EQUINOX

Learning from trial three: Daily evening heat pump flexibility

August 2025

nationalgrid

nationalgrid.co.uk

Version control

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

Publication control

Name	Role
Ewan Steen, Raeesa Hirji, Samantha Piller (Guidehouse)	Author
Stuart Fowler (National Grid) Alex Jakeman, Callum Coghlan (Guidehouse) Jessica Colleen, Tessa Hillen (Octopus Energy)	Reviewer
Stuart Fowler (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
- **Flexibility Service Provider:** An entity that offers or procures flexible energy resources to help balance electricity supply and demand in response to grid needs
- **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
- **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.

[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

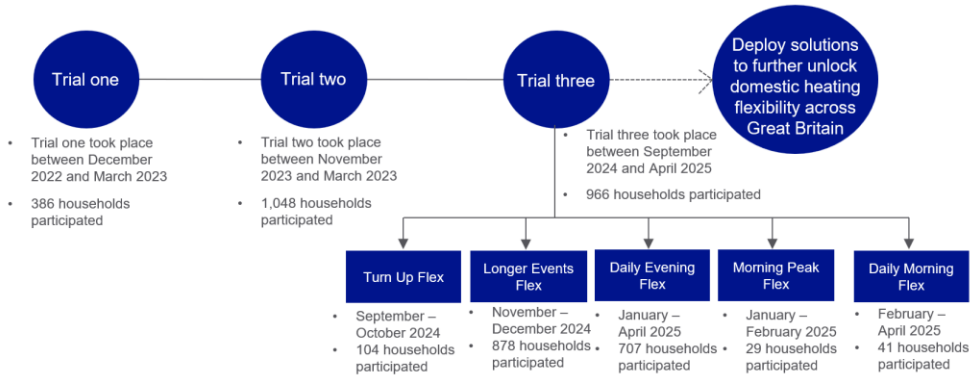

Figure 1: EQUINOX project overview

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:

[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.

- **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 Daily Evening Flex.

3. Daily Evening 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 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].

Feedback from participants in trial two suggested that flexibility 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 turndown daily from

[7] Demand response when more electricity is consumed relative to a baseline.
[8] Demand response when less electricity is consumed relative to a baseline.
[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.

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 first tested 4-hour events less frequently in the Longer Events Flex mini-trial, to establish which customer groups might be comfortable adopting daily 4-hour heat pump flexibility behaviour in the subsequent Daily Evening Flex mini-trial. The Longer Events Flex mini-trial found that participants typically preferred 2-hour turndown events (47% of survey respondents) but also identified a meaningful minority of participants (20% of survey respondents) who preferred 4-hour events or expressed no preference between the event lengths tested.

Building on these results, Daily Evening Flex tested daily turndown events over a longer period, with participants taking part in 2-hour or 4-hour events, according to their preferences reported in the Longer Events Flex end of trial survey[13]. We considered that if customers can consistently participate for 2-hours this would align well with the needs of constraint management, which often requires flexibility over a 4-hour window. This event length discrepancy could be effectively managed by engaging different cohorts of heat pump homes in staggered 2-hour intervals across an event period, or by combining heat pump participation with other flexible assets such as electric vehicle charging. Naturally, if some heat pump homes are capable of responding for longer periods up to 4-hours, this would offer additional flexibility value.

An additional factor that we considered in trial three was facilitating customer participation in multiple, "stacked", flexibility services. Participants in trial two indicated an appetite for being able to choose to access multiple flexibility products where possible, with 69% of survey respondents indicating that they would have liked to have been able to participate in products for NESO's Demand Flexibility Service (DFS)[14], alongside their participation in EQUINOX trials. Beyond the customer benefits of being able to access greater rewards for flexibility participation, there are network benefits to assets being able to provide flexibility for different contexts. For winter 2024-25, the DFS programme rules allowed customers to participate in additional flexibility services while also being enrolled in a customer-facing DFS product. Daily Evening Flex was designed in a way that would permit such stacking, allowing us to conduct a brief observational analysis to understand if participants in EQUINOX would also take part in any DFS events through their energy supplier's DFS product[15].

The Daily Evening Flex mini-trial aimed to assess:

[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".
[13] See full report: "Learning from trial three: Longer Heat Pump Flexibility Events"
[14] A NESO-run programme that enables households and businesses to earn rewards from their registered DFS service provider in return for shifting their electricity usage during DFS events: "Demand Flexibility Service explained".
[15] Octopus Energy's customer-facing product for offering DFS events is "Octoplus Savings Sessions".

- Whether heat pump homes can deliver daily demand response in the evening for 2 hours or for 4-hours, depending on customer preference.
- Whether this daily demand response can be achieved alongside high rates of participant satisfaction and with no or minimal impact on comfort – including for those with potential vulnerabilities.
- The customer experience of stacking daily EQUINOX trial events with ad-hoc NESO DFS events.

3.2 Trial structure

Daily Evening Flex ran for 11 weeks from January to April 2025. Customers who had been part of the previous mini-trial (Longer Events Flex) were invited to take part in daily heat pump turndown events, Monday to Friday each week. Based on their self-reported event length preferences, participants were allocated to either the 2-hour group with events from 5-7pm or to the 4-hour group, with events from 4-8pm. In order to maximise data volume per group, Daily Evening Flex was limited in scope to only consider two event length groups and did not include a 3-hour event length.

Customers in the 2-hour and 4-hour event length groups collectively formed the active trial group (the "treatment group") and were asked to reduce electricity consumption associated with ("turndown") their heat pumps and hot water heating during events. They could achieve this in whatever way made sense for their household, for example by reducing the temperature set point on their thermostat and were not asked to make any other changes to how they used electrical appliances during events. Participants were able to pre-heat their homes before events if they chose to but were not required to do so. There was no penalty if customers chose not to participate in some events, or if they chose to not participate for the full length of an event.

At the beginning of the previous mini-trial, we recruited an additional group of customers for control purposes during data analysis (the "control group"). These customers met the same eligibility criteria as the treatment group customers and then served as the control group for the Daily Evening Flex mini-trial too. 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.

3.3 Participant details

3.3.1 Trial eligibility

All participants in Daily Evening Flex were part of the previous mini-trial, Longer Events Flex. The eligibility criteria were not changed and to be eligible to participate, 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).

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 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. Participants for Daily Evening Flex were recruited exclusively from customers who had been part of the treatment group in the previous mini-trial, Longer Events Flex. Of the 510 customers invited, 353 signed up, a conversion rate of around 70%. The control group from the previous mini-trial was maintained for Daily Evening Flex, with no additional recruitment efforts.

Within the treatment group for Daily Evening Flex, customers were allocated to have either 2-hour events or 4-hour events, based on their self-reported event length preferences at the end of Longer Events Flex. Analysis of the Longer Events Flex mini-trial results showed no strong correlation between participant characteristics and event length preferences, aside from a slight tendency for those with poorer insulation to prefer shorter events. These results did not give any compelling reason to override self-reported preferences from participants. Where participants had previously indicated a preference for 4-hour events or for 2-hour events, they were allocated to those groups. Participants who had indicated a preference for 3-hour events or who had not responded to the survey question were allocated to the 2-hour group. A small group of participants had indicated that they had no preference amongst event lengths, which was contextualized in free text answers as high satisfaction with each, and therefore no preference between them. These participants were allocated to the 4-hour group. All allocations were made ahead of recruitment so that they could be communicated to customers as part of their invitation to take part in the mini-trial. In total, 264 participants were recruited to the 2-hour group and 69 to the 4-hour group.

3.3.3 Demographics

We used self-reported participant data[16] to understand how representative the participants were of the wider UK population. These are important criteria for understanding the wider applicability of the mini-trial results:

- Homeowners were overrepresented among trial participants compared to the UK average. 88% of participants lived in a property they own and 10% pay social rent, compared to the UK average of 64% of individuals owning their home and 17% paying social rent[17]. We saw similar homeownership in trial two.
- High income households tended to be overrepresented: 41% of participants had an income that exceeds £50,000 per year; 8% of participants annual household income was between £20,000 - £24,999. This differed from the UK average with 19% of individuals in the UK having an income that

[16] Customer demographic data was self-reported through an optional start of trial survey administered by Guidehouse.
[17] Office for National Statistics, Census 2021

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 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[18]. 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 trial participants and the general population are in line with our experience in previous EQUINOX trials and based on survey results[19] are thought to reflect the mismatch between current heat pump owners and the wider population. Although Daily Evening Flex is an important early exploration of the potential for heat pump homes to engage in daily demand response for longer periods of time, we recommend caution when applying its trial results to much broader populations that may interact differently with heat pumps than this participant group.
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 Equitable Participation Framework (EPF) [20]. We identified eight self-reported factors[21] 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[22] 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.

Of the 272 participants (80%) who completed the start-of-trial survey, we identified 56% (153 participants) as potentially vulnerable. Figure 2**: Proportion of potentially vulnerable participants in Daily Evening Flex in the 2-hour group (n=211) and the 4-hour group (n=61), based on the start of trial surve** illustrates the distribution of vulnerability factors between the 2-hour and 4-hour event groups. Both groups included a

[18] Office for National Statistics, Census 2021
[19] This is discussed in further detail in "Project Deliverable 5: Learning from Engaging Customers"
[20] See "Project Deliverable 5: Learning from Engaging Customers" for detailed explanation of the customer engagement approach, including how vulnerability factors are considered.
[21] 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.
[22] Households were classified as low income based on their household composition and annual household income, in alignment with the UK minimum income standard.

substantial proportion of potentially vulnerable individuals, with a slightly higher percentage in the 2-hour group (52%) compared to the 4-hour group (43%). The 4-hour group had a higher proportion of participants with a disability (26%) than the 2-hour group (14%). The inclusion of a high proportion of potentially vulnerable participants in Daily Evening Flex provides essential context for interpreting the trial outcomes and evaluating the broader feasibility of implementing longer-duration events.

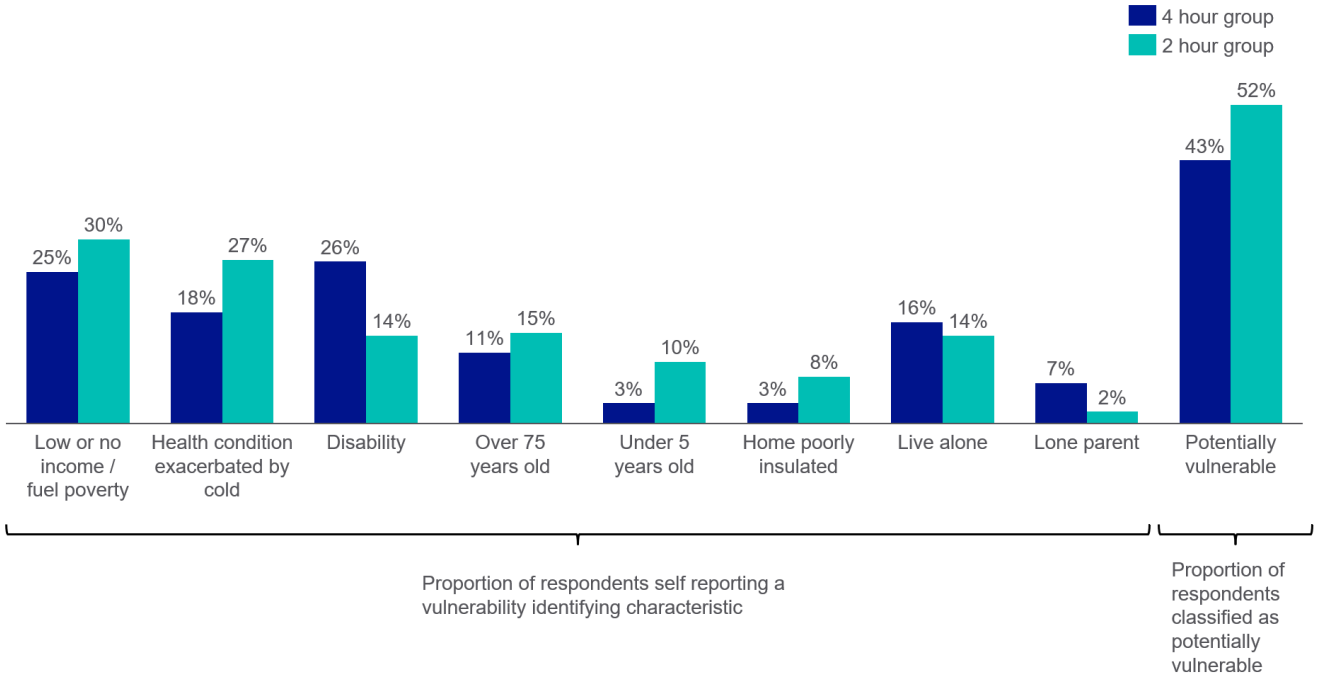

Figure 2: Proportion of potentially vulnerable participants in Daily Evening Flex in the 2-hour group (n=211) and the 4-hour group (n=61), based on the start of trial survey

3.4 Commercial arrangements

3.4.1 Daily Evening Flex settlement

To assess the potential of heat pump homes to participate in daily 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 every weekday for eleven weeks from January to April 2025. Participants were notified of the events structure at the beginning of the trial but were not sent daily or weekly reminders. There was no requirement to opt-in to individual events and participants received the event incentive of £1 per week based on their consumption during events remaining below their personalised trial baseline for four out of five weekdays (see Section 3.4.2 for more details). The event incentive value was based on average market rates offered by NGED for similar flexibility products, and our expectations of how much demand response

participants would deliver over two hours based on the trial two results[23] but does not necessarily represent a business-as-usual customer offering, which may incorporate additional flexibility values.
For Daily Evening Flex, we rewarded participants through a combination of methods, in recognition of the value of their time and participation. Overall, treatment participants could earn up to £25 if they participated in the full trial period and turned down in most weeks. They received a £5 reward for joining the trial, in place of a start-of-trial survey payment, and a £10 reward for completing the end-of-trial survey. Those who took part in focus groups received an additional £25 reward.

Control group participants were not asked to engage in Daily Evening Flex events but were invited to complete an end of trial survey. They could earn up to £20 for participating across trial three, made up of a £10 start of trial survey rewards at the time they joined trial three, and a £10 end of trial survey reward at the end of Daily Evening Flex.

Table 1: Daily Evening Flex commercial arrangements

Item	Treatment participant	Control participant
Payment structure	£5 for joining the mini-trial, in place of a start of trial survey £10 for completion of end of trial survey £1 per week for each week measured as a "success" against a personalised trial baseline (explained in detail in Section 3.4.1) £25 for participating in a focus group (a subset of participants)	£5 for completion of recruitment survey (when joining trial three) £5 for completion of start of trial survey (when joining trial three) £10 for completion of end of trial survey
Control type	Manual and remote customer control	
Notice period	Week 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-hour or 4-hours	
Event timing	2-hours between 5-7pm; 4-hours between 4-8pm	
Event frequency	One event per weekday (Monday-Friday)	

[23] NGED procurement prices for Sustain in 2023/4 range from £0.01/kWh to £4.60/kWh, averaging £0.30/kWh. Based on the average value, and the average performance from trial two of 0.57 kWh per participant over 2 hours.

Supplier notice	Informed by NGED a week before events began

3.4.2 Daily Evening Flex personalised trial baselines

To detect turndown and be able to reward participants with the event incentive, we used a personalised trial baseline approach, adapted from the industry standard p376 baseline method[24]. In previous EQUINOX trials we have used the p376 method to calculate a baseline for each participant. The baseline is refreshed for each turndown event, based on that participant's recent consumption on preceding non-event days. Although this has worked well, it is not a suitable method for daily flexibility events as the baseline would not be updated for each event and would soon lose relevance as a counterfactual for electricity consumption during events. For some commercial flexibility services, a nominated baseline is assigned based on non-personalised factors such as average demand patterns and consumption volumes of homes with the same type of technology (e.g. heat pump or EV charger). For Daily Evening Flex, we used the relatively small-scale nature of the mini-trial to innovate a baseline approach that was both personalised for each participant and did not change from event to event. This allowed us to detect heat pump turndown at the home smart meter level, without specific asset metering, and to reward each participant for their performance in trial events. We do not present this trial baseline methodology as being directly suited to wider commercial use.

Personalised trial baselines were calculated based on each participant's average consumption during their event time (5-7pm for the 2-hour group, 4-8pm for the 4-hour group) on 30 non-event, working days in November-December 2024 and were a set value for each participant that was not updated from week to week or between events. This approach aimed to capture an average of each customer's consumption across a spread of winter temperatures that might be representative of external temperatures across the mini-trial period. To prioritise customer engagement and recognise that this method may under-report turndown on colder event days, we increased each participant's raw baseline by 10% during the mini-trial, making the event incentive more achievable. As illustrated in Figure 3, participants were considered to have achieved turndown in an event if their electricity consumption in that period remained

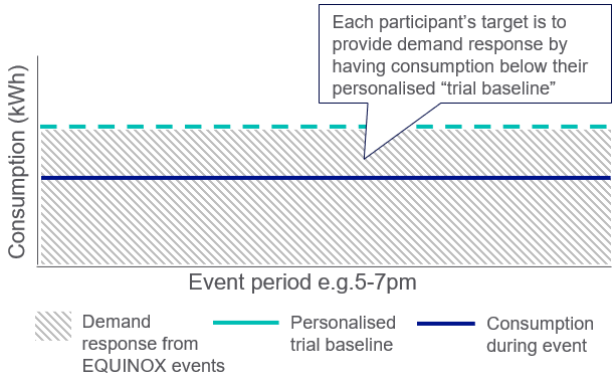

Figure 3: Schematic of Daily Evening Flex settlement approach

[24] For settlement, the p376 method uses the last 10 non-event weekdays to generate an average consumption value per settlement period, resulting in a half-hourly baseline demand profile for an event day. A participant's demand response is calculated by deducting their observed electricity consumption from their baseline consumption. Demand response achieved is then multiplied by their utilisation payment rate for that event to determine the performance incentive owed to that participant. See Elexon report for full details of the p376 method.

below their personalised trial baseline. Eligibility for the event incentive was based on achieving turndown for at least four of five weekdays each week and did not change based on the volume of turndown. The Daily Evening Flex personalised trial baselines were only used for customer settlement in the mini-trial and were not used to evaluate overall demand response.

3.4.3 Stacking

As discussed in Section 3.2, Daily Evening Flex participants were also able to choose to participate in DFS events through their energy supplier. Across winter 2024/25, these events ("DFS events") offered consumers a reward for turndown demand response during specific event periods, often across the same evening peak period (5pm – 7pm) in which participants would be asked to turn down their heat pumps for Daily Evening Flex. We permitted customers to access both programmes if they chose to but did not directly interact with any DFS events, either operationally or for customer settlement.

Our objectives around service stacking were aimed at observing customer appetite for participating in general turndown DFS events alongside heat pump turndown Daily Evening Flex events. EQUINOX trial three was not designed to test the full end to end industry processes for service stacking, or for settlement. Our learnings focus on the customer experience of participating in both services simultaneously, and not on demonstrations of how processes could work for stacking flexibility services more broadly. We will cover aspects of what may follow next in our project "Deliverable 6: Recommended transition of learning to BaU" report (December 2025).

To assess whether Daily Evening Flex participants also engaged with DFS events, we analysed aggregated outcomes for participants who took part in DFS events that occurred concurrently with the Daily Evening Flex trial period (27th January to 4th April). We excluded from analysis any DFS events that occurred in the first two weeks of the mini-trial. This ensured that we only considered customer results from DFS events where their DFS baseline, which is based on each participant's recent average consumption, already reflected heat pump demand response from their ongoing participation in EQUINOX events. Figure 4 illustrates how a participant's average consumption during an event window (e.g., 5–7 pm) may change as a result of participating in EQUINOX and DFS events:

- The light blue line is expected demand if there is not a demand response event (baseline for Daily Evening Flex).
- The red line is demand during an EQUINOX event, based on a request for heat pump turndown.
- The dark blue line is demand during a DFS event, based on a request for turndown from any device.

Customer settlement for Daily Evening Flex was based on whether participants' home consumption during the event window was lower than their average pre-trial consumption (shown as the light blue line in Figure 4). The expected heat pump turndown response is represented by the partially shaded block.

If participants also joined a DFS event that overlapped with a Daily Evening Flex event, they would be rewarded in the DFS product for any additional reduction below their average recent in-trial consumption (the red line). This extra turndown during a stacked event is shown as the fully shaded block.

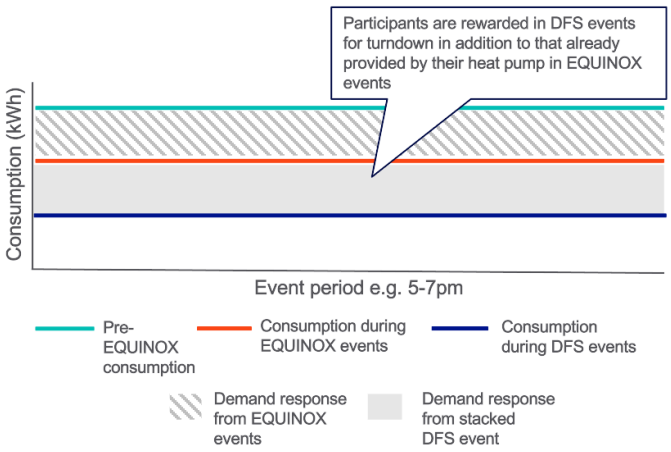

Figure 4: Representation of demand response during a stacked DFS event

To detect if there was an additional turndown response when customers participated in DFS events in addition to Daily Evening Flex events, we looked at the aggregated DFS event outcomes for that subset of trial participants. In addition, trial participants were asked about their engagement with DFS events in the end-of-trial survey for Daily Evening Flex.

3.5 Analysis approach

3.5.1 Difference in Difference approach for Daily Evening Flex

Following industry best practice for evaluating trial programmes, suppliers leveraged a Difference-in-Difference (DiD) approach for calculating demand response. DiD approaches provide an 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 group, per week, and for the full trial duration, though we used the Root Mean Square (RMS) method to calculate aggregate errors[25].

To confirm the control group's ongoing suitability, we compared it with the treatment groups (2-hour and 4-hour event lengths) across key metrics. The 2-hour group and the control group were similar in energy use, heat pump size and ownership of technologies like home batteries, and we deemed these groups to be suitably matched.

The 4-hour group was smaller, only 69 participants, and was not as well aligned to the control group in the key metric of average energy use, particularly in regard to recent average electricity consumption during the relevant 4-8pm period. This was not entirely unexpected as participants with a preference for 4-hour event length may represent a niche group of customers. Within the 4-hour group, we saw that there was an increased level of variation in recent average energy use between customers, often shaped by home battery use and time of use tariffs with overnight off-peak periods. In a few cases, individual participants' recent average consumption over the evening peak period was effectively 0kWh, suggesting that they were relying

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

on stored energy. We interpreted this as an indication that some customers were already engaging in demand flexibility according to some pre-existing motivation.

The specialist nature of the 4-hour group meant that they were not sufficiently matched with the control group and the DiD analysis struggled to detect statistically significant turndown in the latter half of the Daily Evening Flex mini-trial. We speculate that this corresponds with lower demand response volumes as external temperatures increased. This is discussed further in Section 4.2.

3.5.2 Customer experience

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

Throughout the trial, we captured the experiences and perceptions of trial participants through surveys, interviews and focus groups. For Daily Evening Flex, we conducted:

- An end of trial survey, with was administered to participants at the end of the trial to understand the overall experience;
- Three short check-in surveys near the start, middle and end of the trial sent fortnightly to check in on participants comfort and satisfaction; and
- Six focus groups with 22 total participants to gain deeper insights on participant's experience with the processes of the trial.

3.5.3 Stacking

To understand whether the customer appetite for participation in multiple "stacked" services that we observed in trial two would repeat during Daily Evening Flex, we report on the aggregated outcomes from mini-trial participants that also chose to sign up to their energy supplier's DFS product. As discussed in Section 3.4.3, customer outcomes in DFS events that occurred after the first two weeks of the Daily Evening Flex mini-trial would be shaped by these participants' pre-existing flexible behaviours in the mini-trial events.

We also evaluated the participant experience of being able to participate in DFS events in addition to Daily Evening Flex mini-trial, incorporating questions on this topic in the end of trial survey that was sent out to all Daily Evening Flex participants.

4. Demand response results

Section 4 presents our findings for our aim of assessing whether participants could consistently provide statistically significant daily demand response in the evening for 2 hours or for 4 hours, depending on customer preference — confirming the potential for heat pump flexibility to provide value in constraint management flexibility products. Section 4.1 shows week-on-week demand response results for the 2-hour group, and Section 4.2 shows week on week results for the 4-hour group. A brief comparison is made between the groups, in order to consider if the longer events may affect the volume or consistency of the demand response, as compared to the 2-hour events. Average external temperatures are presented but were not used in the DiD calculations.

4.1 Weekly demand response for the 2-hour group

Figure 5 shows average treatment participant demand response and average external temperature (5-7pm) per week for the 2-hour group.

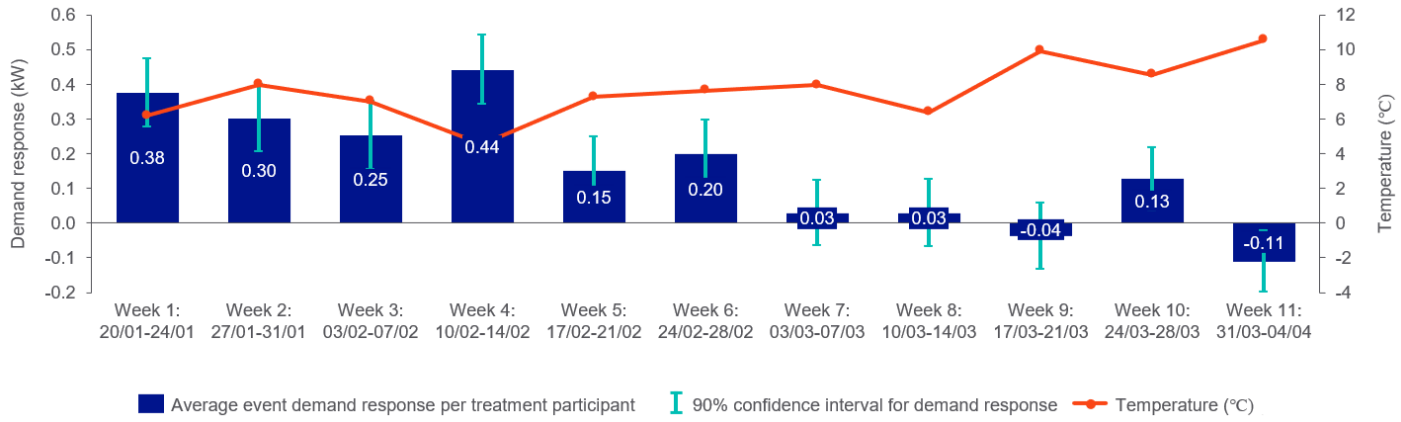

Figure 5: Average week over week demand response per treatment participant for the 2-hour group, and average external temperature during the event time

Figure 5 We observed statistically significant demand response for the 2-hour group in 7 out of the 11 trial weeks. We did not test for statistically significant differences in demand response volume amongst different weeks of the trial, but generally observed a declining trend in the magnitude of the average point estimate detected each week, with the average DiD result per treatment participant declining from 0.38 kW in week 1 to -0.11 kW by week 11, though with week 4 as an notable outlier to that trend. There was also one negative DiD result during week 11, which we suspect was due to ongoing high external temperatures.

The trend of declining average demand response point estimates week on week coincided with a general increase in event time average temperatures, rising from 6.2°C in Week 1 to 10.6°C in Week 11. This pattern aligns with findings from trial two, where we detected greater demand response at lower

temperatures[26]. Notably, the Daily Evening Flex mini-trial period was unseasonably warm. While trial two included event weeks with temperatures as low as -1.4°C, no week during Daily Evening Flex fell below an average of 6.2°C during event times27.

We suspect that high external temperatures reduced heating demand across both treatment and control groups from March onwards, thereby diminishing detectable demand response. This is likely to explain the negative demand response volume in week 11, particularly as consecutive weeks of higher temperatures may lead participants to consider the "heating season" to have ended. Close matching of the 2-hour group to the control group means that if both groups were no longer heating their homes during event times, a slight increase in use of electrical devices by either group could have easily skewed the DiD results in either direction.

Overall, these findings suggest that heat pump homes can contribute to daily 2-hour flexibility. However, the value of this flexibility for DNOs is likely to be lower during milder winter conditions, when heating demand is reduced and there is less potential for abatement.

4.2 Weekly demand response for the 4-hour group

Figure 6 shows average treatment participant demand response and average external temperature (4-8pm) per week for the 4-hour group.

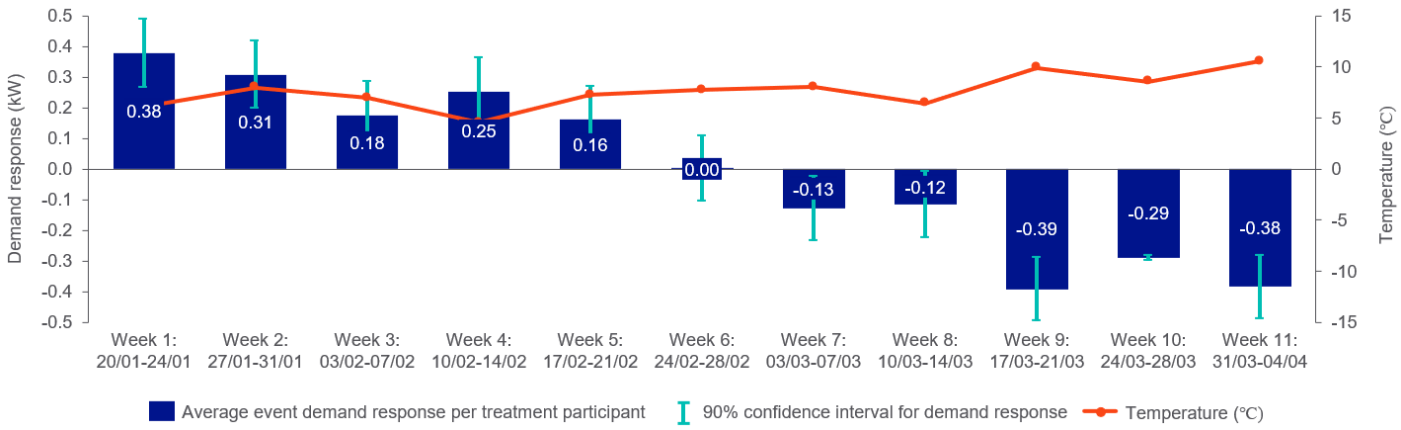

Figure 6: Average week over week demand response per treatment participant for the 4-hour group, and average external temperature during the event time

We detected statistically significant demand response for the 4-hour group in each of the first 5 weeks of the trial. An inconclusive result was detected for week 6, followed by statistically significant negative demand response (an average increase in demand) from weeks 7 to 11. Although we did not test for statistically significant differences in demand response volume amongst different weeks of the trial, we generally

[26] In trial two we observed a trend whereby 0.07 kW greater demand response per treatment participant could be expected per 1°C drop in temperature from the 6.1°C trial average.
[27] According to the Met Office, March temperatures were 1.3°C above the seasonal average, making it the UK's 10th warmest March on record.

observed that the 4-hour group exhibited a similar decline in demand response over time as the 2-hour group, with a more pronounced effect.

We suspect that the negative DiD results in the later weeks were driven by unseasonably warm weather and were exacerbated by the mismatch in energy consumption patterns between the 4-hour treatment group and the control group. As with the 2-hour group, we suspect that warmer conditions from March onwards may have led the 4-hour group to reduce heating use significantly. If the 4-hour group and control group had been perfectly matched in unabated energy consumption, we would have expected a DiD result of 0 kW as the trial progressed. The mismatch between these groups, however, could explain the negative DiD results for week 7 onwards, particularly if the control group had a higher non-heating energy use in this time.

Based on customer feedback in surveys and focus groups and despite these DiD results, we consider it highly likely that the 4-hour group was consistently engaged in the trial for all 11 weeks, and that participants were consistently keeping their heating turned down between 4-8pm.The survey data presented in Section 5.1 examines this customer feedback in more detail.

5. Customer experience results

Section 5 addresses our aim of assessing whether daily evening demand response can be achieved alongside high participant satisfaction and minimal impact on comfort, including for those with potential vulnerabilities. Section 5.1 presents findings from three short check-in surveys sent on 4th February, 28th February, and 4th April. Sections 5.2–5.6 present results from two main surveys: the start-of-trial survey (80% response rate, n=272) and the end-of-trial survey (72% response rate, n=247). These sections cover overall mini-trial satisfaction (5.2), event frequency satisfaction (5.3), experiences by heat pump control method (5.4), event length satisfaction (5.5), and comfort impacts (5.6). Throughout, we compare experiences of 2-hour and 4-hour groups and consider responses from potentially vulnerable participants.

5.1 Participation rate

In the three check-in surveys, we asked treatment participants to self-report their participation in the trial over the previous three weeks28. Their responses are shown in Figure 7. These results show strong and consistent engagement with daily flexibility across both 2-hour and 4-hour groups, with particularly high participation evident in the 4-hour group.

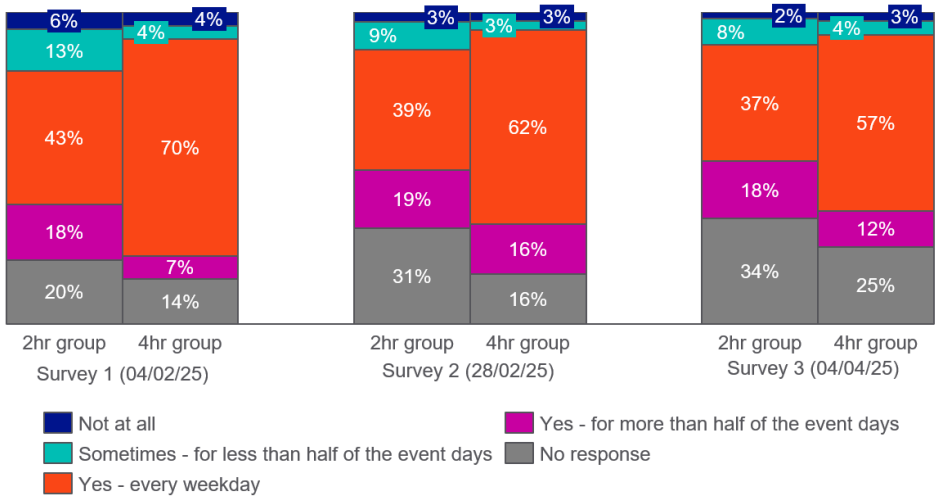

Figure 7: Treatment participants' self reported participation rate in each check in survey, split by 2-hour (n=284) and 4-hour groups (n=69)

Most 2-hour group participants self-reported turning down for some or all the days requested per week throughout the trial. However, there was a general decline in self-reported participation as the trial progressed, which aligns with the temperature increase and decrease demand response volumes seen in Section 4. In Survey 1, 74% of respondents reported turning down, 43% of them every day, with a no response rate of 20%. By Survey 3, this fell to 63% and 37%, and the no response rate rose to 34%.

[28] Participants were asked: "Did you participate in daily flex events in the last three weeks? This could have been by avoiding using your heat pump during events or by using stored energy to run your heat pump during events."

Figure 7 also shows that most 4-hour group participants self-reported turning down for some or all the days requested per week throughout the trial, and self-reported doing so at higher rates than the 2-hour group. Self-reported participation also declined for the 4-hour group as the trial progressed. In Survey 1, 81% of participants reported turning down with 70% of them turning down every day. By Survey 3, 73% of participants reported turning down and 57% did so daily, while the no response rate rose from 14% to 25%. The consistently higher participation rates highlights the strong engagement of the 4-hour group, likely influenced by their self-selection from the Longer Events Flex mini-trial.

These results are highly encouraging, indicating that most participants in both groups were turning down each week. Over a third of the 2-hour group even reported turning down for five days per week, despite only four days being required to qualify for the weekly payment.

The findings are particularly strong for the 4-hour group. More than half reported turning down every day in all three surveys. This consistent engagement highlights the high level of commitment among those willing to participate in longer events. These results also indicate they continued the behaviour of turning down their heat pump, despite the inconclusive findings from the DiD analysis in Section 4.2.

Although self-reported engagement generally declined over the 11-week period, possibly due to participant fatigue, the overall trend remains positive. Additionally, the unseasonably warm weather may have influenced participants' self-reported behaviour, particularly as temperatures rose, since they may have discontinued heating use altogether. Based on the self-reported participation rates, most participants continued to meet or exceed the participation requirements, even under more demanding conditions than in a previous EQUINOX trial.

5.2 Trial satisfaction

Centring customer experience as a key consideration during the trial design phase contributed to high levels of customer satisfaction. Amongst those who completed the end of trial survey, 78% reported that they were satisfied with the trial. This satisfaction rate was similar for both the 2-hour and 4-hour groups, and in alignment with the 79% trial satisfaction rate reported for EQUINOX trial two. This was a positive result given the intensity of the requested behaviour in Daily Evening Flex relative to prior EQUINOX trials. We are encouraged by customer appetite to participate in a range of novel commercial offerings while also being able to maintain high customer satisfaction. However, it is important to point out that the temperatures observed in the trial were unseasonably warm, thus we cannot confirm the response would have been the same if the trial was held during a colder winter.

Event payment satisfaction varied amongst both the 2-hour and 4-hour groups. This was predominantly motivated by whether participants felt as though the event payments received were a fair reflection of their effort to participate, rather than a reflection of any impact of the events on comfort. Although the event incentive in Daily Evening Flex was based on average market rates offered by NGED for similar flexibility products, this does not represent a business-as-usual customer offering, which may incorporate additional flexibility values and may be structured in different ways depending on the preferences of each flexibility service provider. We explore payment satisfaction in more detail in "EQUINOX trial three: Engaging vulnerable customers".

5.3 Event frequency satisfaction

During the Daily Evening Flex trial, participants experienced events on each weekday across 11 weeks. We evaluated participants' views of this frequency[29]. As seen in Figure 8, both 2-hour and 4-hour groups widely viewed daily events were widely viewed positively. 70% of the 2-hour group found daily events either too little or just right. This rose to 84% of the 4-hour group. We observed no meaningful difference in views between potentially vulnerable and non-vulnerable participants. It is important to note that the trial participants were not on tariffs with multiple daytime unit rates and some participants may have intentionally opted out of such tariffs, meaning the outcomes reflect the views of customers who haven't already embraced daily flexibility as part of their routine.

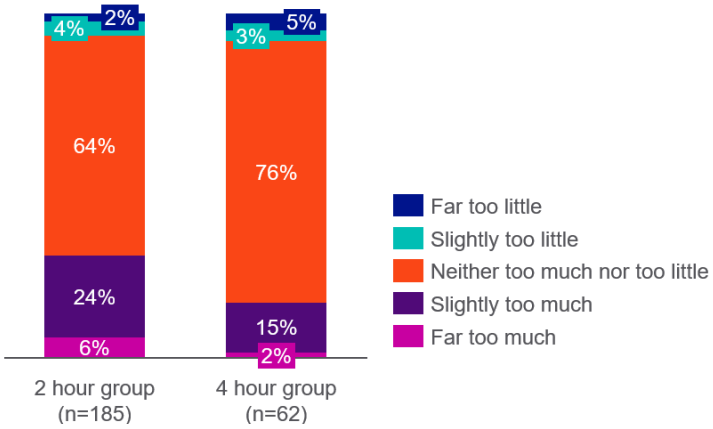

Figure 8: Participants' views on every weekday event frequency, reported in the end of trial survey (n=247)

We also assessed participant views on event frequency in focus groups. Focus group participants were also largely satisfied with the event frequency. Notably, one participant stated: "From our point of view, I think we'd have been happy with it for seven days. In some ways, you are more permanently changing your habits and the way in which you manage your life – to some extent, anyway". Overall, our findings show that daily events were viewed positively by participants, particularly the 4-hour group was particularly engaged in the trial.

5.4 Control type

We also analysed event frequency perspectives based on whether participants needed to be at home in order to control their heat pump. We hypothesised that those that would need to physically adjust their heat pumps settings each day would find the trial more challenging than those who could adjust their heat pump remotely or even set a daily heating schedule that would last across all 11 weeks. We asked participants how they controlled their heat pump during the trial30 and divided them into two groups31:

- **Manual control**: Participants who either turned their heat pump off manually during events or manually adjusted the temperature on their thermostat during events.
- **Remote control**: Participants who either adjusted their home temperature through a mobile app or set a schedule ahead of time.

We segmented participant's views on daily events by this control type, as shown in Figure 9. 83% of remote control participants found events on every weekday just right or too little compared to 63% of manual control participants. When asked whether their control method influenced their ability to participate every weekday

[29] Participants were asked in the end of trial survey: "During the EQUINOX daily flex trial, there was one event per day for 5 consecutive days. What is your view on this frequency?"
[30] Participants were asked: "How did you control your heat pump during the trial?" in the end of trial survey.
[31] 15 participants responded 'Other' and therefore were unable to be placed into a group.

(Figure 10), most (60%) of those using remote control for their heat pump reported that it made participation easier, compared to just 19% of participants with manual control.

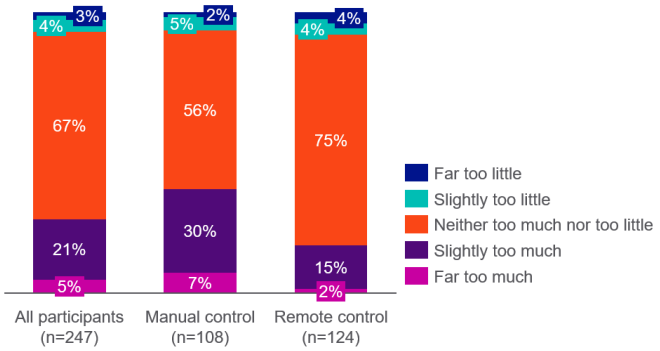

Figure 9: Participants' views on every weekday event frequency segmented by control type, from the end of trial survey

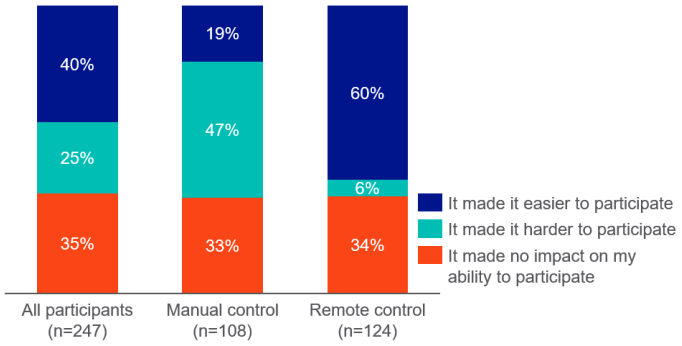

Figure 10: Participants' self-reported impact that their control type had on their ability to participate, segmented by control type , from the end of trial survey (n=247)

These findings are reinforced by insights shared during focus groups. One participant using manual control explained that they could not participate daily due to the nature of their control method: “I don’t mind how often the events are, but I didn’t do them all, because it was so often and because I have to manually [turndown].” In contrast, a participant with remote control noted that the daily frequency enabled them to establish a consistent routine from the start of the trial: “It’s not difficult. You just program the heat pump to do that. It doesn’t matter whether it’s a weekend or a weekday; we’ve just left it the same.” It is noted that participation in each event was entirely voluntary, and customers were encouraged to engage only on the days that suited them.

Overall, these findings indicate that improving accessibility to remote control methods could greatly improve convenience and therefore also increases public willingness to participate in daily flexibility. This represents a relatively simple technical adjustment that could meaningfully increase uptake of flexibility programmes and participation in individual flexibility events.

5.5 Event length satisfaction

At the end of Daily Evening Flex we investigated whether 4-hour group participants were still satisfied with 4-hour event lengths — as they had indicated during Longer Events Flex — or would prefer a 2-hour event duration for further events[32]. Most participants (67%) indicated that they would be neither more nor less

[32] 4-hour group participants were asked: “How likely would you be to participate in everyday events like daily flex trial events if it was 2 hours per day (5-7pm) instead of 4?”

likely to participate if events were 2-hours (Figure 11**: 4-hour group participants' likelihood that they would be willing to participate in sustain events if they were 2 hours in length, from the end of trial survey (n=62)**), suggesting that most participants placed in the 4-hour group were comfortable with the 4-hour duration over the 11-week period. This is an important result, highlighting that 4-hour daily events can be feasible for some customers. Additionally, 30% of participants in the 4-hour group stated that 2-hour event lengths would improve their participation. This emphasises the value of participants being empowered to adjust their participation in flexibility programmes based on their individual circumstances. It also bolsters our findings from the Longer Events Flex trial in indicating that it is best to take a participant led approach to consecutive day flexibility.

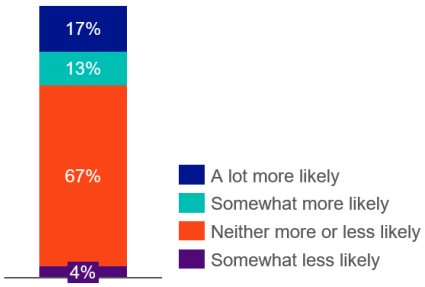

Figure 11: 4-hour group participants' likelihood that they would be willing to participate in sustain events if they were 2 hours in length, from the end of trial survey (n=62)

5.6 Overall Comfort

An important result of the trial was that most participants reported either no change or a minor change in home comfort throughout the trial. Figure 13 shows participant responses on comfort from the end of trial survey33 where 89% of participants reported no change or a slight change in comfort levels. 9% reported the events made their home cool and just 2% (six participants) indicated the events made their home too cold. Of note, 1% of the participants experiencing a slight change indicated that it made their homes warmer. For those who reported a change in comfort, in the end of trial survey and during focus groups, we predominately heard that participants would wear extra layers or use alternative heating to compensate if there were any changes in household comfort levels. These results suggest that everyday heat pump flexibility is generally well tolerated, though comfort levels varied across participants. It is also worth noting that the unseasonably warm winter may have influenced findings.

We asked participants who experienced any change in comfort (49%) how frequently this occurred. Figure 12 shows that 71% felt this changed less than half of the time and only 11% (13 participants) reported that this change occurred every time they participated. Variations in customer comfort are a key consideration for heat flexibility, highlighting the importance of designing programmes with customers in mind from the outset. This includes measures such as clear communication and voluntary participation.

[33] Participants were asked: “Do you think that taking part in the events has impacted the comfort levels in your home?”.

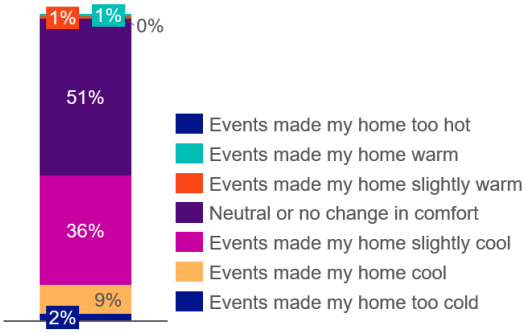

Figure 13: Participants' self-reported impact on comfort levels in the end of trial survey (n=247)

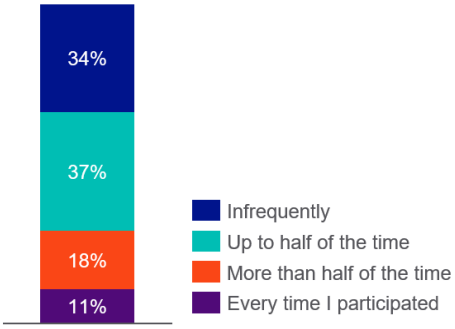

Figure 12: Participants' self reported frequency of change in comfort across the 11 weeks of the trial (n=121)

5.6.2 Comfort by event length group

We also explored whether those participating in the 4-hour group were more likely to feel a change in home comfort than the 2-hour group due to the longer event length. To do so, we segmented responses on comfort, and frequency of changes in comfort by event length.

Figure 14 shows 87% of those who participated in the 2-hour group felt no change or only a slight change in comfort throughout the trial compared to 91% of those in the 4-hour group. Despite the longer event length, the 4-hour group were less likely to feel any change in comfort. This may be attributed to the group being self-selecting, which could suggest that those in the 4-hour group were more likely to have personal circumstances or preferences, such as better home insulation, that made them less sensitive to changes in comfort. While these participants are unlikely to be representative of the wider heat pump population; it is encouraging to see that there are customers for whom 4-hour events are well tolerated. Figure 15 shows both groups experienced a similar frequency of change in comfort.

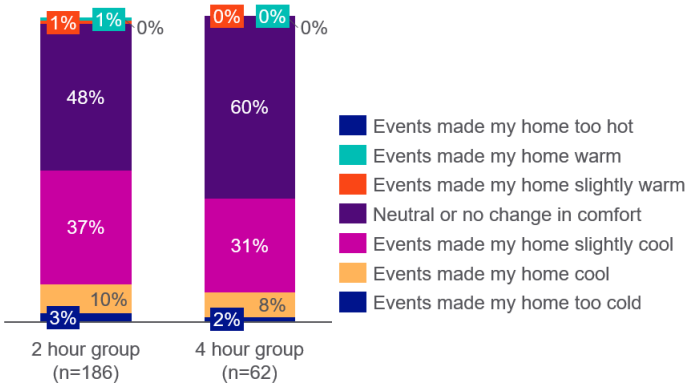

Figure 14: Participants' self-reported impact on comfort levels in the end of trial survey, segmented by event length group (n=248)

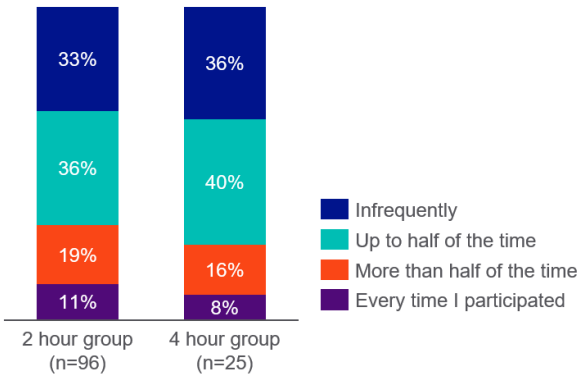

Figure 15: Participants' self reported frequency of change in comfort across the 11 weeks of the trial, segmented by event length group (n=121)

Overall, these findings suggest that event length was not associated with a greater level of reported change in home comfort. Rather, it may have been up to the individual participants making up both groups, and their own individual experiences of flexibility which may have been associated with a variety of factors beyond event length.

5.6.3 Overall comfort of potentially vulnerable participants

We additionally segmented participant responses on comfort, and frequency of change in comfort to check for differences in responses according to whether respondents were categorised as potentially vulnerable or non-vulnerable.

Figure 1617:34 shows that potentially vulnerable participants were somewhat more likely than non-vulnerable participants to report a change in home comfort during events. While 84% of potentially vulnerable participants reported no change or a slight change in comfort, 11% (12 participants) reported that their homes felt cool, and 5% (six participants) indicated their homes felt too cold. Of note, 2% of the participants experiencing a slight change indicated that it made their homes warmer. In comparison, 91% of non-vulnerable participants reported no change or a slight change in comfort, with fewer reporting their home felt cool (9%, 11 participants), and none reporting it felt too cold.

Figure 1716 shows that potentially vulnerable participants reported experiencing these changes with more frequency. 30% felt them infrequently, while 36% felt them up to half the time, and 34% reported feeling them more than half the time or every time. Among non-vulnerable participants, 41% felt changes infrequently, and only 19% experienced them more than half the time or every time.

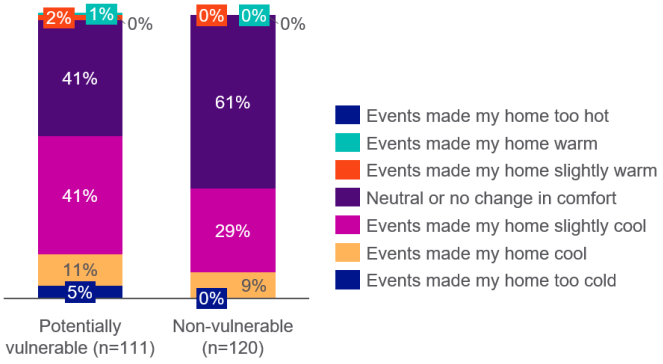

Figure 16: Participant's self reported comfort levels in the end of trial survey, segmented by vulnerability (n=231)

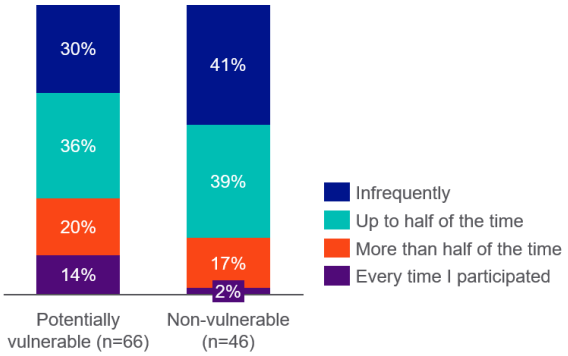

Figure 17: Participants' self reported frequency of change in comfort across the 11 weeks of the trial, segmented by vulnerability (n=112)

These results show that 84% of potentially vulnerable participants reported no change or a slight change in comfort, relative to 91% of non-vulnerable participants. This finding highlights that, with the appropriate

[34] Figures 16 and 17 show responses from fewer total participants than Figures 14 and 15 because segmentation relied on respondents having completed the start of trial survey. This then allowed us to categorise them as potentially vulnerable or non-vulnerable.

considerations, potentially vulnerable customers are interested and can be equitably engaged in heat flexibility offerings. This reinforces that it is critical that flexibility programs have all customers in mind from the outset.

Overall, the results present a nuanced picture of comfort impacts for potentially vulnerable participants. 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 and clear communication to ensure customers are empowered to take part in flexibility programmes.

6. Stacking results

Section 6 presents our findings our aim of assessing the customer experience of stacking daily EQUINOX trial events with ad-hoc NESO DFS events. Section 6.1 presents findings on the amount of additional demand response detected, and Section 6.2 presents findings on customer views of taking part in the DFS events alongside Daily Evening Flex.

6.1 Delivered demand response

As noted in Section 3.4, we excluded from the analysis any DFS events that occurred during the first two weeks of the Daily Evening Flex trial. This ensured that the DFS baseline was based on each participant's recent average consumption, which already reflected heat pump demand response from EQUINOX events.

During the latter nine weeks of the mini-trial, participants who chose to take part in their energy supplier's DFS product took part in seven DFS events. These events ran at various times between 5–7 pm, overlapping with the Daily Evening Flex turndown period.

On average, 71% of treatment group participants opted into each DFS event, achieving an average additional 0.66 kWh of turndown per event per treatment participant. This likely came from reducing the usage of other devices alongside their heat pumps. Across the group, participants earned a total of £96 for their collective extra effort, showing that EQUINOX participants responded to both programmes and increased their turndown when asked to do so.

6.2 Customer experience

We asked both treatment and control group participants in the end of trial survey if they participated in DFS events[35]. 66% of treatment group participant respondents and 68% of control group participant respondents self-reported participating in DFS events during the Daily Evening Flex trial, as shown in Figure 18.

[35] Participants were asked in the end of trial survey: "Did you participate in Octoplus Saving Sessions events during the EQUINOX daily flex trial?"

We asked the self-reported DFS participants what changes they made to be able to participate[36]. As shown in Figure 19, 72% of the control group reported changing how they used other electric appliances, falling to 43% for the treatment group. 39% of the control group also reported using their heat pump or hot water heating due to DFS events, compared to 51% of the treatment group. This discrepancy may be a result of customers in the treatment group being more comfortable with adjusting their heating schedules, as a result of having participated in heat pump flexibility events in Longer Events Flex and in Daily Evening Flex, whereas the control participants had not.

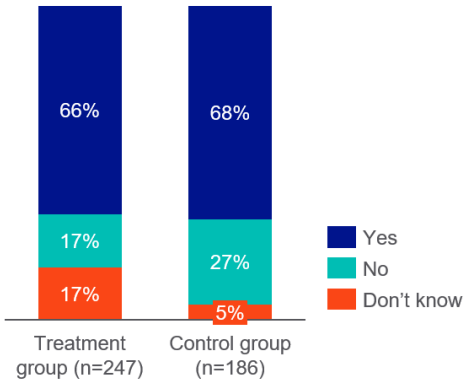

Figure 18: Participants' self-reported participation rate in DFS events in the end of trial

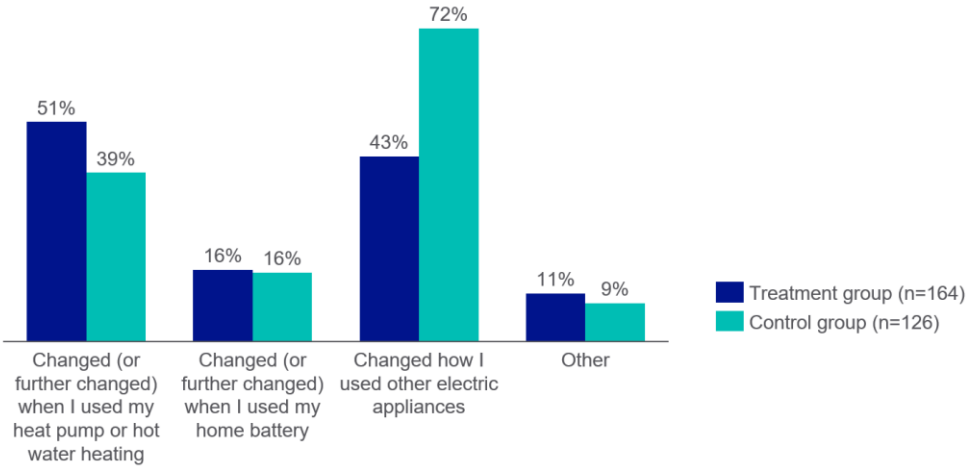

Figure 19: Participants who took part in DFS event's self-reported changes they made during the events in the end of trial survey (n=290)

Overall, treatment participants reported that they were satisfied with participating in both EQUINOX and DFS events. Amongst treatment participants who self-reported participating in these DFS events, 78% (128 out of 164) reported they would recommend the arrangement to others[37]. This suggests that customers who are able to engage in multiple flexibility offerings simultaneously find the experience positive and worthwhile. This is an encouraging result, given that stacking offers the potential to unlock greater rewards for customers while also maximising the engagement of flexible assets to support DNOs and NESO.

[36] Participants were asked in the end of trial survey: "What changes did you make to be able to participate in Octoplus Saving Sessions during the EQUINOX daily flex trial?"
[37] Treatment participants were asked in the end of trial survey: "Would you recommend participating in Octoplus Saving Sessions events alongside EQUINOX events to others?"

7. Summary

The results of the Daily Evening Flex mini-trial indicate strong potential for heat pump homes to participate in daily evening peak flexibility events and demonstrate a positive customer experience of participating in stacked events with NESO's DFS programme. The main takeaways are summarised below:

1. **The 2-hour group participants provided statistically significant average demand response in 7 out of the 11 trial weeks, primarily in the first half of the mini-trial.** Our DiD analysis detected statistically significant demand response primarily in the first half of the mini-trial, across weeks 1-6, and then again in week 10 – with an overall downtrend from 0.38 kW turndown per treatment participant in week 1 to -0.11 kW by week 11. This trend was not explained by participation rates, with more than half of participants continuing to self-report turning down by the final of three check-in surveys. We suspect lower performance in the final weeks was a result of early 2025 being unseasonably warm, with temperatures exceeding 10℃ by the end of the trial. Based on these findings, it's likely that 2-hour participants were successfully carrying out the requested behaviours. The limited impact observed as the trial progressed was likely due to falling heating demand.

2. **The 4-hour group participants provided statistically significant demand response in the first half of the trial**. Our DiD analysis detected statistically significant demand response for trial weeks 1-5, and then a statistically significant increase in demand in the latter half of the trial. This trend cannot be explained by falling participation, with higher self-reported participation for the 4-hour group compared to the 2-hour group. We suspect the 4-hour group's results in the final weeks occurred due to the same high temperatures affecting the 2-hour group, coupled with a general mismatch of unabated consumption of the 4-hour group with the control group, which became more pronounced as both groups progressively heated less due to warmer weather. Based on these findings, it is likely that 4-hour participants were successfully carrying out the requested behaviours and that this may have been detected under different trial conditions.

3. **Overall satisfaction was high, indicating customer willingness to participate in daily flexibility.** 78% of participants reported they were satisfied with the trial – at similar rates across both the 2-hour and 4-hour groups. This result also aligned with the satisfaction rate reported for EQUINOX trial two – which is particularly encouraging given the increased intensity of the requested turndown behaviour in Daily Evening Flex. However, it is important to note that the temperatures observed during the mini-trial were unseasonably warm and we cannot confirm that satisfaction would have been as high if it was held during a colder winter.

4. **Event frequency satisfaction was high, further indicating customer willingness to participate in daily flexibility.** 70% of the 2-hour group reported they were satisfied with the event frequency in the Daily Evening Flex mini-trial, as did 84% of the 4-hour group. This suggests daily flexibility with both event length durations can be suitable for different types of customers.

5. **Control method was highly correlated with event frequency satisfaction**. 83% of participants who could control their heat pump remotely (e.g. by phone app) approved of the daily frequency of events, falling to 63% amongst participants who had to control their heat pump manually. Additionally, 60% of those with remote control reported their control method made participation easier, compared to just 19% with manual control. These findings indicate that improving accessibility to remote control methods could greatly improve rates of public willingness to participate in daily flexibility.

6. **Participants generally maintained comfort during the events, with minimal disruption reported.** 89% of participants noticed no change or a slight change in comfort, while only 2% ever found it too cold. Few participants (5%) reported that they experienced a change in comfort in every event. These results suggest daily flexibility can be achieved with minimal impact on comfort, though we acknowledge that the winter season was unusually warm.

7. **Comfort for potentially vulnerable participants was in line with non-vulnerable participants.** While 91% of non-vulnerable participants reported no or only a slight change in comfort, only slightly fewer (84%) potentially vulnerable participants reported the same. These results present a nuanced picture of comfort impacts for potentially vulnerable participants. They show that potentially vulnerable customers are interested and can be equitably engaged in heat flexibility offerings.

8. **Self-reported change in comfort varied minimally between the 2-hour and 4-hour event length groups.** The proportion of participants indicating they felt any change in comfort, including its magnitude and frequency, barely varied between the 2-hour and 4-hour groups. This suggests experiences of comfort were not correlated with event length but rather due to experiences of individuals making up both groups.

9. **Participants successfully stacked Daily Evening Flex with events in NESO's DFS programme, unlocking greater value for trial participants.** 71% of participants self-reported taking part in DFS events concurrent with the timing of their heat pump turndown events in Daily Evening Flex. By stacking these events and engaging in general turndown in addition to heat pump turndown, these participants collectively earned an additional £96. Customer satisfaction with the experience was high, with 78% of the treatment participants involved reporting that they would recommend the arrangement to others.

In conclusion, the results of the Daily Evening Flex mini-trial demonstrate a strong customer willingness for participation in daily heat pump flexibility events, such as those that could be commercially procured for DNO constraint management flexibility products. Most participants are likely able to do so for just two hours per evening, though a sizable minority can do so for up to four hours per evening. For DNOs, these findings provide encouraging evidence that leveraging heat pump flexibility in this format is achievable. For Flexibility Service Providers (FSPs), these findings provide insights into how to effectively engage customers with heat pumps in flexibility markets. Daily Evening Flex also demonstrated high satisfaction in participation in stacked events, showing the customer appetite to be able to participate in multiple services.

8. Appendix A: Project partners

EQUINOX is led by NGED, 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[38]	
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.

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

9. Appendix B: Difference in Difference approach

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

Equation 1Equation 1 below provides the DiD specification used to calculate an estimate of demand response for each event. Equation 2 details the calculation of uncertainty (standard error). For the trial average analysis, the load profiles run through Equation 1 were representative averages of all the individual load profiles for each day of the trial. For the week-by-week analysis the load profiles run through Equation 1 were a representative average of all the individual load profiles for each weekday in that week. Equation 2 Equation 3. RMS standard error calculationEquation 3details 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[39]

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

Equation 2: DiD standard error calculation[40]

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

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 1Equation 1.

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

[39] 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.

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

Equation 3. RMS standard error calculation[41]

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

[41] 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