Department for Energy Security & Net Zero

Smart Energy Savings Competition (SENS): Energy Saver app

Trial-Level Evaluation Report

June 2023

Acknowledgements

This report was prepared by Ipsos, in partnership with Energy Saving Trust, the University of Edinburgh and Manchester Metropolitan University. The evaluation would not have been possible without GenGame Limited, along with their competition delivery partners Loughborough University Enterprises Limited, Lucid Energy (formerly Intelligent Data Technologies Limited), and SO Energy (sub-contractor), and the cooperation of all of those who participated in interviews, surveys and workshops. The Department for Energy Security and Net Zero (formerly BEIS) would also like to thank the Smart Energy Research Lab (SERL), based at University College London, for their support in the secure collection and provision of energy consumption data, and Charles Michaelis for his expert independent peer review of this report.

Finally, the Department would like to thank all SENS Competition Partners, Ipsos, UCL SERL and AECOM for all their efforts to respond to challenges to deliver SENS throughout the COVID-19 pandemic.

Any enquiries regarding this publication should be sent to us (referencing SENS) at: smartmetering@beis.gov.uk

Contents

Executive Summary

Rationale for and objectives of the SENS Competition

Smart meters are replacing traditional gas and electricity meters in homes and small businesses across Great Britain as part of an important upgrade to the national energy infrastructure, underpinning the cost-effective delivery of Government's Net Zero commitment. They are a critical tool in the transition to a low carbon energy system, for example helping consumers to use energy when renewable generation is available. Prior to the Competition, BEIS found that smart meters would result in average reductions of 3% for electricity customers, 2.2% for gas credit customers, and 0.5% for gas pre-payment customers[1].

Early evaluation and research have shown that these savings are realised through access to near real time feedback (via In-Home Displays (IHDs)), energy efficiency advice at the point of installation, and accurate bills[2]. The Smart Energy Savings Innovation (SENS) Competition was developed on the assumption that more sophisticated uses of energy consumption data can deliver additional savings to those already achieved by having a smart meter installed in the home.

The SENS Competition led by the former Department for Business, Energy and Industrial Strategy (BEIS) committed up to £6.25 million, to support the development, trialling and evaluation of innovative feedback products and services that use smart meter data to help domestic consumers reduce their energy consumption. SENS was launched February 2019, with trials concluding end of March 2022 (extended by one-year due to COVID-19 impacts).

The objectives of the Competition were to:

- Identify innovative products and services using smart meter data that can deliver energy savings in homes, in excess of those currently identified in the smart meter impact assessment, for either the Great Britain population or specific groups within it.
- Ensure that solutions are attractive and valued by consumers and are easily available (using existing technologies and delivery channels or cost-effective new hardware).
- Support the development of a domestic market for energy management products and services, securing investment from technology providers, energy suppliers, and third parties.

[1] https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/831716/smart-meter-roll-out-cost-benefit-analysis-2019.pdf
[2] https://www.gov.uk/government/publications/smart-metering-early-learning-project-and-small-scale-behaviour-trials

Overview of the SENS product

The SENS Energy Saver app project 'Combining Gamification with energy insights to create an energy-saving mobile app' (herein referred to as SENS GenGame), was delivered by GenGame Limited, in partnership with Loughborough University Enterprises Limited, Lucid Energy (formerly Intelligent Data Technologies Limited) and SO Energy.

SENS GenGame was a mobile-only application that aimed to drive household behaviour change by using households' gas and electricity smart meter data to provide a range of energy consumption information and advice, tailored to the individual customer. The app used smart meter data to provide insights to households on their historical consumption patterns, as well as offering forecasted consumption. It also allowed the trialist to track how their energy consumption and expenditure varied. Based on the energy information collected through the smart meter, the app suggested energy saving measures from a built-in database that were bespoke to the trialist's household. The app included a gamification feature to encourage the trialist to engage with it more regularly through features including league tables, prizes, achievements and badges.

Evaluation approach and methodology

The Competition appointed a separate Trial Design and Evaluation Lead (TDEL) team, led by Ipsos, in conjunction with Energy Saving Trust, Manchester Metropolitan University and the University of Edinburgh, to conduct an independent evaluation of the Competition overall, and of each of the individual products and services trialled through the Competition.

This trial-level evaluation sought to test whether the SENS GenGame product was successful in realising its primary objective of reducing energy consumption (gas and electricity) and what features of the app made energy savings more likely to occur.

The trial evaluation employed a Matched Control Design, with a control group who had received the baseline smart meter consumer proposition (i.e., a smart meter installation, access to near real time feedback on gas and electricity used via an In-Home Display (IHD), and energy efficiency advice delivered at the point of installation), and an intervention group, who, in addition to all the above, were offered the SENS intervention. Those in the intervention group who downloaded the SENS GenGame app were then matched (via Propensity Score Matching) to control group trialists using matching variables that were measured pre-intervention and that previous studies had demonstrated to be correlated with the outcome variables being tested.

In total,1,912 trialists were recruited to the intervention group (between April and November 2021), with their smart meter energy consumption data collected from the time of recruitment to end of trial. In addition, 1,068 trialists were recruited to the control group later in the trial (January and February 2022). For the control group, their historic smart meter energy consumption data was collected to cover the same period of time as intervention group trialists, required to evaluate their energy use over the trial (up until end March 2022). Signing up to

take part in a SENS trial was entirely voluntary, and consent could be withdrawn at any time without giving a reason.

After trial completion, the two groups were statistically compared to quantify the effect of the intervention upon energy consumption. Gas and electricity consumption data was used to analyse any changes in gas and electricity consumption before and after the SENS GenGame intervention, using a regression framework including the trial group (control or intervention) as a grouping variable, and with prior consumption as a control variable.

The analysis was supported by a package of wider primary research activities, including baseline and endline quantitative telephone surveys with intervention group trialists, to understand and evaluate their attitudes towards energy, energy usage and management behaviours, uptake of energy efficiency advice and engagement with the intervention.

Finally, in-depth qualitative interviews were conducted with 15 trialists from the intervention group. Recruiters ensured the inclusion of a range of demographics and perspectives, and interviews covered topics including how trialists interacted with the components of the intervention, their initial experiences and behaviour changes, and longer-term impacts.

Outcomes for trialists trialling the product

The energy consumption analysis of SENS GenGame Trial found a statistically significant saving in daily gas use of 4.6% ± 2.0% (95% Confidence Interval, $p<0.001$), for those who used the SENS GenGame Energy Saver app (Treatment on Treated analyses) compared to the control group gas use over the same period. No statistically significant effect on electricity use was identified.

Among the intervention group recruited to the SENS GenGame trial, just over six in ten installed and used the SENS GenGame app (57%). Among these, three quarters logged in at least five times over the trial (78%), with the average number of logins of 26 among those who used the app. Survey evidence indicated that seeing historic gas and electricity use was the feature used by most trialists (over 80% for each fuel), while two thirds reported using tips and advice on using less energy (67%), including reducing heating bills (66%). Slightly fewer used features providing information on the appliances using the most energy (58%), carbon savings by the household (53%), and current energy use (44%). The gamification features appeared to achieve the secondary trial aim of increasing app engagement for some users – of those who used the gamification features, just over half agreed that they made them use the app more than they otherwise would have (56%).

Three quarters of app users surveyed (near to the end of the trial) reported that their understanding of how their household used energy had got a little or a lot better ‘over the last year’ (77%), with over ninety percent of those 77% attributing this at least partly to the app (93%). Over half of the surveyed app users reported improved confidence about knowing which activities or appliances required a lot of energy in their homes (55%), and over two thirds reported more confidence about changes they could make to save the amount of energy used

in their homes (67%). Although this did not appear to translate into changes in the proportions heating their homes when unoccupied 'for a few hours', one of the secondary aims of the trial, there was interview evidence of other heating-related behaviour changes such as wearing jumpers more rather than heating the home.

Finally, there was evidence that a minority of intervention group trialists (just under a quarter of survey respondents who had used the app by the end of the trial) felt they became more able to heat their home to a comfortable level since starting to engage with the app (24%), although just over a third disagreed with this (35%). Around half of that same group of respondents agreed that they found it easier to control how much they spent on energy since starting to engage with the app (49%), with only a fifth disagreeing (21%).

Conclusions

The trial provides robust evidence that the GenGame Energy Saver app achieved the primary aim of reducing energy use for gas consumption. There is potential for this finding to reflect savings for more engaged customers, those that downloaded and used the app at least once (due to the requirement to collect matching data as this point in the customer journey). While evidence suggested trialists engaged with gas and electricity information and advice via the app, there was no statistically significant reduction in electricity use found. This may be because any impact was smaller than the trial was able to detect, because no change in energy use occurred or because there was not sufficient time in the trial for certain changes to have occurred or become embedded, particularly in relation to energy efficiency investment-focused advice that may take longer to act on and implement than behavioural changes.

There was clear evidence that substantial proportions of trialists felt the app's features improved their overall understanding of various aspects of their energy use and drivers of energy consumption. The gamification features were also found to have increased engagement with the app for a substantial proportion of users. The evaluation was unable to identify the specific behaviour changes that led to gas savings, but the research indicated engagement with relevant feedback and tips (e.g., turning off heating at night, closing curtains at night, turning down heating, putting jumpers on, adding more insulation) and instances of heating-related and other energy-related behaviour changes occurring after using the GenGame Energy Saver app.

1 Introduction

The Smart Energy Savings (SENS) Innovation Competition (from here on referred to as 'the Competition') led by the former Department for Business, Energy and Industrial Strategy (BEIS) committed up to £6.25 million, to support the development, trialling and evaluation of innovative feedback products and services that use smart meter data to help domestic customers reduce their energy consumption.

Following a competed application process, eight projects were selected to receive Phase One Competition (matched) grant funding to support the development of their products and/or service. Following a stage-gate review process, five projects were taken through to Phase Two, to trial and evaluate their products and/or services in homes across Great Britain. The Competition was launched February 2019, with trials concluding end of March 2022 (extended by one-year due to COVID-19 impacts).

Ipsos, in partnership with Energy Saving Trust, Manchester Metropolitan University and the University of Edinburgh were commissioned by BEIS as the Trial Design and Evaluation Lead (TDEL), to undertake a robust independent evaluation of the Competition, including separate trial evaluations for each of the individual projects, and to implement a wider package of research. Separately, BEIS awarded a grant to the Smart Energy Research Laboratory (SERL) based at University College London (UCL), for the collection and provision of secure access to energy consumption data from trialists (with their consent) to the TDEL for their analyses. BEIS also appointed an independent Project Management lead, AECOM, to oversee the Competition Partner's project delivery and grant funding milestones.

This report is part of a package of reports published for the Competition, including an overarching competition-level evaluation report, a technical evaluation report and five separate trial-level evaluation reports (including this report).

1.1 Purpose of this report

This report presents the evidence from an evaluation of the SENS GenGame Energy Saver App (hereinafter referred to as SENS GenGame) project that was taken through to Phase Two of the Competition to trial and evaluate their mobile application in real-world households across Great Britain. The report presents the analysis of energy consumption data and other primary and secondary data that were used to answer the primary research question of the SENS GenGame trial, presented in the box below (as well as analysis of other secondary outcomes presented in more detail in chapter five):

What is the added gas and electricity saving achieved from the SENS GenGame Energy Saver app, over and above the baseline smart meter customer proposition (ie. a smart meter, an In-Home display (IHD), and energy efficiency advice provided at install)?

Subsequent chapters of this report provide a summary of the SENS GenGame Energy Saver app and trial design (chapter two) and trial evaluation methodology (chapter three). The overall evaluation findings relating to the primary outcome are presented in chapter four including evidence triangulated across different data sources including energy consumption analysis and quantitative and qualitative research strands. Evidence from the analysis of secondary outcomes is presented in chapter five. Finally, chapter six presents the key conclusions from the trial evaluation.

2 Summary of trial

This chapter provides an *introduction to the SENS GenGame intervention, including its core functionality and mechanisms for behaviour change as presented through its Theory of Change. The core features of the trial design are also presented here.*

2.1 The SENS GenGame intervention

The purpose of the SENS GenGame project was to provide trialists with a mobile application that used smart meter data to provide a range of energy consumption information and advice, tailored to the individual trialist, with the aim to drive household behaviour change and reduce energy consumption.

Table 1 below provides details of the delivery partners for the project.

Table 1: Summary of the project name and delivery partners

Project Title	Competition delivery partner(s)		SENS Product
	Lead	Partner(s)	
Combining Gamification with energy insights to create an energy-saving mobile app	GenGame Limited	Loughborough University Enterprises Limited, Lucid Energy (formerly Intelligent Data Technologies Limited), SO Energy (sub-contractor)	A mobile-only application that used trialists' gas and electricity smart meter data to provide a range of energy consumption information and advice. The product provided insights to trialists on their historical consumption patterns, as well as forecasted consumption. Based on the energy information collected through the smart meter, the product suggested energy saving advice measures from a built-in database that were bespoke to the customer's household. It included a gamification feature to encourage the customer to engage with it more regularly through features including league Tables, prizes, achievements and badges.

2.1.1 Aims of the intervention and how it was expected to achieve these

The intervention aimed to achieve several primary and secondary outcomes, summarised in Table 2 below. These were identified as aims at the outset of the trial. Whether the intervention achieved these primary and secondary outcomes during the trial is evaluated in detail in the rest of this report.

Table 2: Primary and secondary outcomes of the SENS GenGame intervention

Primary/ Secondary	Outcomes to be evaluated
Primary	Reduction in gas and electricity consumption
Secondary	Improved individual perceptions of home comfort
	Reduced unoccupied heating hours
	Increased understanding of energy use and drivers of energy consumption
	Increased customer retention
	Improved household budgeting
	Increase app engagement led via gamification

The SENS GenGame intervention had three component features, described below. The Theory of Change diagram presented in Annex I provides further detail about how these were expected to lead to the intended outcomes.

- **Gamification.** The SENS GenGame app included a gamification feature to encourage the trialist to engage with the app more regularly, through features including challenges, league tables, prizes and rewards, achievements and badges.
- **Energy use and expenditure insight.** The SENS GenGame app enabled households to track how their energy consumption and expenditure varied. The app also drew on smart meter data to provide insights into household's energy consumption in terms of:

  - cumulative history (a household's 'running total' energy usage, for current day, week, month, and year);
  - activity breakdowns (gas and electricity usage, analysed to show what it was used for);
  - consumption characteristics (household's consumption interpreted, showing average usage pattern across day, week, and year);
  - consumption forecasts (household's future energy usage and expenditure, estimated based on their own consumption history);
  - time period comparison (comparison of consumption to similar time period).
- **Energy saving advice.** Based on the energy information collected by the smart meter, the app suggested energy saving measures from a built-in database that were tailored to the trialist. These could include both behaviour-change advice and investment-focused advice, including small investments such as upgrading lighting to more substantial and longer-term investments, for example investing in thermal insulation or replacing the boiler.

Figure 1: Example of features presented through the SENS GenGame app

As the trialists had ready access to tailored energy consumption data/ information and tailored recommendations on how to save energy provided via the app, it was anticipated that they would change their behaviour in the following ways:

- Implement new day-to-day behaviour measures that would reduce energy in the home, such as drying clothes naturally rather than using the tumble dryer;

- Adding thermostatic radiator valves in the house to reduce heat costs;
- Avoid using portable heaters electricity use, and thus cost, will increase;
- Installing a thermostat to reduce energy bills;
- Avoid using the radiator as a mean to dry clothes (as such action results to poorer efficiency);
- Make upgrades to the home that could reduce energy consumption, such as upgrading lighting or installing home insulation.

Such behaviour changes were expected to lead to a reduction in energy consumption (gas and electric).

The routes for these changes to occur are presented in more detail in Annex I, the Theory of Change. As presented there, for these outcomes to materialise, a number of assumptions needed to hold true during the trial period. First, trialists needed to be motivated by a desire to reduce their bills or take low-carbon actions and find ways to maintain thermal comfort[3]. Where this was not the case, the gamification feature of the app needed to motivate trialists in other ways, by promoting engagement with the app and features within it. Secondly, the database of advice needed to be relevant and sufficiently tailored to all trialists, and trialists needed to trust the advice and costing estimates provided. Thirdly, trialists needed to be willing and able to follow these tips, and believe that their households had the potential for further energy-saving improvements. Finally, trialists needed to have access to the necessary forms of personal funding to implement investment-focused advice such as upgrading lighting, insulation or their boiler.

These assumptions have been tested as part of the evaluation of SENS GenGame's contribution to the intended outcomes.

2.2 Design of the SENS GenGame trial

2.2.1 Matched Control Design

It was originally expected that all SENS trials would take the form of Randomised Controlled Trials (RCTs) where possible as this methodology would allow for trialists to be randomly assigned to intervention or control groups, ensuring that any underlying biases in the sample in the distributions of key predictors of the primary outcome (i.e. predictors of energy consumption) were likely to be close to equally distributed between the two groups. However, an RCT design requires a large population available upfront to recruit trialists from, to ensure that sufficient numbers are recruited prior to the start of the trial to enable randomisation to occur at a suitable time-point in the recruitment journey to eliminate selection differences between the intervention and control groups.

[3] Thermal comfort is the condition of mind that expresses satisfaction with the thermal environment.

For SENS GenGame, it was anticipated that SO Energy's population of available customers would not provide the necessary numbers upfront to viably implement an RCT design. Available customers in this case meant those with Data Communications Company (DCC) enrolled (SMETS1 or SMETS2) smart meters, required so that their smart meter data were accessible for the evaluation. Households with SMETS1 meters became available for the trial only at a later stage. Therefore, TDEL recommended a staged matched control design as the most robust viable alternative approach. In this design, the intervention group was recruited first (between April and November 2021), and the control group recruited at a later point in time, towards the end of the trial period (in January and February 2022), and largely from households with SMETS1 DCC enrolled meters. All trialists, as part of joining the trial, consented to give access to their historical smart meter data for the trial evaluation as well as for ongoing smart meter data collection up until end of trial period. As smart meters store over a year of historic energy data, this enabled energy consumption data for the full evaluation period of April 2021 to March 2022 to be collected from both the control group and intervention group trialists, so that data from both groups were from the same period of time.

The risk of the matched control design is that it provides less robust evidence of an intervention effect than an RCT because it results in differences between the control and intervention groups in the distributions of key predictors of the primary outcome. As this cannot be adequately controlled for through the recruitment design, as it would be in an RCT, an alternative method to mitigate this risk was to collect measurements of key predictor variables and control for them during the data analysis phase through a statistical matching process i.e. Propensity Score Matching. For this trial, data were collected and used for several such 'matching variables', described below.

Matching was performed based on seven variables, each of which correlate with energy consumption: prior energy consumption, four attitudinal questions, property type (e.g. flat, terraced house) and geographic region (e.g. South East, Scotland). Prior energy consumption was included as it closely correlates with future energy consumption. The four attitudinal questions were based on literature exploring the links between trialists' values and energy consumption[4], and were selected from a longer list of thematic areas that are strongly associated with energy use as they have been found in previous studies to be the most effective.[5] All attitudinal questions provided respondents with a five-option agreement scale (from strongly agree to strongly disagree), indicating agreement with the following statements:

- Environmentalism: "Climate change is likely to have a big impact on people like me."
- Technological innovation: "I am the type of person who likes to have the newest gadgets in my home."
- Need for comfort: "I am more concerned about having a warm and comfortable home than saving energy."

[4] See Bent, Caitlin & Kmetty, Zoltán. (2017). Intelligent energy feedback: Tailoring advice based on consumer values. 10.5281/zenodo.820511.
[5] See Hatter Kiado (2017), The NATCONSUMERS handbook A guide to introducing ICT tools for customer engagement in energy savings. 10.5281/zenodo.838886

- Need for control: “I feel in control of how much energy I personally use.”

The set of four attitudinal questions and the question on property type were asked to the trialists. To fit in with the SENS GenGame recruitment trialist journey, described in section 2.2.3 below, intervention group trialists (following opt-in consent the trial) were asked the four questions when opening the mobile application for the first time, whilst control group trialists were asked the same questions at the point in which they provided opt-in consent to participate in the trial via SO Energy’s microsite for the trial. This difference in collection of attitudinal data between the intervention group and the control group precluded an Intention To Treat (ITT) analysis in favour of a Treatment on Treated (TOT) approach as described in section 4, as data for matching variables were not collected for those intended to be treated but who did not open the mobile app.

How these matching variables were used in the analyses is described in chapter four.

2.2.2 Eligible trialists

The sampling frame for the trial included all dual fuel households within SO Energy’s customer base from across Great Britain, including urban, suburban, and rural households. SO Energy customers were eligible to participate in the trial if they had a smart meter installed. Customers with a dual-fuel SMETS2 or SMETS1 DCC-enrolled meter were eligible for the intervention group, whilst SMETS1 DCC-enrolled customers were eligible for the control group.[6] Customers were required to be DCC enrolled so that their smart meter data could be collected for trial evaluation purposes.

2.2.3 Recruitment strategy

Recruitment was led by SO Energy. SO Energy was responsible for developing the recruitment materials and the format of the consent form for both intervention and control groups (using standardised consent forms that were General Data Protection Regulation (GDPR) and Smart Energy Code (SEC) compliant, developed by UCL and TDEL).

Signing up to take part in a SENS trial was entirely voluntary, and consent could be withdrawn at any time without giving a reason. To assess the primary aim of this project, to ascertain if the SENS product or service helped trialists to use less energy, trialists gave opt-in consent to provide access to their smart meter data for the evaluation, using a virtual secure lab analysis environment provided by UCL. This smart meter data was used by Ipsos, its approved partners and UCL solely for the purpose of the SENS evaluation. More information on the approach to obtaining consent is provided in the accompanying Technical Report.

To recruit the intervention group, SO Energy sent recruitment emails between April and November 2021 to all eligible smart meter customers to invite them to participate in the SENS trial. SO Energy began with small test batches during the initial recruitment period, before then scaling their recruitment operations and targeting customers through mass recruitment emails.

[6] The timings of smart meter enrolment for SOE meant that a large batch of SMETS1 customers became available after SMETS 2 rollout had started. SMETS2 customers would also have been valid control group trialists but in practice there were not enough.

To maximise sign up to the trial, SO Energy utilised up to three additional reminder emails and offered a £10 voucher incentive to join the SENS GenGame trial.

After signing up to the SENS GenGame trial, intervention group trialists were invited to download the SENS GenGame app. In order to maximise the download rate, GenGame sent reminders to trialists that had not downloaded the app.

To recruit the control group, SO Energy sent recruitment emails to all eligible smart meter customers to invite them to participate in the SENS trial from the end of January 2022 until the end of February 2022. As an incentive, these customers were offered access to the SENS GenGame mobile app once the trial period was over (i.e. after March 2022). Similar to intervention group trialists, they were offered a £10 voucher incentive to take part in the trial.

Both groups completed the attitudinal and property type questions required for data matching. Intervention group trialists were asked to complete these on downloading the SENS GenGame app. Control group trialists meanwhile completed these at the point at which they provided consent to participate in the trial, on SO Energy's microsite, to maximise response rates.

The trialist customer journey for both intervention and control groups are shown in the Figures below.

Figure 2: Trialist customer journey for intervention group

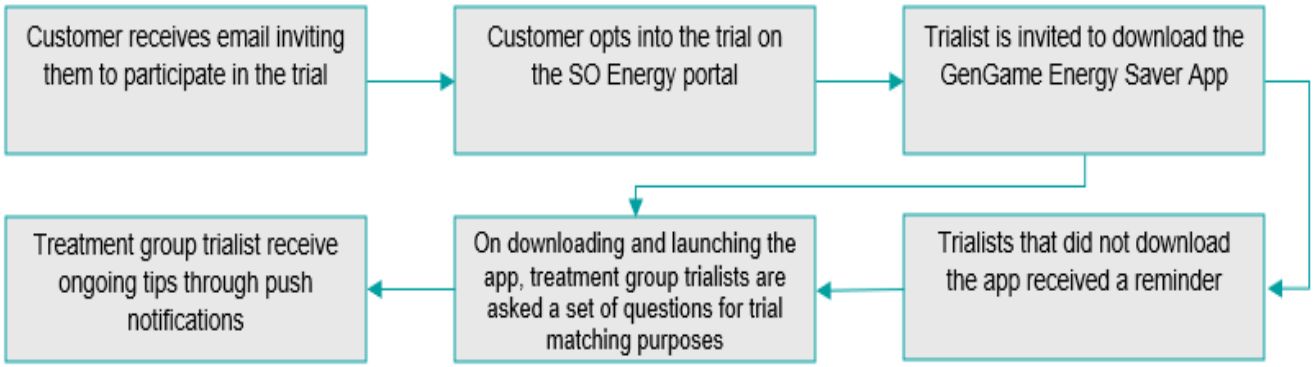
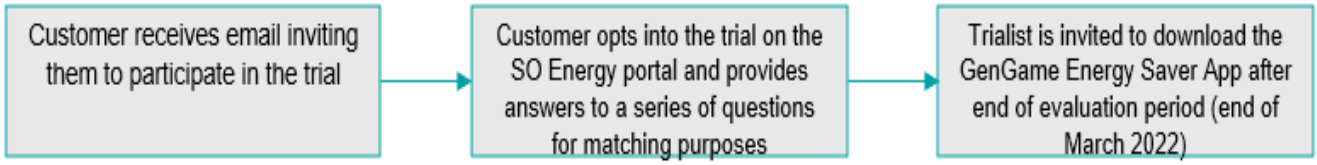


Figure 3: Trialist customer journey for control group



Recruitment targets were initially set by TDEL to achieve the sample sizes needed to detect the expected impact. Based on an anticipated 5% reduction in electricity and gas consumption and the amount of variability in energy consumption that could be explained by pre-trial consumption data, the trial needed to recruit and retain 1,196 trialists in both the intervention and control groups. To account for an assumed 25% drop-out rate (average number of households switching energy supplier or moving home within a 12-month period, as well as

those actively withdrawing from trial), the initial recruitment targets were therefore set at 1,435 in both the intervention and control group.

Table 3 presents these targets and achieved recruitment figures. The recruitment target for intervention group trialists was exceeded, even after drop-out during the trial period was considered, while the target for control group trialists was not reached. However, 766 intervention group trialists consented to participate but did not go on to download the SENS GenGame app, and as such also did not provide responses to the matching questions required for them to be included in the primary analysis of energy consumption. After this and other data quality and other factors were taken into account, 871 intervention group trialists and 915 control group trialists were available for the primary analysis of energy consumption (see Annex II – Trial Overview for more details).

Table 3: Recruitment targets versus achieved recruitment figures

Trial	Recruitment target (intervention / control)	Recruited to trial (intervention/control) initially set out by the TDEL	Onboarded into SERL (intervention/ control)	Final achieved sample (accounting for withdrawals) (intervention/ control)
SENS GenGame	1435/1435	1912/1068	1912/1068	1760/1067

Due to various challenges, including COVID-19 impacts and changes in the wider retail market (see SENS Evaluation Competition Report), the SENS GenGame trial did not recruit the full control group sample but exceeded on recruitment of intervention group sample. The main challenges resulted from a change in energy supplier partner (at start of SENS phase two). This meant that a new pool of eligible SO Energy customers had to be established at short notice before commencing recruitment. Furthermore, the control group was recruited at the end of the trial to maximise the available trialist groups, which could have potentially introduced some form of unobservable bias in the matched design analysis.

3 Methodology

This section describes the methodological approach taken to data collection and analysis of the different data sources, including the Matched Control D*esign method and the statistical methods for the energy consumption analysis. More information is provided in the accompanying* T*echnical* R*eport published alongside this report.*

3.1 Data collection

The evaluation of SENS GenGame utilised a range of data sources to provide evidence against the primary and secondary research questions for the SENS GenGame trial.

3.1.1 Energy Consumption Data

Energy consumption data for evaluation purposes was collected (with consent) to cover two periods[7]:

- During the trial. Gas and electricity consumption data was securely provided to TDEL via the Smart Energy Research Laboratory (responsible for managing the collection and provision of smart meter data from trialists with their consent to TDEL for the purposes of the evaluation) at 30-minute and daily resolution for the trial period.
- Before the trial. Energy consumption data from before the trial ('pre-baseline') were provided by the Competition Partner in the form of estimated annual electricity and gas usage data in kWh (known as EAC, Estimated Annual Consumption, and AQ, Annual Quantity, for electricity and gas respectively) for the participating trialists for the 12 months up to and including 1 April 2021, shortly before the trial began. As described in the analysis section below, these variables were required as matching variables and as control variables in the Energy Consumption Analyses.

3.1.2 Matching Data

As well as the pre-baseline energy data, the following variables were also used to match intervention and control trialists:

- Region (provided by SERL). The region each trialist lived in, e.g., South East, Wales, Scotland.
- Property type. The type of property they lived in, e.g., flat, terraced house.
- Attitudinal questions. Responses by the trialists to each of four questions about attitudes to climate change, gadgets, home comfort and feelings of control over energy use (see Section 2.2.1 for details).

[7] Please refer to the TDEL technical report for further details.

The property type and responses to attitudinal questions were provided by the Competition Partner, who collected this data from trialists. The method of collecting responses to these questions differed between the intervention and control group trialists. Intervention group trialists were asked the four questions only after they had downloaded the SENS GenGame app and opened it for the first time. As such, responses to the questions were not available for intervention group trialists who consented to participate in the trial but did not download or launch the SENS GenGame app. In contrast, control group trialists were asked the same questions at the point in which they provided consent to participate in the trial via SO Energy's microsite, so that their responses are available for every control group trialist.

3.1.3 Engagement data

SO Energy securely collected and provided data to TDEL for analysis on intervention group engagement with the SENS GenGame app (with appropriate consent being in place from the individuals), comprising which trialists had downloaded the mobile application and, during the trial period, how many times a trialist accessed the app, the number of tips accepted, tips marked as completed and tips rejected, and the number of raffle purchases.

3.1.4 Quantitative Telephone Survey with trialists.

Intervention trialists took part in a baseline (August – December 2021) and endline (March 2022) telephone survey to ascertain attitudes to energy, energy usage and management behaviours, uptake of energy efficiency measures, views of smart metering and engagement with the trials and products / services. Of those intervention trialists, 315 took part in the baseline survey and 149 in the endline survey. Of those who completed both surveys, all but 13 had already downloaded the app at baseline. Control group trialists were not able to participate in the survey due to timings of recruitment. More details on the timings and key topics explored are included in the accompanying Technical Report.

One sample t-tests between baseline and endline survey percentages were conducted for the survey findings at the Competition level only (aggregated across all trialists) but not at individual trial level, to determine whether the change was statistically significant at conventional significance levels. Unless explicitly stated, any reported changes (baseline to endline) are indicative only and have either not undergone statistical significance testing or were not found to be statistically significant.

3.1.5 User in-depth interviews

TDEL also conducted in-depth qualitative interviews with 15 consented trialists in the intervention group who had been given access to the SENS GenGame app over the trial period. All apart from one had downloaded the app at baseline. These were recruited from those who completed the endline surveys so there is some overlap with survey responses.

The interviews were semi-structured and typically lasted 45-60 minutes and covered their experiences of use of the intervention and how they found it, as well as more general questions about energy use and household budgeting. A range of quotas across different

demographics and household characteristics were sought, including householder age and property age. Further details of this can be found in chapter seven of the Technical Report.

3.2 Data analysis

3.2.1 Data quality and cleaning

Initial data cleaning was conducted on the data where required, as follows:

- Energy Consumption Data – before the trial. The pre-baseline electricity and gas (EAC and AQ) annual usage estimates provided by the Competition Partner were converted to a daily mean by dividing by 365, to match the units used for the evaluation period energy consumption data.
- Energy Consumption Data – evaluation period. Mean daily estimates of electricity and gas use were calculated for each trialist's participation period using the available smart meter data for their properties. Smart meter data were cleaned and used to produce the estimates following an approach similar to that used by SERL for its data and statistical releases (see Elam, Webborn et al., 2022, and Few, Pullinger et al., 2022[8]). The approach is described in detail in the Technical Report.

3.2.2 Statistical analysis of energy consumption data

After applying the data quality processes above, a total of 871 intervention trialists and 915 control group trialists remained in the dataset for energy consumption analysis. The Trial Overview (Annex II) describes the number of trialists initially recruited and the numbers removed due to different data quality issues. Energy consumption analysis was of 'Treatment on the Treated', i.e. it included those intervention group trialists who consented to participate and subsequently went on to download and use the SENS GenGame app at least once. As noted in section 2.2.1, Intention To Treat (ITT) could not be used because of the way matching variables were collected – the four attitudinal matching variables were not available to those intervention group members who did not download and start up the app: they were collected during the first use of the app.

Matching of intervention and control group trialists

A key challenge of matched control research designs such as the one used for SENS GenGame is that if the control group trialists differ systematically from intervention group trialists (called a 'selection effect') by characteristics that are key predictors of the outcome, then these pre-existing group differences may be the cause of any post-trial group differences

[8] Elam, S., Webborn, E., McKenna, E., Oreszczyn, T., Anderson, B., Few, J., Pullinger, M., European Centre for Medium-Range Weather Forecasts, Ministry of Housing, Communities and Local Government, Royal Mail Group Limited. (2022). *Smart Energy Research Lab Observatory Data, 2019-2021: Secure Access*. [data collection]. *5th Edition.* UK Data Service. SN: 8666, DOI: 10.5255/UKDA-SN-8666-5; Few, Pullinger, McKenna, Elam, Webborn and Oreszczyn (2022) Smart Energy Research Lab: Energy use in GB domestic buildings 2021. Variation in annual, seasonal, and diurnal gas and electricity use with weather, building and occupant characteristics. (SERL Statistical Reports: Volume 1), https://serl.ac.uk/key-documents/reports/.

in the outcome, rather than the intervention (in this case, the SENS GenGame app) being the causal effect.

As described in Chapter 2, data for a range of 'matching variables' were collected from trialists to aid in mitigating this risk. A first stage of analysis was then to apply Propensity Score Matching (PSM), making use of these variables. This aimed to equivalise the control and intervention groups in terms of their characteristics along the matching variables using statistical techniques, so that the control group, after this matching process, provides a more robust counterfactual estimate of the average energy consumption for trialists without the SENS GenGame app.

Before this matching process, the intervention and control groups with full matching data were broadly similar on measures of pre-baseline energy use, but substantially different in terms of the geographic distribution, with far more intervention trialists than control in the east (Greater London, South East and East of England) and more control trialists than intervention in other areas, particularly in Scotland and Northern England. The distribution of responses to certain attitudinal question also varied substantially between the groups, particularly for the questions relating to the latest gadgets and to climate change impacts. These differences would impact on the robustness of final results from analyses of this dataset if matching were not applied.

There are many approaches to PSM, and TDEL explored various methods to identify approaches that led to the optimal matching of the sample's characteristics. The initial large imbalances in certain matching characteristics meant that there were inevitable compromises involved in the matching process, with the approaches either leading to many control group trialists being removed from the analysis (and the remainder, as a result, being matched to multiple intervention group trialists each) or, if matched using a method that did not allow for the removal of trialists, continued substantial imbalances in the control and intervention trialist characteristics. As any approach introduces different impacts on the robustness of the results of analyses, three different approaches were selected. The statistically best matching method is presented in this report, with two other methods used as comparison. As described in the next section, this helps to improve the robustness of the results, by testing the sensitivity of the results.

Annex III (Technical Appendix) presents more details of the three matching approaches selected, and of the distributions of the matching variables for the intervention and control groups before and after matching.

Trial period

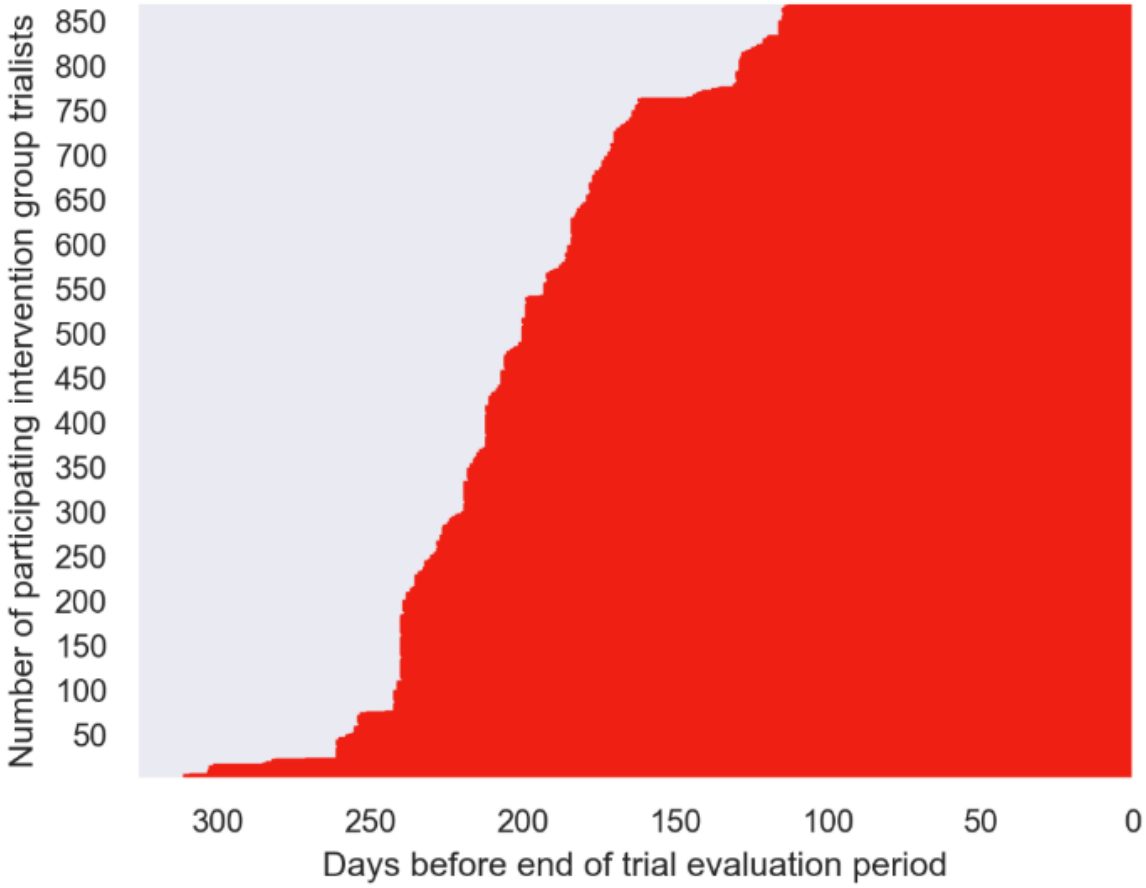
The trial period varied per trialist. Recruitment of intervention group trialists began in April 2021 and concluded in November 2021 (see Annex II for more details). Their trial period began on the day that they provided consent to participate, so varied between April and November 2021, and in all cases ended at the end of March 2022 (at the end of the trial period) to allow sufficient time for analysis. Note that for the energy consumption analysis, the evaluation period is slightly shorter than the trial period, ending 13 March 2022. Of the intervention group trialists that had sufficient data for analysis of their primary outcomes, this evaluation period

was a mean of 203 days (standard deviation 41 days), ranging from a minimum of 110 to a maximum of 311 days.

Control group trialists were recruited in January and February 2022, but their trial period was defined as being the same as the intervention group trialist to which each control trialist was matched (in many-to-one matching where there are multiple 'instances', or copies, of some of the control trialists in the final dataset, each instance was uniquely matched to one intervention group trialist and inherited the trial period of that trialist). As smart meters store historic energy data for 13 months, energy use for the control group trialists could be collected for their evaluation periods even though these periods occurred before recruitment.

Figure 4 below presents the distribution of evaluation periods for the intervention group trialists – note that, by definition, the control group evaluation periods match the intervention groups.

Figure 4: Distribution of lengths of evaluation period for the intervention group



Regression analysis

The primary outcome being tested in the trial was the average consumption difference between the intervention and control groups. Analyses were performed separately for the three outcome measures: daily mean electricity use, daily mean gas use, and daily mean energy use (electricity plus gas).

Testing was conducted using Ordinary Least Squares (OLS) regression. The outcome measure was taken to be the mean daily electricity, gas or energy use during the evaluation period. A binary “trial group” variable was used in the model to distinguish between intervention and control group membership. Prior consumption was included in the model as a control variable. Although there was a degree of skew in the energy consumption measure, and hence non-random distribution of errors in the final results, no transformations of the outcome measure were applied, as this would have impacted on the interpretability of the results.

Once fitted, this model provides a coefficient for the trial group variable that indicates the estimated size of effect of the SENS GenGame app on the primary outcome in terms of average daily energy saving compared to the control group, and the probability of this result being statistically significant, i.e., a true effect rather than a chance result.

In SENS GenGame, as described above, the evaluation period varied substantially between trialists, resulting in the primary outcome measures not being directly comparable between trialists, as they cover different time periods with different external conditions, such as weather. This makes interpreting the meaning of the estimated effect size from the regression model difficult. To account for this, an approximation was calculated of the size of effect that might be expected if the trial had lasted for a full year up until the end of the evaluation period. This was done by dividing the estimated effect size by the mean ‘heating degree day’ (HDD) value for the trialist’s evaluation period and multiplying this per-HDD value by the mean heating degree day value for the trialists for the full year leading up to the end of the evaluation period. Heating degree days are “a measure of the extent to which external temperature over a given period fell below a level below which central heating is assumed to be required (in the UK, commonly taken to be 15.5°C). The heating degree day values are calculated based on the hourly external temperature data.”[9] This partially controls for the variation in average conditions arising from the differing evaluation periods. However, this annualised figure can only be considered indicative, and the robustness of this and the other regression model results are discussed in the results section (Chapter 4).

Matched control designs make it difficult to rule out the effect of differences in the group characteristics on the primary outcome. To partly account for this, the regression analyses were run with the resultant datasets from the statistically best matching approach as well as two other matching approaches, as described in the previous section. This tests the consistency of the estimates between the different methods. Consistency provides some degree of confidence that a stable effect of the intervention exists, where one is identified in the regression results.

These analytical approaches attempt to address the issues arising from the matched control design and the differences in evaluation periods between trialists. Caveats about the

[9] Few, Pullinger, McKenna, Elam, Webborn and Oreszczyn (2022) Smart Energy Research Lab: Energy use in GB domestic buildings 2021. Variation in annual, seasonal, and diurnal gas and electricity use with weather, building and occupant characteristics. (SERL Statistical Reports: Volume 1), https://serl.ac.uk/key-documents/reports/.

robustness of the results arising from these methods are discussed in the results section of the report (Chapter 4).

3.2.3 Secondary analyses

Analyses for the secondary outcomes evaluated in this trial, as well as supplementary analyses for the primary outcome, are based on the survey and interview data collected from a sub-sample of the intervention group trialists. Unless otherwise stated, survey statistics presented in the results sections are based on:

- for baseline results, the full set of responses available (N=315);
- for endline results, the respondents who were surveyed and who had used the SENS GenGame app by the end of the trial (N=113).

As these are results based on intervention group trialists from the start and end of the trial, and the endline results are from a subset of trialists from the baseline results, then care needs to be taken in their interpretation. In particular, there were contextual changes between the baseline and endline that could influence responses and whose effects cannot be excluded, including the fact that the endline was during the heating season whilst many of the baseline surveys were during the non-heating season, and there had been substantial increases in energy prices over the period. The discussions of the survey findings in the results section below highlights these and other factors where relevant.

Interview data has been used to supplement the survey results where relevant, to give a fuller qualitative insight into the thoughts of particular trialists in relation to the secondary outcomes.

Finally, for the secondary outcome “Increased app engagement led by gamification”, summary statistics of trialists’ usage of the SENS GenGame app were used to supplement the survey and interview data. As described earlier, these included data on the numbers of logins and numbers of uses of some of the gamification features within the app.

4 Analysis of primary outcomes

This section describes the extent to which the results of the trial provide evidence that the expected primary outcomes of the SENS GenGame product were achieved, i.e. that the product led to a reduction in average gas and electricity consumption in intervention group trialists, based on a comparison with the control group trialists' consumption. The principal source of evidence comes from an analysis of smart meter energy data from the evaluation period. Survey and interview data provide further context to the results of the energy analyses.

4.1 Energy consumption analysis for the evaluation period

The energy consumption analysis found that the SENS GenGame product supported intervention group trialists who used the app to reduce their gas use by on average around 1.4 to 3.6 kWh/ day, over the evaluation period. This is equivalent to a reduction in daily gas consumption over the evaluation period of 4.6% ± 2.0% (95% Confidence Interval, $p<0.001$) compared to the control group average. Results from the supporting methods were numerically similar and also statistically significant ($p<0.001$) providing evidence that this was a robust result.

However, there was no consumption-based evidence that the product led to reductions in daily mean electricity use. For this metric, neither the primary model nor the two supplementary models produced statistically significant results at below the 5% level.

For overall energy use (combined gas and electricity), estimated savings were 1.4 kWh/day to 3.7 kWh/day, a result that was statistically significant at $p<0.001$. This is consistent with the above results, being predominantly shaped by the influence of the gas data and is equivalent to a reduction of 3.9% ± 1.8% (95% Confidence Interval, $p<0.001$) compared to control group average daily energy use. Again, this result is robust, with the supplementary models producing similar figures, also statistically significant, $p<0.001$.

4.2 Annualised energy consumption analysis

The percentage savings in gas savings in the intervention group compared to the control group can be considered reliable, including the statistical significance of these results. However, the daily mean kWh savings are hard to interpret, as they are based on data from trialists who participated in the study for substantially varying periods of time (between 110 and 311 days, as per Figure 4 on page 24), and hence had differing average climate conditions during the trial, among other changeable factors.

To attempt to produce more interpretable energy savings figures, the results outlined in the previous section were converted into estimates of what the annualised energy savings might be if the trial period had lasted a full year for every participant. These adjusted figures indicate

a saving of 1.1 kWh/day to 2.9 kWh/day for gas, which translates to 398 kWh per year to 1044 kWh per year, with overall energy (gas plus electricity) savings being very similar (as there are no electricity savings indicated by the results).

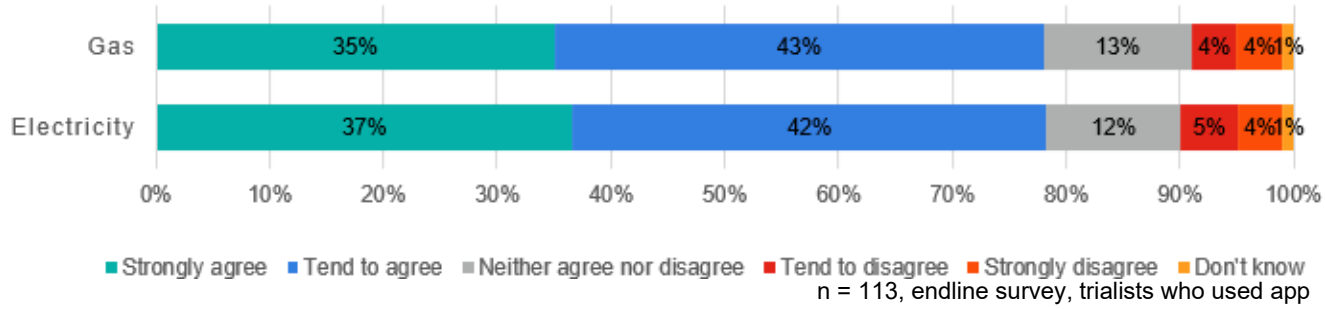
These figures are based on a simple extrapolation of the evaluation period results to a full year, by correcting for differences between the average daily heating degree days over the evaluation period and for the full year ending on 13 March 2022 (the end of the evaluation period). The figures are only estimates, as this extrapolation method assumes that all energy consumption is shaped equally by heating degree days and that the SENS GenGame intervention would continue to have the same average effect on energy consumption over this longer period as it did over the evaluation period, which cannot be verified with the current study data. In reality, engagement with the GenGame feedback and resultant behaviour changes may vary up or down over such a timescale. A longer period of time might also increase the app's achieved average savings by providing more time for trialists to implement advice requiring investments and changes in lighting, insulation, and other modifications in their homes. Nevertheless, in the absence of a full year of data, the figures provide an estimate of the potential average annual savings from the SENS GenGame intervention.

4.3 Survey evidence

Almost 80% of respondents to the endline survey who had used SENS GenGame agreed that they had tried to reduce the amount of both their gas use and their electricity energy use (the survey asked about each separately) since using the app (78% and 79% respectively); only 8% and 9%, respectively, disagreed (see Figure 5 below).

While this suggests a high level of engagement with reducing energy use, the lack of clear effect on electricity use attributable to the app may be partly because, even at the baseline, 87% of intervention group respondents (n=315) agreed with the statement "I have tried to reduce the amount of energy I use at home". Although this had indicatively risen to 90% among the endline respondents who had used the app, it implies some energy saving options open to trialists might already have been taken before the trial began, reducing the scope for further reductions.

Figure 5: Percentage of respondents agreeing with the statement: "I have tried to reduce the amount of [gas or electricity] my households used since using the product".



n = 113, endline survey, trialists who used app

Note that where the main text reports the sum of various response options, rounding of decimal places means that totals may differ slightly from the sum of percentages in the figure below.

The relatively short period of participation for many trialists may also have limited the scope to take action on the investment-related advice within the trial period. There was evidence from the survey that a substantial proportion of the trialists were engaged with considering their energy use – even at the baseline, 57% of respondents disagreed with the statement "I don't spend much time thinking about my energy use", while 31% agreed; by the endline, of those who had used the app., 60% disagreed and 26% agreed.

5 Analysis of secondary outcomes

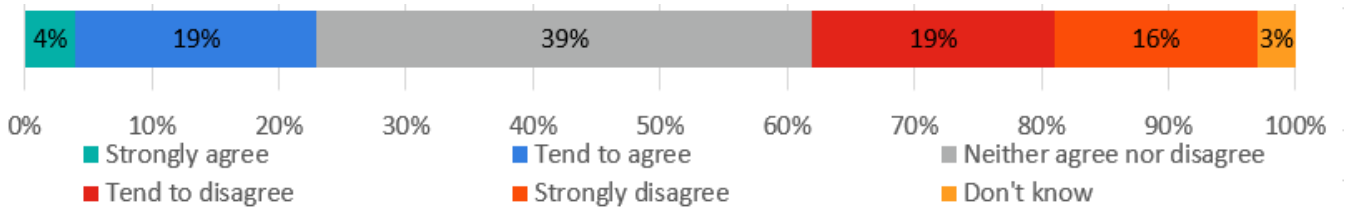
This section reports the results of further analyses of the survey and interview data relating to evaluation of the range of secondary outcomes described in chapter two.

5.1 Improved individual perception of home comfort

There was mixed evidence about the effect of the app on trialists' perceptions of home comfort.

By the end of the trial, almost a quarter (24%) of SENS GenGame users that were surveyed agreed that they found it easier to heat their homes to a comfortable level since starting to engage with the product, while 35% disagreed and 39% neither agreed nor disagreed (see Figure 6 below). This suggests that a minority of users felt they became more able to heat their home to a comfortable level.

Figure 5: Percentage of respondents agreeing with the statement: "I have found it easier to heat my home to a comfortable level since I started to engage with the product".



n = 113, endline survey, trialists who used app

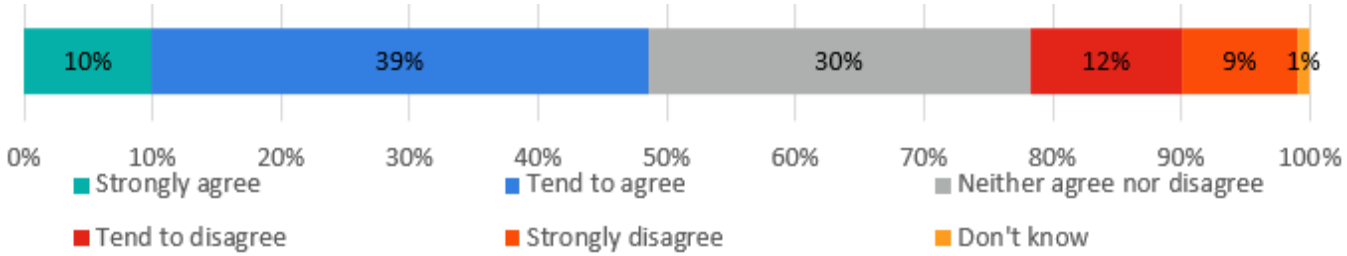
In terms of barriers to achieving home comfort, how well insulated the building was and limitations of heating systems were not an issue for many trialists. At the start of the trial, 90% of survey respondents agreed that they were usually able to keep their homes at a comfortable temperature during the winter, and only 12% reported any difficulty doing so with the heating on. Cost was a barrier for a significant minority of trialists: at the start of the trial, 22% of respondents agreed that it was too expensive to heat their homes to a comfortable temperature. This had indicatively increased slightly to 28% of responding users of the SENS GenGame app at the end. This could be a reflection of wider energy market conditions (such as tariff increases, which occurred for many of the trialists towards end of the evaluation period) and/or the fact that the endline survey was conducted during the heating season (early 2022) rather than an effect of the SENS GenGame app.

5.2 Improved household budgeting

There was evidence that the SENS GenGame app supported users to improve their household budgeting, with about half (49%) of app users agreeing that they found it easier to control how

much they spent on energy since starting to engage with the product (20% disagreed with this statement, and 30% neither agreed nor disagreed – see Figure 7 below).

Figure 6: Percentage of respondents agreeing with the statement: "I have found it easier to control how much I spend on energy since I started to engage with the product".



n = 113, endline survey, trialists who used app

Amongst the same set of respondents (endline survey respondents who had used the app), two thirds (66%) reported that they were a little or much more confident about how much their household spent on energy each month since starting to engage with the app, while 3% were less confident and almost three in ten (28%) felt they were as confident as they were before. Within the same group, three in ten (31%) also responded that they were more confident in knowing which tariff or payment plan they had with their energy supplier than they had been before, around half (54%) said they were as confident as they were before, and 3% were a little less confident (the remaining 12% did not know).

For context, the proportion of respondents who agreed with the statement "I am very conscious about the cost of the energy I am using" indicatively increased between the baseline and endline surveys. Of the 149 trialists who completed both the baseline and endline survey, eight in ten (79%) agreed with this statement at the baseline (46% of those agreed strongly), and only one in ten (11%) disagreed. By the end this had indicatively increased to the large majority (93%) agreeing with this statement (two in three (62%) agreed strongly), while only 5% disagreed. This could well be due to the background context of rising energy prices and potential behaviour changes in response to this towards the end of the trial. Over the same period, for the same group of respondents, when asked to describe how well they and their household were keeping up with their energy bills at the moment, those responding that they were managing very well or quite well indicatively dropped from 86% to 74%; although the number reporting difficulties remained at 3%. This indicative change could also be due to the timing of surveys, with the endline occurring towards the end of the Winter heating period.

Of the 135 trialists who responded to the baseline survey, three quarters (74%) had reported that they would find the ability for their household to monitor what they spend on energy, for example against a household budget, very or fairly useful, while about a quarter (23%) felt it would be not very or not at all useful. Almost six in ten (57%) of 315 respondents reported having monitored what they spend on energy before the trial started.

Among the trialists who took part in the qualitative interviews there were only two who set a household budget, and 13 who did not. Recurrent reasons given for not setting a budget were:

- Already using only what energy is needed/ already reducing what they can, so setting a budget would not change the amount of energy they use

  "Apart from thinking twice about the heating, we need to use what we're using. We need to cook; my daughter can't switch baths to showers."

- Feeling that the cost of energy was outside their control

  "No [don't have a budget], because it's not in my control. It's just, we have to find the money. It's not like we can go to a cheaper tariff, so we just find money for the energy bills."

  "We do not budget because energy prices are a bit out of our control."

- Comfortable financially, so no need to budget.

  "No [there is no] set budget as [I am] financially comfortable – it is a pain to pay more but it will still be affordable."

One interviewee also mentioned that setting a household budget could put them at risk of practising unhealthy behaviours such as not heating the home properly or not cooking proper meals in order to save money on energy.

5.3 Reduced unoccupied heating hours

There was survey evidence that use of the GenGame app did not affect unoccupied heating hours. In both the baseline and endline surveys, almost three in ten respondents (28%) strongly agreed or tended to agree with the statement that they tended to leave the heating on when going out for a few hours, while just over two thirds disagreed (67% at the baseline, and 69% at the endline among those who had used the SENS GenGame app). This indicated that only a minority of surveyed SENS GenGame trialists heated the home while it was unoccupied for substantial periods before the start of the trial, and that this level had not changed by the end of the trial.

Survey data also indicated that levels of another heating energy-saving behaviour, zoning (i.e. heating rooms differently), were not substantially changed by the SENS GenGame app. Responses to a question about agreement that they tend to heat rooms that are not being used to the same temperature as those that are being used similarly changed little between the baseline and endline: 34% agreed at the baseline and 62% disagreed; at the endline, among those who had used the SENS GenGame app, 32% agreed and 63% disagreed.

Six of 15 interviewees mentioned heating-related behaviours and attitudes that they had adopted during the trial, although none discussed their behaviours relating to reducing heating during periods when the home was unoccupied. Rather, they highlighted other measures taken to reduce heating use: putting jumpers on rather than the heating, if they were not too cold;

turning down the central heating and instead using the gas fire in the lounge; turning off radiators in rooms not in use (zoning); turning off heating during the night; closing curtains during the night and generally using heating 'much less' than before.

> *"[I] put the heating on less. We think now, 'do we need the heating on all day'. I am more aware of what I'm using".*

> *"[I] don't fill the kettle anymore after this was one of the challenges, [I] only fill it to what is needed."*

> *"It [the app] makes me think twice 'Do I really need that? Do I need the heating on? Can't I put another jumper on? It makes me think properly."*

> *"[I] will turn everything off if we go away."*

Such behaviour changes may account for some of the gas savings observed in the energy consumption analysis.

5.4 Increased understanding of energy use and drivers of energy consumption

At the end of the trial, among the 113 SENS GenGame app users surveyed, three quarters (77%, or 87 trialists) agreed that their understanding of how their household uses energy had got a little or a lot better 'over the last year'. Nearly all (93%, or 81 trialists) of those 87 attributed this at least partly to information they had received from the SENS GenGame app (63% said it was entirely or mostly because of this information; 30% said it was partly due to this information). In addition, just over half (55%) of the 113 respondents agreed that they were much or a little more confident now about which activities or pieces of equipment required a lot of energy in their homes, and two thirds (67%) agreed they were much or a little more confident now about what changes they could make to save the amount of energy used in their homes.

Furthermore, of 315 intervention group respondents, three quarters (76%) agreed at the baseline that they knew what used the most energy in their homes. By the end of the trial, this had indicatively increased to nine out of ten (91%) of the 113 asked who had used the SENS GenGame app.

As a caveat to these results, six out of ten (62%) of the 315 baseline respondents also agreed with the statement that their understanding of how their household uses energy had got a little or a lot better 'over the last year' (although many had already used the app before this was asked of them).

Among the interviewees, a high proportion mentioned a general increase in awareness of energy use and what different actions and appliances cost, for one or both fuels. The tips, graphical representations of energy use and summaries of energy use and resultant

environmental impacts were all mentioned. A range of specific examples of increased awareness were also highlighted by different trialists in the interviews, including:

- Increased awareness and consideration of impacts on the environment.

  "I was curious about my energy usage, and I thought that gaining better insight, why this use more and at what times, if there is a pattern of energy usage in our household… I was quite curious about that and I thought it might help us, well, reduce our energy consumption, our bills, also go a bit greener."

- The costs of gas and electricity relative to one another.
- The costs of leaving things on standby.

  "It's probably a lot of small appliances contributing to it, that I'm just leaving on when I'm not using them, and I think I just need to develop the habit of turning them off. But it's the things that you know, chargers for laptop, screens on standby mode… those ones are the ones that I wasn't aware. Was I aware of them before? I probably was but I didn't notice the costs."

- The relative costs of different appliances and behaviours, e.g. heating being the most expensive energy use.

  "It [the app] has helped me understand a bit my energy consumption and how it translated to money, it has probably helped me reduce my energy consumption. It's easy to use, not problematic to navigate, not frustrating. It's a been bit fun to use as well."

  "It is much easier to see how whatever it is that you do in the house actually affects your energy consumption and your bill in a graphical way and based on the time. It's easy to track. It has made a difference in connecting the dots between your actions, which would like a simple pressing a button on an appliance or turning on the heating, to how much energy is being used, how your bill goes up and how much more you're destroying the environment."

- Tariff structure (having to pay a standing charge even if not using any energy at all).

Two interviewees mentioned examples of appliance switching that they had implemented as a result of using the app: one switched from using an electric kettle to a gas kettle, another from cooking using the oven to on a hob. Various heating-related behaviour changes that interviewees had implemented are presented in the previous sub-section.

Overall, there is evidence that the SENS GenGame product contributed to increased understanding of energy use and drivers of energy consumption in seven out of ten trialists (72%) who used the app, based on their survey responses[10].

[10] Based on the 81 of 113 SENS GenGame app users surveyed at the endline who agreed that their understanding of how their household uses energy had got a little or a lot better 'over the last year' and that attributed this at least partly to information they had received from the SENS GenGame app.

5.5 Increased customer retention

There was no evidence that the SENS GenGame product increased energy supplier customer retention based on the data collected during the surveys. Several indirect measures that might relate to customer retention were collected via the survey at the baseline and endline, relating to levels of satisfaction with the supplier, the level of trust in them to provide customers with a service that meets their needs, and how easy or difficult they find dealing with their energy supplier. In all cases, levels of satisfaction were high at the baseline and remained essentially unchanged by the endline survey, all at around 80% reporting satisfaction, trust, and ease of dealing with them (responding as fairly or very satisfied, or between 7 and 10 on a 10-point scale). (Percentages for all three questions differed by less than 1% point between the baseline and endline.)

5.6 Increased app engagement led by gamification

Of the 1,912 trialists that signed up to participate in the intervention group, 57% (1092 trialists) went on to download and make use of the app at least once, based on app usage data. 78% of the app users logged in at least five times during the trial period. The average number of logins was 26 among those who used it. The highest was 463 logins.

Survey data suggested that for a slight majority of SENS GenGame app users, once they had installed and made use of the app and the gamification features within it, those gamification features did lead to increased engagement. Seventy-nine respondents who participated in the endline survey reported having used the gamification features. Of those, 56% agreed with the statement “The challenges and rewards system of the Energy Saver App made me use the app more than I otherwise would have” (19% disagreed, 23% neither agreed nor disagreed, and 3% did not know).

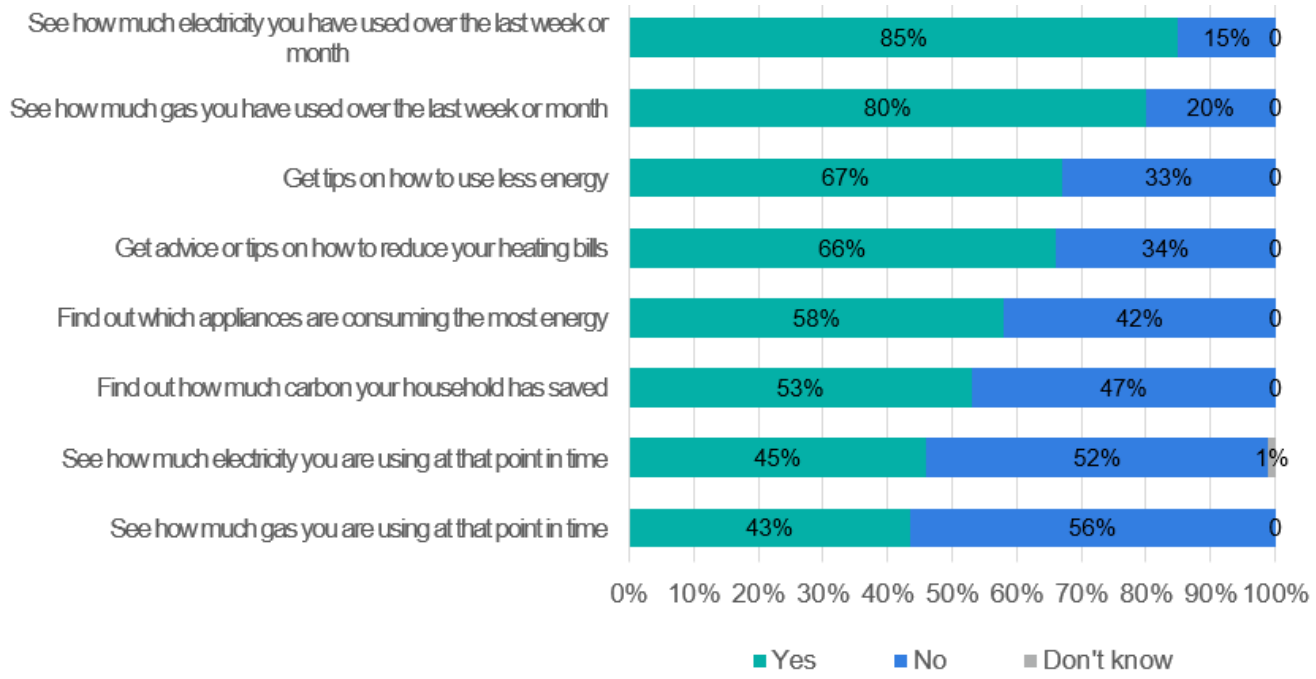
App usage data indicated that three quarters of the trialists (74%, 804 trialists) who logged into the app at least once also made use of the tips and prize draw gamification features at least once. Survey data was consistent with this: 70% of those asked who had also installed the app reported at the endline that they had made use of the gaming feature. There is however no data about whether, for example, trialists’ prior knowledge of the gamification features being available increased the proportion who installed the app or who sought out and tried the gamification features once they had installed the app.

More broadly, survey data gave an indication of which features trialists made use of. Figure 8 overleaf shows responses among endline SENS GenGame app users about which features they had used. The most used features were for seeing how much electricity or gas they had used over the last week or month (85% and 80% of respondents for the electricity and gas features, respectively). Advice and tips on using less energy and on how to reduce heating bills, were the next most frequently used (67% and 66% of respondents, respectively). Finding out which appliances were consuming the most energy was used by 58% of respondents, while finding out how much carbon the household had saved was used by 53% of

respondents. Seeing how much electricity or gas was being used at that point in time was used by 45% and 43% of respondents respectively.

Satisfaction with the product among those who had used it was high in the endline survey. Three quarters (76%) said they were very or fairly satisfied with the app, 19% were neither satisfied nor dissatisfied, and only 5% were fairly or very dissatisfied (1% did not know). Meanwhile, nearly two thirds (64%) said they would recommend it to a friend, colleague or relative (responding between 7 and 10 on a 10-point scale from 1 - would definitely not recommend, to 10 - definitely would recommend).

Figure 8: Distribution of responses to the question: "Have you used SENS GenGame Energy Saver App to do any of the following since you first had access to it?"



n = 113, endline survey, trialists who used app

6 Conclusions

This section discusses the results of the SENS GenGame trial, including the key findings and their implications, the limitations of the trial and options for future work.

The SENS GenGame Energy Saver app aimed primarily to achieve a reduction in gas and electricity consumption in intervention group homes, through providing an app with core components of gamification, energy use and expenditure insights, and both behavioural and investment-focused energy saving advice. Secondary aims included improving individual perceptions of home comfort, reducing unoccupied heating hours, increasing understanding of energy use and the drivers of energy consumption, and increasing app engagement through gamification features.

Despite limitations in the matched control trial design compared to a Randomised Controlled Trial, there was robust evidence that the app achieved the primary outcome for gas. Among the intervention group who made use of the SENS GenGame app (Treatment on the Treated), there was robust and highly statistically significant evidence that gas use over the evaluation period was reduced by 4.6% ± 2.0% (95% Confidence Interval, $p<0.001$) compared to the control group. As discussed in the report, it was not possible to assess the Intention to Treat, those who were allocated to the intervention group but did not download and use the app, due to the point in the customer journey where the matching data was collected. The large variation between trialists in their length of participation in the trial made estimating the typical saving in kilowatt hours problematic, however an approximate annualised figure was 1.1 kWh/day to 2.9 kWh/day, or from 398 kWh to 1044 kWh per year. The trial did not identify a statistically significant difference in electricity use between the app users and the control group.

The evidence did not provide a clear understanding of why a gas saving but not an electricity saving was identified. Possible reasons related to both the trial design and to the intervention. In relation to the trial design, the numbers of trialists available for the energy consumption analysis fell below the originally planned recruitment numbers, reducing the study's power to detect differences in energy use between the groups. The shorter than planned trial period may also have reduced the ability to detect energy savings that the app might have led to over longer periods of time, particularly investment-focused advice that may take longer to act on and implement than behavioural changes.

The survey and interview data provided valuable insights into which features were of most interest to trialists who used the SENS GenGame app, which are potentially applicable to energy feedback and advice services more generally. Viewing the last week's and month's energy usage were the most used features, whilst seeing current energy use was of substantially less interest. Advice and tips on how to reduce energy use, including heating bills, were used by the majority of app users too.

Gamification features boosted engagement among the trialists, at least among those who tried them – just over half of those who used the gamification features reported that they made them use the app more than they otherwise would have. Three quarters of app users surveyed at

the endline were satisfied with the app, and two thirds said they would recommend it to a friend, colleague or relative, indicating a generally positive impression of the app among the trialists who used it.

The data provided some indication as to how this set of app features resulted in energy use outcomes. There was evidence that the SENS GenGame app contributed at least partly to increased understanding of energy use and drivers of energy consumption for over 70% of trialists who used the app, based on their survey responses. Nearly a quarter of surveyed app users also reported finding it easier to heat their homes to a comfortable level after engaging with the app, indicating it may have benefited a significant minority of trialists in this respect. Almost half of surveyed app users reported finding it easier to control how much they spent on energy since starting to engage with the app. There is evidence that use of the SENS GenGame app did not help to reduce unoccupied heating hours, nor increase zoning of heating, however interview evidence indicates it did lead to some heating-related behaviour changes.

Overall, there was evidence that the range of features in the SENS GenGame app, where used, increased engagement with the app and led to increased understanding of energy use and drivers of energy consumption, which in turn led to measurable reductions in gas use among those who used the app compared to the control group, although which behaviour changes and investments led to this result is not clear. When considering the wider potential of such apps for reducing domestic energy use, it is also worth noting that 57% of intervention group trialists who signed up to the trial installed the app and used it at least once. The evaluation results therefore potentially only apply to those more engaged intervention group trialists rather than the full intervention group (especially in the absence of Intention to Treat analyses).

There were several gaps in the evidence that could be addressed in future research. A larger scale trial, over a longer period of time, would be necessary to more robustly assess the impact upon the true mean size of the gas savings over a full year, the impacts upon electricity savings and the range of behavioural changes and investments that arose.

Glossary

ANCOVA	Analysis of Covariance
AQ	Annual Quantity (gas)
ATE	Average Treatment Effect
BAU	Business as Usual
BEAMA	British Electrotechnical and Allied Manufacturers' Association
BEIS	Department for Business, Energy and Industrial Strategy
BIT	Behavioural Insights Team
BST	British Summer Time
CA	Contribution Analysis
CAD	Consumer Access Device
CHP	Combined heat and power
CIC	Community Interest Company
CMO	Context-Mechanism-Outcome
CO2e	Carbon dioxide equivalent
COVID-19	Coronavirus Pandemic
CP	Competition Partner
CRL	Commercial Readiness Level
DCC	Data Communications Company
DESNZ	Department for Energy Security and Net Zero (formerly BEIS)
EAC	Estimated Annual (energy) Consumption
ECA	Energy Consumption Analysis
EL	Energy Local
ELC	(SENS) Energy Local Club

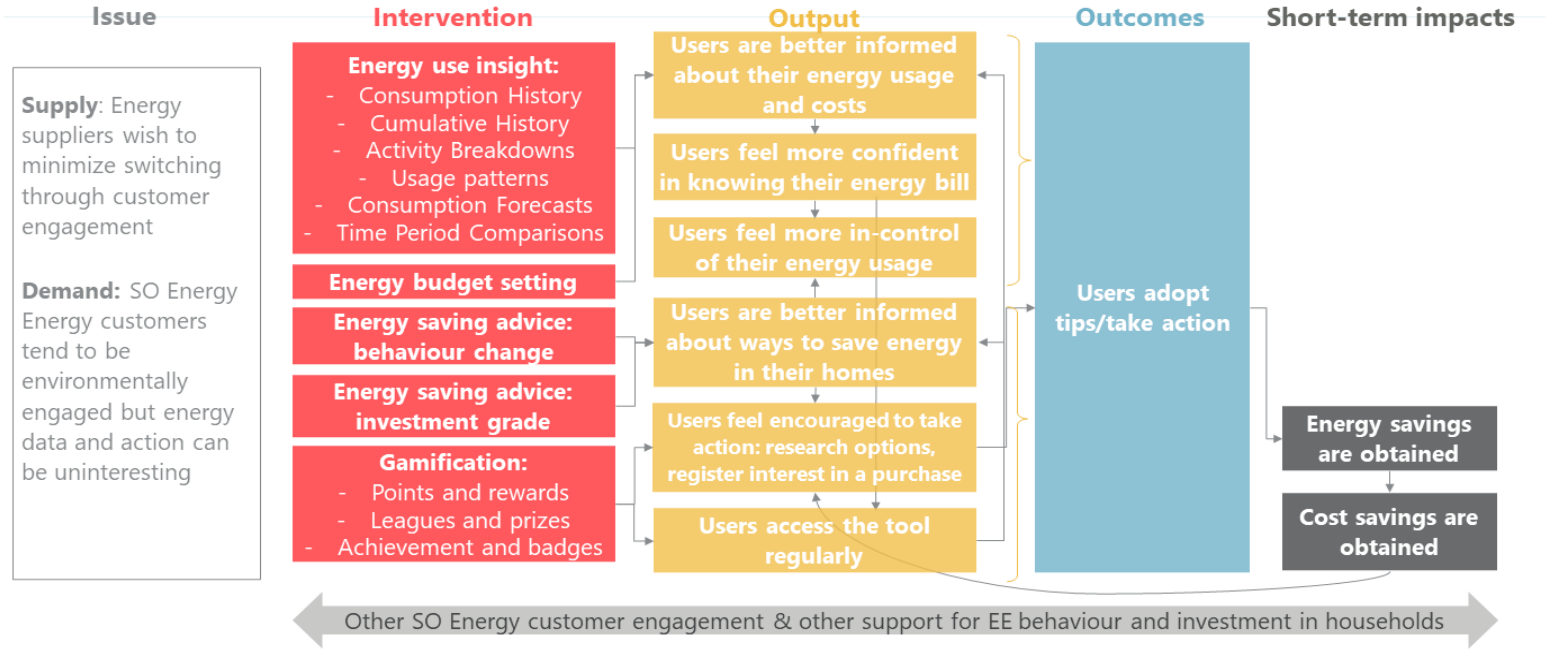
EPC	Energy Performance Certificate
GDPR	General Data Protection Regulation
GEO	Green Energy Options Ltd.
HAN	Home Area Network
HDD	Heating Degree Day
ICE	Igloo Customer Engine
IDEAS	Intelligent Digital Energy Advisory (SENS project)
IHD	In-Home Display
IMD	Index of Multiple Deprivation
ITT	Intention to Treat
KW	Kilowatts
kWh	Kilowatt-hour
M&MH	Me & My Home profile
MDE	Minimum Detectable Effect
MEETS	More Effective and Efficient Thermal comfort with Smart meter data (SENS project)
MI	Monitoring Information
MOP	Meter Operator
MPAN	Meter Point Administration Number
OLS	Ordinary Least Squares
OWL	A brand of electricity monitor used to monitor consumption in Roupell Park
PSM	Propensity Score Matching
RCT	Randomised Controlled Trial
SEC	Smart Energy Code
SECAS	Smart Energy Code Administrator and Secretariat

SENS	Smart Energy Savings Competition
SENS GenGame	SENS GenGame Energy Saver app (SENS project)
SEN-ST	Smart Energy-Smart Thermostat (SENS project)
SERL	Smart Energy Research Laboratory, based at University College London
SM	Smart Meter
SMETER	Smart Meter Enabled Thermal Energy Ratings
SMETS	Smart Metering Equipment Technical Specifications
SMETS1	Smart Metering Equipment Technical Specifications - First Generation
SMETS2	Smart Metering Equipment Technical Specifications - Second Generation
SMS	Smart Metering Services
SoLR	Supplier of Last Resort
TDEL	Trial Design and Evaluation Lead
TOT	Treatment on the Treated
TOU	Time of use
TOUT	Time of Use Tariff
TP	Trial Protocol
TRL	Technology Readiness Level
UCL	University College London
WAN	Wide Area Network

Annex I – GenGame Theory of Change

This section presents the SENS GenGame Theory of Change which sets out the issues the intervention was trying to address, the core components of the intervention itself, the outputs it was expected to deliver, the outcomes to achieve, and ultimately, the impacts of the intervention..

Figure 7: SENS GenGame Theory of Change



Assumptions & Risks:

A1. Users have a desire to reduce bill/ are not comfort seekers and have some intrinsic motivation to take low-carbon action
A2. Tips are relevant to trial households
A3. Users trust the tips and costing estimates and feel able to take action
A4. There is potential to save energy at low cost in trial homes
A5. Following tips leads to savings
A6. Users have access to finance necessary to implement investment level measures

R1. Users are comfort seeking
R2. Users purchase EV as result of increased engagement with energy use
R3. Users access contradictory advice

Annex II – Trial Overview

The table below describes the numbers of households contacted and trialists initially recruited and the numbers removed from the energy consumption analyses due to different data quality issues.

Milestone / stage / sample		Number / count (households)	Date (where applicable, and including start and end date as needed)
Number of households / customers contacted to participate in trial	Intervention	5141 dual-fuel SM2 customers	April 2021 – November 2021
	Control	9,000 duel-fuel SM2	January 2022 – February 2022
Number of households / customers that agreed to participate	Intervention	1912	April 2021 – November 2021
	Control	1068	January 2022 – February 2022
Number of households / customers providing consents to be contacted for TDEL research	Intervention	1912	April 2021 – November 2021
	Control	N/A	N/A
Number of households / customers providing consents to SERL	Intervention	1912	April 2021 – November 2021
	Control	1068	January 2022 – February 2022
Number of trialists onboarded to SERL	Intervention	1912	April 2021 – November 2021
	Control	1068	January 2022 – February 2022
Number of trialists / trialists who went on to download the app and login at least once	Intervention	1092	April 2021 – November 2021
	Control	N/A	N/A

Number of withdrawals over trial period (up to end March 2022)	Change of tenancy	92		April 2021 – March 2022
	Change of supplier	36		
	Withdrawal of consent	18		
	Other (On Hold)	4		
Final achieved sample (Sample at the end of the trial period, accounting for churn of trialists)	Intervention	1760		N/A
	Control	1067		N/A
Final achieved sample for quantitative analysis (i.e. less records excluded for data issues outlined below)	Intervention	871		N/A
	Control	915		N/A
Number of trialists excluded from analysis and reasons: (Intervention group figures on left hand side; control group figures on right hand side. Counts are of trialists excluded from the sample remaining after the previous rounds of exclusions)	Smart meter energy data missing	14	2	N/A
	Trialist has microgeneration (private solar PV or other private power generation)	4	2	N/A
	Did not log into app	766	0	N/A
	Pre-baseline energy data not available (EAC and AQ estimates)	65	7	N/A
	Trialists have multiple MPANs associated with them	3	0	N/A
	Incomplete attitudinal survey responses (for PSM matching)	6	0	N/A
	Less than 50% of smart meter data available for required period (control trialists only)	0	40	N/A
	Less than 50% of smart meter data available after consent_given_date	28	100	N/A

	Pre-baseline energy data (AQ) could not be matched to trialist	3	1	N/A
Baseline survey issued / response rate (intervention group)	No. of contacts available to be contacted[11]	1450		August 2021 – December 2021
	No. of completed interviews	315		
	Completion rate	22%		
Endline survey issued / response rate (intervention group)	No. of contacts available to be contacted	286		March 2022
	No. of completed interviews	149		
	Completion rate	52%		
Qualitative interviews completed with intervention group trialists		15		February 2022 – March 2022

[11] Not all trialists provided supplied contact details.

Annex III – Technical Appendix

This section presents technical figures and data referred to in the main text.

Propensity Score Matching

Section 3 discusses the use of Propensity Score Matching (PSM) in the data analysis for this trial. More details are provided here.

PSM is a statistical method to attempt to better match the distributions of input variables for two groups based on – in this case, for the matching variables described in methodology. Multiple PSM methods and parameters for those methods exist, and a range were tested to find the best performing options, based on standard diagnostic results (primarily the mean within-pair difference of each covariate after matching). Matching was undertaken using the MatchIt library in R.

The matching method selected for use in SENS GenGame was a many-to-one genetic matching algorithm. Many-to-one matching means all treatment group trialists with adequate data were retained, and the best matching control group trialists were matched to each. This allows control group trialists to be potentially matched to multiple treatment group trialists each, while other control group trialists may not be matched to any treatment group trialists, thus omitting them from further analysis. Genetic matching with a many-to-one approach showed significantly better diagnostic results than the other methods so was selected as the primary approach. The results of two other matching methods are also presented here as evidence that the effects found are robust across methods. These two further matching methods were chosen as they achieved the optimum results for different outcomes:

- One-to-one matching – all treatment group trialists with adequate data were retained, and the same number of control group trialists were retained.
- Many-to-one-matching – the method retaining the most control group trialists.

Table 4 summarises the matching parameters and results for the three best performing PSM methods.

After matching, each intervention trialist had one control trialist matched to it (in the case of many-to-one matching, some of the control trialists were matched to multiple intervention trialists, so that there were multiple ‘instances’ of some of them). Control group trialists’ energy consumption data during the trial was then estimated for the same period of time as the participation period of the intervention trialist to which each was matched. In some cases, control group trialists lacked sufficient energy data to estimate one or more of the primary outcome measures. In these cases, these control trialists were dropped from further analysis and the matching process rerun without them, so that the final result was a sample where each intervention trialist with complete data was matched to (an instance of) a control trialist, and all had the complete set of data required for the subsequent regression analyses.

Table 4: MatchIt parameters and results used for the three best performing Propensity Score Matching methods

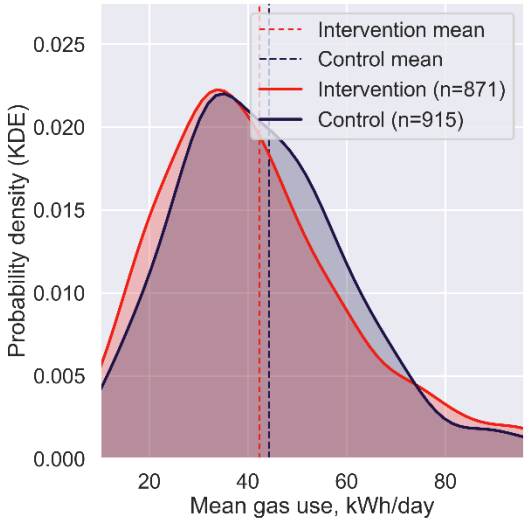
Primary matching method: genetic many-to-one
Method=genetic; distance=glm; link=probit; discard=none; pop.size=10 Outcome: 872 (of 872) intervention trialists matched to 387 (of 915) control trialists Best performing of all the approaches tested
Auxiliary matching method 1: optimal one-to-one
Method=optimal; distance=glm; link=probit; discard=none; pop.size=10 Outcome: 872 (of 872) intervention trialists matched to 872 (of 915) control trialists Best performing one-to-one method approach tested – although it was not possible to obtain high quality matching with ant one-to-one method due to the imbalances in the distributions of the input variables
Auxiliary matching method 2: nearest-neighbour many-to-one
Method=nearest; distance=glm; link=logit; discard=none; pop.size=10 Outcome: 872 (of 872) intervention trialists matched to 411 (of 915) control trialists Many-to-one approach that retained the most control trialists out of those tested

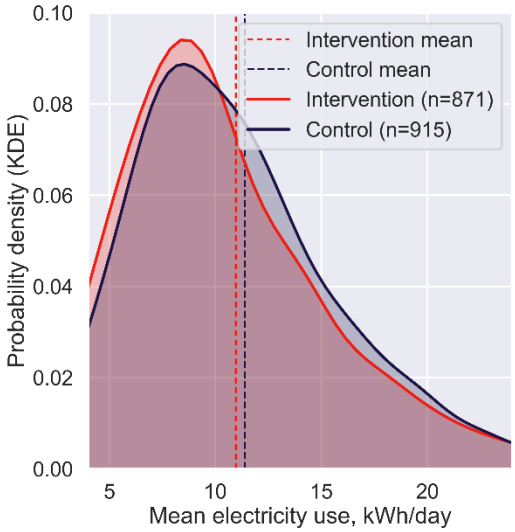
Sample characteristics before and after matching

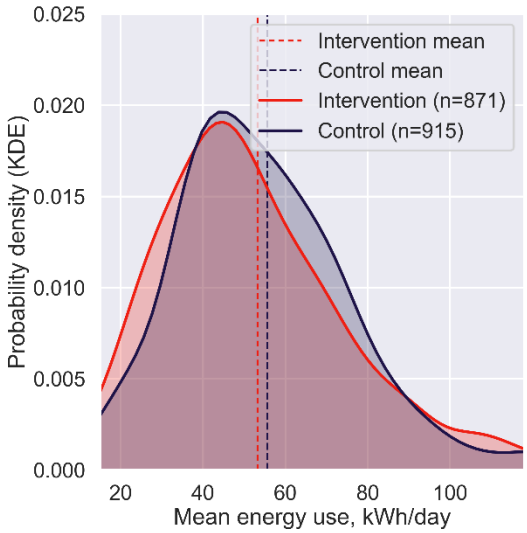
Figure 11 to 14 present the distributions of the matching variables before and after matching, comparing control and intervention groups. Although results are only presented for the primary matching approach (genetic many-to-one), as expected, the many-to-one matching methods resulted in better matching of the values of the variables across the control and intervention trialists, at the cost of omitting the majority of control trialists from the final dataset. The control and intervention trialists are less well balanced after matching when the one-to-one method is applied, but retain most of the control trialists as well as the intervention trialists.

Figure 8: Distribution of pre-baseline electricity, gas and combined energy use for the intervention and control trialists before and after matching (primary matching approach)

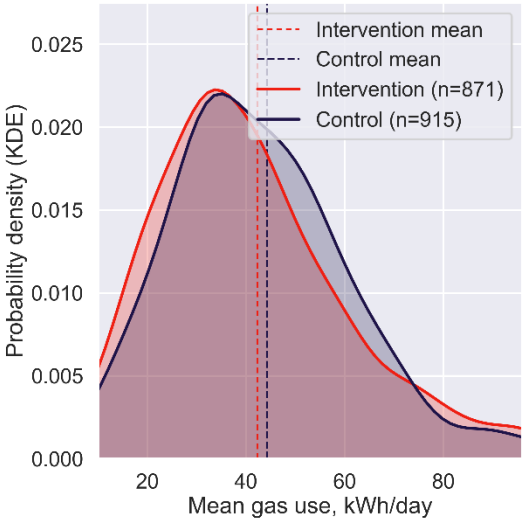
Before matching

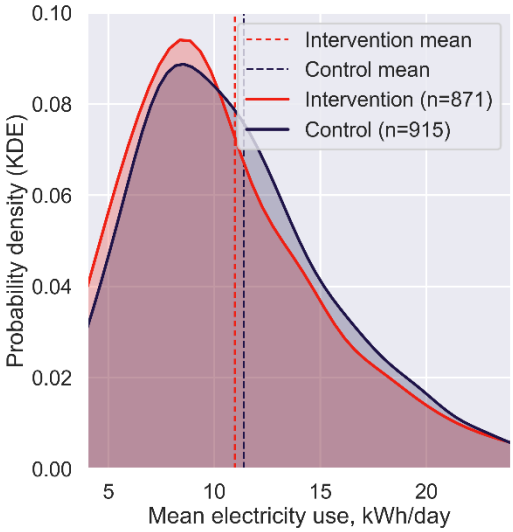






After matching





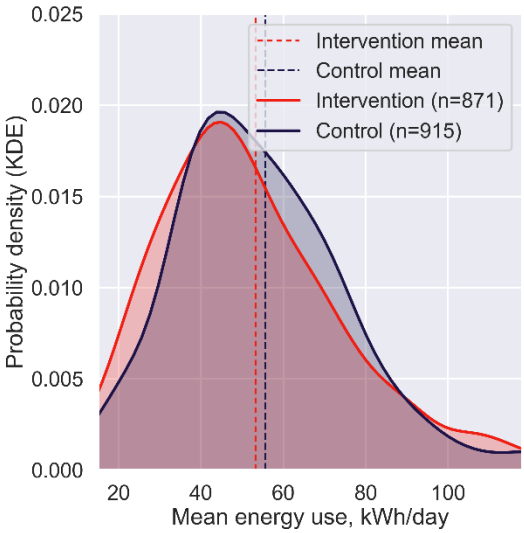
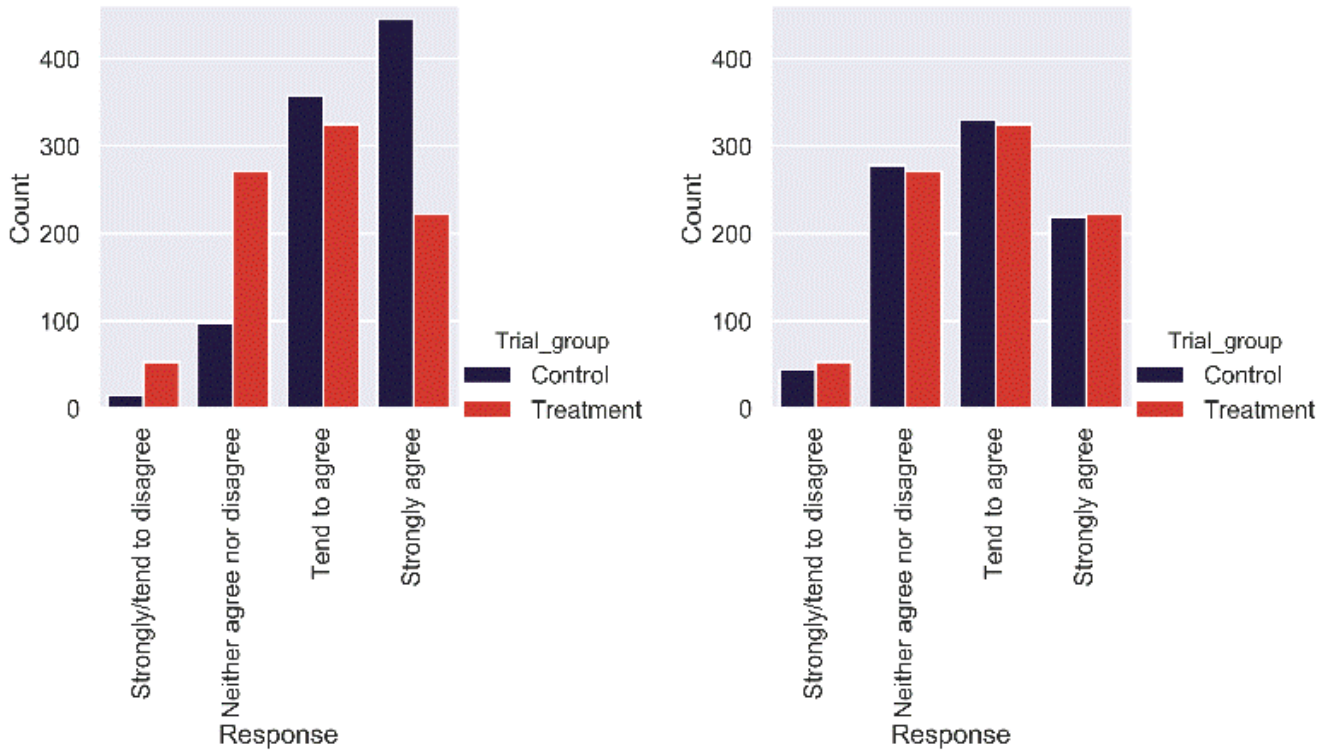


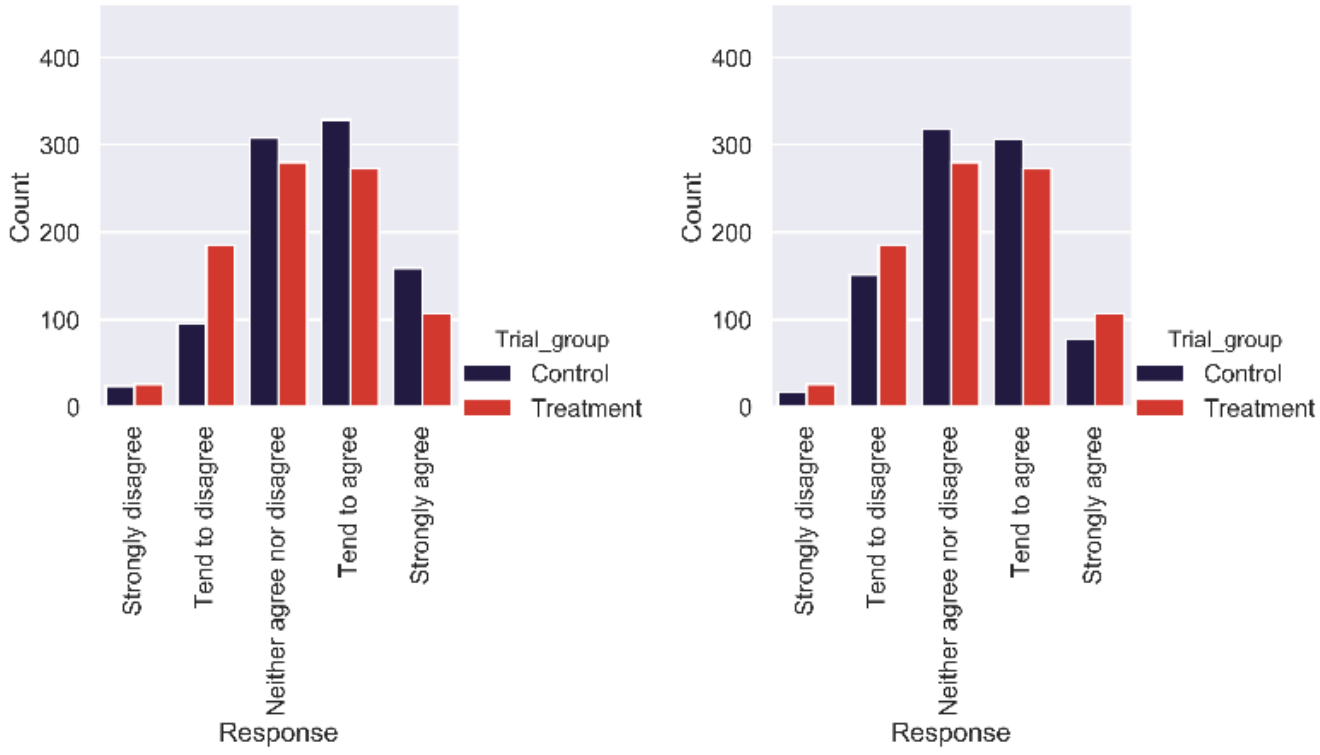
Figure 9: Distribution of survey responses for the intervention and control trialists, before and after matching (primary matching approach)

Before matching	After matching

"Climate change is likely to have a big impact on people like me."

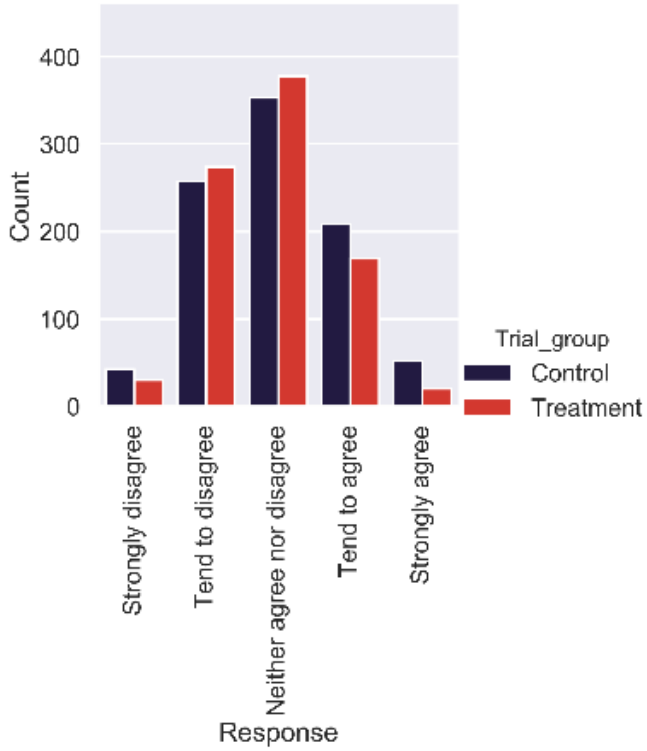


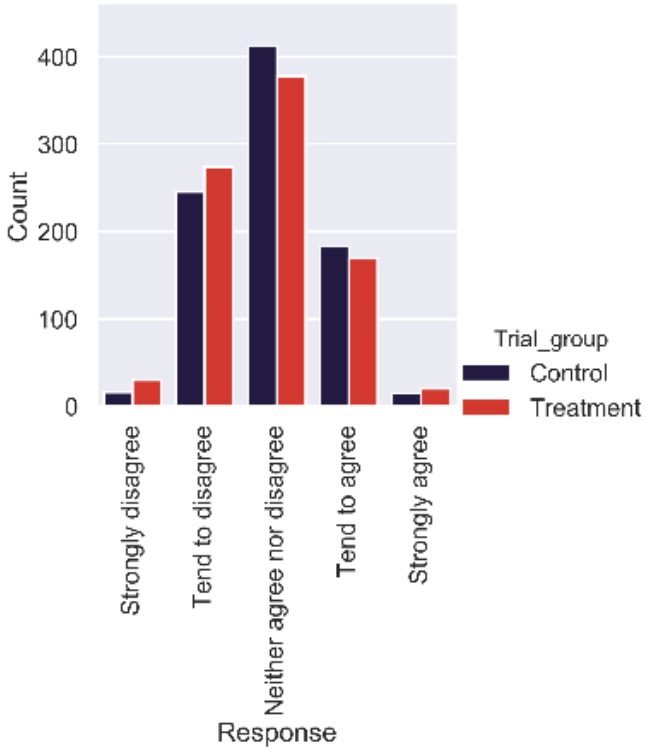
"I am the type of person who likes to have the newest gadgets in my home."



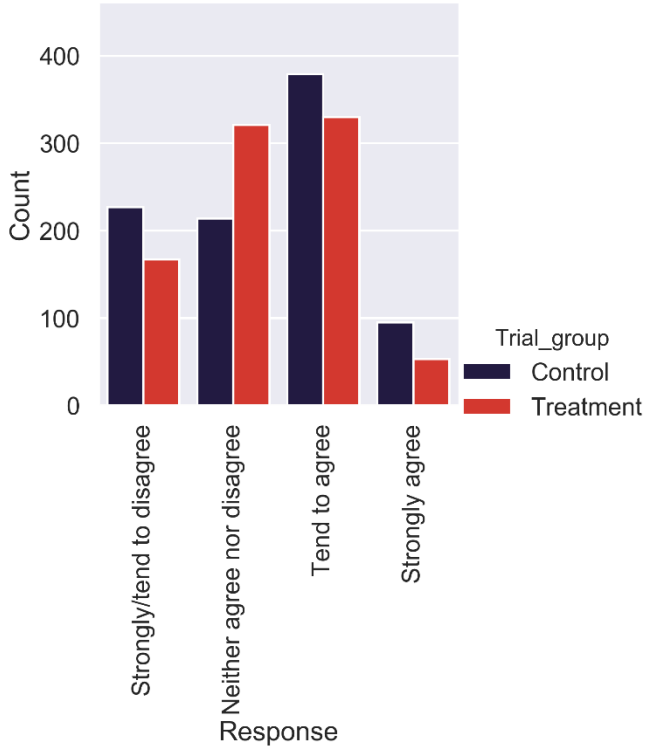
Before matching	After matching

"I am more concerned about having a warm and comfortable home than saving energy."





"I feel in control of how much energy I personally use."



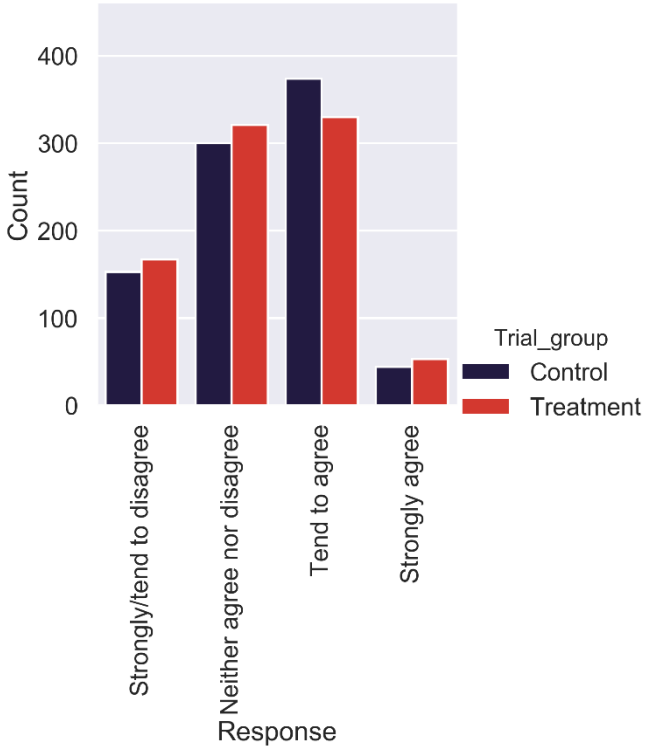
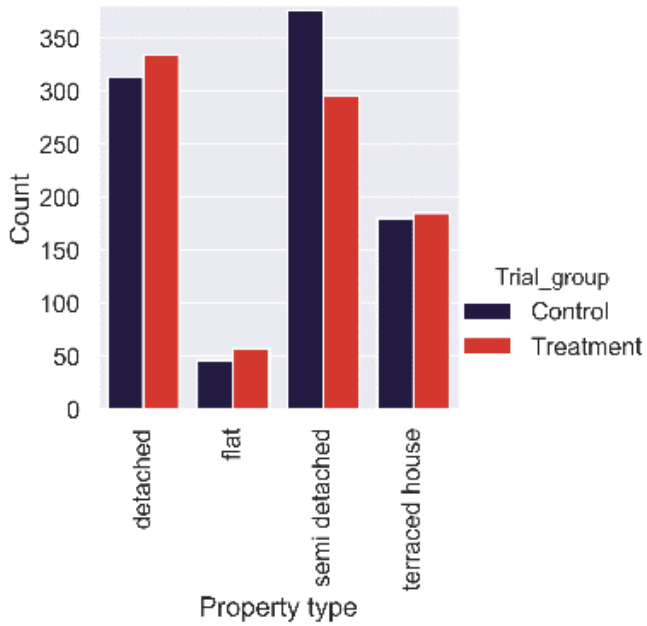
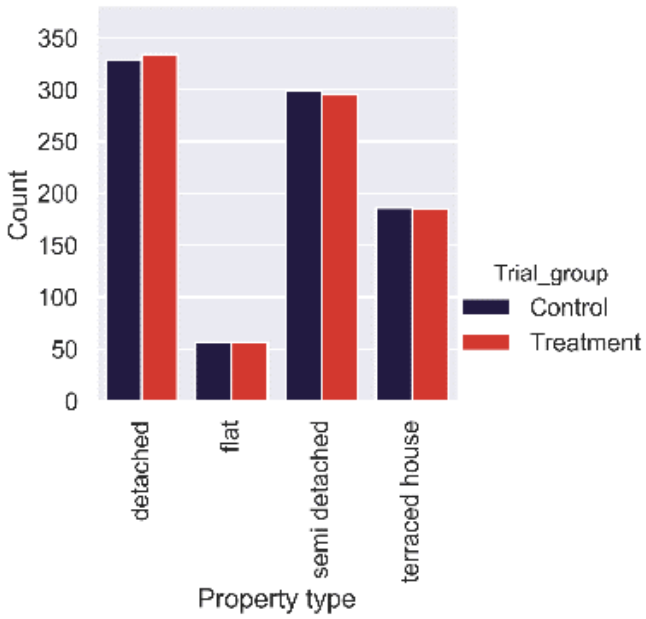


Figure 10: Distribution of property type and region characteristics for the intervention and control trialists, before and after matching (primary matching approach)

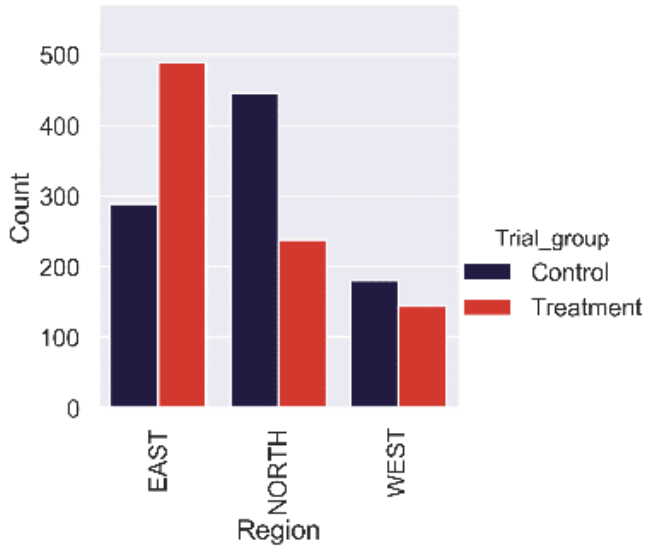
Before matching	After matching

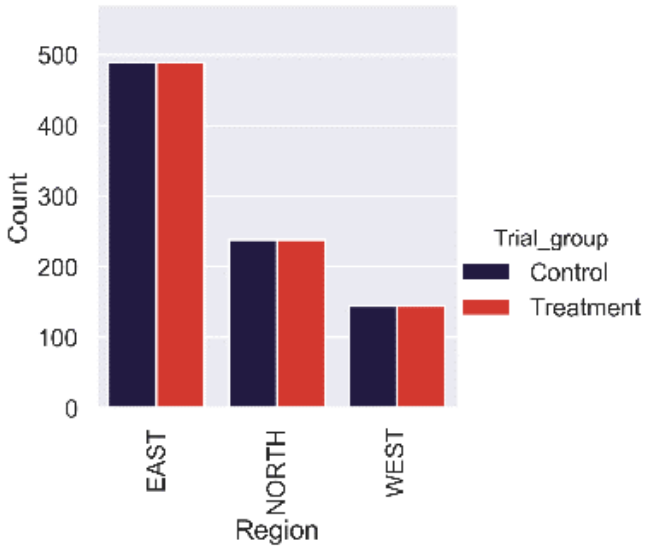
Property type





Region

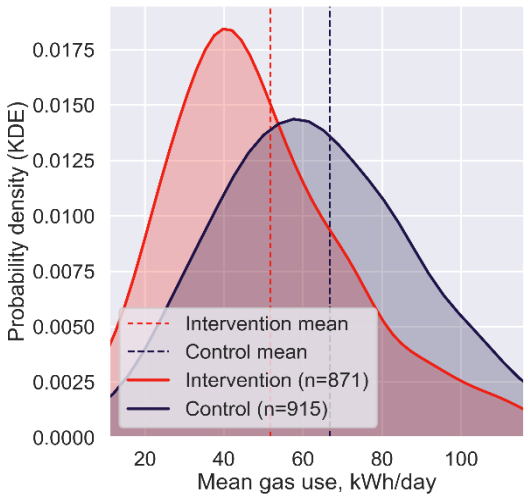


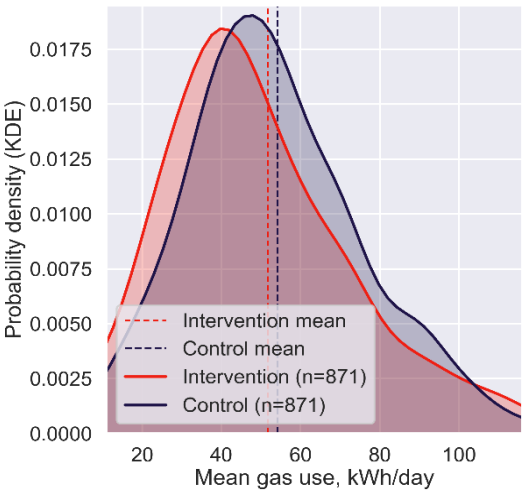


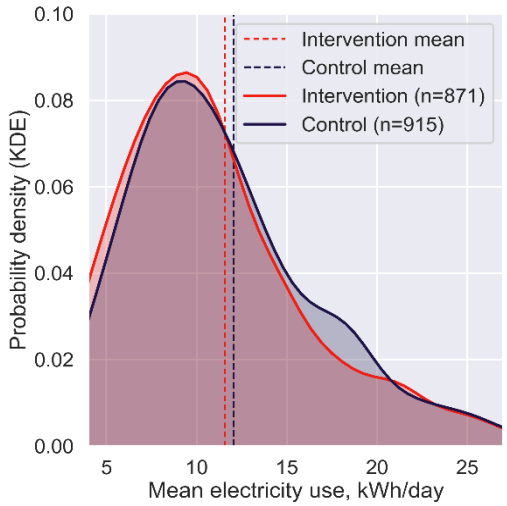
Energy consumption analysis

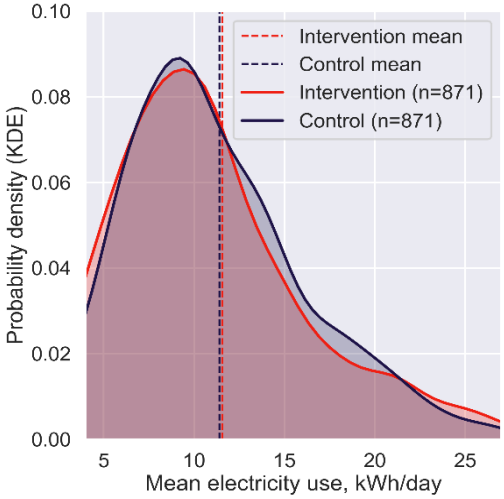
Figure 11: Distribution of mean daily gas, electricity and energy use during the evaluation period, comparing control and intervention groups, before matching and after matching (primary matching approach)

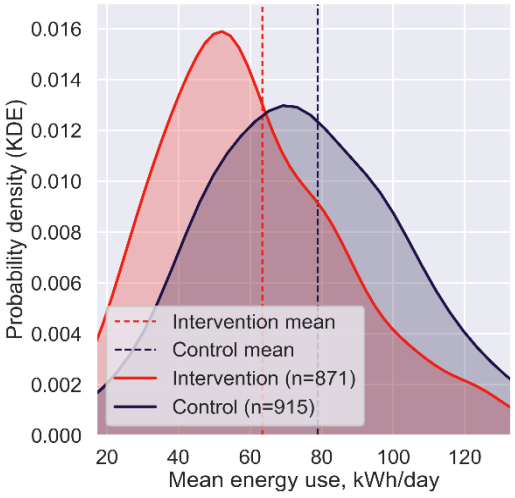
Before matching	After matching

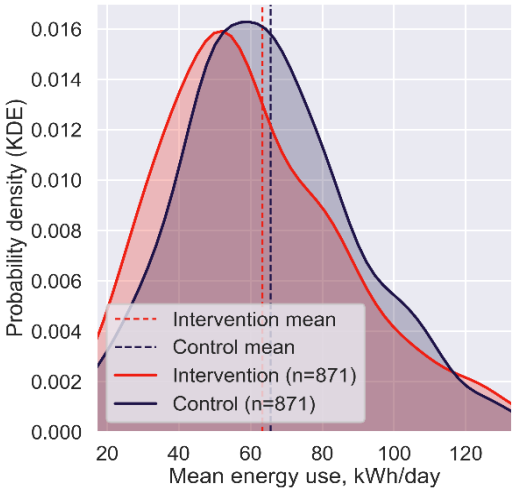












Regression model results

The table below provides the full breakdown of regression model results for the trial group variable, including the kWh / day savings estimates for each fuel and regression model, the percentage equivalent saving compared to the control group mean, and the estimates of the annualised equivalent.

Fuel	Model/ evaluation period	Matching method	Statistical significance level (P> ltl)	Coefficient	95% confidence interval Lower	Upper	Units	Control group mean daily kWh /day, trial period
Gas	Regression model results for trial evaluation period	Genetic (many to one)	0.000	-2.48	-3.59	-1.37	kWh / day	54.19
		Nearest (many to one)	0.000	-2.36	-3.48	-1.25	kWh / day	54.91
		Optimal (one to one)	0.000	-2.33	-3.45	-1.21	kWh / day	55.73
		Midpoint value of statistically significant results		-2.39	-3.51	-1.28	kWh / day	
		Genetic (many to one)	0.000	-4.6%	-6.6%	-2.5%	Percent	

Fuel	Model/ evaluation period	Matching method	Statistical significance level (P> ltl)	Coefficient	95% confidence interval Lower	Upper	Units	Control group mean daily kWh /day, trial period
	Results as percentage of control group mean	Nearest (many to one)	0.000	-4.3%	-6.3%	-2.3%	Percent	
		Optimal (one to one)	0.000	-4.2%	-6.2%	-2.2%	Percent	
		Midpoint value of statistically significant results		-4.4%	-6.4%	-2.3%	Percent	
	Full year ending 13 March 2022*	Genetic (many to one)	0.000	-1.98	-2.86	-1.09	kWh / day	
		Nearest (many to one)	0.000	-1.88	-2.77	-0.99	kWh / day	
		Optimal (one to one)	0.000	-1.86	-2.75	-0.96	kWh / day	
			Daily mean	-1.90	-2.79	-1.02	kWh / day	

Fuel	Model/ evaluation period	Matching method	Statistical significance level (P> ltl)	Coefficient	95% confidence interval Lower	Upper	Units	Control group mean daily kWh /day, trial period
		Midpoint value of statistically significant results	Annual total	-694.7	-1018.7	-370.7	kWh / year	
Electricity	Regression model results for trial evaluation period	Genetic (many to one)	0.511	-0.05	-0.22	0.11	kWh / day	11.4
		Nearest (many to one)	0.868	-0.02	-0.24	0.21	kWh / day	12.08
		Optimal (one to one)	0.635	-0.05	-0.27	0.17	kWh / day	11.96
		Midpoint value of statistically significant results		-	-	-	kWh / day	
	Results as percentage of control	Genetic (many to one)	0.511	-0.5%	-1.9%	0.9%	Percent	
		Nearest (many to one)	0.868	-0.2%	-2.0%	1.7%	Percent	

Fuel	Model/ evaluation period	Matching method	Statistical significance level (P> ltl)	Coefficient	95% confidence interval Lower	Upper	Units	Control group mean daily kWh /day, trial period
	group mean	Optimal (one to one)	0.635	-0.4%	-2.3%	1.4%	Percent	
		Midpoint value of statistically significant results		-	-	-	Percent	
	Full year ending 13 March 2022*	Genetic (many to one)	0.511	-0.04	-0.17	0.09	kWh / day	
		Nearest (many to one)	0.868	-0.02	-0.19	0.16	kWh / day	
		Optimal (one to one)	0.635	-0.04	-0.22	0.13	kWh / day	
		Midpoint value of statistically significant results	Daily mean	-	-	-	kWh / day	
			Annual total	-	-	-	kWh / year	

Fuel	Model/ evaluation period	Matching method	Statistical significance level (P> ltl)	Coefficient	95% confidence interval Lower	Upper	Units	Control group mean daily kWh /day, trial period
Energy (gas + electricity)	Regression model results for trial evaluation period	Genetic (many to one)	0.000	-2.55	-3.71	-1.40	kWh / day	65.58
		Nearest (many to one)	0.000	-2.34	-3.52	-1.17	kWh / day	66.98
		Optimal (one to one)	0.000	-2.36	-3.54	-1.18	kWh / day	67.68
		Midpoint value of statistically significant results		-2.42	-3.59	-1.25	kWh / day	
	Results as percentage of control group mean	Genetic (many to one)	0.000	-3.9%	-5.7%	-2.1%	Percent	
		Nearest (many to one)	0.000	-3.5%	-5.2%	-1.8%	Percent	
		Optimal (one to one)	0.000	-3.5%	-5.2%	-1.7%	Percent	

Fuel	Model/ evaluation period	Matching method	Statistical significance level (P> ltl)	Coefficient	95% confidence interval Lower	Upper	Units	Control group mean daily kWh /day, trial period
		Midpoint value of statistically significant results		-3.6%	-5.4%	-1.9%	Percent	
	Full year ending 13 March 2022*	Genetic (many to one)	0.000	-2.04	-2.96	-1.12	kWh / day	
		Nearest (many to one)	0.000	-1.86	-2.80	-0.93	kWh / day	
		Optimal (one to one)	0.000	-1.87	-2.81	-0.93	kWh / day	
		Midpoint value of statistically significant results	Daily mean	-1.93	-2.86	-0.99	kWh / day	
			Annual total	-702.8	-1042.6	-363.1	kWh / year	

* These figures are based on the evaluation period results, adjusting for the difference in daily mean hdds between the evaluation period and the full year ending 13 March 2022, presented in the table below.

Mean daily heating degree day values - full sample

Matching method	hdd_mean_gas	hdd_mean_elec	hdd-mean_energy	hdd_mean_to_03_22
Genetic (many to one)	6.76	6.74	6.75	5.39
Nearest (many to one)	6.75	6.74	6.75	5.37
Optimal (one to one)	6.85	6.83	6.84	5.45

Detailed regression model results

Tables 5 to 7 present the full OLD regression model results for the primary matched dataset (genetic, many -to-one) for gas, electricity and energy daily usage respectively.

Table 5: Gas results

Dep. Variable:	gas_mean_daily	R-squared:	0.806
Model:	OLS	Adj. R-squared:	0.806
Method:	Least Squares	F-statistic:	3617.
Date:	Fri, 01 Jul 2022	Prob (F-statistic):	0.00
Time:	22:15:36	Log-Likelihood:	-6770.3
No. Observations:	1742	AIC:	1.355e+04
Df Residuals:	1739	BIC:	1.356e+04
Df Model:	2		
Covariance Type:	nonrobust		

	coef	std err	t	P>|t|	[0.025	0.975]
Intercept	3.8375	0.715	5.367	0.000	2.435	5.240
C(Trial_group)[T.Treatment]	-2.4783	0.566	-4.381	0.000	-3.588	-1.369
AQ_daily_mean_kWh	1.1912	0.014	84.940	0.000	1.164	1.219

Omnibus:	505.944	Durbin-Watson:	1.916
Prob(Omnibus):	0.000	Jarque-Bera (JB):	4676.598
Skew:	1.088	Prob(JB):	0.00
Kurtosis:	10.726	Cond. No.	129.

Table 6: Electricity results

Dep. Variable:	electric_mean_daily	R-squared:	0.920
Model:	OLS	Adj. R-squared:	0.920
Method:	Least Squares	F-statistic:	9987.
Date:	Fri, 01 Jul 2022	Prob (F-statistic):	0.00
Time:	22:15:41	Log-Likelihood:	-3414.7
No. Observations:	1742	AIC:	6835.
Df Residuals:	1739	BIC:	6852.
Df Model:	2		
Covariance Type:	nonrobust		

	coef	std err	t	P>|t|	[0.025	0.975]
Intercept	-0.0382	0.100	-0.383	0.702	-0.234	0.157
C(Trial_group)[T.Treatment]	-0.0542	0.082	-0.658	0.511	-0.216	0.107
EAC_daily_mean_kWh	1.0621	0.008	141.315	0.000	1.047	1.077

Omnibus:	1694.725	Durbin-Watson:	1.668
Prob(Omnibus):	0.000	Jarque-Bera (JB):	218561.464
Skew:	4.217	Prob(JB):	0.00
Kurtosis:	57.222	Cond. No.	32.5

Table 7: Energy (gas plus electricity) results

Dep. Variable:	energy_mean_daily	R-squared:	0.830
Model:	OLS	Adj. R-squared:	0.830
Method:	Least Squares	F-statistic:	4250.
Date:	Fri, 01 Jul 2022	Prob (F-statistic):	0.00
Time:	22:15:44	Log-Likelihood:	-6840.2
No. Observations:	1742	AIC:	1.369e+04
Df Residuals:	1739	BIC:	1.370e+04
Df Model:	2		
Covariance Type:	nonrobust		

	coef	std err	t	P>|t|	[0.025	0.975]
Intercept	3.2801	0.794	4.130	0.000	1.722	4.838
C(Trial_group)[T.Treatment]	-2.5547	0.589	-4.339	0.000	-3.710	-1.400
EAC_AQ_daily_mean_kWh	1.1748	0.013	92.117	0.000	1.150	1.200

Omnibus:	515.447	Durbin-Watson:	1.933
Prob(Omnibus):	0.000	Jarque-Bera (JB):	4243.862
Skew:	1.149	Prob(JB):	0.00
Kurtosis:	10.293	Cond. No.	166.

This publication is available from: www.gov.uk/government/publications/smart-energy-savings-sens-competition-evaluation

If you need a version of this document in a more accessible format, please email alt.formats@beis.gov.uk. Please tell us what format you need. It will help us if you say what assistive technology you use.