EQUINOX

# Learning from trial three: Longer heat pump flexibility events

August 2025

nationalgrid

nationalgrid.co.uk

# Version control

| Issue | Date |
|---|---|
| d0.1 | 31/08/2028 |
| d0.2 | 30/09/2025 |
| FINAL | 15/10/2025 |

# Publication control

| Name | Role |
|---|---|
| Ewan Steen, Raeesa Hirji, Samantha Piller, (Guidehouse) | Author |
| Ryan Huxtable (National Grid)<br>Alyssa Fournier, Callum Coghlan, Alex Jakeman (Guidehouse)<br>Jessica Colleen, Tessa Hillen (Octopus Energy) | Reviewer |
| Ryan Huxtable (National Grid) | Approver |

# Contact details

**Email**
For further information, please contact: nged.innovation@nationalgrid.co.uk
**Postal**
Innovation Team
National Grid Electricity Distribution
Pegasus Business Park
Herald Way
Castle Donington
Derbyshire DE74 2TU

# Disclaimer

Neither National Grid, nor any person acting on its behalf, makes any warranty, express or implied, with respect to the use of any information, method or process disclosed in this document or that such use may not infringe the rights of any third party or assumes any liabilities with respect to the se of, or for damage resulting in any way from the use off, any information, apparatus, methos or process disclosed in the document.

# Table of contents

# 1. Glossary of terms

- **Constraint Managed Zones:** Areas of the electricity network where demand or generation is actively managed to avoid network constraints and maintain system stability
- **Customers**: term EQUINOX uses when referring to customers of energy suppliers generally
- **Demand response**: A strategy that adjusts electricity usage by consumers in response to supply conditions, such as price signals or grid needs
- **Demand turn up**: Demand response when more electricity is consumed relative to a baseline
- **Demand turndown**: Demand response when less electricity is consumed relative to a baseline
- **Difference-in-Differences**: A method that compares changes over time between a treatment and control group to estimate the effect of an intervention
- **Distribution Network Operator**: Company licensed to operate the electricity distribution network in a specific region of the UK
- **Dynamic**: an NGED flexibility product procured to support unscheduled maintenance, with a duration from between 30 mins to several hours, on a single day.
- **Equitable novel flexibility exchange (EQUINOX)**: the name of the project
- **Events**: periods of the day when EQUINOX trial participants were asked to provide demand response
- **Heat pump homes**: households with heat pumps
- **Low Carbon Technology**: Innovative technologies that cut or eliminate greenhouse gas emissions
- **National Grid Electricity Distribution (NGED)**: the UK's largest regional Distribution Network Operator
- **Network Innovation Competition**: a programme that funded energy network innovation projects
- **Participants**: term EQUINOX uses when referring to customers enrolled in the EQUINOX trials
- **Root Mean Square**: A supporting method for aggregating demand response standard error impacts by squaring each standard error with a grouping of events, averaging them, and taking the square root
- **Scottish Power Energy Networks**: A UK Distribution Network Operator
- **Secure**: an NGED flexibility product procured for peak load management, with a duration from between 30 mins to several hours, on a single day
- **Scheduled Utilisation:** an NGED flexibility product procured for constraint management, with a variable duration and variable hours daily Monday to Friday across consecutive weeks
- **Sustain**: an NGED flexibility product procured for constraint management, with a duration of 4-hours daily Monday to Friday across consecutive weeks
- **The Office of Gas and Electricity Markets**: the UK's energy regulator responsible for protecting consumers and ensuring a secure, sustainable, and affordable energy system
- **United Kingdom:** The United Kingdom of Great Britain and Northern Ireland

# 2. Context

## 2.1. Introduction to EQUINOX

Equitable Novel Flexibility Exchange (EQUINOX) is a Network Innovation Competition (NIC) project funded by the Office of Gas and Electricity Markets (Ofgem). It is led by National Grid Electricity Distribution (NGED), the UK's largest regional Distribution Network Operator (DNO) and supported by multiple project partners[1]. It is developing, trialling, and where proven, implementing suitable arrangements at scale that can maximise participation of domestic heat pumps in DNO procured flexibility[2] while meeting the needs of all customers, including those with potential vulnerabilities.

Heat pumps are expected to become a mainstream choice to decarbonise home heating in the United Kingdom (UK). The National Energy System Operator (NESO) forecasts an increase in annual heat pump installations from 95,000 in 2024 to more than 1.1 million by 2035[3]. The electrification of heat therefore stands to substantially increase electricity demand. If this new demand coincides with existing demand peaks, demand may more frequently exceed the capacity of DNO infrastructure e.g. substations and cables. Increased demand could exacerbate existing constraints or create new ones. Constraints are ultimately resolved through network reinforcement but can also be managed in the short- to medium- term through procurement of flexibility. In many instances, it is more cost-effective for DNOs to defer reinforcement by procuring flexibility, rather than reinforcing immediately.

EQUINOX is iteratively testing novel commercial arrangements for heat pump flexibility across three trial periods between 2022-25 (Figure 1). The trials measure demand response from heat pumps to better understand the customer experience of heat pump flexibility.

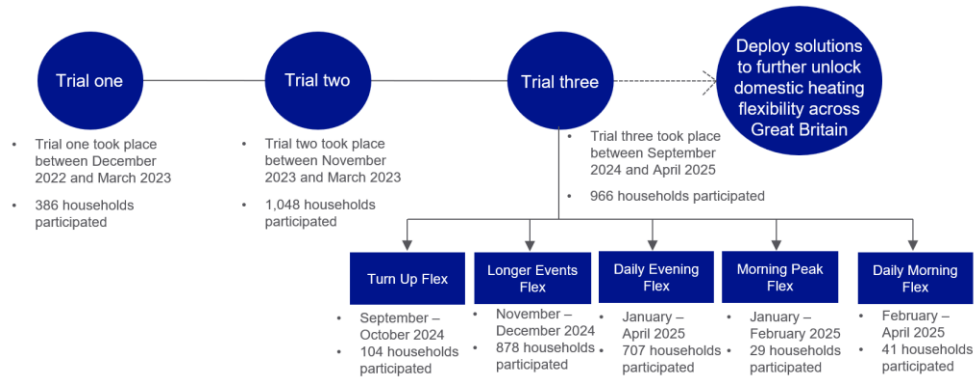


**Figure 1: EQUINOX project overview**

[1] A full list of project partners can be found in Appendix A.

[2] As defined by NGED, flexibility is reducing loads on the network by using customers' ability to change their usage patterns by either reducing consumption, changing their electricity habits, or (at a larger commercial scale) switching on generators. Flex In Five An Overview of Flexibility

[3] NESO records show over 95,000 annual heat pump installations in 2024, rising to 1.13 million to 1.46 million by 2035 across its three net zero compatible Future Energy Scenarios (FES): Hydrogen Evolution, Electric Engagement, and Holistic Transition. Future Energy Scenarios 2025 Data Workbook V001

Trial one[4] ran across winter 2022-23 with almost 400 participating households ("participants") and was a successful proof of concept for heat pump flexibility. Trial two[5] (winter 2023-24) iterated on the results of trial one, increasing the participants pool to over 1,000 customers. This increased scale permitted closer analysis of demand response volumes and used commercial arrangements that more closely mimicked existing DNO flexibility products typically used for unlocking flexibility from electric vehicle charging or other assets. Trial three, held between September 2024 and April 2025, built on the results of trial one and two to investigate additional opportunities for heat pump flexibility to support networks. Specifically, trial three aimed to:

- Test potential for domestic heat pump flexibility during the morning peak as well as the evening peak.
- Uncover the potential for domestic heat pump homes to offer daily flexibility.
- Explore longer heat pump flexibility events, beyond the two-hours tested in trials one and two.
- Test the customer proposition for stacking flexibility from heat pump homes through EQUINOX with an active flexibility product run by NESO. Stacking offers the prospect for customers to unlock greater rewards from flexibility participation and the opportunity for DNOs and NESO to procure flexibility services from the same assets[6].
- Improve understanding of how customer vulnerability should be defined in the context of heating flexibility.

To achieve these aims, we split trial three into five mini-trials each focusing on different aspects of demand response and customer experience.

## 2.2 Trial three mini-trials

Each mini-trial focused on different aspects of heat pump flexibility:

- **Heat pump turn up flexibility ('Turn Up Flex):** demand turn up[7] events on non-consecutive days, between 11am - 1pm.
- **Longer heat pump flexibility events ('Longer Events Flex'):** demand turndown[8] events of differing lengths, on non-consecutive days, between 4-8pm.
- **Daily evening heat pump flexibility ('Daily Evening Flex'):** 'everyday' demand turndown events on weekdays, across consecutive weeks and occurring between 4-8pm.
- **Morning heat pump flexibility events ('Morning Peak Flex'):** demand turndown events on non-consecutive days, between 8-10am.
- **Daily morning heat pump flexibility ('Daily Morning Flex'):** daily demand turndown events on weekdays, across consecutive weeks and between both 8-10am and 5-7pm of the same day.

This report focuses on Longer Events Flex.

[4] Trial one report: Initial Insights on the Effectiveness of Commercial Methods
[5] Trial two report: Learning from trialling novel commercial methods Project deliverable 4
[6] Stacking is when assets provide flexibility to more than one flexibility service. See the Energy Network Association's report: Revenue Stacking Explainer and FAQ's.
[7] Demand response when more electricity is consumed relative to a baseline.
[8] Demand response when less electricity is consumed relative to a baseline.

# 3. Longer Events Flex trial design

## 3.1 Introduction and aims

EQUINOX trial one (winter 2022-3) and trial two (winter 2023-4) demonstrated heat pump flexibility could help reduce electricity demand during the weekday evening during peak times for networks between 4-8pm. However, they primarily tested 2-hour turndown events two to three times per week on non-consecutive weekdays in a format aligned to NGED's Secure and Dynamic flexibility products[9].

A network impact analysis was carried out based on the trial two findings —scaling up the trial two results to estimate the value of heat pump flexibility in resolving a real constraint. It indicated that heat pump homes could offer the most potential to help resolve constraints if they provided flexibility in a format more aligned with NGED's Sustain product. Sustain is a constraint management flexibility product that procures daily flexibility across consecutive weekdays during the entire 4–8pm evening peak[10]. By engaging heat pump homes to participate in line with trial two's format the magnitude of flexibility at any one time would be diluted due to the need to stagger different groups across multiple days and evenings. Instead, if heat pump homes could participate daily for 2-hours they would be better placed to support networks in resolving CMZs – they would be especially impactful if daily participation could be as long as 4-hours for some groups, allowing single groups to cover the entire duration of a CMZ's flexibility need.

Feedback from participants in trial two suggested that events longer than 2-hours would suit many households, with some even open to events lasting up to 4-hours[11]. Following the Week of Consecutive Events mini-trial that we held immediately after trial two, where we asked participants to turn down daily from 6-8pm for one week in April, most participants indicated they would be willing to participate in daily events Monday to Friday[12].

At the time of designing trial three it was therefore a key ambition to test the ability of heat pump homes to participate in daily turndown flexibility in a longer, dedicated trial – including daily 4-hour events where possible. The challenge to implementing this arrangement directly was the lack of quantitative evidence on how participants would respond to the 4-hour daily flexibility request, which represented a substantial step up from what we had tested previously. We hypothesised that such a long duration would only be suitable for some customer groups and

[9] Secure, procured for peak load management, and Dynamic, procured for unscheduled maintenance, are both scheduled from between 30 mins to several hours, on a single day. When trial three was designed these products were still actively procured by NGED. They have now been largely replaced by Operational Utilisation, which procures short-term flexibility lasting for a single day. See NGED's report on the new products formalisation: Distribution Flexibility Services Procurement Statement.

[10] When trial three was designed, Sustain was a commercial flexibility product procured by NGED. Flexibility products have since been updated and NGED's daily flexibility product is now 'Scheduled Utilisation'. The mini-trial designs in trial three were considered in relation to previous products but focused on structural aspects of flexibility that remain relevant to network needs.

[11] Following trial two, 85% of participants (n=541) reported that they would be willing to participate in events up to 3-hours. We also recorded anecdotal feedback from interviews and focus groups of a willingness to participate for up to 4-hours.

[12] For the Week of Consecutive Events mini-trial, 81% of participants (n=439) reported that they would be willing to participate in flexibility offerings asking them to turn down across five consecutive days. See full report: "Learning from trialling novel commercial methods: Week of consecutive events".

therefore sought to first test 4-hour events less frequently in the Longer Events Flex mini-trial. We aimed to establish which customer groups might be comfortable adopting daily 4-hour heat pump flexibility behaviour in the subsequent Daily Evening Flex mini-trial.

In Longer Events Flex we therefore aimed to assess:

- Whether heat pump homes can deliver demand response for 2-hour, 3-hour and 4-hour event lengths during the network evening peak.
- Whether there is consistency in the magnitude of this demand response for different event lengths, and across settlement periods within an event.
- Which event lengths, if any, are preferred by different customer groups based on shared characteristics, particularly those indicating potential vulnerability.

## 3.2. Trial structure

Longer Events Flex was carried out over five weeks during November to December 2024. Eligible customers from Octopus Energy were invited to take part in two 2-hour events, two 3-hour events, and four 4-hour events. We asked these participants, who formed the active trial group (the "treatment group"), to reduce electricity consumption ('turndown') associated with their heat pumps and hot water heating, but not from any other sources. We held all events between 4–8pm on weekdays, which is the time electricity demand typically peaks across NGED's license areas.

We recruited a second group of customers for control purposes (the "control group"). These customers met the same trial eligibility requirements as the treatment customers (detailed below). They took part in a sign-up survey and end of winter survey, and were rewarded for that participation, but were not otherwise contacted during the mini-trial or asked to engage in any behavioural change.

We included two 2-hour events so that the participants would be able to compare across the different event lengths – although we had tested 2-hours events in trials one and two, this had been with different customer cohorts. Our main interest was in obtaining household feedback on the 4-hour duration, though we also tested 3-hour durations given 3-hours of demand response could be sufficient for some CMZs. We had also seen that 3-hour flexibility during trial two, when we observed that participants with tariffs with multiple daytime unit rates[13] were turning down daily regardless of whether an EQUINOX event was called, likely due to their tariff's price signal. We held events on non-consecutive days to avoid any potential issues of demand turndown on an event day subsequently impacting demand turndown activity on the following day.

Since this was the first time EQUINOX participants were asked to reduce usage for more than 2-hours, we designed and ran Longer Events Flex with additional measures to maintain customer comfort as a top priority. Participation in each event remained entirely voluntary, and customers were encouraged to engage only for the duration that suited them. Our messaging made clear that there was no penalty or loss of incentive for partial participation. Instead, incentives were tied to completing the post-event survey, allowing us to gather thoughtful, detailed feedback regardless of how long customers participated in each event.

[13] Tariffs with different prices during the daytime hours of 0800-2200.

## 3.3. Participant details

### 3.3.1 Trial eligibility

To be eligible to participate in Longer Events Flex customers were required to:

- Be a current electricity supply customer of Octopus Energy
- Have a working electricity smart meter that has been successfully sending meter readings on a half-hourly basis for at least 80% of the time in the 15 days prior to being invited to take part in a trial.
- Consent to, and not remove consent for, allowing half-hourly meter reads.
- Opt-in to be part of the EQUINOX trials by completing sign-up surveys as requested and accept terms and conditions of the trial.
- Have and use a heat pump as the primary method of heating their home (and this heat pump must not be shared with any other residence(s)).
- Be resident within the NGED or Scottish Power Energy Networks (SPEN) licenses areas.
- Be using an electricity import tariff ("tariff") that does not have different prices for electricity during the 'day' (between 0800 and 2200) (e.g. a tariff with an off-peak price between 1300 - 1600 was not permitted).[14]

There were no other eligibility requirements related to customers' electricity import tariff, export tariff or to the presence of additional low carbon technologies (LCTs).

### 3.3.2 Recruitment approach

Customer recruitment was a particularly successful aspect of EQUINOX. 878 participants were recruited to as part of the Longer Events Flex mini-trial, 525 in the treatment group and 353 in the control group. Centring customer experience as a key consideration during the trial design phase made the trials an appealing prospect during recruitment.

Customer recruitment was led by Octopus Energy and was conducted via emails which included messaging about the trial objectives, a description of what would be requested of customers participating, and potential financial benefits of taking part. Customers received an invitation to either the treatment group or to the control group, and were required to agree to trial Terms and Conditions and to complete a recruitment survey to become participants in the trial.

### 3.3.3 Demographics

We used self-reported participant data[15] to understand how representative the participants are of the wider UK population, which are important criteria for understanding the wider applicability of the mini-trial results:

[14] Customers were eligible if on a tariff with a fixed price across a 24-hour period, or if on a tariff with a fixed daytime price but reduced or variable prices overnight. We included this eligibility criteria based on our findings with trial two, when we saw participants with multiple day time rate tariffs turn down regardless of whether there was an EQUINOX event. For Dynamic Longer we wanted participants to not have a pre-existing incentive to turn down, so any turndown they did achieve could be attributed to the mini-trial.

[15] Customer demographic data was self-reported through an optional start of trial survey administered by Guidehouse. 566 participants (64% of enrolled treatment and control customers) completed it.

- Homeowners were overrepresented among trial participants compared to the UK average. 88% of participants lived in a property they owned and 9% paid social rent, compared to the UK average of 64% of individuals owning their home and 17% payment social rent[16]. We saw similar homeownership in trial two.
- High income households tended to be overrepresented: 56% of participants had an income that exceeded £50,000 per year; 6% of participants annual household income was between £20,000 - £24,999. This differs from the UK average with 19% of individuals in the UK having an income that exceeds £50,000 per year and 27% of individuals in the UK having an income between £20,000 - £30,000. We saw similar income level distribution in trial two.
- Participants' homes tended to be more energy efficient, with 72% living in a home with an Energy Performance Certificate (EPC) rating of A, B, or C. This is higher than the average EPC rating for a home in the UK at an EPC rating of D[17]. It may reflect both the general recommendations that a home should have higher insulation standards to get the best efficiency from using a heat pump, and previous requirements of government grant schemes that specified that homes must meet certain insulation requirements in order to access funding for a heat pump installation. Similarly, the ECO4 (Energy Company Obligation) Scottish Government Obligation incentivised potentially vulnerable participants whose houses would be suitable from an energy efficiency perspective to install heat pumps.

These differences between the mini-trial participants and the general population are in line with our experience in previous EQUINOX trials[18]. Moreover, recruiting a customer pool representative of the UK population was difficult as heat pumps are not yet widespread across all demographics. This is an important consideration because over time, heat pumps will become the main heating technology. As heat pumps continue to be installed in more homes, the population of heat pump users changes and will begin to more closely match the UK population. If we can understand how well our participants represent the general UK population, then we have some measure of assurance that our results will or will not continue to be relevant as DNOs undertake future planning activities.

Although the Longer Events Flex mini-trial is an important early exploration of the potential for heat pump homes to engage in demand response for longer periods of time, we stress that heating is highly individualised and our customer experience results reflect this group of customers. Any wider application of those findings should be undertaken cautiously. Customer safety should be put first (for example through maintaining the voluntary nature of events) and should expect that new or different trends may emerge as greater numbers of customers participate in heat flexibility.

For trial three we refined our approach to categorise, identify, and account for potential vulnerabilities. This involved considering a higher number of individual and combined potential vulnerability factors, identified through a bespoke

[16] Office for National Statistics, Census 2021.
[17] Office for National Statistics, Census 2021.
[18] This is discussed in further detail "Project Deliverable 5: Learning from Engaging Customers"

Equitable Participation Framework (EPF)[19]. We identified eight self-reported factors[20] that would contribute to whether a participating household was considered potentially vulnerable or non-vulnerable.

We classified participants as potentially vulnerable only if they met one of the following three circumstances:

- The household was defined as having a low income[21] according to their household composition. Households were also removed from the vulnerability segmentation if they were classified as low income, but self-reported that they were able to regularly afford their energy bills.
- Someone within the household self-reported having a health condition exacerbated by the cold; or
- The household met at least two of the self-reported EPF vulnerability factors such as having individuals in the home of a certain age, home being poorly insulated or single occupancy.

Figure 2**Figure 2** shows that amongst the 69% (576 participants) of participants who completed the start of trial survey, 46% (266 participants) could be classed as potentially vulnerable. This is consistent across both the treatment and control groups, showing the degree of similarity across both groups which is ideal for a treatment-control trial. Additionally, Longer Events Flex's representation of potentially vulnerable participants itself provides important context for interpreting the mini-trial results and assessing the broader applicability of longer events.

[19] See "Project Deliverable 5: Learning from Engaging Customers" for detailed explanation of the customer engagement approach, including how vulnerability factors are considered.
[20] The eight vulnerability factors included: Having a health condition exacerbated by the cold, having a disability or long-term health condition, having children under 5 years old in the home, having individuals over the age of 75 in the home, home being poorly insulated, living alone, being a lone parent, and meeting the low-income threshold.
[21] Households were classified as low income based on their household composition and annual household income, in alignment with the UK minimum income standard.

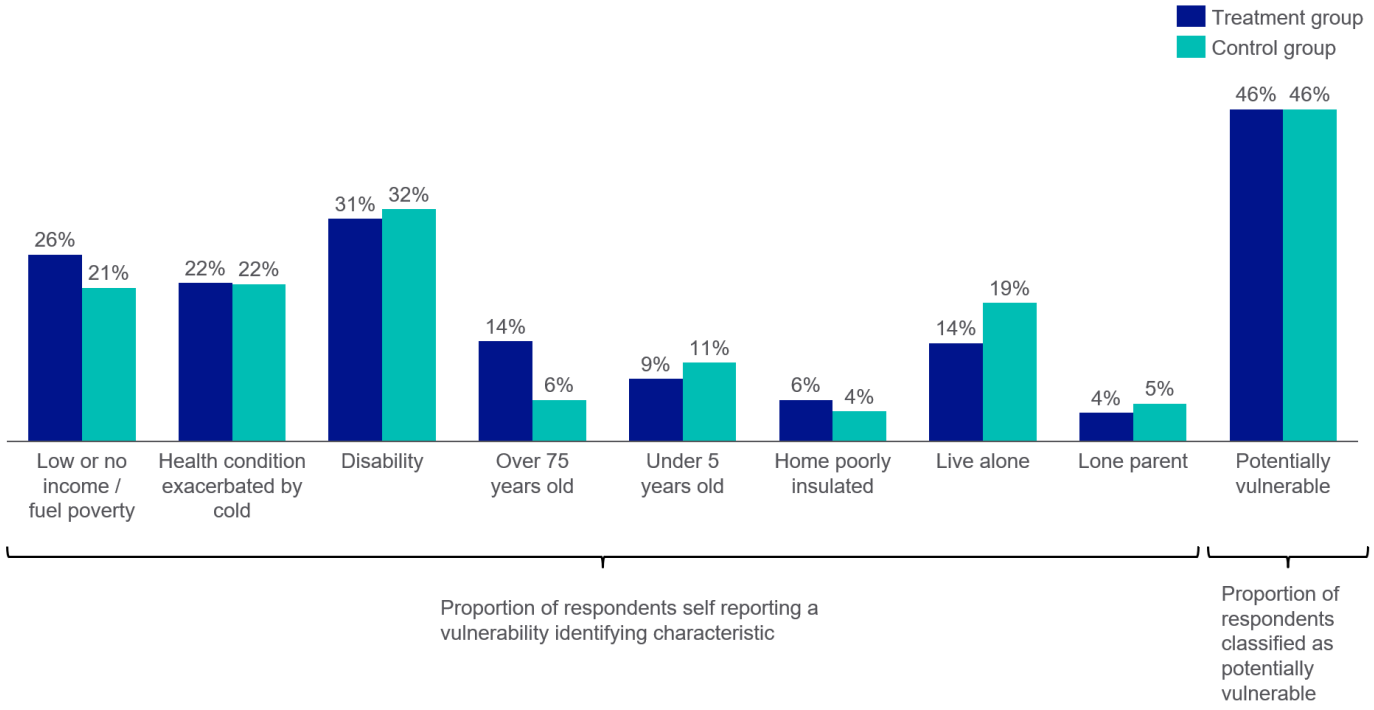


**Figure 2: Proportion of potentially vulnerable participants in Longer Events Flex in the treatment (n=383) and control group (n=193), based on start of trial survey**

## 3.4. Commercial arrangements for Longer Events Flex

To assess the potential of heat pump homes to participate in longer flexibility events in this mini-trial, commercial arrangements and event design emulated expected network flexibility needs where possible and otherwise prioritised simplicity and customer convenience. Design of the commercial arrangements and events are summarised in Table 1 (below) for treatment group and control group participants.

Events were held at three different times at a rate of zero to two events per week on weekdays. We gave participants day ahead notice of events, which was based on positive feedback on this notice period during both trial two and Turn Up Flex. Octopus Energy was notified of events on the preceding Wednesday to simplify trial operations.

Participants were not required to opt-in to individual events and received the £2 per event incentive if they completed the post-event survey. Treatment group participants were expected to earn over £25 if they completed the recruitment survey, start of trial survey, end of trial survey, and all eight post-event surveys.

Participants in the control group provided a comparison for both demand response volumes and for 'typical' energy consumption behaviours without the influence of EQUINOX events. To reflect that value, we incentivised control group participation via an initial incentive of up to £10 for surveys upon sign up and a further £10 incentive in Spring 2025 for customers who remained a part of the control group during Daily Evening Flex.

Table 1: Longer Events Flex commercial arrangements

| Item | Treatment participant | Control participant |
|---|---|---|
| Payment structure | £5 for completion of recruitment survey[22]<br>£5 for completion of start of trial survey<br>£5 for completion of end of trial survey<br>£2 per event upon completion of each post-event survey | £5 for completion of recruitment survey<br>£5 for completion of start of trial survey<br>Further £10 incentive available in Spring 2025 for customers who remained a part of the control group by the end of Daily Evening Flex |
| Control type | Manual and remote customer control | |
| Notice period | Day ahead | |
| Eligible supplier tariffs | No tariff with multiple unit rates between 8am and 10pm – which might incentivise load shifting during the daytime. Any other tariff accepted | |
| Event duration | 2 hours, 3 hours, 4 hours | |
| Event timing | 2 hours between 5-7pm; 3 hours between 5-8pm; 4 hours between 4-8pm | |
| Event frequency | 0[23] to 2 events per week, on non-consecutive days | |
| Supplier notice | Informed by NGED on Wednesday of the preceding week | |

## 3.5. Analysis approach

[22] The recruitment survey was only sent to treatment and control participants that had not previously participated as treatment or control in the Turn Up Flex mini-trial. The start of trial survey was only sent to control participants that had not previously participated as control in Turn Up Flex.
[23] Events were suspended for one week during the mini-trial, due to storms in the NGED area that created a risk of power outages for some participants.

Following industry best practice for evaluating trial programmes, suppliers leveraged a Difference-in-Difference (DiD) approach for calculating demand response. DiD approaches provide an unbiased estimate of programme demand response when energy consumption from the entire treatment group is compared to the entire control group for each event. We used a simple average to develop average point estimates of demand response per event length (2-hour, 3-hour, 4-hour), and we used the Root Mean Square (RMS) method to calculate aggregate errors[24].

To confirm the control group's suitability, we compared it with the treatment group across key metrics. Both groups were similar in energy use, heat pump size, and ownership of technologies like home batteries. These similarities supported the use of the DiD approach.

As well as demand response, customer experience was a key focus. Throughout all the mini-trials including, Longer Events Flex, we prioritised equitable access and participation for all customers, including those experiencing vulnerability or fuel poverty. We employed a mix-methods approach, combining quantitative surveys with qualitative interviews and focus groups to explore customer willingness to engage in heat pump flexibility and understand the experiences and barriers for potentially vulnerable customers.

We evaluated the participant experience in Longer Events Flex via several surveys which captured self-reported behaviour change during events and satisfaction with various event design elements. These surveys were:

- Eight short post-event surveys sent out at the end of each event.
- An end of trial survey.

[24] The full DiD and RMS approach is provided in Appendix B.

# 4. Demand response results

Section 4 summarises the findings for our aim of determining whether participants could achieve statistically significant demand response during 2-hour, 3-hour and 4-hour events, and whether this demand response would remain consistent across event length and settlement periods, which are crucial criteria for determining whether heat pump homes can support networks when participating in longer events. Section 4.1 Demand response by event length compares response magnitudes across event durations, Section 4.2 Demand response by event settlement period assesses consistency across settlement periods, and Section 4.3 Demand Profiles presents average daily demand profiles for an alternative view of the results.

## 4.1 Demand response by event length

Participants achieved statistically significant demand response for the 2-hour and 4-hour event lengths, though there was no statistically significant difference in demand response magnitude between them. There was no statistically significant demand response for the 3-hour event lengths.

Figure 3 shows that each treatment participant provided an average:

- 0.42 kW per 2-hour event (0.16 kW to 0.68 kW 90% confidence interval)
- 0.22 kW per 3-hour event (-0.10 kW to 0.44 kW 90% confidence interval)
- 0.27 kW per 4-hour event (0.10 kW to 0.44 kW 90% confidence interval)

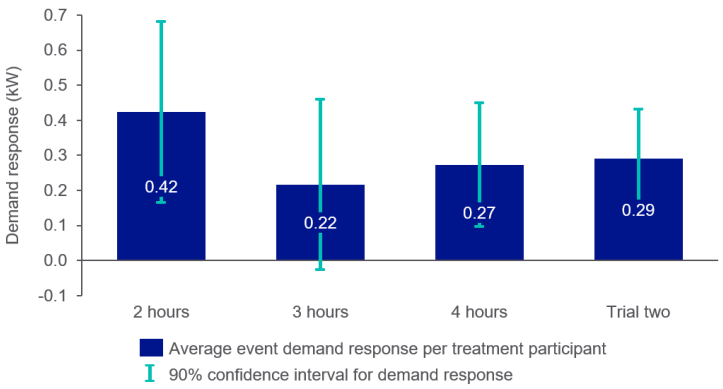


**Figure 3: Average demand response per treatment participant by event length**

The overlapping confidence intervals of the 2- and 4-hour event lengths means they are not statistically different from one another. Since the lower confidence interval for the 3-hour events crosses the 0 kW line, we cannot conclude that 3-hour events resulted in demand response greater than 0 kW. We cannot explain this outcome, though suggest it could be due to data limitations since we held just two 3-hour events and there was more variation in external temperature between them than we saw for other event lengths. We consider it likely that demand response was achieved during the 3-hour events, since it was conclusively achieved for the longer 4-hour duration events. Additionally, there is indication participants were turning down during the 3-hour events based on observational analysis of participants' demand profiles (Section 4.3), participant survey responses (Section 5.25), and that in trial two we had already observed that participants with multiple daytime rate tariffs could provide demand response for 3-hours.

Figure 3 does show wide confidence intervals for each event length, which are in line with expectations for a small trial over a low number of events. Figure 3 also shows that demand response volumes in Longer Events Flex DiD results are comparable to those from EQUINOX trial two, which was carried out at a larger scale. It provides

reassurance that Longer Events Flex's results sit within the bounds of what was detected during trial two, despite varying factors and event numbers between the two trials.

## 4.2 Demand response by event settlement period

We also used a DiD analysis to investigate whether the magnitude of demand response varied by settlement period, to see if participants responded more strongly during certain periods of an event—especially during longer 3- and 4-hour events. Our focus was particularly on whether the demand response weakened in the later settlement periods, which could reduce the overall benefit to the network at those times. However, we did not detect any differences in effect across the different settlement periods.

Figure 4, Figure 5 and Figure 6, show the average demand response detected per settlement period for the 2-hour, 3-hour and 4-hour event lengths, respectively. There were no statistically significant differences in demand response between settlement periods for any event length, as shown by the overlapping 90% confidence intervals for all point estimates in all charts. However, we note the comparatively lower average point estimates in the last half hour of the 2-hour events (Figure 4) and the last hour of the 4-hour events (Figure 6). It is uncertain whether these differences would persist in a larger trial with more events, or whether they would diminish as the number of events increases.

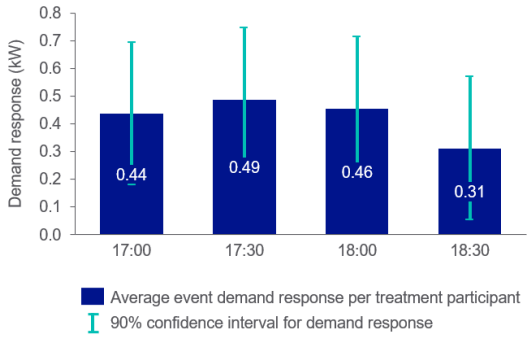


**Figure 4: Average demand reponse per treatment participant for 2-hour events across each settlement period**

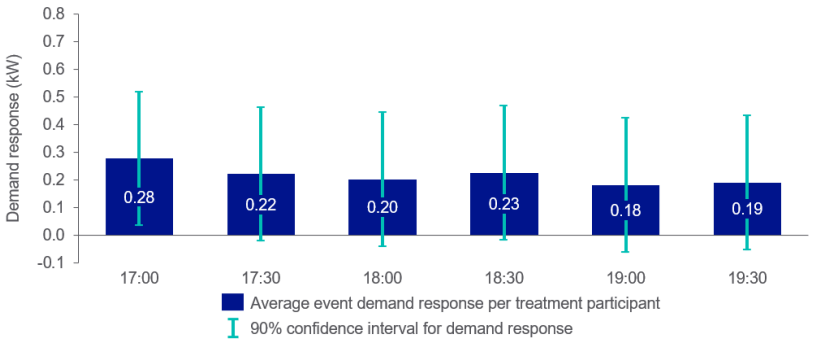


**Figure 5: Average demand reponse per treatment participant for 3-hour events across each settlement period**

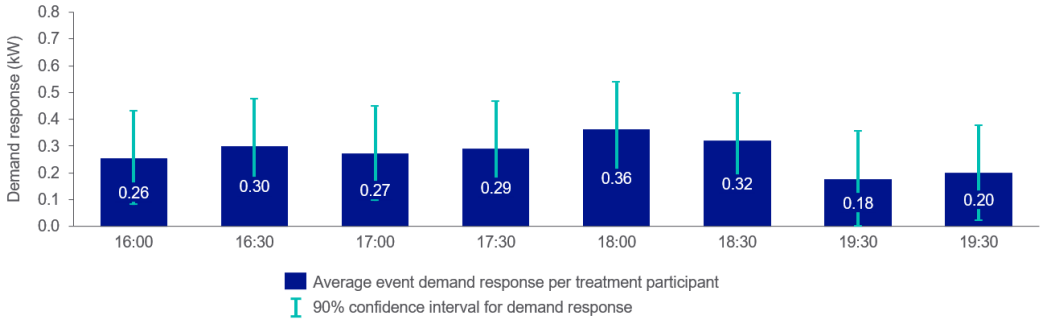


**Figure 6: Average demand response per treatment participant for 4-hour events across each settlement period**

## 4.3 Demand Profiles

To provide an alternative view of these results, Figure 7Figure 7, Figure 8Figure 8 andFigure 9 Figure 9 show the average half hourly demand profile for each of the treatment and control groups (solid lines), for 2-hour, 3-hour and 4-hour event days, respectively[25]. The dotted lines represent a historical baseline for each group, constructed from the p376 method[26]. The 2-hour event window is highlighted in blue. This comparison shows dips in consumption among treatment participants for all three event lengths.

These patterns align with the DiD findings for the 2- and 4-hour events, supporting the conclusion that demand response was achieved. While the DiD results for the 3-hour events were inconclusive, the observed dipFigure 8 suggests a possible response. However, this visual trend should be interpreted with caution given the inconclusive result for 3-hour events per Figure 3 and Figure 5.

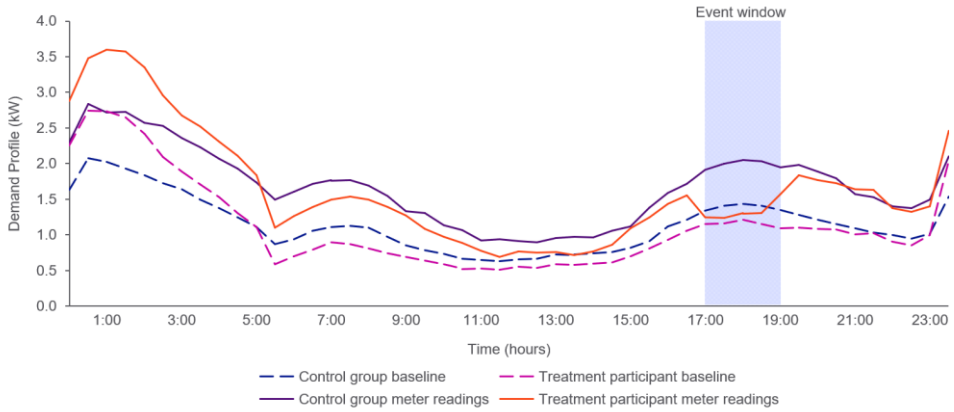


**Figure 7: Average baselines and meter readings for treatment and control Longer Events Flex participants across 2-hour event days (5-7pm)**

[25] Note the high overnight peak on all three charts. This is likely the result of many treatment and control group participants using EV-chargers. We are not concerned by this trend since it is rather the time of the day when aggregate demand of all assets on the network peaks, at 4-8pm, which is the key time to reduce demand from a network perspective. Nevertheless, we recognise that in future flexibility strategies may need to evolve with the deployment of greater numbers of EV-chargers.

[26] The p376 baseline method was used in EQUINOX trial two and constructs a personalised historic baseline for each participant based on the average of their consumption on the last 10 similar non-event days (for weekdays) or the last 4 similar non-event days (for weekends and holidays): See Elexon report for full details of the p376 method.

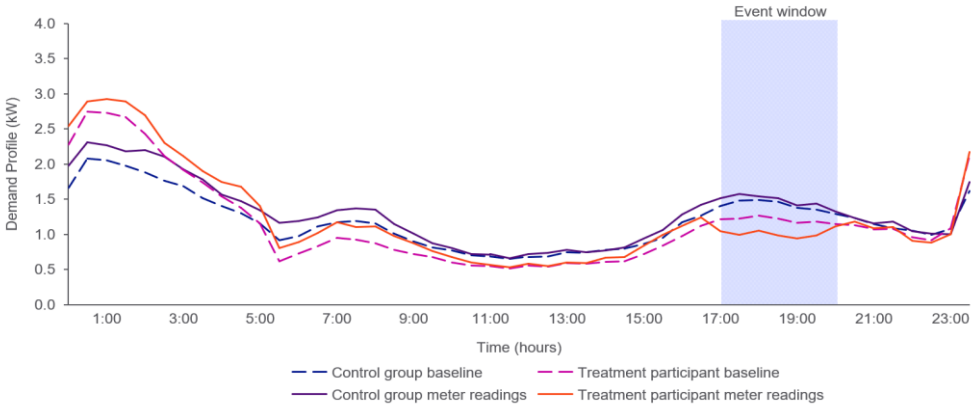


**Figure 8: Average baseline and meter readings for treatment and control Longer Events Flex participants across 3-hour event days (5-8pm)**

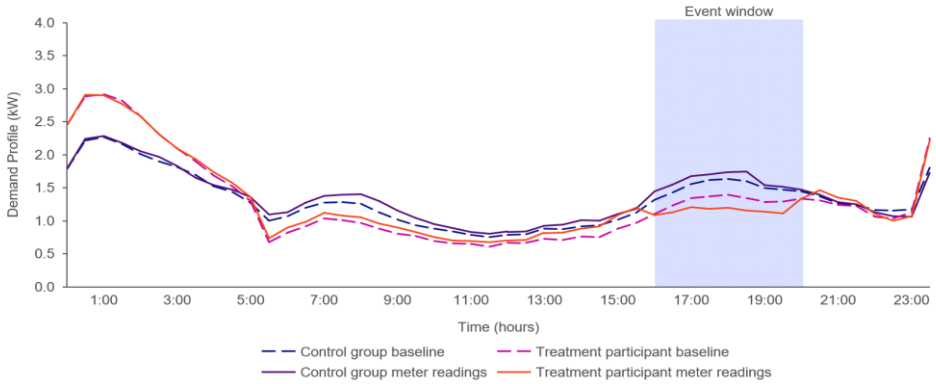


**Figure 9: Average baseline and meter readings for treatment and control Longer Events Flex participants across 4-hour event days (4-8pm)**

Overall, based on these results we are confident that Longer Events Flex participants achieved demand response during the 2-hour and 4-hour events, and have no reason to conclude that the magnitude of response differed across the two event lengths or across individual settlement periods. We also suspect that demand response was achieved for the 3-hour event lengths, though cannot prove this based on the findings.

# 5. Customer experience

Section 5 addresses our aim of determining which event lengths may suit different customer groups, particularly those experiencing vulnerability or fuel poverty. These findings are primarily based on survey analysis, including a start-of-trial survey, post-event surveys issued immediately after events, and an end-of-trial survey[27]. Section 5.1 presents findings on satisfaction, Section 5.2 on self-reported participation rates, Section 5.3 on views of event length, Section 5.4 on comfort, and Section 5.5 on notification periods and the value of reminders sent when events ended. Throughout, we carefully consider the experiences of customers with differing shared characteristics, particularly those that are potentially vulnerable.

## 5.1 Satisfaction

Centring customer experience as a key consideration during the trial design phase contributed to high levels of customer satisfaction. 78% of survey respondents indicated that they were satisfied with Longer Events Flex. The satisfaction rate was similar between potentially vulnerable participants and non-vulnerable participants. This was encouragingly consistent with the overall 79% satisfaction rate seen in EQUINOX trial two, where events were 2-hours. We are encouraged by customer appetite to participate in a range of novel commercial offerings while also being able to maintain high customer satisfaction. This highlights the success of our customer-centric trial design approach. These insights will directly inform our transition planning efforts.

## 5.2 Participation rate

### 5.2.1 Overall participation rate

Figure 10 shows that participation was consistently high across all three event lengths, with 76-78% self-reporting participating for the whole or part of the event time. Though, the proportion of participants reporting they turned down for the full event was slightly lower for the 4-hour events (60%) than the 2-hour (67%) and 3-hour (68%) events. Additionally, 8-11% of participants reported they did not participate. We assumed that non-respondents generally did not take part, likely due to disengagement. Overall, the consistently high participation rates reflect strong customer engagement and a clear willingness from participants to try longer event durations.

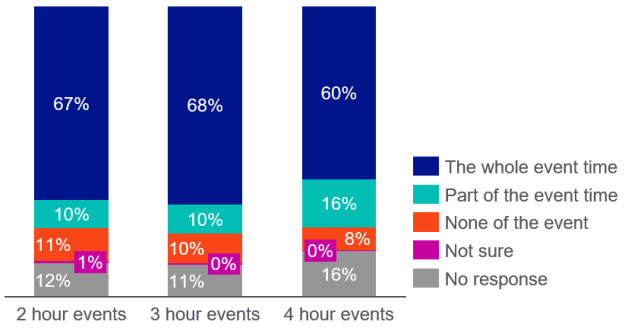


**Figure 10: Self-reported event participation rate in post-event surveys (n=515)**

Within participation, we found that more customers reported participating for only part of the event in 4-hour events. This shows that customer engagement was consistent at 2- and 3-hour event lengths but decreased slightly in 4-hour events. Nonetheless, 60% of respondents on average reported taking part in all four hours. It is important to note that there was no penalty for not taking part in events or for participating in only part of an event.

[27] 74% of participants completed the start of trial survey, 86% of participants completed the post-event surveys on average, and 77% completed the end of trial survey.

Participants were encouraged to only participate in events for as long as it made sense to them based on individual preferences and circumstances including comfort.

For the 10-16% of customers who only participated for part of the event, we asked why this was the case and they reported that comfort was the leading driver. As shown in Figure 11, 42%, 47% and 56% (21, 24, and 48 participants) of those who reported participating in part of the event reported either their home (or someone in it) was getting too cool or that they knew participating for the full event would make their home feel too cold, across 2-hour, 3-hour and 4-hour event lengths, respectively. This is an important finding as it suggests concerns over the cold were highest during the 4-hour events, though notably other factors unrelated to cold remained important.

In addition to comfort-related reasons, 58–44% (29-37 participants) of these partial event participants cited other factors for not completing the full event. These included being away from home, forgetting the event had started, using hot water, or the thermostat activating before the event ended. Notably, 19–24% (12-16 participants) of partial event participants mentioned that their thermostat kicked in before the event concluded, not because they chose to end participation early, but because their heat pump automatically maintained a participant's pre-set comfort threshold. Overall, comfort was the main driver of participants choosing to participate in only part of the event. This highlights the benefit of clear communication and not penalising customer opt-out to ensure participants are safeguarded against feeling discomforted due to events. Participant's experiences on comfort across various event lengths are explored further in greater detail in Section 5.4 Comfort.

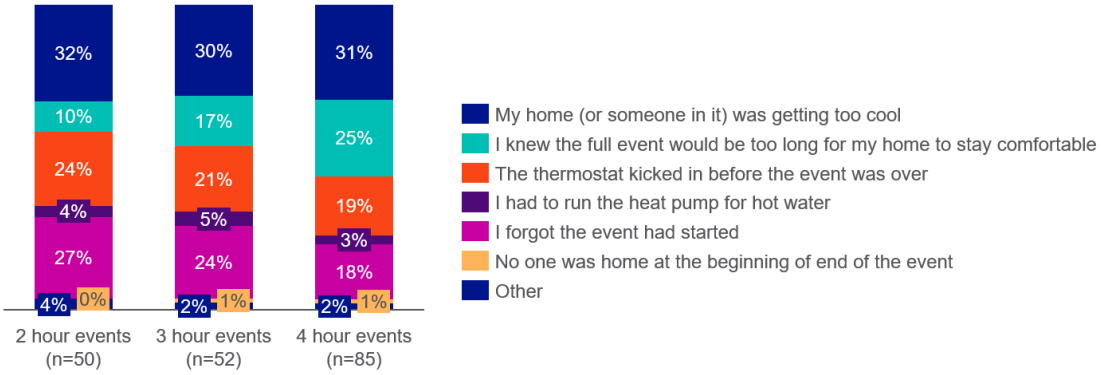


**Figure 11: Self-reported reasons for participating in part of the event reported in post-event survey, by participants who participated in part of an event (n=50-85)**

### 5.2.2 Participation rate of potentially vulnerable participants

We prioritised ensuring that all customers including potentially vulnerable customers and fuel poor customers could participate and benefit equitability from the trials. Enabling all customers to benefit from flexibility underpinned trial design factors. This approach valued customers' time and efforts within the trial structure, without losing sight of how these factors could later be maintained in enduring flexibility services. Methods for engaging potentially vulnerable and fuel poor customers evolved across the three trials based on lessons learnt.

To understand the experiences and perceptions of all participants, including potentially vulnerable participants, we disaggregated the mini-trial results into two categories: all participants and potentially vulnerable participants. We then compared Figure 10 and **Error! Reference source not found.** above, across participants we had identified as potentially vulnerable and non-vulnerable per Figure 2. We observed minimal differences in participation rate and reasons for participating. This lack of variation in results was an indication that potential vulnerability itself was likely not correlated with event length preference. Rather, individual preferences of participants would be the main influencing factors. This is explored further in Section 5.3 Views on event length and 5.4 Comfort.

## 5.3 Views on event length

### 5.3.1 Event length acceptance and preference

In order to gauge the holistic customer experience of different event lengths, the end-of-trial survey gauged suitability of different event durations for participants by asking them about their preference for and experience of each event length [28]. Figure 12 shows that on average 74% found 2-hour events about right, falling to 57% for 3-hour events, and 27% for 4-hour events. The overall trend is a clear preference towards 2-hour events, yet it is notable that over half of participants indicated acceptance of 3-hour events, while a sizable minority of more than 1 in 4 participants indicated acceptance of 4-hour events.

When asked to explicitly state their preferred event length[29], Figure 13 shows 46% of participants preferred 2-hour events, 32% preferred 3-hour events, 16% had no preference and just 6% selected 4-hour events. Considering 27% previously indicated they could tolerate 4-hour durations (Figure 12), it is possible that those reporting no preference could accept longer events, even if they did not prefer them. In fact, of those who reported no event length preference, 82% indicated that they could accept 4-hour event durations – significantly higher than the overall average (Figure 12). This suggests that those who selected no event length preference are more likely to accept 4-hour events.

It was valuable asking both questions: the first revealed what participants could reasonably tolerate, while the second revealed what they would ideally choose. While most participants favour 2-hour durations to support networks, a sizable number could potentially engage in longer events—offering still greater flexibility to support network needs.

[28] End of trial survey question stated: "Of the two-, three- and four-hour event lengths, which event length did you prefer?"
[29] End of trial survey question stated: "During the trial, EQUINOX events were two, three, or four hours long. What did you think about the two-, three-, and four-hour duration of EQUINOX events?"

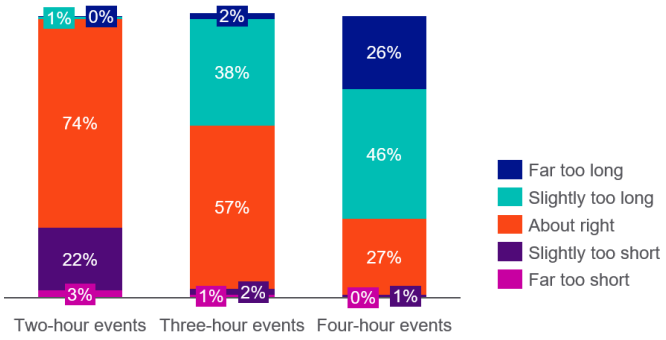


Figure 12: Participant views on each event length in end of trial survey (n=396)

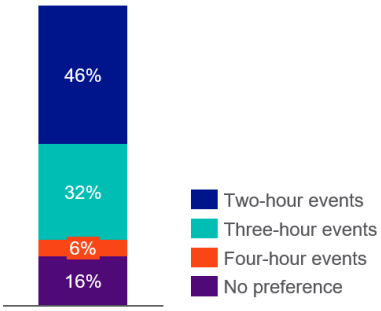


Figure 13: Participants' self-reported event length preference in the end of trial survey (n=396)

### 5.3.2 Reasons for event length preference

To explore the reasons behind participants' event length preferences, we analysed their responses based on household characteristics we considered might affect their ability to participate in longer events. We collected information on these characteristics in the start-of-trial survey and analysed it against preferred event lengths reported at the end of the mini-trial. The characteristics were:

- **Insulation levels**[30]**:** to determine whether poor insulation could reduce preference for longer events due to faster home heat loss. We decided to look at self-reported insultation levels as it provides a more direct measurement of how well a participant perceives their home maintains comfortable temperatures. In previous EQUINOX trials we analysed self-reported EPC ratings but while the ratings are useful for assessing a home's energy efficiency, they don't always capture the full picture of household comfort and are an indirect proxy for insulation quality.
- **Additional heating:** to determine whether the presence of additional heating could increase preference for longer events by providing an additional way for participants to stay warm without their heat pump being turned on.
- **Home batteries:** to determine whether the presence of home batteries could increase preference for longer events by enabling continued use of their heat pump without drawing electricity from the grid.
- **Potential vulnerability:** to determine whether potentially vulnerable customers could be less tolerant of longer events due to higher sensitivity to cold.

Figures 12 and 13 presents the proportion of respondents preferring each event length across the household characteristics. Since they were not mutually exclusive, participants could be grouped under multiple characteristics. The key findings were:

[30] Start of trial survey question stated: "How would you describe the quality of your home insulation?" If they selected "My home is not very well insulated and impacts the way I heat my home", they would be classified as "poorly insulated".

- **Insulation levels**: better insulation was correlated with longer events and conversely, self-reported poor insulation correlated with a preference for shorter events. 90% of participants with poor insulation preferred 2- or 3-hour events. None preferred 4-hour events. This suggests that such households may be better suited to participating in shorter events, though the small sample size of poorly insulated homes (n=20) limits the strength of this conclusion.
- **Battery ownership, additional heating, potential vulnerability**: none of these characteristics were correlated with a substantial deviation from the all-participant average. This suggests whether a participant had a home battery, had access to additional heating, or were potentially vulnerable was not a predicator to defining event length preference across the mini-trial cohort.

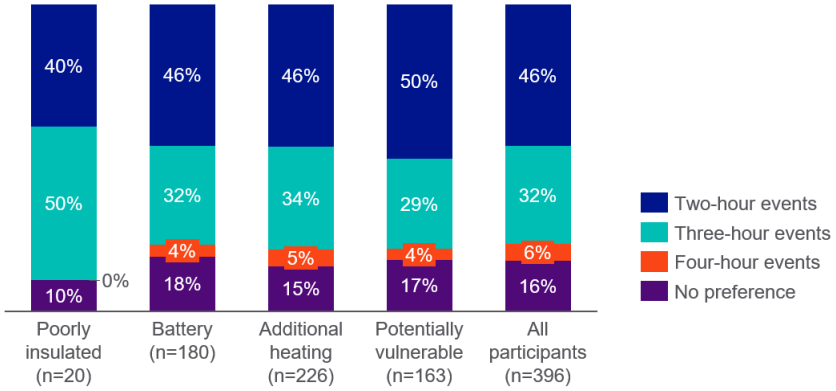


**Figure 14: Participants' self-reported event length preference in end of trial survey, segmented by household factors in start of trial survey (n=396)**

Overall, these findings suggest that households with good insulation may be better suited for 4-hour events. However, other characteristics we investigated did not strongly predict event length preference. This suggests that except for insulation, which may impact a participant's event length preference, any other variation in event length preference is likely due to individual customers' personal preferences rather than driven by specific household characteristics. It may be the case that the best approach to allocating participants to a given event length is to allocate them based on their self-disclosed preference.

## 5.4 Comfort

### 5.4.1 Overall comfort and comfort by event length

Comfort was a key consideration when testing longer event lengths during Longer Events Flex. To safeguard against potential underheating due to longer event duration, there was no penalty for not taking part in events or for participating in only part of an event.

We were encouraged that few participants reported a change in comfort[31]. As shown in Figure 15, 89% indicated that the events caused no change or a slight change in overall comfort. Only 8% responded that events made their home "cool" and 2% (8 participants) indicated that the events had made their home "too cold".

Figure 16Figure 15 shows responses to a follow-up question sent to participants who had reported a change in comfort. We asked which event durations were associated with that change. Half (50%) reported experiencing a change in comfort only during the 3- and 4-hour events, while 28% attributed it solely to the 4-hour events. A smaller group (14%) reported changes in comfort across all event lengths. Interestingly, 8% stated that event duration did not affect their comfort, despite having earlier indicated that they noticed a change[32].

Out of the 396 participants who answered the initial question on comfort, 234 reported experiencing some level of change. Among these, 189 described their home as feeling only "slightly cool," and just 8 participants said it felt "too cold." This indicates that the majority of comfort changes discussed in Figure 16 & Figure 15 were relatively minor, suggesting that while participants noticed a difference, it typically did not result in significant discomfort.

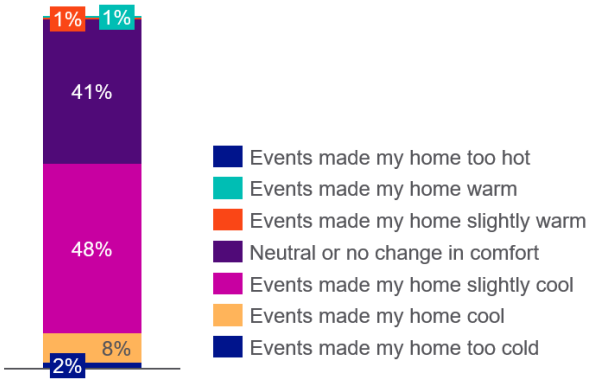


Figure 15: Participants' self-reported impact on comfort levels in end of trial survey (n=396)


Figure 16: Participants' self-reported impact of event length on their comfort level in end of trial survey (n=234)

Taken together, these findings on comfort align with the self-disclosed event length preferences. The combined results suggest that 2-hour events are likely to be the most widely accepted and represent a strong baseline for large-scale heat pump flexibility. A sizable proportion of participants may be able to participate in 3-hour events without a significant change in comfort – aligning with our findings on the EQUINOX trial two participants with tariff types that have multiple daytime prices. Meanwhile, a smaller but still relevant proportion appear capable of participating in 4-hour events. This supports the case for a tiered approach to event design, offering shorter events as

[31] End of trial survey question stated: "How much, if at all, did the Longer Events Flex events impact the comfort levels of your home?"

[32] From discussions with participants through interviews and focus groups we understand that comfort can be highly personal and not only impacted by the characteristics of one's home (e.g., home insulation level, etc.), but also personal preference and circumstances.

standard, with longer durations targeted to participants whose homes are more resilient or who have indicated comfort with extended turndown events.

### 5.4.2 Comfort of potentially vulnerable participants

We additionally segmented the responses on comfort according to whether respondents were potentially vulnerable or non-vulnerable, to identify any important trends in the preference for different event lengths by potentially vulnerable participants. This is shown in Figure 17 [33].

Among participants who reported a change in comfort during the mini-trial, Figure 17 shows potentially vulnerable individuals were slightly more likely to experience a change in comfort during shorter events. 19% of potentially vulnerable participants reported a comfort change during all events. 52% experienced changes during both 3-hour and 4-hour events, while 20% reported a change only during 4-hour events. In comparison, a lower 9% of non-vulnerable participants reported a comfort change during 2-hour events. A slightly lower 46% experienced changes during both 3-hour and 4-hour events, and a higher 38% reported a change only during 4-hour events.

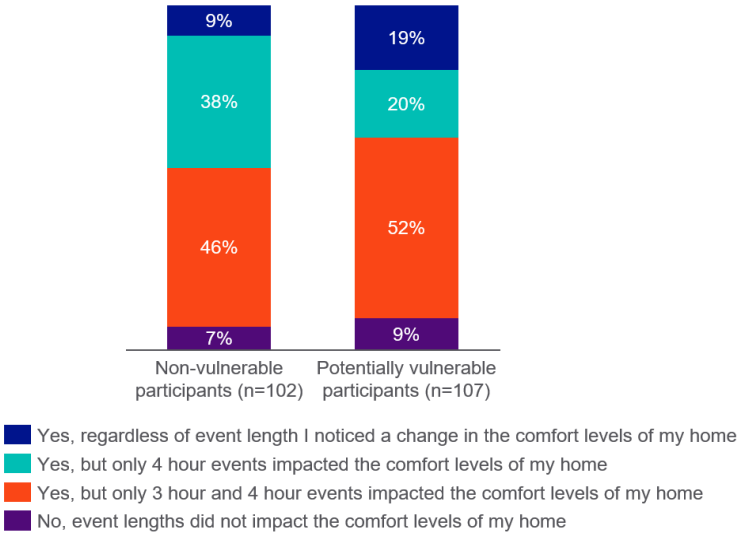


**Figure 17: Participants' self-reported impact of event length on their comfort level in end of trial survey, segmented by potential vulnerability (n= 209)**

This comparison suggests potentially vulnerable participants were more likely to experience comfort changes across all event durations, whereas non-vulnerable participants were more likely to feel an impact during the longest events. However, it is important to note the segmentation did not include responses from the 39% of potentially vulnerable participants that recorded no change in comfort, showing that despite the trend observable in the subset of participants in **Error! Reference source not found.**, a significant proportion of potentially vulnerable participants could participate in longer events without issue. These findings show that there is an interest and ability from potentially vulnerable customers to take part, but that comfort is highly personal and depends on both the household and the home itself. This makes it critical that flexibility programmes have customers in mind from the offset and use simple measures like voluntary participation, clear communication and not penalising customer opt-out to ensure customers are empowered to take part in flexibility programmes. With such measures in place, direct participation in heat pump flexibility should be an option for all customers, including those with potential vulnerabilities.

[33] Note, for Figure 17 n=209 because 25 of the participants that indicated they had felt a change in comfort as per Figure 16 did not complete the start of trial survey. Therefore, it was unclear if they were potentially vulnerable.

Overall, the results present a nuanced picture of comfort impacts for potentially vulnerable participants. Since the definition of potentially vulnerable encompasses a wide range of individual characteristics and circumstances, the most appropriate approach to determining participation in longer events is to rely on individual customers' self-reported preferences. The exception may be poor insulation as results suggest that homes with this characteristic may be better suited to shorter duration flexibility events.

## 5.5 Event reminders and notice period

Participants received a notification to complete the post-event survey at the end of each event. While we wanted to encourage participants to complete the post-event survey, we also sent this survey as a reminder to participants that the event had ended and to encourage them to turn their heating back on. Figure 18 **Figure 18**shows 42% of non-vulnerable participants found the reminder helpful for turning their heating back on and 49% of potentially vulnerable participants found the reminder helpful. This indicates the importance of sending a reminder at the end of events when implementing domestic flexibility services for heat pump homes, particularly to potentially vulnerable households.

Participants were notified of Longer Events Flex events a day prior to the event via an email from their energy supplier. To assess if day ahead notification is sufficient for customers to provide flexibility for all event lengths, participants were asked about their satisfaction with this notice period in the end of trial survey[34]. As shown in Figure 19 **Figure 19**, 91% of participants were satisfied with the day ahead notice period, with no significant difference observed between potentially vulnerable and non-vulnerable participants. These results suggest that day ahead notice periods are sufficient for most customers.

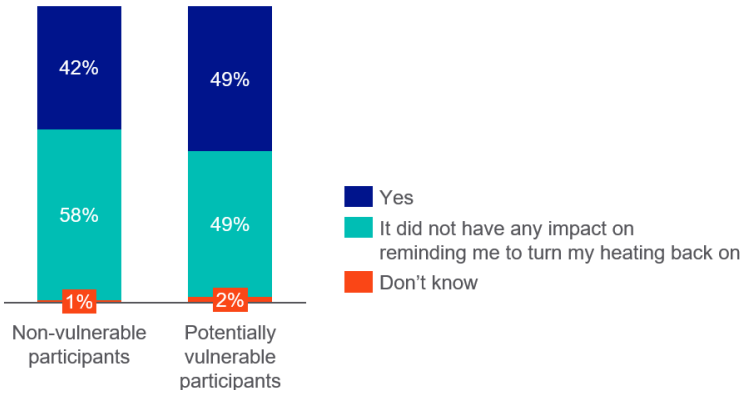


**Figure 18: Non-vulnerable participants responses on whether the post-event survey email helped remind them to turn their heating back on in end of trial survey (n=232) compared to potentially vulnerable participants (n=163)**

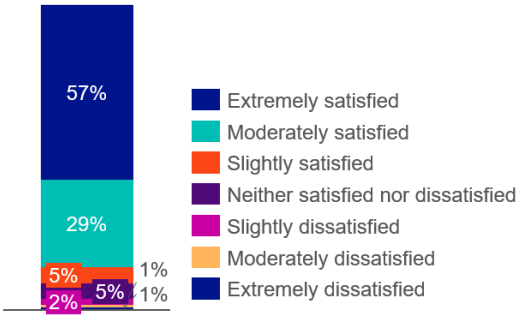


**Figure 19: Participants' satisfaction with the day-ahead notice period in end of trial survey (n=396)**

[34] End of trial survey question stated: "How satisfied were you with the day-ahead notice periods for the EQUINOX events?"

# 6. Summary

We are encouraged by the results of Longer Events Flex, and the potential for heat pump homes to participate in longer duration evening peak flexibility events. The key takeaways are:

1. **Heat pump homes can achieve demand response during the evening peak for up to 4-hours.** Our DiD analysis found statistically significant demand response volumes for 2-hour events (0.42 kW average, within a 90% CI of 0.16 - 0.68 kW) and for 4-hour events (0.27 kW average, within a 90% CI of 0.10 - 0.44 kW). These volumes are not statistically different from one another in this mini-trial and are in line with our expectations for heat pump turndown, based on trial two analysis. The 3-hour events did not yield statistically significant results, but observational analysis and survey responses suggest that demand response likely occurred.

2. **We observed that heat pump homes maintained the same magnitude of demand response across the full event window, including for events up to 4-hours**. Our DiD analysis detected statistically significant demand response in every settlement period across 2-hour and 4-hour event lengths, without any statistically significant difference in magnitude across settlement periods. This mini-trial finding suggests that heat pump homes can provide consistent demand response during events up to 4-hours long.

3. **Half of trial participants preferred 2-hour events, though large minorities preferred 3-hour and 4-hour events.** 46% preferred 2-hour events, 32% preferred 3-hour events, and 22% either preferred 4-hour events or had no preference. This suggests that large minorities of heat pump homes can support 3- or 4-hour durations to help networks better manage constraints – even if 2-hours is the preferred length for many.

4. **Poor insulation was the only identifiable determinant of event length preference.** Amongst participants with poor insulation, none preferred 4-hour events and only 10% reported no preference. Across other factors such as home battery ownership, access to additional heating and potential vulnerability, there was no deviation from the overall treatment group on event length preference. This suggests that except for poor insulation, which may predicate a customer to prefer shorter events, any other variation in event length preference could be due to individual customers' personal preferences rather than specific characteristics.

5. **Most participants did not experience a meaningful change in comfort due to events. For those who reported any change in comfort, the change was overwhelmingly mild.** Of all participants, 89% reported no change or slight change in comfort throughout the trial. Of the 59% who experienced a change, only 3% (8 participants) said it became "too cold." Taken together, findings on comfort aligned with participants' self-disclosed event length preferences and indicated that comfort was the most common motivator for event length preference.

6. **There was little variation in participant comfort among potentially vulnerable and non-vulnerable participants.** 39% of potentially vulnerable participants reported feeling no change in home comfort at all, and in general across all participants the defined change in comfort was overwhelmingly mild in characterisation, with 89% of participants reporting no change or a slight change in comfort. The 61% of

potentially vulnerable participants who reported feeling a change in home comfort were slightly more likely to report that the change in comfort occurred in all events compared with non-vulnerable participants, who skewed slightly more towards longer events.

7. **Trial participants expressed strong satisfaction with day ahead event notification and end-of-event notifications.** 91% of participants were satisfied with day ahead event notice periods. 42% of non-vulnerable participants found the reminder to turn their heating back on alongside the post-event survey helpful, compared to 49% of potentially vulnerable participants. This suggests day ahead notification is an effective design approach for domestic flexibility procurement.

8. **Trial satisfaction was high, suggesting willingness to participant in longer duration flexibility events.** 78% of participants were satisfied with the trial. Satisfaction levels aligned with those in EQUINOX trial two at 79%, when just 2-hour events were tested. This indicates that heat pump homes are open to considering longer events.

In conclusion, the results of the Longer Events Flex mini-trial provide a strong foundation for exploring customer willingness to provide flexibility on a more structural basis. We found that half of participants preferred 2-hour events, but sizeable minorities were open to 3- or 4-hour durations. Self-reported poor insulation emerged as the only clear factor influencing event length preference. These key findings informed our design for the Daily Evening Flex mini-trial, where we ultimately decided to divide participants into different event length groups defined by their own self-declared event length preferences.

# 7. Appendix A: Project Partners

EQUINOX is led by National Grid, along with multiple project partners and collaborators, as detailed in Table 2.

**Table 2: List of EQUINOX partners and collaborators**

| Name | Project function | Role |
|---|---|---|
| NGED | DNO | Project lead. Responsible for running the technical integration, trial design, and project management and knowledge workstreams. |
| Guidehouse | Consultancy | Partner. Responsible for supporting the commercial arrangement design and customer engagement workstreams. Supporting on trial design, data analysis, project management, and knowledge dissemination. |
| Octopus Energy | Energy supplier | Partners. Responsible for planning and delivering EQUINOX trials with participants from their customer base. Supporting all project workstreams as commercial flexibility service providers and customer experts. |
| Sero | Energy supplier[35] | |
| ScottishPower | Energy supplier | |
| Passiv UK | Smart technology company | Partner. Responsible for simulating the flexibility impacts for different intervention strategies and household archetypes. |
| West Midlands Combined Authority | Local government | Partner. Responsible for coordinating a social housing heat pump installation programme which can contribute customers to trials two and three. Also advising on equitable participation. |
| Welsh Government | Government | Partner. Responsible for running a social housing heat pump installation programme which can contribute customers to trial three. |
| National Energy Action | Charity | Collaborator. Responsible for running participant focus groups to understand trial perceptions. NEA will ensure that the needs of customers with vulnerabilities are accounted for in the trial design. |
| SP Energy Networks | DNO | Partner. A DNO brought on board to ensure that the design is interoperable for all DNOs. SPEN's license areas will join trial three. |
| National Energy System Operator | NESO | Collaborator. Responsible for sharing learnings between EQUINOX and other NESO flexibility programmes. |

[35] Sero is not an energy supplier but assuming the role for the purpose of trials one and two.

# 8. Appendix B: Difference in Difference approach

As detailed in Section 3.5, the DiD approach was used to determine the treatment effect during Longer Events Flex.

**Equation 1 below provides the DiD specification used to calculate an estimate of demand response for each event. Equation 1**

Equation 2 details the calculation of uncertainty (standard error). Equation 2Equation 3. RMS standard error calculationEquation 3 details the calculation we used to aggregate the standard errors to generate an overall average demand response impact across multiple events.

### Equation 1: DiD demand response calculation[36]

$$Demand\ response = \left[mean(observed\ demand_{treatment,\ event}) - mean(observed\ demand_{control,\ event})\right] - \left[mean(observed\ demand_{treatment,\ non-event}) - mean(observed\ demand_{control,\ non-event})\right]$$

### Equation 2: DiD standard error calculation[37]

$$Demand\ response\ standard\ error = \sqrt{\frac{Variance(observed\ demand_{treatment,\ event})}{Customer\ count_{treatment,event}} + \frac{Variance(observed\ demand_{control,\ event})}{Customer\ count_{control,event}} + \frac{Variance(observed\ demand_{treatment,\ non-event})}{Customer\ count_{treatment,non-event}} + \frac{Variance(observed\ demand_{control,\ non-event})}{Customer\ count_{control,non-event}}}$$

The treatment group participants were deemed to have delivered demand response for an event if they increased their energy consumption during the event as per Equation 1.

As detailed in Section 3.5, the RMS approach was used to determine the aggregated standard error during Longer Events Flex. Equation 3Equation 3 below provides the RMS specification used to calculate an estimated standard error grouped across events.

[36] Observed Demand = household consumption in kWh; Treatment = group of customers called to participate during an event; Control = group of customers not called to participate; Event = time frame during which consumption was averaged across the treatment or control group of customers on event days; Non-event = time frame during which consumption was averaged across the treatment or control group of customers on non-event days.

[37] Variance = statistical measure quantifying estimate uncertainty; Customer count = number of customers called during event.

### Equation 3. RMS standard error calculation[38]

$$Root\ Mean\ Square\ (RMS) = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(SE_i)^2}$$

[38] $SE_i$ = standard error from event *i*; *N* = total number of events.

National Grid Electricity Distribution plc
Avonbank
Feeder Road
Bristol BS2 0TB
United Kingdom

nationalgrid.co.uk

Month YEAR. Reference Number