716 (df = 3842)	0.693 (df = 3815)	0.708 (df = 3785)
F Statistic	1.055 (df = 3; 3550)	0.826 (df = 3; 3712)	1.342 (df = 3; 3759)	1.289 (df = 3; 3829)	1.592 (df = 3; 3842)	1.065 (df = 3; 3815)	3.183* (df = 3; 3785)

A.3.3.2 Regression model results: household attribute model

Due to the size of the tabular results, only the Week 1 model results are included here (Table 45).

Table 45 4-8 messaging regression model results: Week 1

	Dependent variable:	
	logMeanWhPerWeek	modelDef
	Model 1	Model 2
	(1)	(2)
int4_8Online and postal	0.018 (-0.051, 0.088)	-0.304 (-0.802, 0.194)
int4_8Online only	-0.014 (-0.080, 0.052)	-0.085 (-0.559, 0.389)
int4_8Postal only	-0.043 (-0.111, 0.025)	-0.089 (-0.609, 0.431)
storage heaters		0.207 (-0.017, 0.430)
biomas, other		0.392*** (0.198, 0.587)
heat pumps		0.691** (0.194, 1.188)
ba_presenceChildren.latest1+ child		-0.146 (-0.299, 0.006)
ba_Q2_npeople_reduced.latest2		0.501*** (0.364, 0.638)
ba_Q2_npeople_reduced.latest3		0.755*** (0.580, 0.930)
ba_Q2_npeople_reduced.latest4		0.827*** (0.614, 1.040)
ba_Q2_npeople_reduced.latest5+		1.095*** (0.853, 1.336)
ba_censusTenurePrivate rent		-0.169 (-0.352, 0.014)
ba_censusTenureSocial rent		-0.136 (-0.292, 0.019)
ba_censusTenurex.Refused/dk/Other		-0.240 (-0.670, 0.191)
tmpDwellingFlat/Other		-0.513*** (-0.754, -0.273)
tmpDwellingSemi		-0.160** (-0.273, -0.047)
tmpDwellingTerrace		-0.215*** (-0.342, -0.087)
ba_Q8_27_Income£10-£20k		-0.038 (-0.276, 0.201)
ba_Q8_27_Income£20-£30k		-0.014 (-0.265, 0.236)
ba_Q8_27_Income£30-£40k		0.036 (-0.226, 0.297)
ba_Q8_27_Income£40-£50k		0.070 (-0.207, 0.347)
ba_Q8_27_Income£50-60k		0.100 (-0.192, 0.391)
ba_Q8_27_Income£60-80k		0.126 (-0.154, 0.407)
ba_Q8_27_Income> £80k		0.272 (-0.018, 0.562)
ba_Q8_27_IncomeDon't know		0.092 (-0.159, 0.344)
ba_Q8_27_IncomeRefused		0.118 (-0.108, 0.345)
ba_Q22_HRPethnicityMixed		-0.192 (-0.879, 0.494)
ba_Q22_HRPethnicityAsian/Asian British		-0.165 (-0.484, 0.154)
ba_Q22_HRPethnicityBlack/Black British		0.017 (-0.440, 0.475)
ba_Q22_HRPethnicityOther		0.198 (-0.496, 0.892)
ba_Q22_HRPethnicityRefused		-0.019 (-0.450, 0.412)
ba_Q2D_HRPemplType.latestHRP in part-time employment (8-29 hours/week)		-0.048 (-0.203, 0.108)
ba_Q2D_HRPemplType.latestHRP retired		-0.176** (-0.308, -0.043)
ba_Q2D_HRPemplType.latestHRP self-employed (unkown hours)		-0.209* (-0.398, -0.021)
ba_Q2D_HRPemplType.latestOther		0.080 (-0.106, 0.266)
ba_Q2D_HRPemplType.latestUnemployed		0.013 (-0.354, 0.380)
ba_Q7_2_ecoMean		-0.111** (-0.180, -0.042)
storage heaters		0.564** (0.167, 0.961)
storage heaters		0.263 (-0.040, 0.565)
storage heaters		0.223 (-0.177, 0.623)
biomas, other		0.045 (-0.244, 0.335)
biomas, other		-0.098 (-0.376, 0.180)
biomas, other		0.214 (-0.079, 0.506)
heat pumps		0.038 (-0.945, 1.022)
heat pumps		-0.318 (-1.291, 0.655)
heat pumps		0.518 (-0.271, 1.308)
int4_8Online and postal:ba_presenceChildren.latest1+ child		-0.055 (-0.281, 0.170)
int4_8Online only:ba_presenceChildren.latest1+ child		0.023 (-0.185, 0.231)
int4_8Postal only:ba_presenceChildren.latest1+ child		-0.049 (-0.277, 0.179)
int4_8Online and postal:ba_Q2_npeople_reduced.latest2		-0.022 (-0.229, 0.184)
int4_8Online only:ba_Q2_npeople_reduced.latest2		-0.084 (-0.278, 0.109)
int4_8Postal only:ba_Q2_npeople_reduced.latest2		0.080 (-0.119, 0.279)
int4_8Online and postal:ba_Q2_npeople_reduced.latest3		-0.007 (-0.279, 0.265)
int4_8Online only:ba_Q2_npeople_reduced.latest3		-0.037 (-0.283, 0.210)
int4_8Postal only:ba_Q2_npeople_reduced.latest3		0.187 (-0.071, 0.446)
int4_8Online and postal:ba_Q2_npeople_reduced.latest4		0.119 (-0.193, 0.430)
int4_8Online only:ba_Q2_npeople_reduced.latest4		0.087 (-0.200, 0.373)

int4_8Postal only:ba_Q2_npeople_reduced.latest4	0.306* (0.001, 0.612)
int4_8Online and postal:ba_Q2_npeople_reduced.latest5+	0.060 (-0.308, 0.429)
int4_8Online only:ba_Q2_npeople_reduced.latest5+	-0.037 (-0.378, 0.304)
int4_8Postal only:ba_Q2_npeople_reduced.latest5+	0.277 (-0.078, 0.631)
int4_8Online and postal:ba_censusTenurePrivate rent	0.015 (-0.263, 0.294)
int4_8Online only:ba_censusTenurePrivate rent	-0.068 (-0.325, 0.188)
int4_8Postal only:ba_censusTenurePrivate rent	-0.012 (-0.297, 0.273)
int4_8Online and postal:ba_censusTenureSocial rent	-0.118 (-0.359, 0.124)
int4_8Online only:ba_censusTenureSocial rent	0.137 (-0.077, 0.352)
int4_8Postal only:ba_censusTenureSocial rent	0.084 (-0.151, 0.319)
int4_8Online and postal:ba_censusTenurex.Refused/dk/Other	-0.328 (-1.028, 0.372)
int4_8Online only:ba_censusTenurex.Refused/dk/Other	0.294 (-0.272, 0.860)
int4_8Postal only:ba_censusTenurex.Refused/dk/Other	0.549 (-0.224, 1.322)
int4_8Online and postal:tmpDwellingFlat/Other	0.412* (0.055, 0.769)
int4_8Online only:tmpDwellingFlat/Other	0.083 (-0.244, 0.410)
int4_8Postal only:tmpDwellingFlat/Other	0.341 (-0.025, 0.707)
int4_8Online and postal:tmpDwellingSemi	0.065 (-0.103, 0.234)
int4_8Online only:tmpDwellingSemi	0.065 (-0.093, 0.224)
int4_8Postal only:tmpDwellingSemi	-0.003 (-0.168, 0.163)
int4_8Online and postal:tmpDwellingTerrace	0.089 (-0.099, 0.278)
int4_8Online only:tmpDwellingTerrace	0.001 (-0.175, 0.177)
int4_8Postal only:tmpDwellingTerrace	-0.013 (-0.196, 0.169)
int4_8Online and postal:ba_Q8_27_Income£10-£20k	0.304 (-0.065, 0.673)
int4_8Online only:ba_Q8_27_Income£10-£20k	0.073 (-0.263, 0.408)
int4_8Postal only:ba_Q8_27_Income£10-£20k	-0.064 (-0.420, 0.292)
int4_8Online and postal:ba_Q8_27_Income£20-£30k	0.254 (-0.129, 0.638)
int4_8Online only:ba_Q8_27_Income£20-£30k	0.052 (-0.293, 0.397)
int4_8Postal only:ba_Q8_27_Income£20-£30k	0.032 (-0.335, 0.400)
int4_8Online and postal:ba_Q8_27_Income£30-£40k	0.153 (-0.250, 0.555)
int4_8Online only:ba_Q8_27_Income£30-£40k	-0.001 (-0.367, 0.364)
int4_8Postal only:ba_Q8_27_Income£30-£40k	0.050 (-0.337, 0.437)
int4_8Online and postal:ba_Q8_27_Income£40-£50k	0.132 (-0.283, 0.547)
int4_8Online only:ba_Q8_27_Income£40-£50k	-0.015 (-0.404, 0.373)
int4_8Postal only:ba_Q8_27_Income£40-£50k	0.035 (-0.366, 0.437)
int4_8Online and postal:ba_Q8_27_Income£50-60k	0.174 (-0.262, 0.609)
int4_8Online only:ba_Q8_27_Income£50-60k	-0.086 (-0.511, 0.340)
int4_8Postal only:ba_Q8_27_Income£50-60k	0.025 (-0.416, 0.467)
int4_8Online and postal:ba_Q8_27_Income£60-80k	0.204 (-0.223, 0.630)
int4_8Online only:ba_Q8_27_Income£60-80k	-0.051 (-0.445, 0.342)
int4_8Postal only:ba_Q8_27_Income£60-80k	-0.047 (-0.463, 0.370)
int4_8Online and postal:ba_Q8_27_Income> £80k	0.260 (-0.171, 0.692)
int4_8Online only:ba_Q8_27_Income> £80k	-0.107 (-0.516, 0.301)
int4_8Postal only:ba_Q8_27_Income> £80k	-0.154 (-0.588, 0.280)
int4_8Online and postal:ba_Q8_27_IncomeDon't know	0.207 (-0.182, 0.595)
int4_8Online only:ba_Q8_27_IncomeDon't know	-0.013 (-0.367, 0.342)
int4_8Postal only:ba_Q8_27_IncomeDon't know	-0.258 (-0.637, 0.121)
int4_8Online and postal:ba_Q8_27_IncomeRefused	0.286 (-0.068, 0.641)
int4_8Online only:ba_Q8_27_IncomeRefused	-0.077 (-0.398, 0.245)
int4_8Postal only:ba_Q8_27_IncomeRefused	-0.131 (-0.473, 0.212)
int4_8Online and postal:ba_Q22_HRPethnicityMixed	-0.483 (-1.581, 0.616)
int4_8Online only:ba_Q22_HRPethnicityMixed	0.140 (-0.666, 0.946)
int4_8Postal only:ba_Q22_HRPethnicityMixed	
int4_8Online and postal:ba_Q22_HRPethnicityAsian/Asian British	-0.306 (-0.806, 0.195)
int4_8Online only:ba_Q22_HRPethnicityAsian/Asian British	-0.362 (-0.844, 0.120)
int4_8Postal only:ba_Q22_HRPethnicityAsian/Asian British	0.036 (-0.444, 0.515)
int4_8Online and postal:ba_Q22_HRPethnicityBlack/Black British	-0.116 (-0.792, 0.560)
int4_8Online only:ba_Q22_HRPethnicityBlack/Black British	0.214 (-0.542, 0.969)
int4_8Postal only:ba_Q22_HRPethnicityBlack/Black British	-0.641 (-1.663, 0.380)
int4_8Online and postal:ba_Q22_HRPethnicityOther	-0.608 (-1.744, 0.527)
int4_8Online only:ba_Q22_HRPethnicityOther	-0.217 (-1.135, 0.700)
int4_8Postal only:ba_Q22_HRPethnicityOther	-0.467 (-1.479, 0.544)
int4_8Online and postal:ba_Q22_HRPethnicityRefused	-0.310 (-0.896, 0.276)
int4_8Online only:ba_Q22_HRPethnicityRefused	-0.025 (-0.538, 0.489)
int4_8Postal only:ba_Q22_HRPethnicityRefused	-0.066 (-0.632, 0.499)

int4_8Online and postal:ba_Q2D_HRPemplType.latestHRP in part-time employment (8-29 hours/week)		0.064 (-0.158, 0.287)
int4_8Online only:ba_Q2D_HRPemplType.latestHRP in part-time employment (8-29 hours/week)		0.067 (-0.153, 0.286)
int4_8Postal only:ba_Q2D_HRPemplType.latestHRP in part-time employment (8-29 hours/week)		0.193 (-0.026, 0.412)
int4_8Online and postal:ba_Q2D_HRPemplType.latestHRP retired		0.153 (-0.046, 0.352)
int4_8Online only:ba_Q2D_HRPemplType.latestHRP retired		0.202* (0.017, 0.386)
int4_8Postal only:ba_Q2D_HRPemplType.latestHRP retired		0.239* (0.044, 0.434)
int4_8Online and postal:ba_Q2D_HRPemplType.latestHRP self-employed (unkown hours)		0.403** (0.114, 0.692)
int4_8Online only:ba_Q2D_HRPemplType.latestHRP self-employed (unkown hours)		0.431** (0.170, 0.693)
int4_8Postal only:ba_Q2D_HRPemplType.latestHRP self-employed (unkown hours)		0.085 (-0.209, 0.379)
int4_8Online and postal:ba_Q2D_HRPemplType.latestOther		-0.011 (-0.302, 0.280)
int4_8Online only:ba_Q2D_HRPemplType.latestOther		-0.142 (-0.393, 0.109)
int4_8Postal only:ba_Q2D_HRPemplType.latestOther		-0.051 (-0.315, 0.212)
int4_8Online and postal:ba_Q2D_HRPemplType.latestUnemployed		0.062 (-0.530, 0.655)
int4_8Online only:ba_Q2D_HRPemplType.latestUnemployed		-0.099 (-0.553, 0.356)
int4_8Postal only:ba_Q2D_HRPemplType.latestUnemployed		0.256 (-0.240, 0.752)
int4_8Online and postal:ba_Q7_2_ecoMean		-0.003 (-0.105, 0.100)
int4_8Online only:ba_Q7_2_ecoMean		0.001 (-0.097, 0.098)
int4_8Postal only:ba_Q7_2_ecoMean		-0.050 (-0.153, 0.053)
Constant	5.162*** (5.114, 5.209)	5.068*** (4.731, 5.404)
Observations	3,554	2,643
R^2	0.001	0.338
Adjusted R^2	0.00005	0.302
Residual Std. Error	0.732 (df = 3550)	0.597 (df = 2504)
F Statistic	1.055 (df = 3; 3550)	9.278*** (df = 138; 2504)
Note:		
p<0.05; ***p<0.01***; p<0.001		

A.3.3.3 Regression model results: interpretation of other household attributes

This section presents visualisations of sections of the second regression model (household attributes).

The first presents the results for household size. The lower set of bars represent the ‘main’ (non-interaction) effect. The others represent the effects of each household size interacted with the intervention.

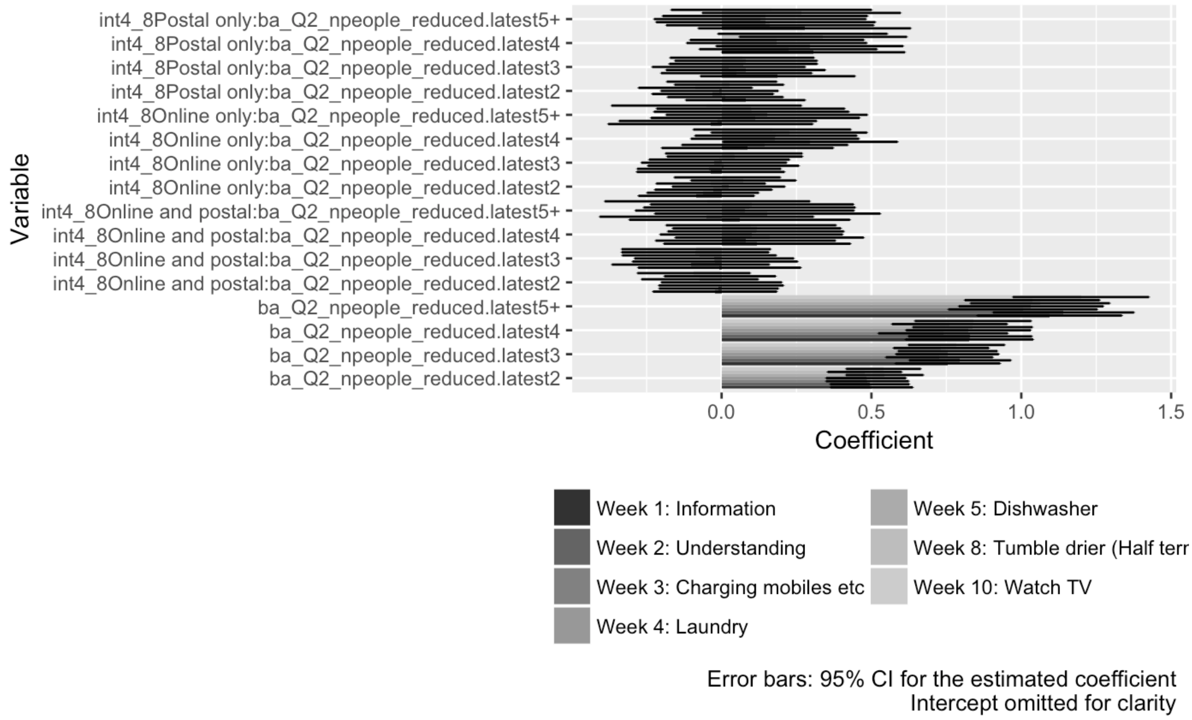

Figure 45 Regression model results for intervention: household size interaction

The next example examines dwelling type. Again, the top set of bars represent the 'main' effect. The others represent the effects of each dwelling type interacted with the intervention.

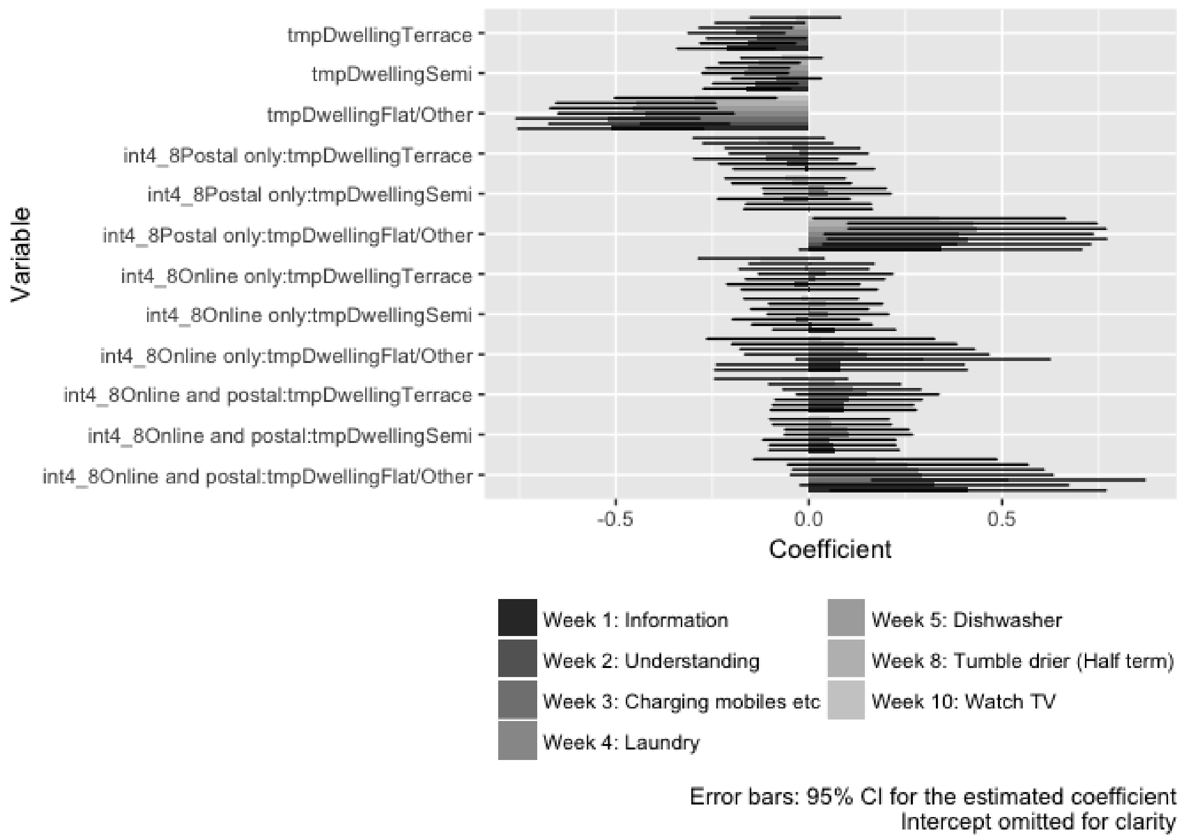

Figure 46 Regression model results for intervention: dwelling type interaction

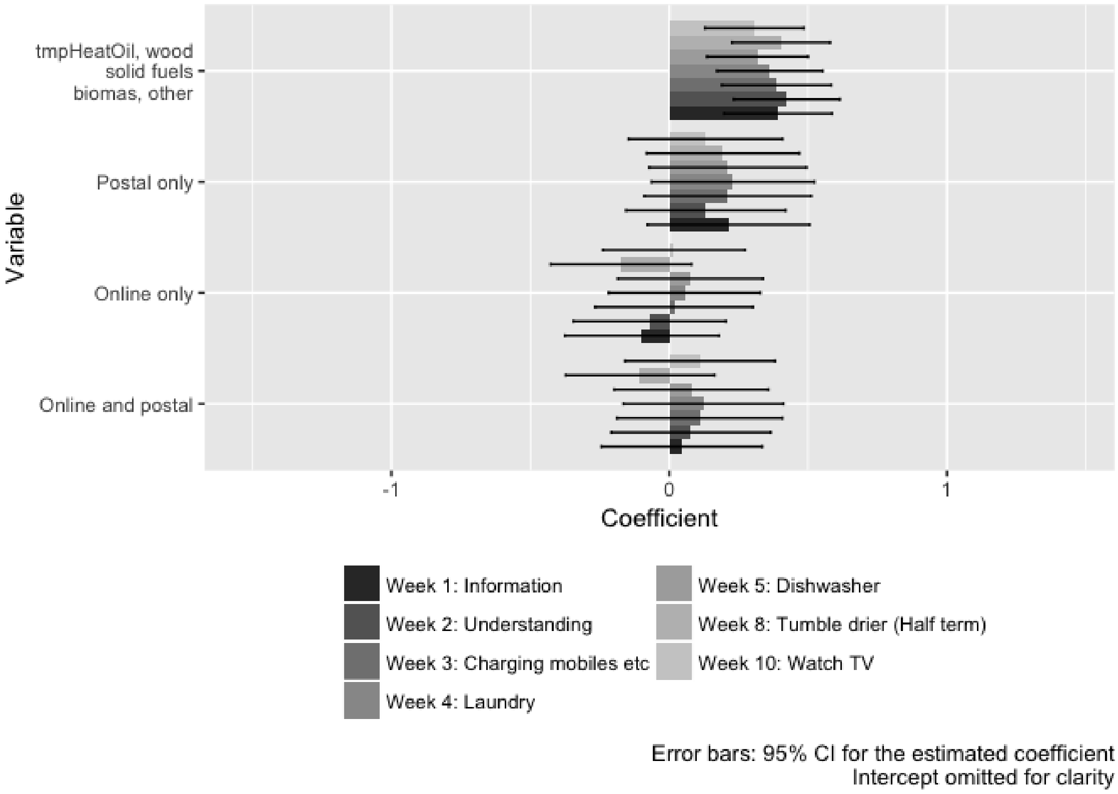

Figure 47 Regression model results for intervention: oil heating interaction

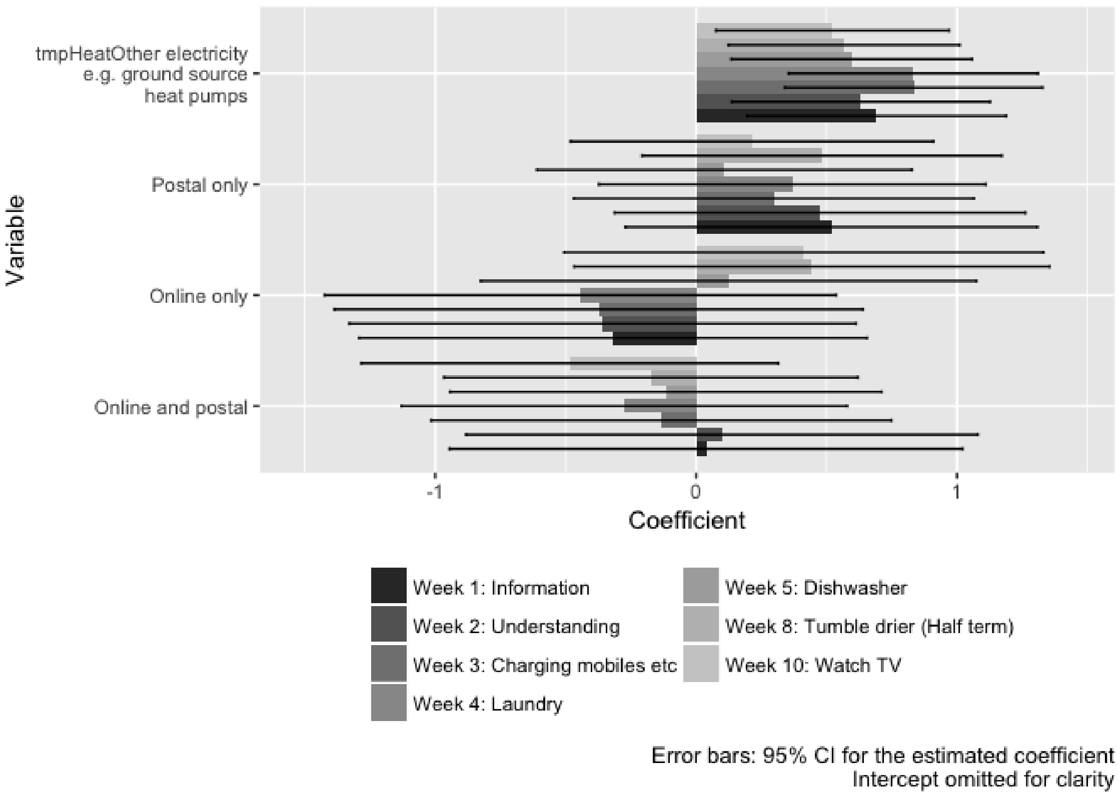

Figure 48 Regression model results for intervention: other electric heating interaction

Figure 49 examines the results for the 'eco mean' variable, the bottom set of bars show the main effect coefficient.

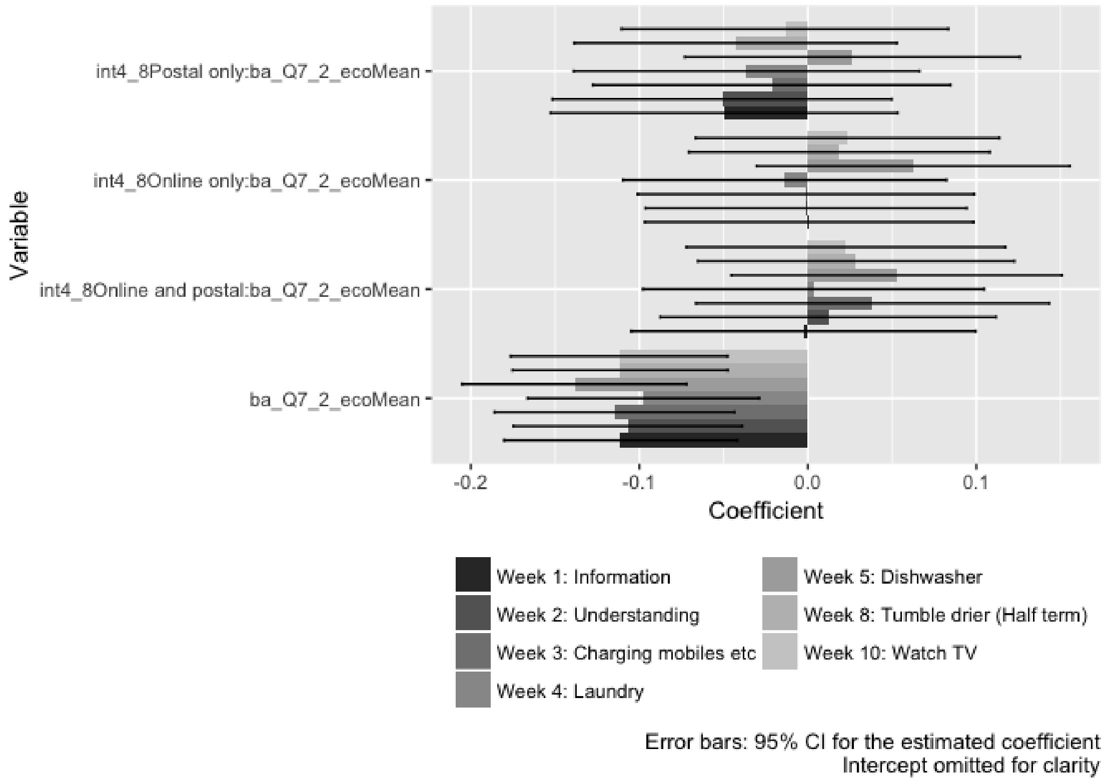

Figure 49 Regression model results for intervention: eco_mean interaction

Figure 50 examines the results for the income > £80k variable.

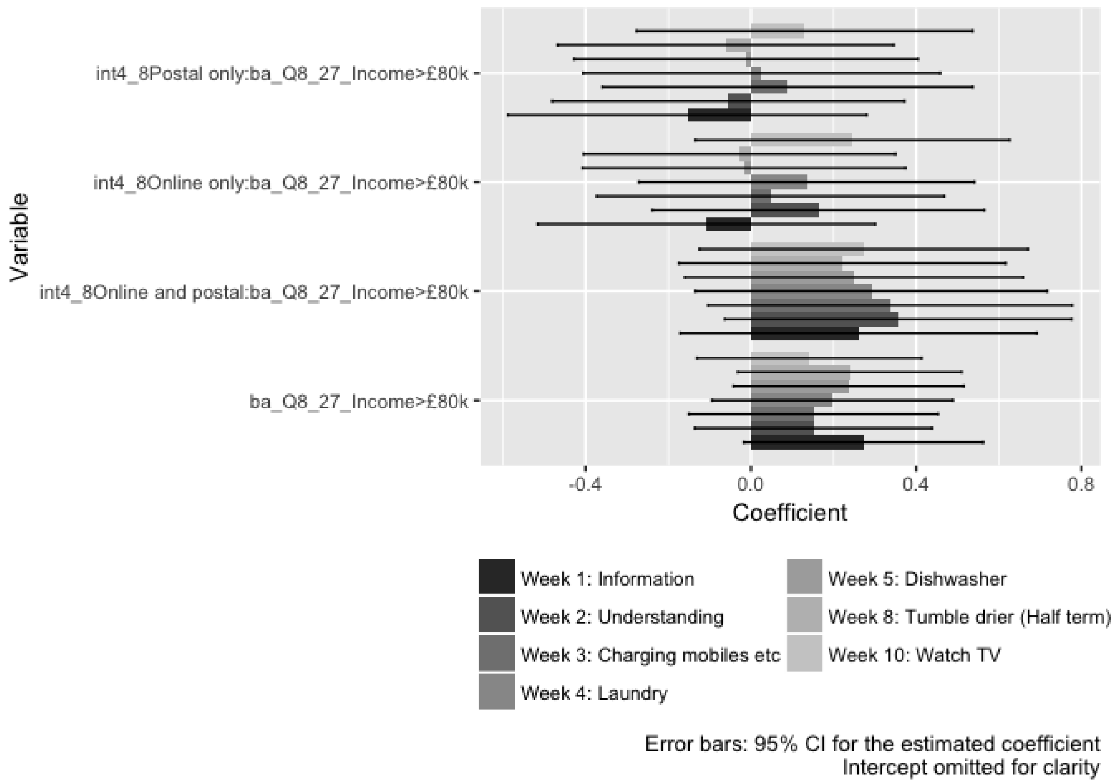

Figure 50 Regression model results for intervention: income>£80k interaction

Figure 51 and Figure 52 examine the results for the two ethnicity sub-groups that appear to show a response.

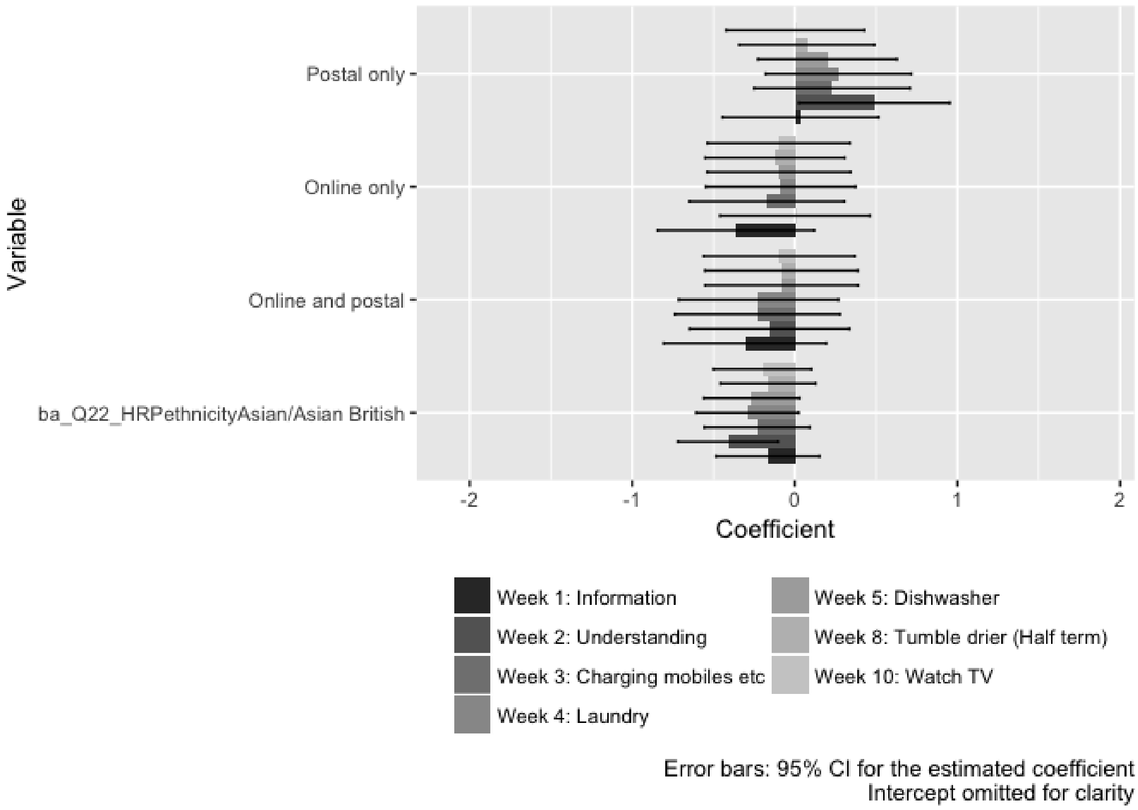

Figure 51 Regression model results for intervention: ethnicity_asian interaction

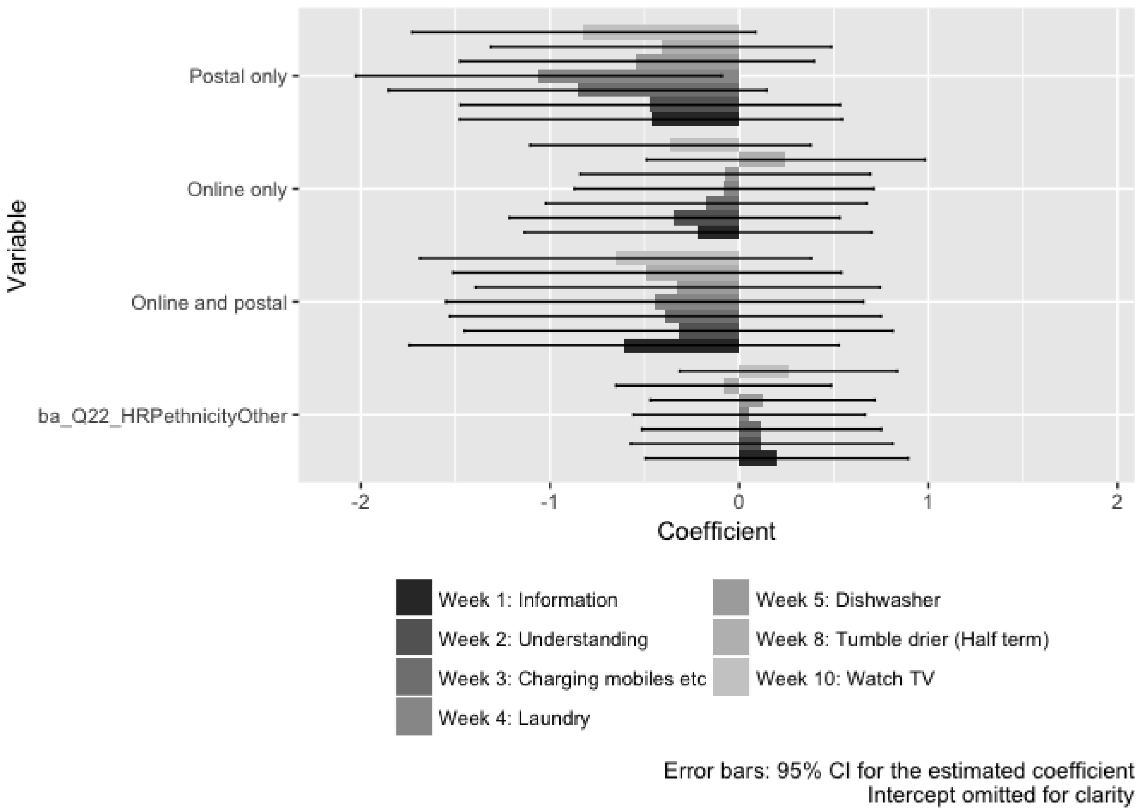

Figure 52 Regression model results for intervention: ethnicity_other interaction

Figure 53 examines the results for the presence of children variable.

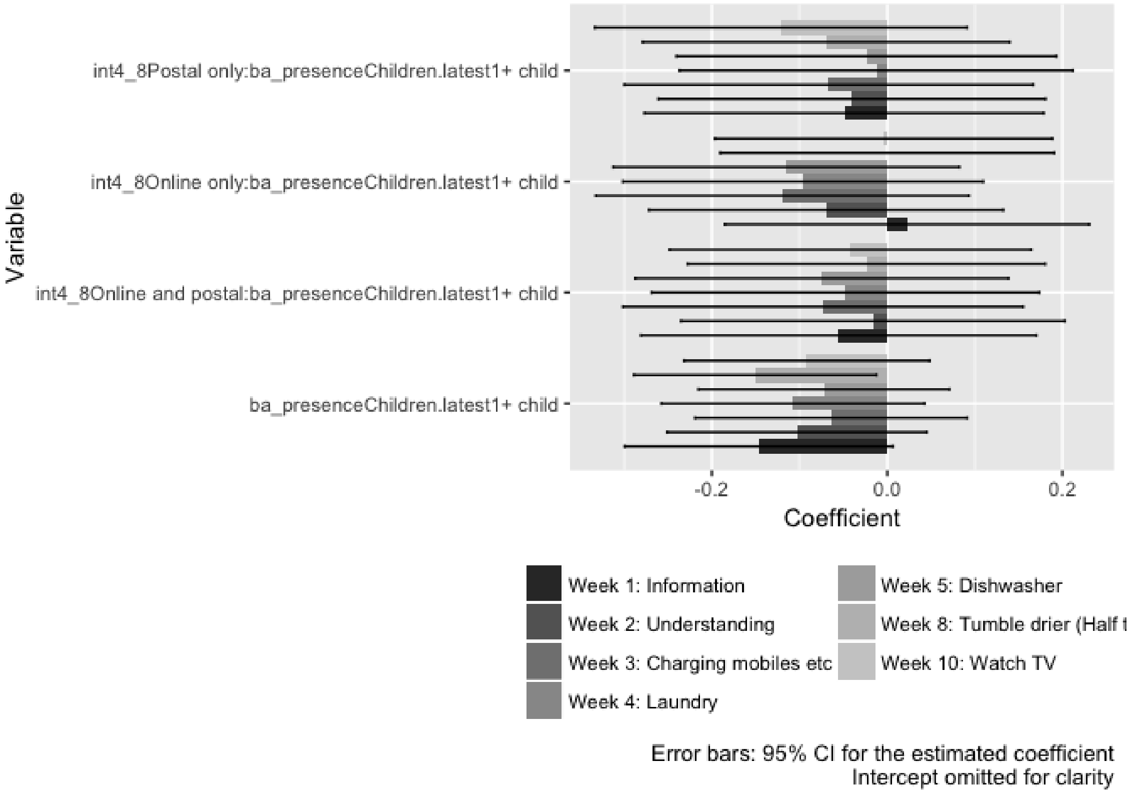

Figure 53 Regression model results for intervention: presence of children interaction

A.3.4 Impact of event day interventions

Table 46 shows the number of observations and households used in the event day analysis (after the removal of zero Wh observations).

Table 46 Number of observations and clamps on the event day and each day before/after

Trial Group	Day	N observations	N households
BMG Group 1: Control	Tue 14 Mar 2017	91,575	968
BMG Group 1: Control	Wed 15 Mar 2017	91,612	967
BMG Group 1: Control	Thu 16 Mar 2017	91,435	965
BMG Group 2	Tue 14 Mar 2017	97,605	1024
BMG Group 2	Wed 15 Mar 2017	97,613	1024
BMG Group 2	Thu 16 Mar 2017	97,431	1023
BMG Group 3	Tue 14 Mar 2017	84,082	880
BMG Group 3	Wed 15 Mar 2017	84,087	880
BMG Group 3	Thu 16 Mar 2017	83,958	880

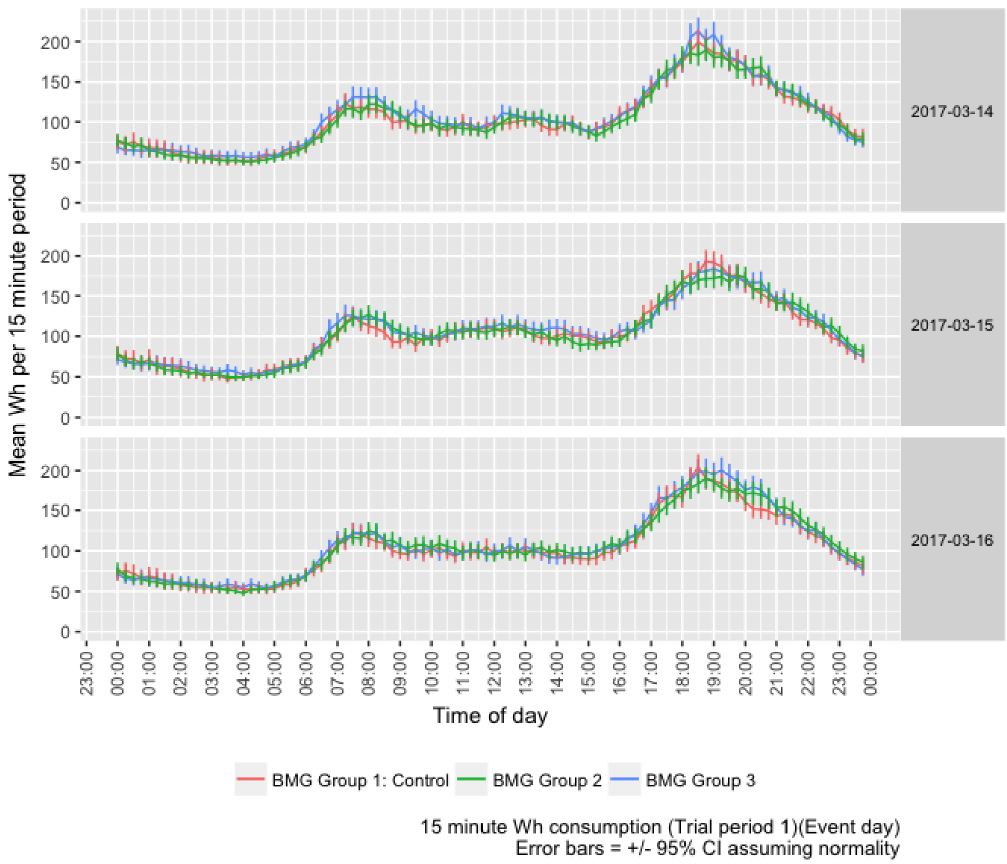

Figure 54 Mean 15 minute Wh consumption profile by trial group – including error bars

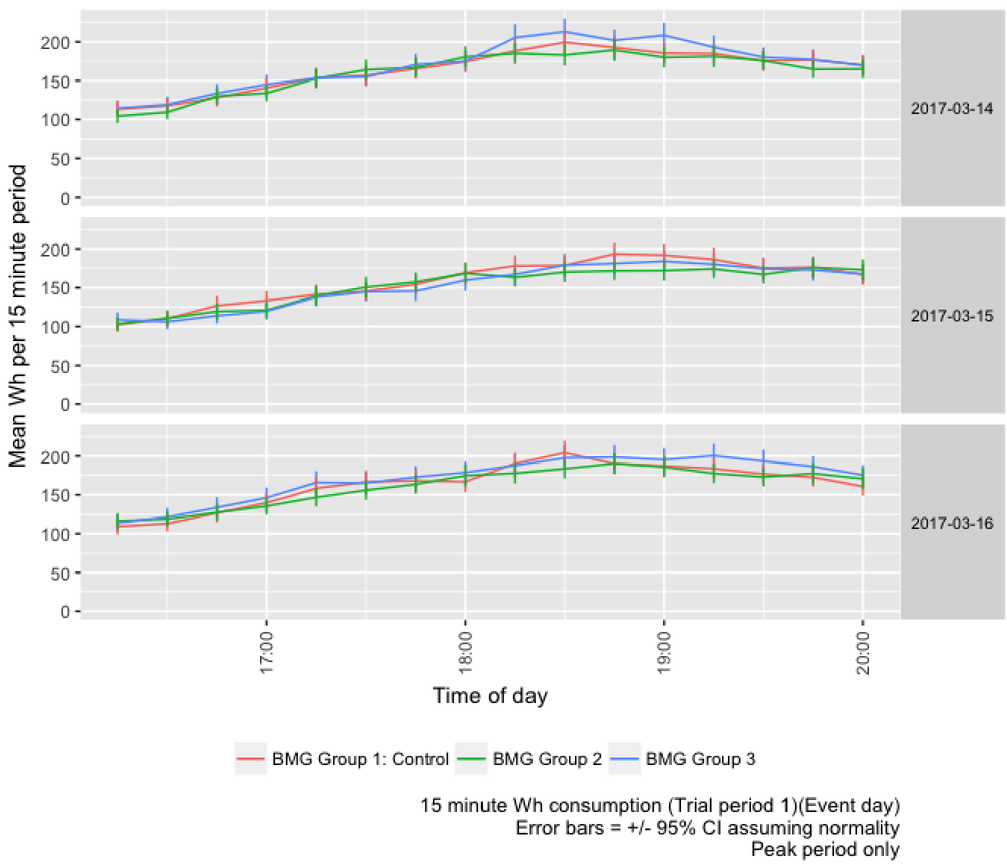

Figure 55 Mean 15 minute Wh peak hours profile by trial group – including error bars

A.3.5 Event day regression models

The following tables report the results of the regression models for each period. As a guide, any co-efficient that has more than one * in the table is statistically significant at the 95% level or greater.

A.3.5.1 Pre-peak model results

Table 47 Event day pre-peak period models results

	Dependent variable:			
	logMeanWh			
	Model 1	Model 2	Model 3	Model 4
Group 2	-0.017 (-0.099, 0.065)	-0.025 (-0.115, 0.065)	0.027 (-0.393, 0.447)	0.069 (-0.357, 0.494)
Group 3	0.067 (-0.018, 0.152)	0.060 (-0.033, 0.152)	-0.091 (-0.529, 0.346)	-0.095 (-0.541, 0.350)
openedEdEmail Not applicable				
openedEdEmail Yes		0.019 (-0.066, 0.104)	0.022 (-0.074, 0.118)	0.021 (-0.075, 0.117)
ecoMean			-0.151** (-0.248, -0.053)	-0.150** (-0.247, -0.052)
Group 2: ecoMean			-0.021 (-0.158, 0.117)	-0.022 (-0.160, 0.115)
Group 3: ecoMean			0.059 (-0.086, 0.204)	0.062 (-0.083, 0.207)
presenceChildren 1+ child				0.102 (-0.046, 0.250)
Group 2: presenceChildren 1+ child				-0.130 (-0.336, 0.077)
Group 3: presenceChildren 1+ child				-0.023 (-0.239, 0.193)
Constant	4.184*** (4.125, 4.243)	4.184*** (4.125, 4.243)	4.645*** (4.351, 4.938)	4.613*** (4.316, 4.910)
Observations	2,859	2,859	2,199	2,199
R^2	0.001	0.002	0.014	0.015
Adjusted R^2	0.001	0.0005	0.011	0.011
Residual Std. Error	0.933 (df = 2856)	0.933 (df = 2855)	0.927 (df = 2192)	0.927 (df = 2189)
F Statistic	2.088 (df = 2; 2856)	1.453 (df = 3; 2855)	5.148*** (df = 6; 2192)	3.759*** (df = 9; 2189)
Note:	*p<0.05;* ***p<0.01;*** p<0.001			

A.3.5.2 Peak model results I: 16:00 – 20:00

Table 48 Event day peak period models results

	Dependent variable:			
	logMeanWh			
	Model 1	Model 2	Model 3	Model 4
	(1)	(2)	(3)	(4)
Group 2	-0.036 (-0.110, 0.038)	-0.030 (-0.111, 0.051)	-0.084 (-0.448, 0.280)	-0.072 (-0.435, 0.291)
Group 3	-0.019 (-0.096, 0.057)	-0.014 (-0.096, 0.069)	-0.160 (-0.539, 0.219)	-0.189 (-0.569, 0.192)
openedEdEmail Not applicable				
openedEdEmail Yes		-0.014 (-0.090, 0.062)	-0.001 (-0.085, 0.082)	-0.004 (-0.086, 0.078)

ecoMean			-0.230*** (-0.315, -0.146)	-0.227*** (-0.310, -0.143)
Group 2: ecoMean			0.016 (-0.103, 0.136)	0.019 (-0.099, 0.137)
Group 3: ecoMean			0.047 (-0.079, 0.172)	0.059 (-0.065, 0.183)
presenceChildren 1+ child				0.339*** (0.213, 0.466)
Group 2: presenceChildren 1+ child				-0.063 (-0.239, 0.114)
Group 3: presenceChildren 1+ child				-0.042 (-0.227, 0.142)
Constant	4.759*** (4.706, 4.812)	4.759*** (4.706, 4.812)	5.424*** (5.170, 5.678)	5.320*** (5.066, 5.574)
Observations	2,859	2,859	2,197	2,197
R^2	0.0003	0.0004	0.031	0.059
Adjusted R^2	-0.0004	-0.001	0.028	0.055
Residual Std. Error	0.836 (df = 2856)	0.836 (df = 2855)	0.804 (df = 2190)	0.792 (df = 2187)
F Statistic	0.460 (df = 2; 2856)	0.351 (df = 3; 2855)	11.596*** (df = 6; 2190)	15.275*** (df = 9; 2187)
Note:	*p<0.05;* ***p<0.01;*** p<0.001			

A.3.5.3 Peak model results II: 18:00 – 20:00

Table 49 Event day peak period models results: 18:00 – 20:00

	Dependent variable:		
	logMeanWh		
	Model 1a	Model 2a	Model 3a
	(1)	(2)	(3)
Group 2	-0.055 (-0.134, 0.025)	-0.040 (-0.127, 0.047)	-0.140 (-0.534, 0.254)
Group 3	-0.027 (-0.110, 0.056)	-0.014 (-0.103, 0.075)	-0.082 (-0.492, 0.329)
openedEdEmail Not applicable			
openedEdEmail Yes		-0.033 (-0.116, 0.049)	-0.014 (-0.105, 0.076)
ecoMean			-0.247*** (-0.338, -0.155)
Group 2: ecoMean			0.036 (-0.093, 0.165)
Group 3: ecoMean			0.021 (-0.115, 0.157)
Constant	4.847*** (4.790, 4.904)	4.847*** (4.790, 4.904)	5.543*** (5.268, 5.818)
Observations	2,858	2,858	2,196
R^2	0.001	0.001	0.031
Adjusted R^2	-0.0001	-0.0002	0.028
Residual Std. Error	0.902 (df = 2855)	0.902 (df = 2854)	0.870 (df = 2189)
F Statistic	0.910 (df = 2; 2855)	0.818 (df = 3; 2854)	11.525*** (df = 6; 2189)
Note:	*p<0.05;* ***p<0.01;*** p<0.001		

A.3.5.4 Post-peak model results

Table 50 Event day post-peak period models results

	Dependent variable:		
	logMeanWh		
	Model 1	Model 2	Model 3
	(1)	(2)	(3)
Group 2	0.037 (-0.038, 0.112)	-0.040 (-0.127, 0.047)	0.169 (-0.207, 0.545)
Group 3	0.066 (-0.012, 0.144)	-0.014 (-0.103, 0.075)	0.137 (-0.254, 0.529)
openedEdEmail Not applicable			
openedEdEmail Yes		-0.033 (-0.116, 0.049)	-0.0005 (-0.087, 0.086)
ecoMean			-0.186*** (-0.273, -0.099)
Group 2: ecoMean			-0.045 (-0.168, 0.078)
Group 3: ecoMean			-0.028 (-0.158, 0.102)
Constant	4.480*** (4.426, 4.533)	4.847*** (4.790, 4.904)	5.013*** (4.751, 5.275)
Observations	2,860	2,858	2,200
R^2	0.001	0.001	0.029
Adjusted R^2	0.0003	-0.0002	0.026
Residual Std. Error	0.849 (df = 2857)	0.902 (df = 2854)	0.834 (df = 2193)
F Statistic	1.396 (df = 2; 2857)	0.818 (df = 3; 2854)	10.888*** (df = 6; 2193)
Note:	*p<0.05;* ***p<0.01;*** p<0.001		

A.3.6 Time-use diary analysis

Table 51 Summary of 'energy acts' (at home)

period16_20	Trial group	sum Act	Mean HeActs	Sd HeActs	Obs	nDiaries	ci_upper	ci_lower
04:00 - 12:00	Group 1	435	0.6016598	0.4898951	1573	314	0.6258526	0.5774669
04:00 - 12:00	Group 3	377	0.5808937	0.4937934	1381	269	0.6069169	0.5548704
04:00 - 12:00	Group 4	423	0.5708502	0.4952891	1523	298	0.5957081	0.5459923
12:00 - 16:00	Group 1	137	0.4724138	0.5001014	628	219	0.5114598	0.4333678
12:00 - 16:00	Group 3	83	0.3807339	0.4866848	499	162	0.4233376	0.3381303
12:00 - 16:00	Group 4	127	0.4409722	0.4973677	625	206	0.4798974	0.4020471
16:00 - 20:00	Group 1	263	0.5680346	0.4958855	1059	292	0.5978686	0.5382005
16:00 - 20:00	Group 3	210	0.5198020	0.5002272	846	249	0.5534680	0.4861359
16:00 - 20:00	Group 4	266	0.5518672	0.4978192	1005	280	0.5826130	0.5211214
20:00 - 00:00	Group 1	88	0.4782609	0.5008902	622	290	0.5175174	0.4390043
20:00 - 00:00	Group 3	70	0.4294479	0.4965228	582	257	0.4696630	0.3892327
20:00 - 00:00	Group 4	70	0.3954802	0.4903407	641	282	0.4333320	0.3576284
Other	Group 1	4	0.4000000	0.5163978	77	59	0.5094229	0.2905771
Other	Group 3	1	0.1428571	0.3779645	43	34	0.2474475	0.0382668
Other	Group 4	3	0.3750000	0.5175492	47	34	0.5134060	0.2365940

First, the distributions of the underlying % of households with n electricity acts is tested in each period at household level to assess which modelling approach to use. The distribution of electricity-using acts at home is in the form of a Pareto curve, as shown in Figure 56.

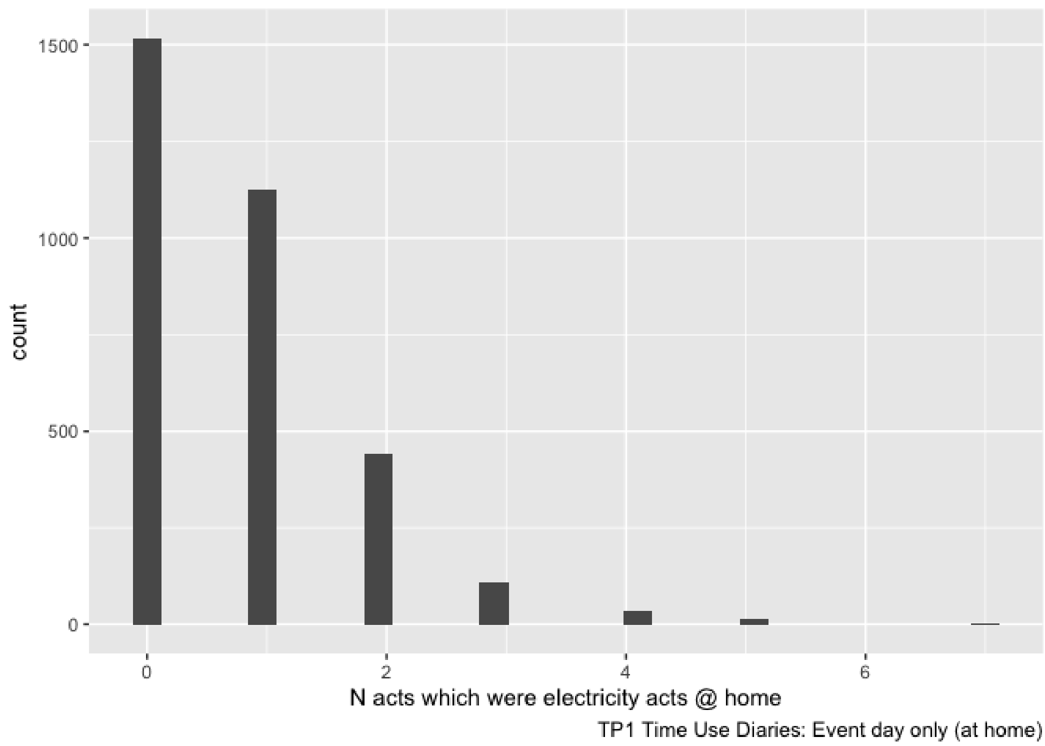

Figure 56 Distribution of electricity-using acts at home

Following the format of the analysis using the event day consumption data, two Poisson regression models are used to test for treatment effects in the pre-peak, peak and post-peak hours on the event day. Only the trial groups are included as the predictor variable in the first model, whilst the second included a variable for the event day notification email. The results are presented in Table 52 and Figure 57 below.

A.3.7 Time-use diary regression model results

A.3.7.1 All 'electricity acts'

Table 52 SAVE TP1 Event Day Time Use Diary Poisson model: all 'electricity acts'

	Dependent variable: sumAct					
	Pre Peak Model		Peak Model		Post Peak Model	
	12:00 - 16:00 (1)	12:00 - 16:00 (2)	16:00 - 20:00 (1)	16:00 - 20:00 (2)	20:00 - 00:00 (1)	20:00 - 00:00 (2)
Group 3	-0.165 (-0.635, 0.305)	-0.453 (-1.088, 0.183)	-0.066 (-0.247, 0.116)	0.026 (-0.181, 0.234)	-0.108 (-0.422, 0.206)	-0.308 (-0.708, 0.092)
Group 4	0.059 (-0.349, 0.468)	0.059 (-0.349, 0.468)	0.053 (-0.117, 0.224)	0.053 (-0.117, 0.224)	-0.201 (-0.515, 0.113)	-0.201 (-0.515, 0.113)
openedEdEmail NA						
openedEdEmail Yes		0.575 (-0.173, 1.324)		-0.239 (-0.521, 0.043)		0.420 (-0.049, 0.890)
Constant	-1.221*** (-1.510, -0.932)	-1.221*** (-1.510, -0.932)	-0.105 (-0.225, 0.016)	-0.105 (-0.225, 0.016)	-1.193*** (-1.401, -0.984)	-1.193*** (-1.401, -0.984)
Observations	415	415	821	821	829	829
Log	-278.256	-277.106	-961.215	-959.811	-543.246	-541.705

Likelihood						
Akaike Inf. Crit.	562.513	562.211	1,928.431	1,927.621	1,092.493	1,091.410
Note:						*p<0.05;* ***p<0.01;*** p<0.001

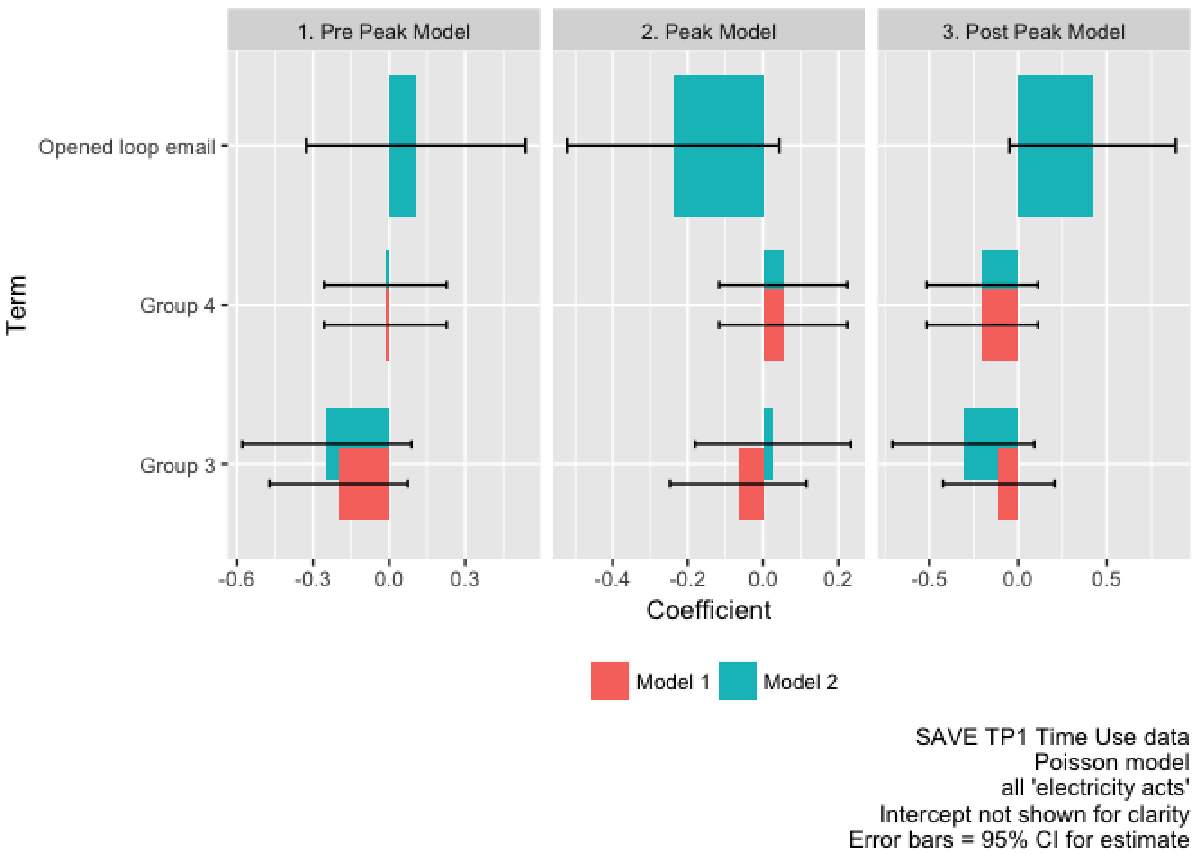

Figure 57 Event Day Time Use Diary Poisson model results: all 'electricity acts'

A.3.7.2 Acts involving cooking

Table 53 Event Day Time Use Diary Poisson model results: Acts involving cooking

	Dependent variable: sumAct					
	Pre Peak Model 12:00 - 16:00 (1)	12:00 - 16:00 (2)	Peak Model 16:00 - 20:00 (1)	16:00 - 20:00 (2)	Post Peak Model 20:00 - 00:00 (1)	20:00 - 00:00 (2)
Group 3	-0.069 (-0.380, 0.243)	-0.091 (-0.466, 0.284)	-0.005 (-0.207, 0.197)	0.015 (-0.221, 0.250)	-0.245 (-0.682, 0.193)	-0.511 (-1.089, 0.067)
Group 4	0.008 (-0.277, 0.293)	0.008 (-0.277, 0.293)	0.144 (-0.045, 0.333)	0.144 (-0.045, 0.333)	-0.175 (-0.593, 0.243)	-0.175 (-0.593, 0.243)
openedEdEmail NA						
openedEdEmail Yes		0.053 (-0.430, 0.537)		-0.047 (-0.350, 0.256)		0.542 (-0.135, 1.219)
Constant	-0.814*** (-1.013, -0.615)	-0.814*** (-1.013, -0.615)	-0.354*** (-0.491, -0.217)	-0.354*** (-0.491, -0.217)	-1.778*** (-2.058, -1.498)	-1.778*** (-2.058, -1.498)
Observations	587	587	821	821	829	829
Log Likelihood	-490.606	-490.583	-855.526	-855.479	-359.098	-357.854
Akaike Inf. Crit.	987.212	989.165	1,717.051	1,718.958	724.196	723.708

Note: *p<0.05;* ***p<0.01;*** p<0.001

Table 54 Summary of n acts by time period: acts involving cooking

nActs	12:00 – 16:00	16:00 – 20:00	20:00 – 00:00	Other
0	361	306	709	123
1	199	433	117	4
2	25	74	3	0
3	1	7	0	0
4	1	1	0	0
5	0	0	0	0

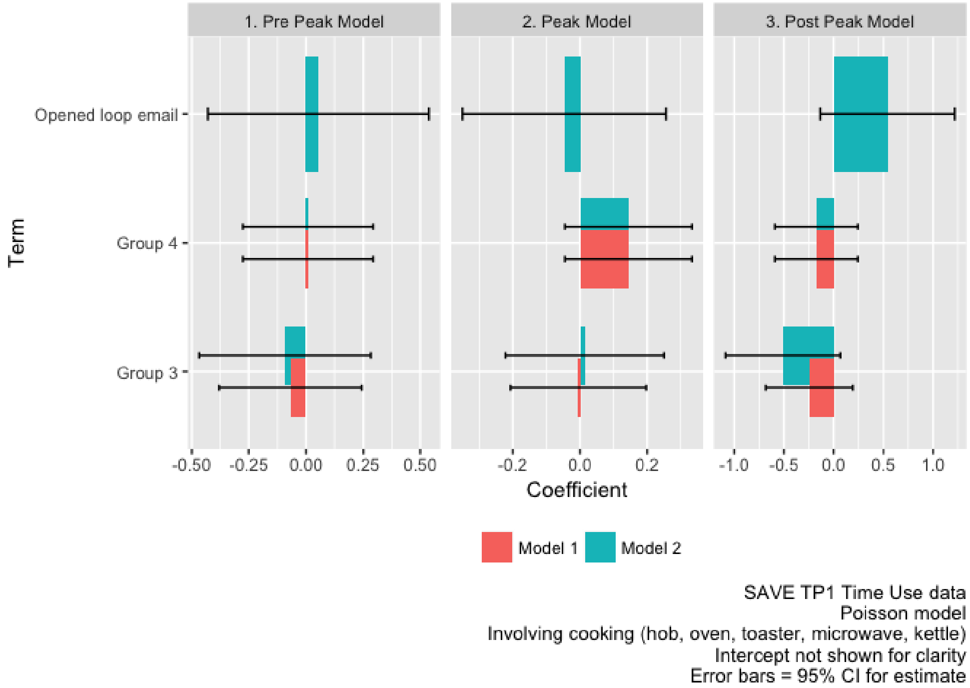

Figure 58 Event Day Time Use Diary Poisson model results: acts involving cooking

A.3.7.3 Dishwasher

Table 55 Event Day Time Use Diary Poisson model results: Using dishwasher

	Dependent variable: sumAct					
	Pre Peak Model		Peak Model		Post Peak Model	
	12:00 - 16:00 (1)	12:00 - 16:00 (2)	16:00 - 20:00 (1)	16:00 - 20:00 (2)	20:00 - 00:00 (1)	20:00 - 00:00 (2)
Group 3	0.301 (-2.470, 3.073)	0.846 (-1.926, 3.618)	-0.870 (-1.891, 0.151)	-0.336 (-1.358, 0.685)	-0.572 (-1.372, 0.228)	-1.119 (-2.341, 0.103)
Group 4	0.754 (-1.646, 3.155)	0.754 (-1.646, 3.155)	-0.295 (-1.106, 0.517)	-0.295 (-1.106, 0.517)	-0.783 (-1.616, 0.050)	-0.783 (-1.616, 0.050)
openedEdEmail NA						
openedEdEmail Yes		-15.759 (-3,709.758, 3,678.240)		-15.928 (-1,836.406, 1,804.549)		0.999 (-0.387, 2.385)
Constant	-5.389*** (-7.349, -	-5.389*** (-7.349, -	-3.038*** (-3.562, -	-3.038*** (-3.562, -	-2.780*** (-3.241, -	-2.780*** (-3.241, -

	3.429)	3.429)	2.514)	2.514)	2.318)	2.318)
Observations	587	587	821	821	829	829
Log Likelihood	-23.746	-23.202	-124.390	-121.721	-143.698	-142.624
Akaike Inf. Crit.	53.492	54.404	254.780	251.441	293.395	293.248
Note:						*p<0.05*; ***p<0.01***; p<0.001

Table 56 Summary of n acts by time period: Using dishwasher

nActs	04:00 - 12:00	12:00 - 16:00	16:00 - 20:00	20:00 - 00:00	Other
0	869	583	792	794	125
1	12	4	29	35	2

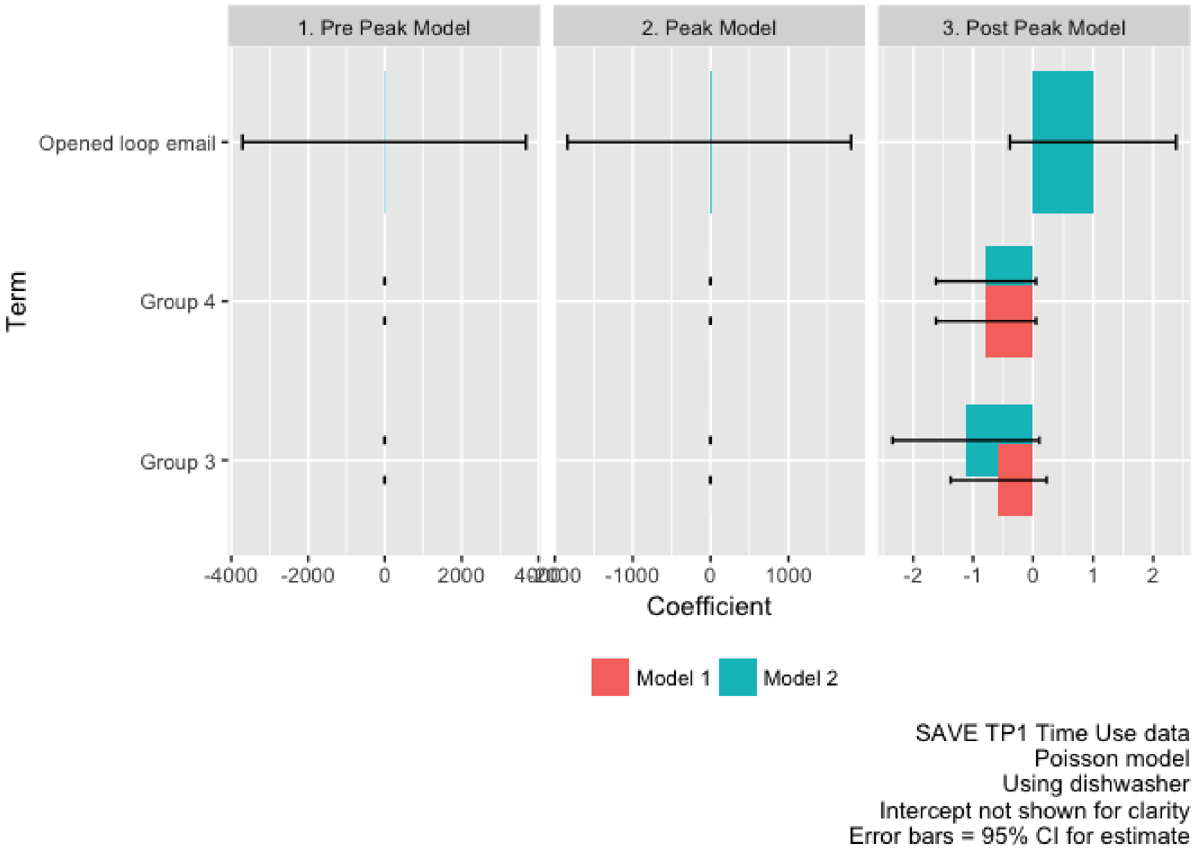

Figure 59 Event Day Time Use Diary Poisson model results: Using dishwasher

A.3.7.4 Laundry: washing machine

Table 57 Event Day Time Use Diary Poisson model results: laundry by machine

	Dependent variable: sumAct					
	Pre Peak Model		Peak Model		Post Peak Model	
	12:00 - 16:00 (1)	12:00 - 16:00 (2)	16:00 - 20:00 (1)	16:00 - 20:00 (2)	20:00 - 00:00 (1)	20:00 - 00:00 (2)
Group 3	-0.710 (-1.854, 0.434)	-0.859 (-2.366, 0.648)	-0.534 (-1.607, 0.540)	-0.223 (-1.383, 0.936)	17.852 (-4,844.503, 4,880.207)	18.404 (-4,843.951, 4,880.759)
Group 4	-0.391 (-1.338, 0.557)	-0.391 (-1.338, 0.557)	-1.162 (-2.452, 0.128)	-1.162 (-2.452, 0.128)	17.354 (-4,845.001, 4,879.709)	17.354 (-4,845.001, 4,879.709)
openedEdEmail NA						
openedEdEmail Yes		0.324 (-1.636, 2.284)		-1.037 (-3.228, 1.153)		-18.404 (-7,949.486, 7,912.678)
Constant	-2.991*** (-3.582, -2.400)	-2.991*** (-3.582, -2.400)	-3.374*** (-3.994, -2.754)	-3.374*** (-3.994, -2.754)	-22.303 (-4,884.658, 4,840.052)	-22.303 (-4,884.658, 4,840.052)
Observations	587	587	821	821	829	829
Log Likelihood	-93.382	-93.330	-84.890	-84.374	-28.249	-26.593
Akaike Inf. Crit.	192.764	194.659	175.781	176.748	62.498	61.187
Note:						*p<0.05;* **p<0.01;** p<0.001

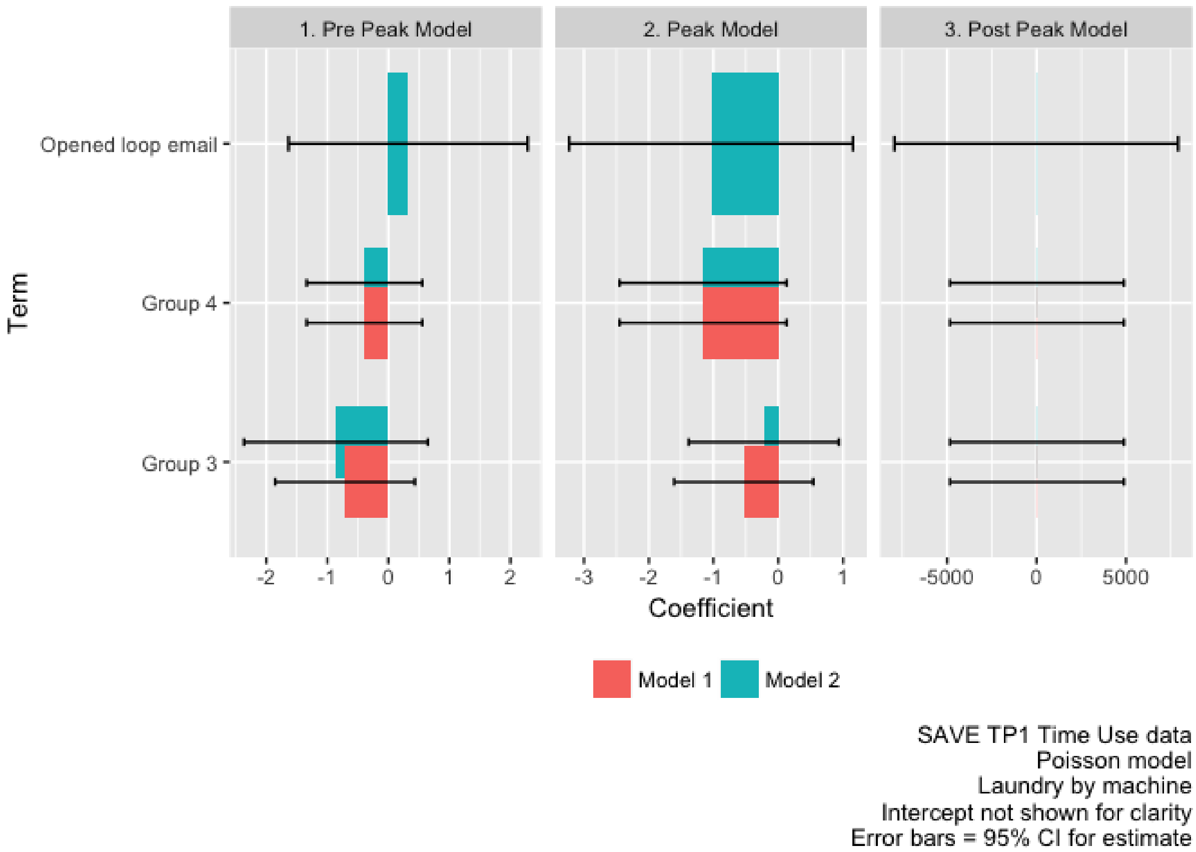

Figure 60 Event Day Time Use Diary Poisson model results: laundry by machine

Table 58 Summary of n acts by time period: laundry by machine

nActs	04:00 - 12:00	12:00 - 16:00	16:00 - 20:00	20:00 - 00:00	Other
0	820	565	803	824	127
1	60	22	18	5	0
2	1	0	0	0	0

A.3.7.5 Tumble drying

Table 59 Event Day Time Use Diary Poisson model results: tumble drying

	Dependent variable: sumAct					
	Pre Peak Model		Peak Model		Post Peak Model	
	12:00 - 16:00 (1)	12:00 - 16:00 (2)	16:00 - 20:00 (1)	16:00 - 20:00 (2)	20:00 - 00:00 (1)	20:00 - 00:00 (2)
Group 3	-16.523 (-2,409.804, 2,376.758)	-16.523 (-3,158.386, 3,125.340)	-17.235 (-3,199.958, 3,165.487)	-17.235 (-4,173.680, 4,139.209)	0.121 (-1.839, 2.081)	-0.020 (-2.421, 2.380)
Group 4	0.531 (-0.586, 1.649)	0.531 (-0.586, 1.649)	-0.874 (-2.514, 0.765)	-0.874 (-2.514, 0.765)	-17.326 (-4,948.168, 4,913.517)	-17.326 (-4,948.168, 4,913.517)
openedEdEmail NA						
openedEdEmail Yes		-0.000 (-4,849.426, 4,849.426)		-0.000 (-6,462.537, 6,462.537)		0.306 (-2.466, 3.078)
Constant	-3.780*** (-4.656, -2.903)	-3.780*** (-4.656, -2.903)	-4.067*** (-4.944, -3.191)	-4.067*** (-4.944, -3.191)	-4.977*** (-6.363, -3.591)	-4.977*** (-6.363, -3.591)
Observations	587	587	821	821	829	829
Log Likelihood	-57.886	-57.886	-37.220	-37.220	-23.665	-23.642
Akaike Inf. Crit.	121.771	123.771	80.440	82.440	53.331	55.284
Note:						*p<0.05;* ***p<0.01;*** p<0.001

Table 60 Summary of n acts by time period: Tumble drying

nActs	12:00 – 16:00	16:00 – 20:00	20:00 – 00:00	Other
0	574	814	825	127
1	13	7	4	0
2	0	0	0	0

A.3.7.6 Acts at home

Table 61 SAVE TP1 Event Day Time Use Diary Poisson model: n acts at home

	Dependent variable:					
	sumAct					
	14:00 - 16:00	14:00 - 16:00	16:00 - 20:00	16:00 - 20:00	20:00 - 00:00	20:00 - 00:00
	(1)	(2)	(3)	(4)	(5)	(6)
Group 3	0.061 (-0.155, 0.277)	-0.052 (-0.321, 0.217)	-0.111* (-0.214, -0.009)	-0.048 (-0.167, 0.071)	0.008 (-0.112, 0.128)	0.008 (-0.134, 0.149)
Group 4	-0.011 (-0.216, 0.193)	-0.011 (-0.216, 0.193)	-0.028 (-0.126, 0.069)	-0.028 (-0.126, 0.069)	0.037 (-0.080, 0.153)	0.037 (-0.080, 0.153)
openedEdEmail NA						
openedEdEmail Yes		0.246 (-0.079, 0.572)		-0.162* (-0.322, -0.002)		0.001 (-0.176, 0.178)
Constant	0.197** (0.055, 0.339)	0.197** (0.055, 0.339)	1.055*** (0.988, 1.123)	1.055*** (0.988, 1.123)	0.662*** (0.579, 0.744)	0.662*** (0.579, 0.744)
Observations	415	415	821	821	829	829
Log Likelihood	-555.240	-554.145	-1,597.645	-1,595.653	-1,314.144	-1,314.144
Akaike Inf. Crit.	1,116.479	1,116.291	3,201.290	3,199.306	2,634.288	2,636.288
Note:	*p<0.05;* ***p<0.01;*** p<0.001					

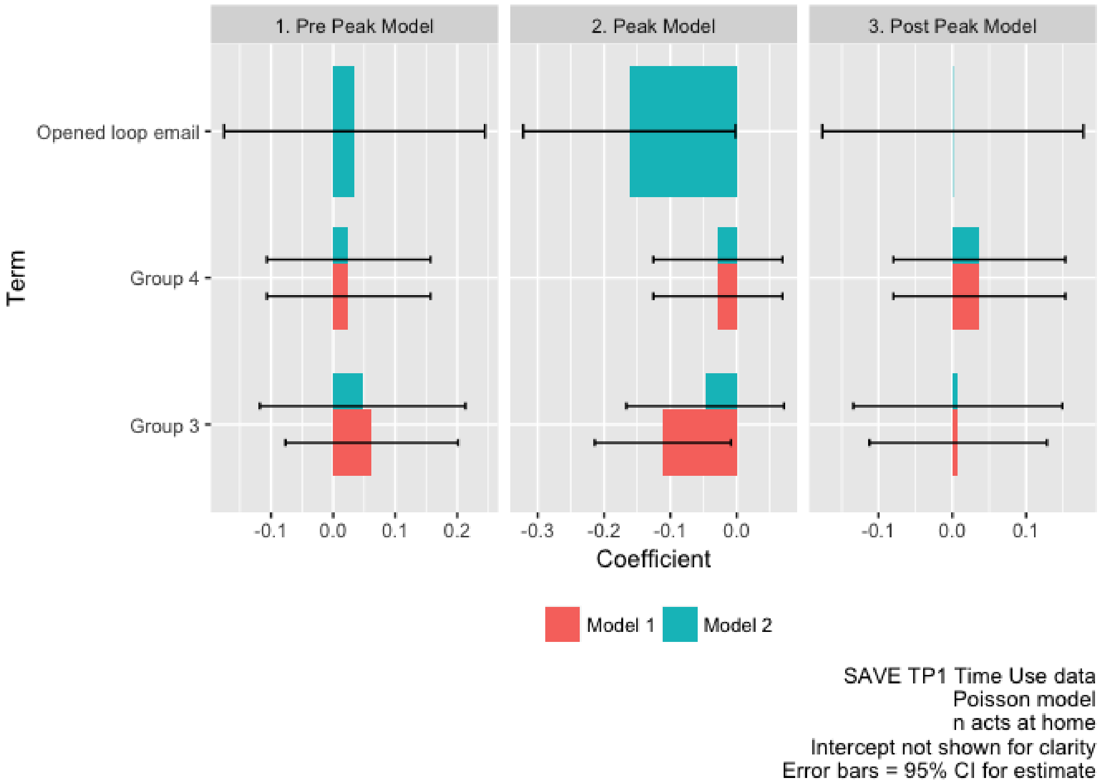

Figure 61 Event Day Time Use Diary Poisson model coefficientsl: number of acts at home, where bar colour refers to model number and term denotes trial groups and dummy variable for households opening event notification email

A.4 Spatial microsimulation

A.4.1 Select constraint variables: consumption histograms

Histograms of mean half-hourly consumption are produced to check the distribution (see below).

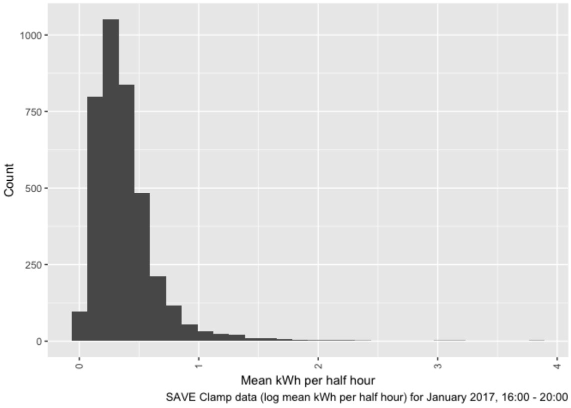

Figure 62 Histogram of mean values for half-hourly consumption (mean Wh)

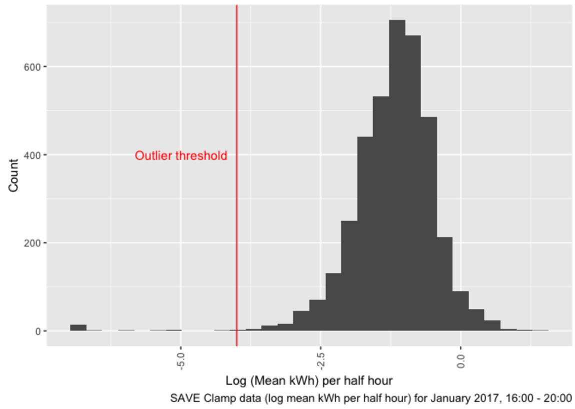

Figure 63 Histogram of log-transformed half-hourly consumption (log mean Wh)

Clearly the untransformed consumption values are extremely skewed, therefore the log(mean) Wh values are used to avoid the problem of a non-normal dependent variable. The transformed values are also cut at -4 (see histogram above) to avoid the outliers with non-zero but extremely low consumption values.

A.4.1.1 Regression model results

Table 62 IPF constraint regression model results

	Dependent variable:	
	mod1	mod2
	(1)	(2)
ba_censusNpeople.latest2	0.424*** (0.360, 0.488)	0.435*** (0.370, 0.500)
ba_censusNpeople.latest3	0.668*** (0.583, 0.754)	0.665*** (0.578, 0.752)
ba_censusNpeople.latest4	0.803*** (0.705, 0.900)	0.784*** (0.684, 0.883)
ba_censusNpeople.latest5	0.996*** (0.872, 1.120)	1.000*** (0.873, 1.127)
ba_censusNpeople.latest6	1.063*** (0.884, 1.243)	1.080*** (0.896, 1.264)
ba_censusNpeople.latest7	0.993*** (0.706, 1.279)	0.936*** (0.640, 1.232)
ba_censusNpeople.latest8+	0.826*** (0.412, 1.241)	0.715*** (0.291, 1.139)
ba_presenceChildren.latest1+ child	-0.081** (-0.155, -0.008)	-0.074* (-0.150, 0.001)
ba_heatSourceReducedGas boiler	-0.332*** (-0.463, -0.201)	
ba_heatSourceReducedOil/wood/solid fuels/biomass/other	-0.091 (-0.229, 0.047)	
ba_heatSourceReducedOther electricity (e.g. heat pump)	0.324** (0.044, 0.603)	
ba_Q3_6_mainsGasLPG gas	0.625** (0.112, 1.139)	
ba_Q3_6_mainsGasMains gas	0.427* (-0.024, 0.878)	
ba_Q3_6_mainsGasNo gas	0.687*** (0.218, 1.157)	
ba_Q8_2_dwellingRecodedSemi-detached	-0.049* (-0.102, 0.004)	-0.060** (-0.114, -0.005)
ba_Q8_2_dwellingRecodedTerrace or end terrace	-0.059* (-0.119, 0.0004)	-0.062** (-0.123, -0.001)
ba_Q8_2_dwellingRecodedFlat	-0.106* (-0.223, 0.011)	-0.016 (-0.132, 0.101)
ba_Q8_2_dwellingRecodedCommercial	-0.367 (-0.850, 0.117)	-0.249 (-0.748, 0.250)
ba_Q8_2_dwellingRecodedCaravan etc	-0.175 (-0.580, 0.229)	0.244 (-0.150, 0.639)
ba_Q8_2_dwellingRecodedRefused	-0.414** (-0.825, -0.003)	-0.425* (-0.850, 0.0003)
ba_censusTenurePrivate rent	-0.079* (-0.166, 0.008)	-0.048 (-0.137, 0.041)
ba_censusTenureSocial rent	-0.038 (-0.113, 0.037)	-0.062 (-0.137, 0.014)
ba_censusTenurex.Refused/dk/Other	-0.175* (-0.371, 0.021)	-0.055 (-0.253, 0.143)
ba_censusNcars1	-0.015 (-0.110, 0.079)	-0.001 (-0.098, 0.095)
ba_censusNcars2	0.095* (-0.006, 0.197)	0.118** (0.014, 0.222)

ba_censusNcars3	0.133** (0.015, 0.251)	0.163*** (0.042, 0.283)
ba_censusNcars4+	0.099 (-0.029, 0.226)	0.170** (0.040, 0.300)
ba_Q8_27_Income£10-£20k	0.106* (-0.002, 0.215)	
ba_Q8_27_Income£20-£30k	0.114** (0.002, 0.226)	
ba_Q8_27_Income£30-£40k	0.104* (-0.015, 0.223)	
ba_Q8_27_Income£40-£50k	0.118* (-0.007, 0.243)	
ba_Q8_27_Income£50-60k	0.085 (-0.048, 0.218)	
ba_Q8_27_Income£60-80k	0.135** (0.008, 0.262)	
ba_Q8_27_Income> £80k	0.254*** (0.122, 0.386)	
ba_Q8_27_IncomeDon't know	0.138** (0.023, 0.252)	
ba_Q8_27_IncomeRefused	0.156*** (0.051, 0.261)	
ba_Q22_HRPethnicityMixed	-0.217 (-0.491, 0.056)	-0.215 (-0.489, 0.059)
ba_Q22_HRPethnicityAsian/Asian British	-0.275*** (-0.423, -0.127)	-0.265*** (-0.417, -0.114)
ba_Q22_HRPethnicityBlack/Black British	0.137 (-0.112, 0.386)	0.086 (-0.167, 0.339)
ba_Q22_HRPethnicityOther	-0.152 (-0.430, 0.125)	-0.177 (-0.464, 0.110)
ba_Q22_HRPethnicityRefused	0.006 (-0.143, 0.154)	0.027 (-0.124, 0.179)
ba_Q2D_HRPemplType.latestHRP in part-time employment (8-29 hours/week)	-0.006 (-0.073, 0.060)	-0.015 (-0.082, 0.053)
ba_Q2D_HRPemplType.latestHRP retired	-0.068* (-0.147, 0.011)	-0.119*** (-0.198, -0.039)
ba_Q2D_HRPemplType.latestHRP self-employed (unkown hours)	-0.024 (-0.110, 0.063)	0.003 (-0.085, 0.092)
ba_Q2D_HRPemplType.latestOther	-0.020 (-0.101, 0.061)	-0.039 (-0.119, 0.042)
ba_Q2D_HRPemplType.latestUnemployed	0.076 (-0.070, 0.223)	0.055 (-0.096, 0.206)
ba_Q2B_HRPage.fullSurvey25 - 34	0.018 (-0.178, 0.213)	-0.024 (-0.222, 0.175)
ba_Q2B_HRPage.fullSurvey35 - 44	0.149 (-0.042, 0.341)	0.089 (-0.106, 0.283)
ba_Q2B_HRPage.fullSurvey45 - 54	0.100 (-0.088, 0.288)	0.050 (-0.141, 0.241)
ba_Q2B_HRPage.fullSurvey55 - 64	0.171* (-0.020, 0.361)	0.135 (-0.058, 0.328)
ba_Q2B_HRPage.fullSurvey65 - 74	0.175* (-0.024, 0.374)	0.155 (-0.047, 0.358)
ba_Q2B_HRPage.fullSurvey75+	0.138 (-0.068, 0.345)	0.134 (-0.077, 0.344)
ba_Q2B_HRPage.fullSurveyRefused	0.060 (-0.389, 0.509)	-0.084 (-0.518, 0.350)
Constant	-1.872*** (-2.781, -0.964)	-1.441*** (-2.246, -0.636)
Observations	2,801	2,816
R^2	0.352	0.302
Adjusted R^2	0.338	0.290
Residual Std. Error	0.534 (df = 2739)	0.554 (df = 2769)
F Statistic	24.401*** (df = 61; 2739)	26.008*** (df = 46; 2769)
Note:	*p<0.1;* ***p<0.05;*** p<0.01	

A.4.1.2 Stepwise regression results: Model 1 – all constraints

```
## Start:  AIC=-3450.12
## log(meanHalfHourkWh) ~ ba_censusNpeople.latest + 
ba_presenceChildren.latest + 
##     ba_nrooms + ba_heatSourceReduced + ba_Q3_6_mainsGas + 
ba_Q8_2_dwellingRecoded + 
##     ba_censusTenure + ba_censusNcars + ba_Q8_27_Income + 
ba_Q22_HRPethnicity + 
##     ba_Q2D_HRPemplType.latest + ba_Q2B_HRPage.fullSurvey
## 
##                              Df Sum of Sq    RSS     AIC
## - ba_Q2D_HRPemplType.latest   5     1.337 783.23 -3455.3
## - ba_Q8_2_dwellingRecoded     6     2.978 784.87 -3451.5
## <none>                                    781.89 -3450.1
## - ba_censusTenure             3     1.715 783.61 -3450.0
## - ba_Q8_27_Income             9     5.486 787.38 -3448.5
## - ba_presenceChildren.latest  1     1.337 783.23 -3447.3
## - ba_Q2B_HRPage.fullSurvey    7     4.719 786.61 -3447.3
## - ba_Q22_HRPethnicity         5     5.132 787.02 -3441.8
## - ba_Q3_6_mainsGas            3     5.467 787.36 -3436.6
## - ba_censusNcars              4     6.558 788.45 -3434.7
## - ba_heatSourceReduced        3    12.058 793.95 -3413.3
## - ba_nrooms                   8    17.508 799.40 -3404.1
## - ba_censusNpeople.latest     7    99.986 881.88 -3127.1
## 
## Step:  AIC=-3455.34
```

```
## log(meanHalfHourkWh) ~ ba_censusNpeople.latest +
ba_presenceChildren.latest +
##     ba_nrooms + ba_heatSourceReduced + ba_Q3_6_mainsGas +
ba_Q8_2_dwellingRecoded +
##     ba_censusTenure + ba_censusNcars + ba_Q8_27_Income +
ba_Q22_HRPethnicity +
##     ba_Q2B_HRPage.fullSurvey
##
##                              Df Sum of Sq    RSS     AIC
## - ba_Q8_2_dwellingRecoded     6     2.932 786.16 -3456.9
## - ba_censusTenure             3     1.673 784.90 -3455.4
## <none>                                    783.23 -3455.3
## - ba_Q2B_HRPage.fullSurvey    7     4.262 787.49 -3454.1
## - ba_presenceChildren.latest  1     1.331 784.56 -3452.6
## - ba_Q8_27_Income             9     5.848 789.08 -3452.5
## + ba_Q2D_HRPemplType.latest   5     1.337 781.89 -3450.1
## - ba_Q22_HRPethnicity         5     5.108 788.34 -3447.1
## - ba_Q3_6_mainsGas            3     5.601 788.83 -3441.4
## - ba_censusNcars              4     6.851 790.08 -3438.9
## - ba_heatSourceReduced        3    12.047 795.27 -3418.6
## - ba_nrooms                   8    17.370 800.60 -3409.9
## - ba_censusNpeople.latest     7    99.852 883.08 -3133.2
##
## Step:  AIC=-3456.88
## log(meanHalfHourkWh) ~ ba_censusNpeople.latest +
ba_presenceChildren.latest +
##     ba_nrooms + ba_heatSourceReduced + ba_Q3_6_mainsGas + ba_censusTenure
+
##     ba_censusNcars + ba_Q8_27_Income + ba_Q22_HRPethnicity +
##     ba_Q2B_HRPage.fullSurvey
##
##                              Df Sum of Sq    RSS     AIC
## <none>                                    786.16 -3456.9
## + ba_Q8_2_dwellingRecoded     6     2.932 783.23 -3455.3
## - ba_Q2B_HRPage.fullSurvey    7     4.384 790.54 -3455.3
## - ba_censusTenure             3     2.190 788.35 -3455.1
## - ba_presenceChildren.latest  1     1.345 787.50 -3454.1
## - ba_Q8_27_Income             9     6.520 792.68 -3451.7
## + ba_Q2D_HRPemplType.latest   5     1.290 784.87 -3451.5
## - ba_Q22_HRPethnicity         5     5.162 791.32 -3448.5
## - ba_Q3_6_mainsGas            3     5.521 791.68 -3443.3
## - ba_censusNcars              4     7.728 793.89 -3437.5
## - ba_heatSourceReduced        3    11.546 797.71 -3422.0
## - ba_nrooms                   8    23.817 809.98 -3389.3
## - ba_censusNpeople.latest     7   100.765 886.92 -3133.1
```

```
## Stepwise Model Path
## Analysis of Deviance Table
##
## Initial Model:
## log(meanHalfHourkWh) ~ ba_censusNpeople.latest +
ba_presenceChildren.latest +
##     ba_nrooms + ba_heatSourceReduced + ba_Q3_6_mainsGas +
ba_Q8_2_dwellingRecoded +
##     ba_censusTenure + ba_censusNcars + ba_Q8_27_Income +
ba_Q22_HRPethnicity +
##     ba_Q2D_HRPemplType.latest + ba_Q2B_HRPage.fullSurvey
##
## Final Model:
```

```
## log(meanHalfHourkWh) ~ ba_censusNpeople.latest +
ba_presenceChildren.latest +
##     ba_nrooms + ba_heatSourceReduced + ba_Q3_6_mainsGas + ba_censusTenure
+
##     ba_censusNcars + ba_Q8_27_Income + ba_Q22_HRPethnicity +
##     ba_Q2B_HRPage.fullSurvey
##
##
##                          Step Df Deviance Resid. Df Resid. Dev       AIC
## 1                                              2739   781.8902 -3450.125
## 2 - ba_Q2D_HRPemplType.latest  5 1.336698      2744   783.2269 -3455.340
## 3   - ba_Q8_2_dwellingRecoded  6 2.931736      2750   786.1587 -3456.875
```

A.4.1.3 Stepwise regression results: Model 2 – OA level constraints only

```
## Start:  AIC=-3283.65
## log(meanHalfHourkWh) ~ ba_censusNpeople.latest +
ba_presenceChildren.latest +
##     ba_nrooms + ba_Q8_2_dwellingRecoded + ba_censusTenure + ba_censusNcars
+
##     ba_Q22_HRPethnicity + ba_Q2D_HRPemplType.latest +
ba_Q2B_HRPage.fullSurvey
##
##                              Df Sum of Sq    RSS     AIC
## - ba_censusTenure             3     0.954 849.58 -3286.5
## <none>                                    848.63 -3283.7
## - ba_Q8_2_dwellingRecoded     6     3.673 852.30 -3283.5
## - ba_Q2D_HRPemplType.latest   5     3.279 851.91 -3282.8
## - ba_presenceChildren.latest  1     1.138 849.77 -3281.9
## - ba_Q2B_HRPage.fullSurvey    7     4.823 853.45 -3281.7
## - ba_Q22_HRPethnicity         5     4.911 853.54 -3277.4
## - ba_censusNcars              4     9.036 857.67 -3261.8
## - ba_nrooms                   8    17.447 866.08 -3242.3
## - ba_censusNpeople.latest     7   103.205 951.84 -2974.5
##
## Step:  AIC=-3286.49
## log(meanHalfHourkWh) ~ ba_censusNpeople.latest +
ba_presenceChildren.latest +
##     ba_nrooms + ba_Q8_2_dwellingRecoded + ba_censusNcars +
ba_Q22_HRPethnicity +
##     ba_Q2D_HRPemplType.latest + ba_Q2B_HRPage.fullSurvey
##
##                              Df Sum of Sq    RSS     AIC
## <none>                                    849.58 -3286.5
## - ba_Q8_2_dwellingRecoded     6     3.966 853.55 -3285.4
## - ba_Q2D_HRPemplType.latest   5     3.456 853.04 -3285.1
## - ba_presenceChildren.latest  1     1.225 850.81 -3284.4
## - ba_Q2B_HRPage.fullSurvey    7     5.017 854.60 -3283.9
## + ba_censusTenure             3     0.954 848.63 -3283.7
## - ba_Q22_HRPethnicity         5     4.775 854.36 -3280.7
## - ba_censusNcars              4     9.746 859.33 -3262.4
## - ba_nrooms                   8    20.167 869.75 -3236.4
## - ba_censusNpeople.latest     7   102.622 952.21 -2979.4
## Stepwise Model Path
## Analysis of Deviance Table
##
## Initial Model:
## log(meanHalfHourkWh) ~ ba_censusNpeople.latest +
ba_presenceChildren.latest +
##     ba_nrooms + ba_Q8_2_dwellingRecoded + ba_censusTenure + ba_censusNcars
+
```

```
##     ba_Q22_HRPethnicity + ba_Q2D_HRPemplType.latest + 
ba_Q2B_HRPage.fullSurvey
##
## Final Model:
## log(meanHalfHourkWh) ~ ba_censusNpeople.latest + 
ba_presenceChildren.latest + 
##     ba_nrooms + ba_Q8_2_dwellingRecoded + ba_censusNcars + 
ba_Q22_HRPethnicity + 
##     ba_Q2D_HRPemplType.latest + ba_Q2B_HRPage.fullSurvey
##
##
##                Step Df  Deviance Resid. Df Resid. Dev       AIC
## 1                                     2769   848.6298 -3283.650
## 2 - ba_censusTenure  3 0.9540999      2772   849.5839 -3286.486
```

A.4.2 Constraint data setup

First, the constraint categories are prepared which are created from the aggregate level spatial (Census OA) data. The relevant OA level tables were downloaded from nomisweb (https://www.nomisweb.co.uk/census/2011/data_finder) and the key columns selected and saved to a new (clean) .csv file with simple column headings. The individual files are loaded and merged keeping only the OAs in the study area. The assembled constraint data table contains 6,136 rows corresponding to the individual OAs within the SAVE study area, with constraint sub-categories containing a minimum of 40 and a maximum of 314 households.

Checks for the correct number of households in the data are made by summing the counts within the variable categories (e.g. sumKids) and checking against the number of households expected in each OA (e.g. nHhsChildren). The check column (e.g. checkDiffKids in Table 63) reveals quickly if counts differ from those expected.

Table 63 Census data: presence of children

oaCode	nHsChildren	nchild_0	nchild_1m	sumKids	checkDiffKids
E00080972	113	75	38	113	0
E00080973	128	82	46	128	0
E00080974	132	90	42	132	0
E00080975	132	99	33	132	0
E00080976	132	71	61	132	0
E00080977	126	86	40	126	0

The process is repeated for each of the household and household representative person (HRP) attributes to be used as constraint variables. Finally, the overall household counts of all constraint variables are cross-referenced to check for anomalies. Figure 64 contains a matrix where each plot shows the household counts of the OAs for a constraint variable plotted against the household counts for each of the other constraints.

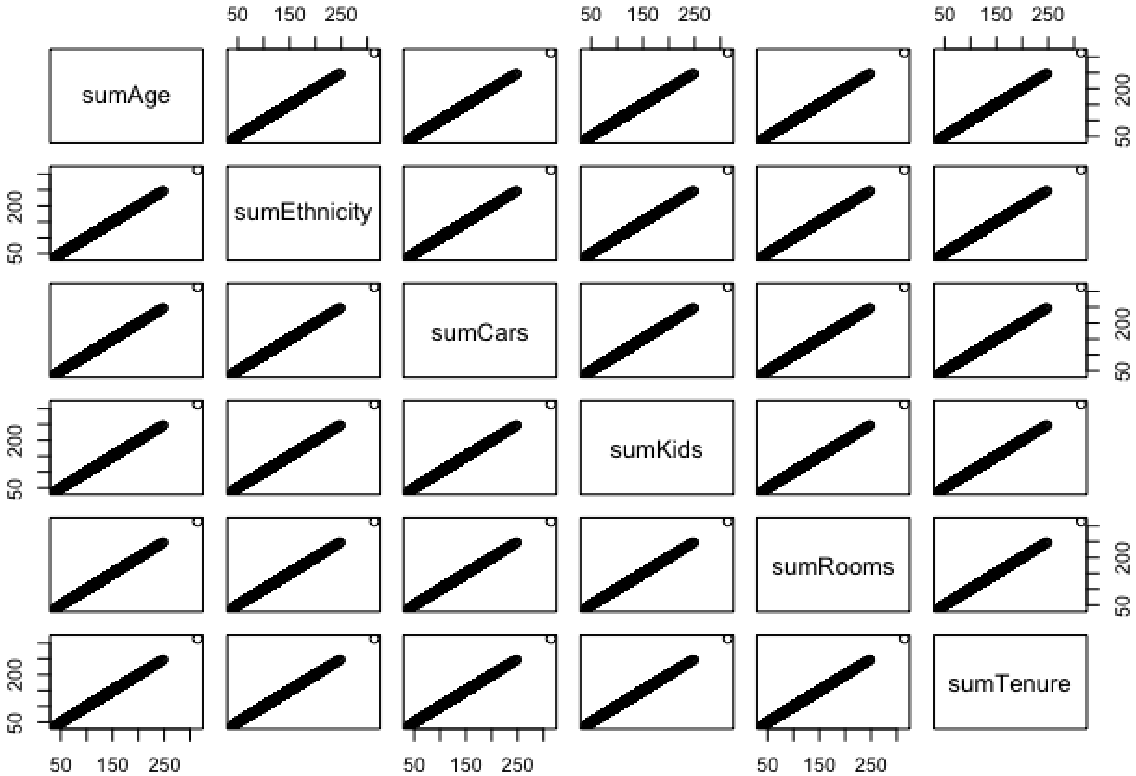

Figure 64 Household counts for constraint variables

As can be seen, the constraints always add up to the correct number of households (maximum difference in cell counts are reported as +/- 0/0).

A.4.3 IPF results

A.4.3.1 Baseline model

To produce correctly weighted household counts the weights are used as part of a survey (weighted) table calculation. As an example, Table 64 shows the weighted ethnic distributions by household size for the entire SAVE study region.

Table 64 Ethnicity vs n people (all households, all OAs, SAVE study area, weighted IPF results rounded to 2 d.p.)

	1	2	3	4	5	6	7	8+
White British/Irish	220422.12	277938.74	113785.08	98841.07	30294.45	8771.82	2244.16	901.03
Mixed	2042.17	2033.40	1950.90	755.06	0.00	0.00	0.00	0.00
Asian/Asian British	1022.96	4576.95	4909.48	3476.08	4257.12	1901.20	256.84	0.00
Black/Black British	619.04	1466.37	632.27	2127.88	841.03	403.98	0.00	406.97
Other	2185.72	332.55	273.27	174.90	221.40	0.00	0.00	0.00
Refused	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

Note that fractional counts result due to the weighted data. Based on this example table, the simulated household population is 790,066 households compared to the overall Census count of 790,066 households – an exact match.

The full results of the systematic validation tests, which calculate the sum of the differences between the weighted SAVE sample and Census data for each constraint sub-category, are reported in Table 65 and Table 66 below. Each row summarises the total absolute error for each of the constraint sub-categories (across all of the OAs); that is, the difference between the total simulated household count and the 'real' household count from the Census data for each OA.

The weighted counts are calculated for each OA for each constraint used and compared with the original constraint counts sourced from the Census.

This involves:

- creating a matrix of the weighted counts (for all areas) that result from the IPF
- joining this to the Census constraint data
- calculating the matrix difference

The errors for both the HRP (Table 65) and household attributes (Table 66).

Table 65 Summary of error table per HRP constraint category (all households, all OAs)

Constraint category	Min.	Mean	Max.
hrpWhite	Min. :-0.12817	Mean : 0.03577	Max. :64.02147
hrpMixed	Min. :-6.781432	Mean :-0.003664	Max. : 0.476174
hrpAsianBrit	Min. :-44.43772	Mean : -0.01571	Max. : 2.35721
hrpBlackBrit	Min. :-13.704297	Mean : -0.010019	Max. : 0.294116
hrpOtherEthnic	Min. :-12.165550	Mean : -0.006381	Max. : 0.060824
hrp16_24	Min. :-114.16045	Mean : -0.03452	Max. : 7.37570
hrp25_34	Min. :-9.27395	Mean : 0.01353	Max. :75.25399
hrp35_64	Min. :-2.109865	Mean : 0.019090	Max. :29.890387
hrp65plus	Min. :-0.889220	Mean : 0.001898	Max. : 9.016069
econActEmplFt	Min. :-1.772414	Mean : 0.007716	Max. :20.102178
econActEmplPt	Min. :-1.986239	Mean : 0.002069	Max. : 7.941040
econActSelfEmpl	Min. :-0.656312	Mean : 0.001961	Max. : 4.894664
econActUnemp	Min. :-1.959539	Mean : 0.001362	Max. : 4.604880
econActInactiveRetired	Min. :-0.371236	Mean : 0.000769	Max. : 4.603922
econActInactiveOther	Min. :-22.846876	Mean : -0.013877	Max. : 1.406713

Table 66 Summary of error table per household attribute constraint category (all households, all OAs)

Constraint category	Min.	Mean	Max.
tenureOwned	Min. :-2.092420	Mean : 0.006113	Max. :26.065992
tenureSocialRented	Min. :-0.36375	Mean : 0.01824	Max. :63.11844
tenurePrivateRented	Min. :-89.18443	Mean : -0.02436	Max. : 0.21777
cars_0	Min. :-4.106125	Mean : 0.007123	Max. :26.742084
cars_1	Min. :-21.087969	Mean : -0.008183	Max. : 4.224824
cars_2	Min. :-4.084654	Mean : 0.000082	Max. : 4.001130
cars_3	Min. :-3.755817	Mean :-0.000996	Max. : 0.256623
cars_4m	Min. :-0.002159	Mean : 0.001974	Max. : 5.520720
nchild_0	Min. :-3.788008	Mean : 0.002820	Max. : 4.498700
nchild_1m	Min. :-4.498700	Mean :-0.002820	Max. : 3.788008
npeople_1	Min. :-6.253e-13	Mean :-1.360e-15	Max. : 5.826e-13
npeople_2	Min. :-1.137e-12	Mean : 1.341e-15	Max. : 8.669e-13
npeople_3	Min. :-1.208e-13	Mean : 8.596e-16	Max. : 1.421e-13
npeople_4	Min. :-2.629e-13	Mean :-2.369e-16	Max. : 2.274e-13
npeople_5	Min. :-4.974e-14	Mean :-6.622e-17	Max. : 5.684e-14
npeople_6	Min. :-7.105e-15	Mean : 1.800e-20	Max. : 1.421e-14

npeople_7	Min. :-1.776e-15	Mean :-1.013e-18	Max. : 1.776e-15
npeople_8m	Min. :-1.776e-15	Mean :-2.262e-18	Max. : 1.776e-15

A.4.3.2 Model 1: Trial Group 1 (Control)

This model repeats the process using a reduced household survey dataset: only those households in trial group 1 (control group). This run of IPF uses 730 households for 6136 areas. The first few rows of the output weights file are shown in Table 67.

Table 67 First few rows/columns of IPF weights table (Trial Group 1: Control)

bmg_id	oaCode	ipfWeight
956600058	E00085919	0.1808724
956600128	E00085919	0.1024742
956600237	E00085919	0.3753534
956610383	E00085919	0.1624927
956610449	E00085919	0.5617075
956610474	E00085919	0.0315803

A check for the number of households in each group is conducted, comparing for the original survey data and the unweighted IPF results. The first step is to check the unweighted household counts in the SAVE sample 'pool' before (Table 68) and after (Table 69) the IPF process:

Table 68 Household counts unweighted survey (household size x ethnicity): TG1 households

	1	2	3	4	5	6	7	8+	NA
White British/Irish	121	289	124	108	42	8	5	1	0
Mixed	1	1	3	0	0	0	0	0	0
Asian/Asian British	0	3	5	2	3	3	0	0	0
Black/Black British	0	1	1	2	1	1	0	1	0
Other	0	1	2	1	0	0	0	0	0
Refused	0	0	0	0	0	0	0	0	0
NA	0	0	0	0	0	0	0	0	0

Table 69 Household counts unweighted IPF results (household size x ethnicity): TG1 households, E00085919

	1	2	3	4	5	6	7	NA
White British/Irish	121	289	124	108	42	8	5	0
Mixed	1	1	3	0	0	0	0	0
Asian/Asian British	0	3	5	2	3	3	0	0
Black/Black British	0	0	0	0	0	0	0	0
Other	0	0	0	0	0	0	0	0
Refused	0	0	0	0	0	0	0	0
NA	0	0	0	0	0	0	0	0

The household counts match, except for the loss of *Black/Black British, Other* and households with 8+ people in the IPF results table due to zero weights. Survey data is added to the weighted households to test aggregate household counts for all OAs by ethnicity and household size (Table 70). In this example, the total simulated population is 783,631 households compared to the overall Census count of 790,066 households; indicating an imperfect match potentially due to the smaller household sample used. However, this still represents an under-estimate of less than 1%.

Table 70 Ethnicity vs n people (TG1 households, all OAs, SAVE study area, weighted IPF results rounded to 2 d.p.)

	1	2	3	4	5	6	7	8+
White British/Irish	222287.28	279351.36	107784.92	96714.31	31904.66	6843.71	2475	797.17
Mixed	2606.72	808.04	3226.34	0.00	0.00	0.00	0.00	0.00
Asian/Asian British	0.00	2351.64	7075.51	4561.19	2174.91	3510.61	0.00	0.00
Black/Black British	0.00	486.26	1237.28	2199.06	1197.43	640.68	0.00	449.83
Other	0.00	778.71	1249.95	918.43	0.00	0.00	0.00	0.00
Refused	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

The systematic validation tests of the errors within constraint categories are reported in Table 71 and Table 72. The mean error is still within +/-1 but in some cases the min/max errors indicate larger disparities caused by the use of a smaller household sample.

Table 71 Summary of error table per HRP constraint category (TG1 households, all OAs)

Constraint category	Min.	Mean	Max.
hrpWhite	Min. :-152.00000	Mean : -0.78562	Max. : 67.07444
hrpMixed	Min. :-9.22871	Mean :-0.02655	Max. : 3.25928
hrpAsianBrit	Min. :-45.53255	Mean : -0.13415	Max. : 0.73864
hrpBlackBrit	Min. :-25.964324	Mean : -0.056789	Max. : 4.262102
hrpOtherEthnic	Min. :-28.607987	Mean : -0.045617	Max. : 0.009946
hrp16_24	Min. :-134.63650	Mean : -0.10575	Max. : 11.70962
hrp25_34	Min. :-33.00000	Mean : -0.03189	Max. : 74.24652
hrp35_64	Min. :-109.00000	Mean : -0.60248	Max. : 55.09966
hrp65plus	Min. :-98.00000	Mean : -0.30861	Max. : 16.90294
econActEmplFt	Min. :-97.00000	Mean : -0.49425	Max. : 26.17118
econActEmplPt	Min. :-18.194870	Mean : -0.072005	Max. : 12.104468
econActSelfEmpl	Min. :-38.00000	Mean : -0.17716	Max. : 11.69707
econActUnemp	Min. :-34.56160	Mean : -0.07233	Max. : 9.96471
econActInactiveRetired	Min. :-106.00000	Mean : -0.27866	Max. : 15.39447
econActInactiveOther	Min. :-53.00000	Mean : 0.04567	Max. : 49.08052

Table 72 Summary of error table per household attribute constraint category (TG1 households, all OAs)

Constraint category	Min.	Mean	Max.
tenureOwned	Min. :-138.00000	Mean : -0.90054	Max. : 29.36504
tenureSocialRented	Min. :-107.00000	Mean : -0.00156	Max. : 51.75413
tenurePrivateRented	Min. :-78.00000	Mean : -0.14663	Max. : 15.53047
cars_0	Min. :-86.00000	Mean : -0.10738	Max. : 35.66555
cars_1	Min. :-74.00000	Mean : -0.37333	Max. : 21.94196
cars_2	Min. :-81.00000	Mean : -0.40632	Max. : 11.59724
cars_3	Min. :-39.00000	Mean : -0.11283	Max. : 10.96796
cars_4m	Min. :-20.000000	Mean : -0.048875	Max. : 5.906950
nchild_0	Min. :-124.00000	Mean : -0.75386	Max. : 12.20723
nchild_1m	Min. :-74.00000	Mean : -0.29487	Max. : 20.05183
npeople_1	Min. :-101.0000	Mean : -0.2278	Max. : 0.0000
npeople_2	Min. :-87.0000	Mean : -0.4192	Max. : 0.0000
npeople_3	Min. :-31.0000	Mean : -0.1592	Max. : 0.0000
npeople_4	Min. :-44.00	Mean : -0.16	Max. : 0.00
npeople_5	Min. :-19.00000	Mean : -0.05492	Max. : 0.00000
npeople_6	Min. :-12.00000	Mean : -0.01336	Max. : 0.00000
npeople_7	Min. :-8.000000	Mean :-0.004237	Max. : 0.000000
npeople_8m	Min. :-5.000000	Mean :-0.009941	Max. : 0.000000

A.4.3.3 Model 2: Trial group 2

Model 2 includes only those households in Trial Group 2 (who received online communications including event day notifications in TP1). This IPF run uses 759 households for 6,136 areas. The first few rows of the output are shown in Table 73.

Table 73 First few rows/columns of IPF weights table (Trial Group 2)

bmg_id	oaCode	ipfWeight
956600179	E00085919	0.1069915
956600197	E00085919	0.5651163
956610355	E00085919	0.0250599
956610483	E00085919	0.2567679
956610552	E00085919	0.0754967
956610603	E00085919	0.1105885

The check is repeated for the number of unweighted households in each group for the original survey data (Table 74) and following the IPF process after removal of households with zero weight (Table 75).

Table 74 Household counts unweighted survey (household size x ethnicity): TG2 households

	1	2	3	4	5	6	7	8+	NA
White British/Irish	120	321	115	121	34	12	1	2	0
Mixed	1	4	1	3	0	0	0	0	0
Asian/Asian British	0	4	0	3	5	1	1	0	0
Black/Black British	0	1	0	2	0	1	0	0	0
Other	1	1	1	1	2	0	0	0	0
Refused	0	0	0	0	0	0	0	0	0
NA	0	0	0	0	0	0	0	0	0

Table 75 Household counts unweighted IPF results (household size x ethnicity): TG2 households, E00085919

	1	2	3	4	5	6	7	NA
White British/Irish	120	321	115	121	34	12	1	0
Mixed	1	4	1	3	0	0	0	0
Asian/Asian British	0	4	0	3	5	1	1	0
Black/Black British	0	0	0	0	0	0	0	0
Other	0	0	0	0	0	0	0	0
Refused	0	0	0	0	0	0	0	0
NA	0	0	0	0	0	0	0	0

The household counts match, except for *Black/Black British, Other* and households with 8+ people as before.

The test of aggregate household counts for all OAs by ethnicity and household size (Table 76) results in a simulated household population of 782,590 households compared to the overall Census count of 790,066 households – a similar underestimate to the Trial Group 1 model.

Table 76 Ethnicity vs n people (TG2 households, all OAs, SAVE study area, weighted IPF results rounded to 2 d.p.)

	1	2	3	4	5	6	7	8+
White British/Irish	221751.69	270394.18	118954.78	97766.30	27114.96	9225.84	1380.97	1270
Mixed	2116.33	2643.42	683.50	1201.70	0.00	0.00	0.00	0.00

Asian/Asian British	0.00	6622.91	0.00	3165.04	7267.46	1056.67	1077.03	0.00
Black/Black British	0.00	3279.22	0.00	2034.23	0.00	717.49	0.00	0.00
Other	601.98	269.27	899.72	192.73	902.58	0.00	0.00	0.00
Refused	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

The systematic validation tests of the errors within constraint categories are reported in Table 77 and Table 78, and are comparable to the previous model results, giving no cause for concern.

Table 77 Summary of error table per HRP constraint category (TG2 households, all OAs)

Constraint category	Min.	Mean	Max.
hrpWhite	Min. :-156.00000	Mean : -0.83447	Max. : 69.12137
hrpMixed	Min. :-8.142202	Mean :-0.025921	Max. : 4.447151
hrpAsianBrit	Min. :-51.01780	Mean : -0.21315	Max. : 1.00000
hrpBlackBrit	Min. :-20.920980	Mean : -0.086059	Max. : 0.100085
hrpOtherEthnic	Min. :-29.84770	Mean : -0.05879	Max. : 1.36810
hrp16_24	Min. :-140.07340	Mean : -0.20665	Max. : 15.81488
hrp25_34	Min. :-45.39576	Mean : -0.28516	Max. : 4.76766
hrp35_64	Min. :-113.00000	Mean : -0.34231	Max. : 137.71606
hrp65plus	Min. :-101.00000	Mean : -0.38426	Max. : 28.24077
econActEmplFt	Min. :-112.00000	Mean : -0.66334	Max. : 13.22792
econActEmplPt	Min. :-16.000000	Mean : -0.059572	Max. : 18.313557
econActSelfEmpl	Min. :-38.00000	Mean : -0.18422	Max. : 4.65555
econActUnemp	Min. :-5.00000	Mean : 0.03173	Max. :63.99998
econActInactiveRetired	Min. :-106.00000	Mean : -0.34303	Max. : 28.72253
econActInactiveOther	Min. :-53.00000	Mean : 0.00004	Max. : 50.28673

Table 78 Summary of error table per household attribute constraint category (TG2 households, all OAs)

Constraint category	Min.	Mean	Max.
tenureOwned	Min. :-145.00000	Mean : -0.91872	Max. : 45.90416
tenureSocialRented	Min. :-107.00000	Mean : 0.16991	Max. : 122.32555
tenurePrivateRented	Min. :-110.11696	Mean : -0.46957	Max. : 1.12766
cars_0	Min. :-86.00000	Mean : -0.16919	Max. : 37.01832
cars_1	Min. :-71.00000	Mean : -0.42339	Max. : 21.13476
cars_2	Min. :-82.00000	Mean : -0.47003	Max. : 8.06580
cars_3	Min. :-26.000000	Mean : -0.119283	Max. : 4.539850
cars_4m	Min. :-17.000000	Mean : -0.036491	Max. : 8.694317
nchild_0	Min. :-136.00000	Mean : -0.95485	Max. : 10.25816
nchild_1m	Min. :-71.00000	Mean : -0.26353	Max. : 15.55973
npeople_1	Min. :-101.0000	Mean : -0.2969	Max. : 0.0000
npeople_2	Min. :-89.0000	Mean : -0.5116	Max. : 0.0000
npeople_3	Min. :-31.0000	Mean : -0.1651	Max. : 0.0000
npeople_4	Min. :-44.0000	Mean : -0.1654	Max. : 0.0000
npeople_5	Min. :-26.00000	Mean : -0.05362	Max. : 0.00000
npeople_6	Min. :-7.00000	Mean :-0.01255	Max. : 0.00000
npeople_7	Min. :-3.000000	Mean :-0.007008	Max. : 0.000000
npeople_8m	Min. :-2.000000	Mean :-0.006193	Max. : 0.000000

A.4.3.4 Model 3: Trial group 3

Model 3 includes just households in Trial Group 3 (who received both online and postal communications, and price signalling including event day notifications in TP1). The IPF process is run using 623 households for 6,136 areas. The first few rows of the resulting weights table are shown in Table 79.

Table 79 First few rows/columns of IPF weights table (Trial Group 3)

bmg_id	oaCode	ipfWeight
956600053	E00085919	0.4836127
956600056	E00085919	0.3661462
956600093	E00085919	0.0924718
956600200	E00085919	0.0424994
956610266	E00085919	0.0197768
956610335	E00085919	0.2250730

The check for number of households in each group is repeated for the original survey data (Table 80) and the unweighted IPF results (Table 81).

Table 80 Household counts unweighted survey (household size x ethnicity): TG3 households

	1	2	3	4	5	6	7	8+	NA
White British/Irish	97	258	106	97	32	9	3	1	0
Mixed	1	0	0	1	0	0	0	0	0
Asian/Asian British	0	3	1	7	0	0	0	0	0
Black/Black British	1	0	1	2	2	0	0	0	0
Other	1	0	0	0	0	0	0	0	0
Refused	0	0	0	0	0	0	0	0	0
NA	0	0	0	0	0	0	0	0	0

Table 81 Household counts unweighted IPF results (household size x ethnicity): TG3 households, E00085919

	1	2	3	4	5	6	7	NA
White British/Irish	97	258	106	97	32	9	3	0
Mixed	1	0	0	1	0	0	0	0
Asian/Asian British	0	3	1	7	0	0	0	0
Black/Black British	0	0	0	0	0	0	0	0
Other	0	0	0	0	0	0	0	0
Refused	0	0	0	0	0	0	0	0
NA	0	0	0	0	0	0	0	0

Inspection of the household counts in the SAVE survey 'pool' is as expected and provides no cause for concern. Finally, the test of aggregate household counts for all OAs by ethnicity and household size is repeated in Table 82. The resulting simulated household population of 788,457 households represents a smaller underestimate than the previous 2 models.

Table 82 Ethnicity vs n people (TG3 households, all OAs, SAVE study area, weighted IPF results rounded to 2 d.p.)

	1	2	3	4	5	6	7	8+
White British/Irish	216996.93	282349.18	111661.58	93696.71	34285.95	11037	2492	1299
Mixed	4657.89	0.00	0.00	1929.82	0.00	0.00	0.00	0.00
Asian/Asian British	0.00	3515.82	8674.80	6784.63	0.00	0.00	0.00	0.00
Black/Black British	1157.91	0.00	1038.62	2760.84	1264.05	0.00	0.00	0.00
Other	2854.27	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Refused	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

The systematic error tests (Table 85 and Table 86) are similar to those for Group 1 and 2 with larger errors than for the initial all households model, but still with a small mean error of +/- 1 household per OA.

Table 83 Summary of error table per HRP constraint category (TG3 households, all OAs)

Constraint category	Min.	Mean	Max.
hrpWhite	-146.00000	0.13679	69.07300
hrpMixed	-11.000000	-0.035249	1.313647
hrpAsianBrit	-52.53068	-0.24800	1.89257
hrpBlackBrit	-18.751567	-0.055018	2.817556
hrpOtherEthnic	-29.32199	-0.06074	1.42933
hrp16_24	-134.99608	-0.36727	5.06028
hrp25_34	-59.00000	-0.34365	9.01709
hrp35_64	-81.00000	0.34922	154.17747
hrp65plus	-80.00000	0.09948	36.24795
econActEmplFt	-86.00000	-0.36262	8.08467
econActEmplPt	-20.000000	0.045723	14.047955
econActSelfEmpl	-33.00000	-0.02558	7.19613
econActUnemp	-36.00000	-0.05433	3.26481
econActInactiveRetired	-77.00000	-0.02864	10.15221
econActInactiveOther	-130.00000	0.16323	29.17884

Table 84 Summary of error table per household attribute constraint category (TG3 households, all OAs)

Constraint category	Min.	Mean	Max.
tenureOwned	Min. :-122.00000	Mean : -0.06139	Max. : 28.07115
tenureSocialRented	Min. :-88.00000	Mean : 0.03721	Max. : 31.52087
tenurePrivateRented	Min. :-153.00000	Mean : -0.23804	Max. : 18.01908
cars_0	Min. :-111.00000	Mean : -0.07978	Max. : 20.25904
cars_1	Min. :-70.00000	Mean : -0.15697	Max. : 28.88694
cars_2	Min. :-51.00000	Mean : -0.03781	Max. : 15.43485
cars_3	Min. :-23.000000	Mean : 0.010101	Max. : 13.230168
cars_4m	Min. :-10.000000	Mean : 0.002235	Max. : 2.819485
nchild_0	Min. :-173.00000	Mean : -0.26008	Max. : 10.02674
nchild_1m	Min. :-76.00000	Mean : -0.00214	Max. : 27.96273
npeople_1	Min. :-113.0000	Mean : -0.1019	Max. : 0.0000
npeople_2	Min. :-67.00000	Mean : -0.07872	Max. : 0.00000
npeople_3	Min. :-23.00000	Mean : -0.02868	Max. : 0.00000
npeople_4	Min. :-38.00000	Mean : -0.03308	Max. : 0.00000
npeople_5	Min. :-12.00000	Mean : -0.01043	Max. : 0.00000
npeople_6	Min. :-13.000000	Mean : -0.006519	Max. : 0.000000
npeople_7	Min. :-3.000000	Mean :-0.001467	Max. : 0.000000
npeople_8m	Min. :-2.000000	Mean :-0.001467	Max. : 0.000000