Virtual Energy System

Powered by National Grid ESO

CrowdFlex: Alpha

January 2023

D7.1 - Trial Specification and Delivery Approach

Strategic Innovation Fund

Authors:

Centre for Net Zero

Powered by Octopus Energy

ESO

1. Purpose of this document

The trial protocol (TP) completed in Beta will describe in detail what we plan to do in CrowdFlex and how we will do it. Researchers write trial protocols for three main reasons:

1. **To avoid future problems.** Writing each section of the trial protocol forces researchers and delivery partners to consider things that could go wrong in the trial.
2. **To get everyone on the same page.** CrowdFlex will last for approximately three years, and different team members may come and go during that time. Trial protocols serve as a single authoritative document that people can refer back to, thus facilitating smoother hand-offs and greater continuity.
3. **To ensure our analysis is robust.** TPs pre-specify our analysis**,** saying in advance exactly how we will analyse the data**.** Social science has experienced a 'replication crisis' in recent years with many findings failing to replicate. Part of the cause of the crisis is post hoc analysis of datasets and testing many hypotheses on the same set of data. For this reason, a growing number of academic journals now encourage pre-registered analysis plans to safe-guard against pressures to 'fish' for results later. Trial protocols facilitate this pre-registration.

This document is a trial protocol outline. Here, we scope the key components that will need to be written and pre-specified in advance of CrowdFlex. The trial protocol will be finalised in the Beta phase.

2. Contents

3. One-page PICOS statement

	Utilisation payment trial	Availability payment trial
What is the problem?	Encouraging residential energy customers to turn-down or up their energy consumption during flexibility response events, as well as making their appliances available to flexibility service providers to remotely control for flexibility both inside *and* outside of specific events.	
Population	Customers who (1) have a smart meter and (2) opt in to participate in CrowdFlex.	Customers who opt in to participate in CrowdFlex. Specifically, those who own assets that can be controlled by the flexibility service provider for the availability payment trial, in particular a controllable heat pump or electric vehicle charger.
Intervention	Differing financial incentives for flexibility (turn-down or turn-up) during events. Note: other interventions are in scope for other mini-trials within CrowdFlex.	Differing financial incentives for suppliers (or other parties) to remotely automate customer assets (such as EV chargers or heat pumps) to provide flexibility services.
Comparison	This sub-trial will involve a Randomised Control Trial (RCT). Randomisation will involve two components: • Wait-list randomisation: We will waitlist some customers who opt into the CrowdFlex programme, leaving them as a 'control' group until the final events. • Individual-level randomisation within a mini-trial. The comparison condition thus comes from the waitlist: not receiving the intervention during all events except the final ones.	This sub-trial will involve a Randomised Control Trial (RCT). Randomisation will involve customer-level randomisation.
Outcome(s)	**Primary outcome:** Half-hourly energy consumption during the event window (from smart meter data). **Secondary outcome:** Energy consumption 24 hours before and after the event window (from smart meter data). **Exploratory outcomes:** Survey responses: customer experience, comfort, understanding of the flexibility event logistics, understanding of the purpose of flexibility, views on and satisfaction with event logistics, and behaviours customers report employing to shift their demand.	**Primary outcome:** Whether customers opt in to allow flexibility service providers to control their appliance AND being available (e.g., having the charger plugged in during a pre-specified timeframe)
Setting	The trial will take place in Great Britain (Wales, Scotland, and England).	

4. Recommendations and summary

Section 5 explains the policy area and trial rationale, Section 6 outlines roles and responsibilities, and Section 7 outlines the research aims. We recommend a bifurcated approach to CrowdFlex, with a trial focusing on availability payments for remote control of EV chargers and heat pumps and, separately, a trial on utilisation payments for flexibility events.

Section 8 recommends the interventions we propose testing in the first iteration of the trial:

	Group	Availability payments	Utilisation payments - turn-down	Utilisation payments - turn-up
Incentive (based on CrowdFlex Deliverable D7.2)	Low	£2 - £4 / asset / month	£0.05 / kWh	
	Medium	£6 - £8 / asset / month	£0.40 / kWh	
	High	£12 - £14 / asset / month	£1.25 / kWh	
	Control	No incentive, but invited[1]	No communications at all	

We suggest an iterative and learning approach, where we will alter the treatments in later phases of CrowdFlex.

Section 9 recommends a series of two availability payments trials and three mini-trials for utilisation payments. We propose mini-trials of 60 turn-down and 20 turn-up events.

Section 10, Section 11 and Section 12 outline the trial procedure. The primary outcome measure for the availability payment trials will be the proportion of customers that make their assets available to be remotely controlled. The primary outcome measure for the utilisation payment trials will be half-hourly consumption during events.

Section 13 recommends sample sizes to detect meaningful effects while balancing budget considerations:

1. **Availability payment trials**
   - At least 2,500 customers in each incentive variation.
   - Two trials, each of 12 months duration.

   Under these conditions, we would be confident in our ability to detect a 2.5 percentage point difference (for each trial) between the lowest incentive group and customers who do not receive any incentive (against a baseline of 10%). We estimate the incentives to then cost between **£135,000 and £270,000** per trial, and between **£270,000 and £540,000 across the two availability trials**.
2. **Utilisation payment trials**
   - 60 turn-down events and 20 turn-up events per mini-trial.
   - 4 batches of events: 3 mini-trials + one 'final' batch of events that is more open to non-consortium flexibility service providers.
   - 8,200 customers per treatment arm in each turn-down event.
   - 4,200 customers/treatment arm in each turn-up event.

   Under these conditions, we would be confident in our ability to detect a 0.56% difference in consumption between treatment and control for turn-down events, and a 1.42% difference in consumption between treatment and control for turn-up events.

[1] We may also include a "no invitation" group as a backup.

We estimate the incentives to then cost between **£416,640 and £762,720** per mini-trial, and between **£1,666,560 and £3,050,880 across the four batches of events**.

Combining these costs (availability payments, turn-down utilisation payments across three mini-trials plus a fourth batch of events, and turn-up utilisation payments across three mini-trials plus a fourth batch of events), we estimate **the total cost for incentives to be £1,936,560 to £3,590,880**.

Section 14 outlines our analytical approach, Section 15 discusses ethical issues, and Section 16 discusses risks and mitigations.

5. Policy area, trial rationale & challenges

5.1. Social impact objective

As more electricity is generated by intermittent renewable generation, the electricity grid needs to become more responsive to possible changes in both demand and supply to avoid, to the extent possible, electricity curtailment related to the variability of renewable sources.[2,3] This responsiveness ensures the grid can maximise the benefits of renewables and reduce carbon emissions by displacing gas and coal generation.[4]

Traditionally, generation has been controlled to 'match' demand, but with increasing renewable penetration and increasing electricity demand, other ways are being considered to avoid electricity curtailment. 'Intelligent demand' involves embedding the concept of domestic demand flexibility into the energy market, i.e., moving demand to periods when there is plenty of cheap, renewable electricity generation.[5] This involves assets like electric vehicles (EVs), heat pumps (HPs) and potentially other home appliances adjusting their demand profiles.

By harnessing flexibility in all parts of the power system, both generation *and* demand, the UK can more cost-effectively transform its energy system towards a cleaner one dominated by renewable sources. The introduction of flexibility to fully utilise renewable generation could potentially save as much as £17-40 billion cumulative to 2050 compared to a business as usual (BAU) scenario, where flexibility is not considered.[6]

5.2. The challenge

As noted in CrowdFlex deliverable D2.1, different methods for influencing patterns of electricity consumption have been employed for several decades, such as the introduction of the UK Economy 7 tariff in 1987.[7] [8] [9]

These methods are helpful at shifting demand. However, CrowdFlex investigates novel services that can exploit domestic flexibility using existing technologies and seeks to help

[2] L. V. Villamor, V. Avagyan, and H. Chalmers, 'Opportunities for reducing curtailment of wind energy in the future electricity systems: Insights from modelling analysis of Great Britain', Energy, vol. 195, p. 116777, Mar. 2020, doi: 10.1016/j.energy.2019.116777.
[3] A. Fragaki, T. Markvart, and G. Laskos, 'All UK electricity supplied by wind and photovoltaics – The 30–30 rule', Energy, vol. 169, pp. 228–237, Feb. 2019, doi: 10.1016/j.energy.2018.11.151.
[4] F. Teng, M. Aunedi, and G. Strbac, 'Benefits of flexibility from smart electrified transportation and heating in the future UK electricity system', Applied Energy, vol. 167, pp. 420–431, Apr. 2016, doi: 10.1016/j.apenergy.2015.10.028.
[5] Moving from 'Demand Side Response' to 'Intelligent Demand'. Octopus Energy. https://octopus.energy/blog/intelligent-demand/ (accessed Sep. 1, 2022)
[6] An analysis of electricity system flexibility for Great Britain. GOV.UK. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/568982/An_analysis_of_electricity_flexibility_for_Great_Britain.pdf (accessed Aug. 31, 2022)
[7] J. Crawley, C. Johnson, P. Calver, and M. Fell, 'Demand response beyond the numbers: A critical reappraisal of flexibility in two United Kingdom field trials', Energy Research & Social Science, vol. 75, p. 102032, May 2021, doi: 10.1016/j.erss.2021.102032.
[8] The Electricity Council, 'Electricity Supply in the UK: A chronology'. Fourth edition, London, UK, 1987
[9] W. L. Kidd, 'Development, design and use of ripple control', Proceedings of the Institution of Electrical Engineers, vol. 122, no. 10R, pp. 993–1008, Oct. 1975, doi: 10.1049/piee.1975.0260.

industry participants and policymakers better understand how to best incentivise and reward domestic assets that provide this flexibility.[10]

In addition, most of the consumer-side flexibility from static time-of-use (ToU) tariffs, such as Economy 7, provides what National Grid Electricity System Operator (ESO) calls ‘routine’ flexibility. This flexibility reduces peak consumption on a constant, ongoing basis, but it is not useful for acute supply shortfalls or excesses relative to demand. Traditionally, ESO has relied on non-domestic solutions for this ‘response’ (acute) flexibility, such as power plants that can turn generation up or on at short notice or industrial manufacturers that can turn production up or down at short notice. A core focus of CrowdFlex is investigating the extent to which domestic consumers can provide response flexibility too.

5.3. Barriers

5.3.1. Consumer side

Flexibility comes from ‘manual’ or ‘automated’ dispatch. Manual dispatch includes manually setting delays on electrical appliances such as dishwashers, or plugging one’s electric vehicle into a charger at a different time than one usually does. Automatic dispatch involves automatic control of appliances (such as a dishwasher or EV charger) to turn on at different times.

Consumer side barriers include:

- Consumer disengagement with energy industry.
- Inconvenience and hassles associated with manual dispatch.
- Discomfort (from some consumers) with automated dispatch (i.e., due to a perceived loss of autonomy).
- Difficulties manually changing or automating flexibility from appliances that customers rely on (especially where customers use an appliance at unpredictable times and cadences).

5.3.2. Technical side

In addition to the aforementioned consumer side barriers, there are technical barriers to overcome. They include managing the data flows between ESO, suppliers (and other flexibility service providers), and individual households – all of which must work smoothly to scale up flexibility from many individual households; the technical limitations to the response assets can provide, both in terms of kW and duration of response; and the need to aggregate domestic participants’ flexibility response to ensure that market participants can use their flexibility. [This section will be expanded in beta; also, see CrowdFlex Deliverable D5.1 for further details.]

5.4. How to operationalise domestic response flexibility

A separate piece of work as part of this project, conducted by Element Energy in collaboration with ESO (CrowdFlex deliverable D4.1), examined how ESO might ‘operationalise’ domestic response flexibility once it is proven and reliable enough to rely on as a ‘business-as-usual’ source of flexibility. We show here a summary of the ‘flexibility

[10] ‘CrowdFlex Discovery Phase’. National Grid ESO. https://www.nationalgrideso.com/future-energy/virtual-energy-system/crowdflex (accessed Aug. 31, 2022)

services' that Element Energy and ESO believe that consumers may (one day) provide, as noted in CrowdFlex deliverable D4.1.

Table 1: Summary of the types of 'services' in scope for CrowdFlex (adapted from CrowdFlex deliverable D4.1)

<table>
<tr><th>Trial Service</th><th>System Operator</th><th>Local / National Response</th><th>System Need</th><th>Dispatch Notice & Duration</th><th>Response type</th></tr>
<tr><td>1. Balancing Mechanism</td><td>ESO</td><td>National</td><td>ESO Frequency</td><td><1-hour
~30 mins</td><td>Automated turn-up / turn-down</td></tr>
<tr><td colspan="6">Consumer proposition: Short-timescale request from ESO for flexibility exclusively from automated appliances like smart EV chargers and heat pumps. Make your EV / HP available for control of for x hours and receive £y availability payment.</td></tr>
<tr><td>2. Thermal Constraint Management</td><td>ESO</td><td>Local</td><td>ESO Thermal</td><td>Day Ahead
<4 hours</td><td>Manual & automated turn-up / down</td></tr>
<tr><td colspan="6">Consumer proposition: Provide demand turn-up/down on either side of constraint by increasing / reducing consumption during notified time period (vs baseline as calculated by flexibility service provider) and receive £y / kWh.</td></tr>
<tr><td>3. Demand Flexibility Service[11]</td><td>ESO</td><td>National</td><td>ESO Adequacy</td><td>Day-Ahead
1-2 hours</td><td>Manual & automated turn-up / down</td></tr>
<tr><td colspan="6">Consumer proposition: The nature of this service means it will always be more beneficial for consumers to participate in than other services. The aim of trialling will be to see how other flexibility (and DSO-level services) can operate in concert with an evolved DFS.</td></tr>
<tr><td>4. Sustain-H</td><td>DSO</td><td>Local</td><td>DSO Constraint Management</td><td>Contract Stage
4 hours</td><td>Manual & automated turn-down</td></tr>
<tr><td colspan="6">Consumer proposition: Sign-up service for consumers to reduce consumption during a notified time period (vs baseline) and receive £y / kWh utilisation payment – this is like Service 2, but only certain areas of the high and low voltage networks experience transmission and/or distribution constraints, likely on a repeated basis, but over a short period of time. CrowdFlex will investigate primacy rules to prioritise DSO services and how households can still participate in other flexibility events.</td></tr>
<tr><td>5. LMA Secure</td><td>DSO</td><td>Local</td><td>DSO Constraint Management</td><td>Day Ahead
24 hours</td><td>Manual & automated demand diversification</td></tr>
<tr><td colspan="6">Some distribution system operators (DSOs) have the ability to control households' electric (storage) heating to avert destabilising spikes in demand if electric heating turns on across many households in an area at the same time. DSOs mitigate this problem using constraint management zones, where they utilise radio tele-switching (RTS) to introduce diversity in turn-on time. It has been agreed that, as part of moving to Smart meters, this capability will be withdrawn; however, as households modernise their heating, DSOs must maintain the ability to automatically introduce this diversity. They are interested in a trial with just heating to reintroduce diversity provided via suppliers, rather than directly with consumers.</td></tr>
</table>

[11] Our proposal is that CrowdFlex does not attempt to design a service that provides demand turn-down during system stress events, as the Demand Flexibility Service does, but stack the other CrowdFlex trial services alongside the output of the DFS in Winter 23/24.

Understanding customers' price sensitivity to incentives to provide flexibility is a question that cuts across all five of these 'trial services' – as noted in CrowdFlex deliverable D6.2, it is of critical importance for predicting and reducing the uncertainty around domestic demand flexibility.

There are two types of prices that are relevant to these trial services:

1. **Utilisation payments** – payments to households per kWh for flexibility (turn-down or turn-up compared to an industry-agreed calculation of the customer's counterfactual consumption (their 'baseline')) during **flexibility events**. These are relevant to services 2, 3, 4 and potentially 5.
2. **Availability payments** – payments to households with assets such as electric heating systems or EV chargers to allow third parties (such as energy suppliers, other flexibility service providers, and/or DSOs or ESO directly) to operate them. This is relevant to service 1, and potentially 5. (EV chargers will likely be more prominent in (1), and heat pumps and other electric heating systems in (5), but our understanding is that both services will involve both types of assets.)

There is thus a clear bifurcation in the research questions of interest to CrowdFlex. **This trial protocol outline reflects this bifurcation, listing separate research questions, analytical strategies, and power calculations for: utilisation payments, on the one hand; and availability payments, on the other.**

6. Roles and responsibilities in this trial

Table 2: High-level roles and responsibilities

Who is responsible for…	Organisation name	Person responsible
Data sharing agreements	ESO	
Collecting outcome data	Energy supplier(s)	
Randomisation	CNZ, in participation with Flexibility service providers	
Delivering the intervention	Flexibility service providers – energy supplier(s), aggregator(s) and technology/asset provider(s)	
Analysis	CNZ, in participation with flexibility service provider(s), ESO, Element Energy	

7. Research aims, question and hypotheses

7.1. Research questions regarding availability payments

7.1.1. Technical research questions

Are these events technically feasible with our current technology? If not, what additional technology is required, and how is the industry addressing these needs? [This section will be expanded in beta.]

7.1.2. Social research questions

Primary: How much do the treatments – specifically, different levels of availability payments (in £ per asset per month) cause domestic customers to allow their assets to be controlled remotely for the purpose of allowing flexibility service providers[12] to use them to participate in the Balancing Mechanism (and/or diversity in domestic appliances in Load Managed Areas[13]) by providing flexibility on a seconds- or minutes-ahead basis? Concretely, how does the availability rate differ by treatment arm?

Note that we assume no *social* research questions about the actual performance (i.e., revenue earned, kWh balanced, etc.) of assets participating in the balancing mechanism. Instead, the research questions focus on participation only. Note, though, that: 1) payments will be scaled to realistic levels, 2) flexibility service providers may classify customers as ineligible for payment unless certain conditions are met, and 3) performance will be examined as a 'technical' research question (i.e., looking at the capacity of response throughout the year and how it varies based on external factors (e.g., as we approach system peak).

Exploratory: We will also examine how the primary outcome above varies by subgroup, e.g., tariff type, household size and type, location in Great Britain, automatic vs manual dispatch, etc.

7.2. Research questions regarding utilisation payments (for flexibility events)

7.2.1. Technical research questions

Are these events technically feasible with our current technology? If not, what additional technology is required, and how is the industry addressing these needs? What would be needed to deliver the services effectively, particularly at a consumer engagement level?

Also in scope: how much response is generated, whether repeated calls for the service reduces the response (as the person or the asset gets exhausted), how the response varies through the year, and the level of response at/close to the system peaks.

[This section will be expanded in beta.]

---

[12] Flexibility service providers include suppliers, but also remote managers of EV charging systems (such as Ohme), and perhaps remote managers of heat pumps and other relevant assets.

[13] To avoid repetition, we will usually simply reference entry into the Balancing Mechanism, but note that availability payments are *always* treated as being for entry into the Balancing Mechanism, provision of diversity in domestic appliance consumption in Load Managed Areas, or both.

7.2.2. Social research questions

Primary: How much do the treatments – specifically, different incentives (£ per kWh flexibility) – cause domestic customers to turn-down or up their energy consumption during event windows?

- For clarity, the policy-relevant outcome measure we will use is *half-hourly* consumption during event windows – we believe this is a better measure than aggregate consumption during the event because it allows comparisons of flexibility per half hour between events of different duration.
- We note turn-down *or up* because some events will involve encouragement to turn-up electricity consumption to 'soak up' excess electricity supply at specific times and/or in specific areas of the country. Sometimes, events will involve certain regions turning up and other regions turning down. We generally recommend treating turn-up as distinct in analysis from turn-down, even if both happen (in different regions) in the same event because we expect responses to be asymmetric and to have qualitatively different underlying behavioural mechanisms to achieve the flexibility.

Currently, the priority is to understand the causal effect of monetary incentives. However, other treatments are in scope for later mini-trials in CrowdFlex. These include varying notice period (hours announcement before event commences), frequency of events, opt-in vs opt-out journey to participate in events, and randomised incentives for technologies that automate flexibility response.

Secondary: How much does being treated (as opposed to still being on the waitlist) cause domestic customers to turn-down or up their energy consumption *outside* event windows, e.g., in other hours during the same day as the event, and/or in days leading up to or following the event? Note that this spillover may be positive (where turn-down begets more turn-down, perhaps due to learning or habit changes) or negative (where turn-down during a window begets turn-up outside the window, as people *shift* their consumption around, potentially causing a secondary peak).

- For clarity, the policy-relevant outcome measure we will use is *half-hourly* consumption in the 24 hours before and the 24 hours after events.

Exploratory:

- We will examine how the above outcomes (consumption during flexibility window; and consumption outside the flexibility window) vary by event parameters which will differ across events (but not randomly). These include: day of week,[14] season of year, days since the most recent event, and a marker of frequency (e.g., number of events that week).
- We will also examine how the above outcomes vary by subgroup, e.g., tariff type, household size and type, location in Great Britain, automatic vs manual dispatch, etc.
- We will also survey customers in the treatment group about their satisfaction, comfort, and other sentiments. We will analyse survey responses, examining how they differ by subgroup.

[14] Note that events will occur at times dictated mostly by ESO, at times that mimic when the grid will have excess supply or demand. Thus these timings are non-random.

- What are the longer-term effects of incentives on demand flexibility? For example:
  - Do customers exhibit fatigue?
  - Do customers form habits?
  - Are there carryover effects across different levels of incentives (e.g., anchoring)?
- What are the behavioural mechanisms of behavioural change? (e.g., manual dispatch, automatic dispatch, technology adoption)
- Does the treatment cause backfiring, as indicated by certain survey responses? (For example, does demand response lead to self-rationing or customers who otherwise regret actions they took to achieve flexibility?)

8. Intervention(s) being tested in this trial

8.1. Intervention for availability payments

What is the intervention: Different levels of availability payments (see table below) to customers to allow flexibility service providers to use their asset portfolio to earn revenue through the Balancing Mechanism (and/or through provision of diversity in domestic appliance consumption in Load Managed Areas). To be eligible, customers who opted in are required to make their asset (typically a heat pump / electric vehicle in CrowdFlex) available for control by the flexibility service provider for at least [threshold to be agreed, but likely 4 days per week].

What resources are needed to deliver the intervention: Emails or other outreach to customers informing them of the opportunity to obtain availability payments to allow their assets to be remotely controlled to earn Balancing Mechanism revenue for the flexibility service provider. (The flexibility service provider must also have the ability to track and verify availability, but the analyst does not necessarily need this capability – for the purposes of the analyses we pre-specify in this trial protocol, the trial analyst will only need to know from the flexibility service provider whether the customer was available – as defined and verified by the flexibility service provider – or not.)

What are the practical steps taken for intervention delivery? We (or, specifically, flexibility service providers involved in CrowdFlex) need to communicate the opportunity to earn the availability payments to participants, and then control their assets during the highest-Balancing-Mechanism-value half hours.

Who will deliver the intervention? Flexibility service providers involved in CrowdFlex – suppliers, remote managers of EV charging systems (such as Ohme), and perhaps remote managers of heat pumps and other relevant assets.

How long will the intervention last? At least six months, but up to 18 months.

Where will it take place? All supply regions of the participating energy suppliers in Great Britain – across England, Scotland, and Wales.

Control/business as usual (BAU): Customers in the control condition will not receive an opportunity to earn availability payments from their asset.

Table 3: Availability payment trial treatments

	Group	Availability Payments
Incentive (based on CrowdFlex deliverable D7.2)	Low	£24 - £48 / asset / year (£2 - £4 / asset / month)
	Medium	£72 - £96 / asset / year (£6 - £8 / asset / month)
	High	£144 - £168 / asset / year (£12 - £14 / asset / month)
	Control	No incentive, but invited[15]

[15] We may also include a “no invitation” group as a backup.

8.2. Intervention for utilisation payments (for flexibility events)

What is the intervention: Different levels of the primary incentive mechanism (£/kWh flexibility) and notice period (hours announcement before event commences).

- Incentive: Low (£0.05 per kWh), Medium (£0.40 per kWh), High (£1.25 per kWh), or Mixed (low or high within a mini-trial, details in Section 9.3)[16]. This is the primary focus in the first 'mini-trial'.
- Notice period: 24 hours, 12 hours, 8 hours, 4 hours[17] – this is currently in scope for later mini-trials.

Table 4: Utilisation payment treatments (for mini-trial 1)

	Group	Utilisation payments - turn-down	Utilisation payments - turn-up
Incentive (based on CrowdFlex Deliverable D7.2)	**Low**	£0.05 / kWh	£0.05 / kWh
	Medium	£0.40 / kWh	£0.40 / kWh
	High	£1.25 / kWh	£1.25 / kWh
	Control	no communications at all	

What resources are needed to deliver the intervention: We need to communicate the intervention to participants in the trial. 1 to 24 hours before the event, we will notify them of:

1. The timing of the event.
2. The incentive for participating in the event.

What are the practical steps taken for intervention delivery? We will give customers a notice 1 to 24 hours in advance and ask customers to turn their consumption down or up. Flexibility service providers will notify customers through:

1. Email.
2. Push notifications from the apps (and potentially other devices) of participating flexibility service providers.

Who will deliver the intervention? Participating flexibility service providers, with CNZ undertaking analysis and trial implementation oversight.

How long will the intervention last? There will be 60 turn-down and 20 turn-up events over 6 months, during the first mini-trial. Each event will be 1 to 4 hours long.

Where will it take place? All supply regions of the participating energy suppliers in Great Britain – across England, Scotland, and Wales.

Control: Customers in the waitlisted control group will receive no notifications about events, nor receive any incentives for reducing their energy consumption.

[16] When we discuss incentives 'per kWh' of flexibility, we mean incentive per kWh that customers reduce their consumption below (or increase their consumption above) a baseline consumption during the flexibility window calculated by the energy suppliers in accordance with standards set by National Grid ESO on 'baselining'.
[17] Note that in all notice periods, we expect customers to receive a *separate* reminder of the event approximately one hour before the event commences.

8.2.1. Event specifications

We will vary many more parameters at the 'event' level (rather than at the 'treatment group' level). Parameters that might vary by event include:

- Event duration (e.g., 1 hour vs 2 hours vs 4 hours)
- Daily/seasonal timing of event
- How much time passes between events (i.e., how many hours or days since the most recent event an event occurs).[18]
- Is the event turn-up or turn-down? (Sometimes, events will involve certain regions turning up and other regions turning down. Note that we will treat turn-up as distinct in analysis from turn-down, even if both happen (in different regions) in the same event. As explained above, this is because we expect responses to be asymmetric and to have qualitatively different underlying behavioural mechanisms to achieve the flexibility.)

The consortium is interested in testing events that involve certain regions turning up and others turning down, in order to improve understanding of demand side flexibility to manage thermal constraints – both testing the social research questions outlined above *and* testing technical feasibility questions related to these sorts of simultaneous turn-up *and* turn-down events. This is further discussed in Section 9.5. We show a high-level visualisation of this setup in the table below.

Table 5: Visualising how events may vary from each other

Event	1	2	3	4	5	6	7	8	9	10
Turn up or down	Down	Down	Down	Down	Down	Up	Down	Up and down	Up	Down
Event duration (hours)	2	2	1	1	4	3	2	2	etc.	
Day / season	June	June	July	July	July	August	etc.			
Days since previous event	5	1	1	4	5	20	etc.			
Location	All regions	All regions	Just 2 regions	etc.				Scot: up Eng: down		

See Appendix A for a more in-depth example event schedule.

[18] Note that the consortium is especially interested in testing repeated calls and how the flexibility deteriorates as a result – e.g., a minimum of 3 events over 5 working days.

9. Design

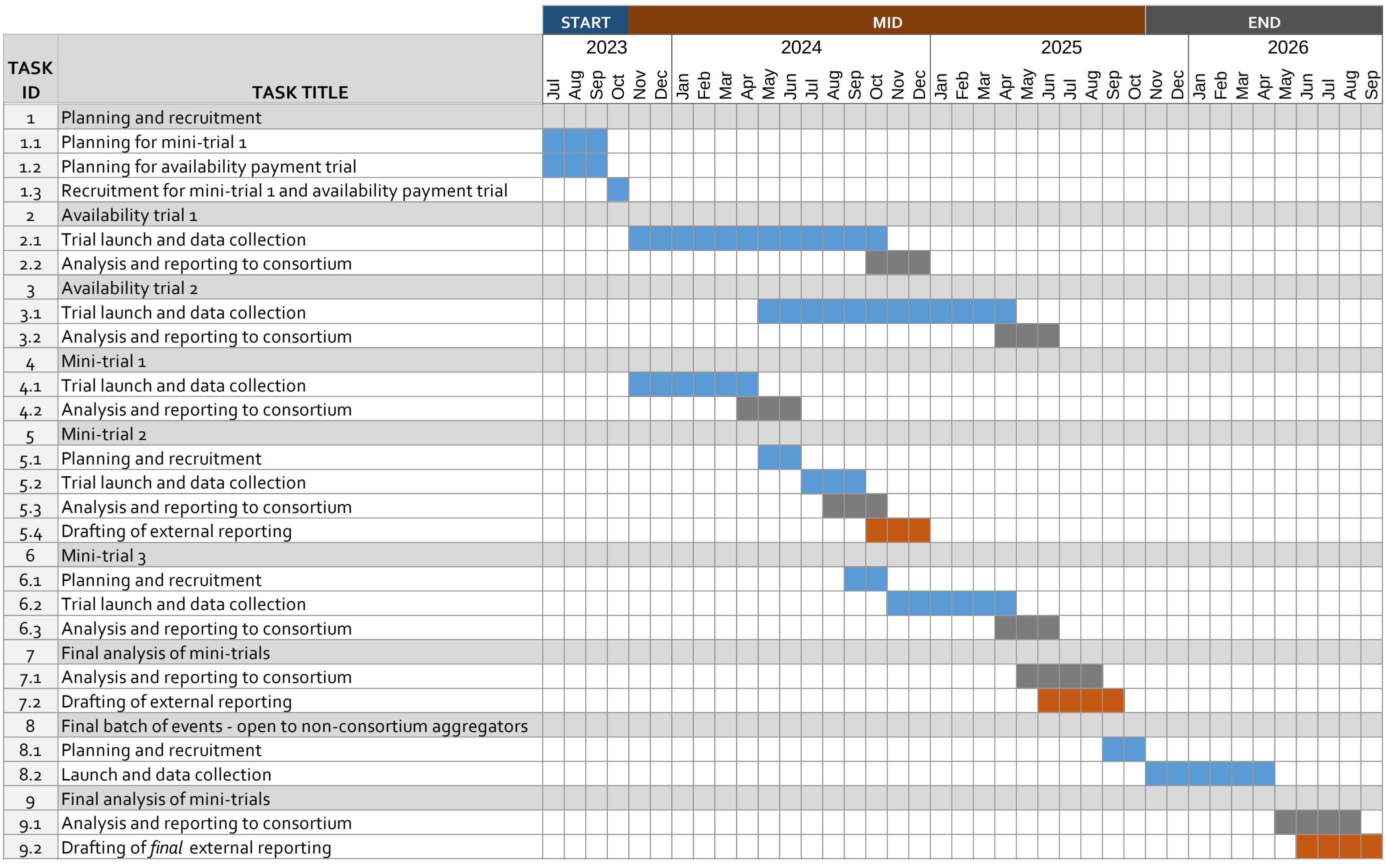
9.1. High-level design of CrowdFlex

Within this trial, we recommend three phases over three years. The goal of this iterative approach is to make operational improvements, and also gives an opportunity to adapt what we test (for example, if we feel that we've answered the original research questions of interest earlier than expected).

Table 6: The three proposed CrowdFlex 'phases'

Start: **Jul 23 - Oct 23**	Test messaging and communications with customers via surveys and qualitative research (such as focus groups). In addition, test programme logistics, accessibility, and any new technology that CrowdFlex will rely on. **Deliverable: Publish analysis of this initial testing.** Refine outstanding questions from this trial protocol outline; complete a full trial protocol; pre-register the trial protocol. **Deliverable: Publish and pre-register trial protocol.** Recruit participants into CrowdFlex, including an optional survey as part of opt-in about their household (discussed below).
Mid: **Nov 23 - Nov 25**	Conduct the availability payment trial: flexibility service providers contact customers with relevant assets about the availability payments offer (where these customers are identified, at least partially, through the optional survey from the Start phase). Run 3-5 CrowdFlex 'mini-trials' for the utilisation payment trial, encompassing 20-60 events over 3-6 months. After each 'mini-trial', CNZ and the consortium will conduct interim analyses and provide recommendations on changes to the design of the next sub-trial. These changes may involve operational improvements, new ways of testing research questions, new onboarding and recruitment, or new research questions to test. **Deliverable: Publish interim analyses.** This 'batched' design will allow CrowdFlex to be agile, respond to new information, and adapt to changing contexts.
End: **Nov 25 - Sept 26**	CNZ and the consortium will write up final analysis findings. **Deliverable: Publish final analysis.** A final batch of 20-60 events with all customers, including the waitlisted customers (discussed in more depth below, in Section 9.5.2) from the participating flexibility service providers. In addition, other suppliers will be invited to participate.

Figure 1: Gantt chart showing the three proposed CrowdFlex 'phases' and mini-trials within the 'Mid' phase



In the following subsections, we recommend the design of the first mini-trial (the first 'batch' of approximately 60 turn-down and 20 turn-up events).

9.2. Start phase

Before conducting the RCTs, the start phase of Beta will:

- Test messaging and communications with customers via surveys and qualitative research (such as focus groups).
- Answer technical questions: Test programme logistics, accessibility, and any new technology that CrowdFlex will rely on.
- Refine outstanding questions from trial protocol outline.
- Complete a full trial protocol; pre-register the trial protocol.
- Recruit participants into CrowdFlex, including an optional survey as part of opt-in about their household.

At the end of the Start phase, suppliers (and potentially other flexibility service providers) will send an email to approximately 300,000 randomly chosen smart-meter customers asking them to opt in to CrowdFlex, expecting approximately 150,000 to respond affirmatively. (Details on this are to be decided including: 1) how many 'batches' of invitations, in order to avoid obtaining an unnecessarily large sample size; 2) whether to have an open invitation as well.)

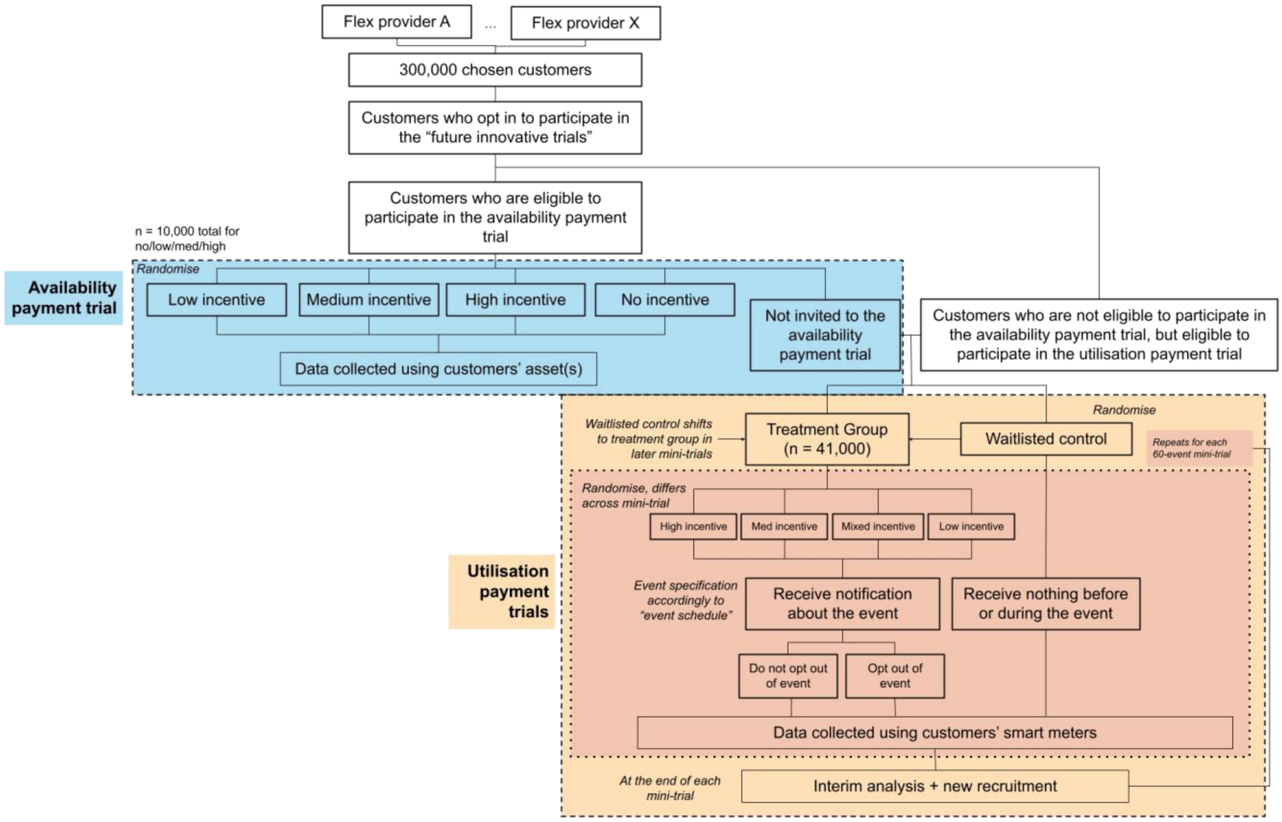
In the invitation to opt in to CrowdFlex, suppliers and other flexibility service providers will be transparent that incentives for turn-down (and turn-up) will vary by customer and by event, sometimes randomly. Based on conversations with Octopus Energy's marketing and communications team, we recommend *also* noting that customers who opt in to CrowdFlex will be entered into a prize draw with attractive prizes.[19]

Once a customer opts in, we will assume participation for subsequent events – e.g., an 'opt-out' framework (see Section 9.5.1 below).

9.3. Mid and end phase

The mid-phase will involve running two trials: (i) the availability payment trial and (ii) the utilisation payment trial. We illustrate this in the figure below.

Figure 2: Trial randomisation and delivery flow-chart



9.4. Design for trial dealing with availability payments

9.4.1. Sample selection

Our sample (before attrition) is customers of consortium members (e.g., Octopus Energy, Ohme, and potentially other relevant flexibility service providers who join the consortium) who satisfy the following conditions:

[19] This prize draw will serve at least two purposes. First, we hope that it will increase the overall opt-in rate. Second, we hope that it will reduce disappointment for customers who are placed in the waitlisted control or in the lower-incentive treatment arms.

- Have a smart meter – or some other means of recording avoided or additional electricity consumption (e.g., a 'qualifying meter').
- Have been customers of the relevant flexibility service provider for at least 6 months.
- Opt in to participate in CrowdFlex.
- Not on an exporting tariff.
- Have a controllable asset (e.g., a controllable EV charger or heat pump) as already identified by flexibility service providers or through the recruitment survey.
- Customers who have not already agreed to allow their assets to be remotely controlled, such as those on the Intelligent Octopus tariff.[20]

9.4.2. Randomisation

For the part of CrowdFlex associated with availability payments, we recommend a cluster randomisation design (clustered at the customer level as customers may have multiple assets[21]), with flexibility service providers across the consortium randomising their customers into four groups.

Table 7: Randomisation for availability payments

	Group	Flexibility service provider 1	Flexibility service provider 2	Flexibility service provider 3
Group allocation (incentives based on CrowdFlex deliverable D7.2)	Low	£24 - £48 / asset / year (£2 - £4 / asset / month)		
	Medium	£72 - £96 / asset / year (£6 - £8 / asset / month)		
	High	£144 - £168 / asset / year (£12 - £14 / asset / month)		
	Control	No incentive, but invited.		

Note that this sample selection strategy assumes that flexibility service providers know which of their customers have assets that could be remotely controlled – and thus would be eligible to receive availability payments.

- Energy suppliers know this about customers from customers' asset-specific choice of tariff (for example, for Octopus Energy customers, the choice of an EV-specific tariff).
- Other flexibility service providers (such as Ohme) would know that customers are eligible based on the basic customer proposition of providing them with an eligible asset.
- Flexibility service providers may also find who is eligible for availability payments based on customers' responses to the optional survey filled out during opt-in to CrowdFlex, discussed in Section 12.1.

Customers' response to the question of whether they would be willing to opt in to managed charging (or equivalent for heat pumps) is then half of the core outcome measure for the availability payment trial, and the key hypothesis is that their response will depend on the

[20] For customers with the flexibility service provider Octopus Energy, availability payments will be an extra incentive offered to customers to switch to the Intelligent Octopus tariff. Octopus Energy has explained during the drafting of this trial protocol outline that their vision is for Intelligent Octopus to be the single product for remote-control of EV chargers (and later heat pumps).
[21] This may mean linking customers across multiple flexibility service providers, if they are receiving availability payments for one asset from one flexibility service provider and from another asset from another flexibility service provider.

payment level being offered. (The other 'half' of the outcome measure is that the customer actually follows through, making their asset sufficiently available to receive the payments.)

Flexibility service providers will ask about opt-in to the availability payments through a separate email, app interaction, or other customer journey; not via the optional survey discussed in Section 12.1. Thus, the survey's only connection to the availability payments trial is that it may surface some customers who are eligible for availability payments that the flexibility service provider(s) did not previously know about. However, to emphasise: A customer who fails to fill out the survey may still be known to the flexibility service provider to be eligible.

9.5. Design for trial dealing with utilisation payments (for flexibility events)

Customers who are invited in the availability payment trial will not be invited to any flexibility events.

9.5.1. Opt-in and opt-out for utilisation payments (flexibility events)

A key part of the trial procedure involves recruitment and opt-in. We have provided our recommendations on recruitment in CrowdFlex deliverable D2.2, including recommending the following opt-in / opt-out framework:

- Customers ***opt in*** to the CrowdFlex *programme*, before the start of each "mini-trial".
- Customers may ***opt out*** of *specific events* – so, customers will be included in all events by default if they have opted in to the CrowdFlex program. They can choose to say 'no, I would not like to participate', opting out of any specific event. They can also opt out of the programme entirely at any time. Note: even if customers opt out of receiving emails and other event notifications, we will be able to observe attriters' energy consumption through their smart meter (see Section 10.4 for details on attrition).

This would make the opt-in / opt-out system more like Ecobee's pilot program with San Diego Gas & Electric (SDGE)[22] than the framework used in CrowdFlex NIA.

9.5.2. Randomisation for utilisation payments (flexibility events)

We recommend a wait-list randomisation to create a pure control group, with the rest of the customers randomly allocated into one of 5 groups.

From a marketing and comms perspective, if customers are opting into the CrowdFlex programme, they would expect to take part in at least *some* events.

We recommend inviting people to opt into CrowdFlex but note in the invitation that 'we will be onboarding customers in batches, due to operational constraints'.

Then, we will split opted-in customers into a larger randomised subset invited to CrowdFlex events, and a smaller randomised subset not invited to any CrowdFlex events until its very final months. We let the latter group know that they weren't picked for the first series of events – but that they'll participate in later ones.

CrowdFlex deliverable D2.2 recommended that recruitment marketing be direct and transparent. We believe that this approach meets those standards, yet also have given

[22] See: https://support.ecobee.com/s/articles/eco-Frequently-Asked-Questions and https://www.ecobee.com/en-us/newsroom/press-releases/ecobee-launches-pilot-program-with-california-utility-to-help-prevent-power/

ourselves a pure control group for all but the last few months of CrowdFlex. With that said, we will liaise with all CrowdFlex partners about the marketing, communications, and customer satisfaction implications of this approach.[23]

We illustrate this randomisation approach in the following diagram:

Table 8: Utilisation payments trials randomisation approach

		Mid-phase			**End phase**
		Mini-trial 1: Months 1-6	**Mini-trial 2: Months 7-12**	**Mini-trial 3: Months 13-18**	**Months: 19-24**
Group 1	**Low**	£0.05 / kWh	To be confirmed, based on consortium discussion and priorities and interim analysis after the first mini-trial. The randomisation approach for these customers will be finalised in Beta.		The most cost-effective incentive, based on results from the previous mini-trials.
Group 2	**Medium**	£0.40 / kWh			
Group 3	**High**	£1.25 / kWh			
Group 4	**Mixed**	Low, then High			
Group 5	**Mixed**	High, then Low			
Group 6	**Control**	Waitlisted			

In future mini-trials, regardless of the treatment, we recommend withholding a group of customers whom we don't expose to any treatment until the 'end phase'. However, some of them could be allocated to treatments in mini-trials 2 and 3. Indeed, if the waitlisted control group is large enough, each mini-trial could theoretically have its own sample, such that customers are never in treatment groups for multiple mini-trials. As noted in the diagram above, this approach needs to be finalised in Beta.

Table 9: How the mixed incentives arms will work over the course of mini-trial 1

	Mini-trial 1					
Turn-down events	**1**	**…**	**30**	**31**	**…**	**60**
Turn-up events	**1**	**…**	**10**	**11**	**…**	**20**
Group 1	Low					
Group 2	Medium					
Group 3	High					
Group 4	Low			High		
Group 5	High			Low		
Group 6	Waitlisted control					

[23] We recommend the trial designer does so with any other suppliers involved in the beta phase of this project, too.

Number of arms: Within the first mini-trial, there are 6 variations:

1. Waitlisted control
2. Persistent Low
3. Persistent Medium
4. Persistent High
5. Initial Low, Subsequent High (Group 4)
6. Initial High, Subsequent Low (Group 5)

We suggest equal sized Groups between Group 1 - 5 (the treatment groups) with at least 8,200 customers in each turn-down event 4,200 in each turn-up event per arm (see Section 13.2.2 for details on balancing power and budget for the utilisation payment trials), while reserving a majority of the customers in the waitlisted control for future mini-trials.

What is the unit of randomisation? Household[24]

What is the unit of analysis? Household * half-hour combination

Spillover / contamination risks: We rate the risk of spillover between treatment arms to be low in this trial, as the incentive to turn-up/turn-down is specific to the household. That said, there is some chance that neighbours or friends discuss their differing incentives with each other, which could affect how customers perceive their incentive levels.

We rate higher risks of contamination *within-participants, between mini-trials.* Specifically, there may be *carryover* effects across mini-trials for treated customers. The anticipated effect of the carryover is ambiguous. For example, during mini-trial 1, participants compensated at £1.25/kWh might be incentivised to *learn* how to use their energy more efficiently. This might cause them to subsequently use *less* energy in mini-trial 2 (even if they are allocated to a smaller incentive £0.40/kWh group, where they would not have originally been sufficiently incentivised to learn how to use their energy more efficiently). We might incorrectly attribute their turn-down to the lower incentive in this case. To a certain extent, we try to understand this carryover effect using the ‘mixed incentive’ arm. We discuss our approach in Section 14.2.

9.5.3. Non-randomised changes in flexibility event parameters

In addition to the randomisation suggested above, we then suggest varying many more parameters on an ‘event’ level (rather than on a ‘treatment group’ level), as discussed in Section 8.2.1 (Event specifications). We suggest analysing the impact of these changes descriptively in exploratory analysis (details in the analytical strategy section).

9.5.4. Flexibility events for subsets of customers

Note that there may be some events for subsets of customers, e.g., an event only for customers with automation (such as customers with Ohme chargers). These might operate a little differently from the above events open to all customers, and we have not fully spelled out how they would work here. To be determined how these events interact with the 'availability payments' for automated response.

[24] Note that we use the term ‘household’ and ‘customer’ interchangeably in this document. Where a customer has multiple properties, we assume that they will need to choose a specific property to participate.

9.5.5. Simultaneous turn-up and turn-down flexibility events

Note that within a given event, depending on the region of the country, some customers may be asked to 'turn-up' (increase) their consumption and others to 'turn-down' (reduce) their consumption. The reason for this is that electricity can be abundant in certain areas of the country (e.g., ones with high levels of wind generation and low population density) and scarce in others (e.g., ones with high population density and relatively less local wind generation) during the same half-hour. Increasing consumption in electricity-abundant areas reduces curtailment costs, while decreasing it in electricity-scarce areas reduces system costs associated with drawing on or generating additional electricity.

We are proceeding on the assumption that these turn-up-and-turn-down events would be 'phased in' over the course of CrowdFlex, as ESO, DSOs, and suppliers learn about user engagement and refine event logistics. We will work with the consortium to define this schedule as much as possible in the full trial protocol.

In addition, note that there may be a seasonal component to when the grid needs more 'turning up' versus when it needs more 'turning down': the summer will likely involve more cases where generation is abundant but demand is low and 'turning up' is called for, whereas the winter will likely involve more cases where demand is high and supply is constrained and 'turning down' is useful.

We believe that the general analysis framework still applies to both turn-up-and-turn-down events and in understanding customer behaviour given the seasonal turn-up / turn-down pattern. Treatment allocation (£ incentive per kWh) will have proportionately the same number of customers in different regions (and of course seasons). So, in exploratory analysis, we will be able to:

- Examine how random treatment allocation affects customers' flexibility response.
- Examine descriptively the impact of specific event parameters (such as whether an event is a 'turn-up-and-turn-down' event.

However, note that most of this trial protocol outline focuses on turn-down events, where customers' unit rates (e.g., currently 34p/kWh consumption) create the same directional incentive as the turn-down incentive (e.g., £0.40/kWh turn-down, which would lead to a 'total' incentive of £0.74/kWh benefit from avoiding consumption, assuming the demand is 'destroyed' rather than shifted). Turn-up events are a little more complex, and further thought may be required as to how best to vary incentives per kWh of turn-up.

10. Data collection and outcome measures

10.1. Outcome measures for availability payments

Our primary outcome will be whether a customer's asset is **available** to be controlled remotely for the purpose of allowing flexibility service providers[25] to participate in the Balancing Mechanism (and/or diversity in domestic appliance in Load Managed Areas) by providing flexibility on a seconds- or minutes-ahead basis. That is, whether the participant opts in to allow flexibility service providers to control their asset remotely and the flexibility service provider being able to control it in pre-specified timeframes.

Note that we assume no *social* research questions about the actual performance (i.e., revenue earned, kWh balanced, etc.) of assets participating in the balancing mechanism. Instead, the research questions focus on participation only. Thus, we do not consider revenue earned or kWh balanced to be outcome measures of interest for analysis of availability payments in pre-specified trial analysis, though these measures *will* be collected and analysed to assess the performance and viability of this 'mode' of flexibility service.[26]

Table 10: Outcome measure for availability payments trial

Outcome Measures	Data to be collected	Point of Collection
Primary: Whether the customer's asset is available to participate in the Balancing Mechanism	Whether the customer opts-in to allow their asset to be controlled remotely (customer interaction data) Whether the customers' asset is available to be controlled remotely • EV: Charger plugin time and whether the customer overrides • Heat pump: TBC	• Before Interim-analysis • End of the trial

10.2. Outcome measures for utilisation payments (for flexibility events)

Our primary outcome of interest is customers' electricity consumption during the event window. This will primarily be measured using customers' smart meters, which record electricity consumption in 30-minute intervals, though asset level data may also be utilised where available.

Our secondary outcome of interest is customers' electricity consumption outside the event window (24 hours before and after the event).[27] This is to understand whether customers *substitute* their energy consumption to half hours outside the event window versus whether the energy demand is *destroyed*. This will also be measured using smart meters and asset data where available.

[25] Flexibility service providers include suppliers, but also perhaps smart charging management or other asset management systems.
[26] Note, also, that: 1) payments will be scaled to realistic levels, 2) flexibility service providers may classify customers as ineligible for payment unless certain conditions are met, and 3) performance will be examined as a 'technical' research question (i.e., looking at the capacity of response throughout the year and how it varies based on external factors (e.g., as we approach system peak).
[27] We will exclude events that are less than 48 hours apart.

Exploratory outcomes of interest are:

- Customers' comfort and satisfaction during or after events – this is to understand whether these events change comfort and satisfaction, and will be measured using customer surveys during the trial.
- Opt-outs from an event or the programme overall – collected from customer relationship management data.
- Home heating behaviours, predominantly via automated systems, where available.
- Perhaps data on pre- or post-event EV driving distances, collected from Ohme (and other participating companies') smart chargers – the idea here is to help answer the exploratory research question of whether EV drivers drive less (or more) in anticipation of an event or following an event.

Table 11: Outcome measures for utilisation payments trials

Outcome Measures	Data to be collected	Point of Collection
Primary: Electricity consumption during the event window	• Flexibility service providers' customer energy consumption data • Smart-meter dataset	This is recorded in near real-time, but we collect this data before interim analyses and before the end-of-trial analysis.
Secondary: Electricity consumption 24 hrs before/after the event window (event window specified in the event timetable)	• Flexibility service providers' customer energy consumption data • Smart-meter dataset	This is recorded in near real-time, but we collect this data before the interim analysis and before the end-of-trial analysis.
Exploratory: Subgroup analysis for primary outcome measure: • Participant's appliance data (including EV charger, heat pump, solar panels, battery, etc.) • Location • Number of events invited	• Flexibility service providers' customer energy consumption data • Customers' appliance data • Ohme charging data	Pre-event surveying and the same frequency/data as the primary and secondary measures.
Exploratory: The effect of different event parameters: • Event duration • Hours of the event • Season of the event • Day of the week	• Flexibility service providers' customer energy consumption data • Suppliers' customer energy consumption data • Smart-meter dataset	Event design (pre-trial) and the same frequency/data as the primary and secondary measures.
Exploratory: Opt-out rate	• Customers' interaction with the flexibility service provider	This is recorded in near real-time, but we will be sent this data before the interim analysis and before the end-of-trial analysis. Event-by-event communication.
Exploratory: Self-reported comfort, satisfaction	• Customer's qualitative feedback survey	The survey will be sent to a random sample of participants, periodically. We collect this data before the interim analysis and before the end-of-trial analysis. [We will refine the survey plan in beta.]

10.3. Attrition

Based on a previous flexibility trial, we expect the proportion of customers opting out from receiving event notifications to be <2%. However, this is not attrition in the classic sense of the word; even if customers opt out of receiving emails and other event notifications, we will be able to observe attriters' energy consumption through their smart meter (our primary and secondary outcome).

Another type of attrition is customer switching and customers moving home during the trial. We will likely lose the ability to track these customers' energy consumption once they switch or move home. Switching rates are lower than Great Britain's historical average[28] due to the current energy crisis, but they may increase to their traditional levels during CrowdFlex, which is a multi-year study. It is theoretically possible that this type of attrition could differ by treatment allocation, but we consider this to be unlikely. Hence, we will use data up to the point when the customer attrits (e.g., moves out, opts out from the trial etc.).

We also propose rounds of additional recruitment before each mini-trial to replace attriters and to allow the option to refine the sub-groups.

10.4. Data gathering

Table 12: Data gathering for availability payment trials

Data	Collection Point	Source
Participant property reference number (we assume a 1:1 relationship between energy meter and property reference number, and we treat this as our 'id' variable)[29]	• Before randomisation	Energy supplier(s) and other flexibility service providers
Participants' randomisation outcome	• After randomisation	Flexibility service provider(s)
Customer interaction data with the aggregator (e.g., whether the participant opts in to having their asset be controlled by the aggregator).	• Before interim analysis • Before endline analysis	Flexibility service provider(s)
EV charging data (performance and whether the participant opts out from automation, whether their chargers are plugged in, for example)	• Before interim analysis • Before endline analysis	Flexibility service provider for EVs
Customers' heat pump and/or thermostat data (performance and whether the participant opts out from automation, whether the customer overrides automation, for example)	• Before interim analysis • Before endline analysis	Flexibility service provider for heat pumps

[28] Less than 100k customers switching per month, versus a historic average of approximately 0.5m switching per month, according to: https://www.gov.uk/government/statistical-data-sets/quarterly-domestic-energy-switching-statistics
[29] It is not yet decided how we will manage customers with multiple properties for the availability payment trial. Theoretically they could be paid for each of their properties (and indeed receive payments for separate assets at each property.)

Table 13: Data gathering for utilisation payment trials

Data	Collection Point	Source
Participant property reference number (we assume a 1:1 relationship between energy meter and property reference number, and we treat this as our 'id' variable)[30]	• Before randomisation	Energy supplier(s) and other flexibility service providers
Participants' randomisation outcome	• After randomisation	Flexibility service provider(s)
Participants' energy consumption data	• Before randomisation • Before interim analysis • Before endline analysis	Energy supplier(s) – and perhaps flexibility service providers with access to qualifying meters that are not the customer's smart meter
Participants' opt-out data	• Before interim analysis • Before endline analysis	Flexibility service provider(s)
Customer information (e.g., location, whether they own a LCT, property EPC)	• Before randomisation • Before interim analysis • Before endline analysis	Flexibility service provider(s)
Participants' comfort and satisfaction	• Before interim analysis • Before endline analysis	Flexibility service provider(s)

10.5. Data storage and transmission

Data will be anonymised and stored in project folders with access restricted to the project team only. Data will not be transmitted to third parties, except where this is appropriate under the conditions of appropriate data sharing agreements.

[30] It is not yet decided how we will manage customers with multiple properties. For the time being, we assume that they must choose *one* property to participate in CrowdFlex events – so no customer will have multiple properties in our analyses datasets.

11. Stratifying randomisation

11.1. Availability payment trial

For the randomisation for the availability payments, we suggest stratifying on flexibility service providers and GSP group (a high-level location category).[31]

11.2. Utilisation payment trial

For the randomisation for utilisation payments for flexibility events – both for the waitlist and treatment randomisations (both described in Section 9.5.2 above), we suggest using the following variables for stratification:

- Average household-level electricity consumption for the 6 months before randomisation (exact categories to be defined).
- Tariff type (three categories: dynamic ToU, other ToU, no ToU).
- GSP Group (a high-level location category).

[31] There may be insufficient customers in all GSP groups to use GSP group as a stratification variable – we will assess this in more detail in beta.

12. Trial procedure

12.1. Survey during opt-in for flexibility events

During the opt-in stage for participating in CrowdFlex flexibility events, we suggest that customers fill in an optional initial survey about:

- What technology they have in their home – in particular what low-carbon technologies:
  - EVs, what sort of EV charger (if any), what alternative transport options they have (e.g., a second vehicle, reliable public transport)
  - What sort of heating system they have – and, if a heat pump, what type of heat pump, and whether it is the sole source of heating versus one of multiple heating options
  - Whether they have an in-home battery
  - Whether they have solar panels
- Characteristics about their home – e.g., number of bedrooms, type of building, EPC rating
- Characteristics about the occupants – e.g., number of occupants, working patterns, and attitudinal questions)

The purpose of this survey is to allow much richer subgroup analysis. ESO and other consortium members are interested in understanding the impact of household and occupant characteristics on flexibility provision. The models to be developed in CrowdFlex for flexibility provision are based in part on these characteristics; however, flexibility service providers do not necessarily know these characteristics about customers before the trial – they may know some of them, for some customers, but coverage is not complete. A baseline survey conducted during opt-in to CrowdFlex would allow us to have better coverage on these characteristics of interest across our full sample (including the pure control group), greatly enhancing exploratory subgroup analysis and predictive modelling.

Note that this survey would be optional – to avoid creating a new friction to involvement that reduces the overall rate of opt-in to CrowdFlex. However, we believe that even the partial completion of this survey will enhance the insights we can obtain from subgroup analyses.

12.2. Availability payment trial implementation

We then suggest that randomisation of offers of availability payments be chosen from the group of customers who have opted in to the flexibility events, many of whom will have already filled out the optional pre-trial survey. In this way, **survey completion rates will (by randomisation) be approximately the same across control and treatment trial arms, for both flexibility events and availability payments.**

12.3. Utilisation payment trial implementation

See Section 9.1 for high-level trial design. Details below to be filled in during Beta.

Table 14: Trial delivery roles and responsibilities

Timing	Owner	Action
		Recruitment
		Randomisation
	N/A	Trial start (refer to 'event timetable')
		Interim analysis
		End-of-trial analysis

Table 15: Example event timetable

Event	Date / Time	Action	Owner
1		Calculate customers' 'baseline' energy consumption	Flexibility Provider(s)
		Send customers day-ahead notification	Flexibility Provider (s)
		Send customers reminder notification	Flexibility Provider (s)
		Event window	Event parameters specified by ESO (& DSOs)
		Send customers feedback	Flexibility Provider (s)
		Provide compensation for participation (e.g., discount to customers' bills)	Flexibility Provider (s)
2		Calculate customers' 'baseline' energy consumption	Flexibility Provider(s)
		Send customers day-ahead notification	Flexibility Provider (s)
		Send customers reminder notification	Flexibility Provider (s)
		Event window	Event parameters specified by ESO (& DSOs)
		Send customers feedback	Flexibility Provider (s)
		Provide compensation for participation (e.g., discount to customers' bills)	Flexibility Provider (s)
3		Calculate customers' 'baseline' energy consumption	Flexibility Provider(s)
		Send customers day-ahead notification	Flexibility Provider (s)
		Send customers reminder notification	Flexibility Provider (s)
		Event window	Event parameters specified by ESO (& DSOs)
		Send customers feedback	Flexibility Provider (s)
		Provide compensation for participation (e.g., discount to customers' bills)	Flexibility Provider (s)

4	Interim survey	Flexibility Provider(s)
	Calculate customers' 'baseline' energy consumption	Flexibility Provider (s)
	Send customers day-ahead notification	Flexibility Provider (s)
	Send customers reminder notification	Event parameters specified by ESO (& DSOs)
	Event window	Flexibility Provider (s)
	Send customers feedback	Flexibility Provider (s)
	Provide compensation for participation (e.g., discount to customers' bills)	Flexibility Provider(s)
	…	
	Final survey	Flexibility service provider(s)

12.4. Interim and final surveys

As we outlined in CrowdFlex deliverable D2.2 recommendations on customer recruitment and engagement, **discomfort and excessive self-rationing** are possibilities that raise ethical concerns in this trial. Citizens Advice emphasised that the impact on different vulnerable consumer groups needs to be considered in flexibility trials,[32] first to ensure that they are not negatively affected by demand-side response, and second to empower them to share in its benefits. Vulnerable customers can be negatively impacted by DSR as demand response can potentially lead to self-rationing, which can affect their physical and mental health. However, outcomes such as the customers' subjective comfort and stress cannot be directly observed from smart meter data. Customers may receive financial incentives at the cost of their comfort and health due to self-rationing. Hence, **we recommended conducting interim surveys, ideally oversampling on customers with pre-identified vulnerabilities.** These surveys will collect important qualitative outcomes such as participants' subjective comfort during the event. We will collect the data on customer experience, comfort, understanding of the flexibility event logistics, understanding of the purpose of flexibility, views on and satisfaction with event logistics, and the behaviours customers report employing to shift their demand.

We suggest staggering the surveys, sending out X surveys after each event, where the survey will be sent at various points during the trial to a random subset of customers who are not in the control group.

A complementary idea we will explore in beta involves sending short-survey text messages immediately after the end of the events to a random subsample of participants. The short-survey could include the highest-priority questions about customer welfare.

[32] Citizens Advice (2014): Take a walk on the demand side, Retrieved from: https://www.citizensadvice.org.uk/about-us/our-work/policy/policy-research-topics/energy-policy-research-and-consultation-responses/energy-policy-research/take-a-walk-on-the-demand-side/

12.5. Intervention adaptation rules

If we plan to change the intervention in response to new information, we should define rules by which we will do so in this section. We have left this section blank for now, with the intention to fill it out in the Beta phase of this project.

Table 16: Stopping and trial adaptation rules

Stopping or adaptation rule	Rule is that…	Who is responsible?
Rule 1	E.g., If carryover estimates are statistically significant, use persistent incentives across mini-trials	
Rule 2	E.g., Harm detected to low-income, vulnerable, or other consumers or assets.	
	Etc.	

13. Power

A power calculation estimates the minimum detectable effect size (MDES) given key assumptions – in particular, sample size. We conduct a power calculation for both trials. Our primary outcomes are opt-in rates for the availability payment trial and energy consumption during an event for the utilisation payment trial. Since these are the main objectives of the trial, we want to be confident that we can detect economically meaningful differences, if they exist.

13.1. Power calculations

13.1.1. Power for availability payment trial

Table 17: Availability payments trial power calculation assumptions and inputs

Alpha (significance level)	5%
Power	80%
Total planned sample size	2,500 per arm, 10,000 in total.
Clustered trial?	No
Number of trial arms	4
Base rate	10%
Intracluster correlation coefficient (ICC)	Not applicable
Proportion of outcome variation explained by pre-intervention outcome	0
What is the planned MDES for this trial?	2.5 percentage points
Anticipated substantive effect of the intervention?	> 4 percentage points
Is the planned MDES the same as or smaller than the anticipated effect of the intervention?	Yes
Have you corrected for multiple comparisons?	No – CNZ policy is to not correct for multiple comparisons if the comparisons are pre-registered in the trial protocol.

The following table describes the minimum detectable effect size for the primary analysis - the opt-in rate, in percentage points, varying the number of participants.

We assume:

- 3 levels of incentives + control group[33]
- Equal number of participants in each trial arm
- Individual-level randomisation
- The unit of analysis is the individual[34]
- No clustering
- We do not adjust for multiple comparisons, as is CNZ's policy when the comparisons are pre-registered in the trial protocol
- A baseline opt-in rate of 10%

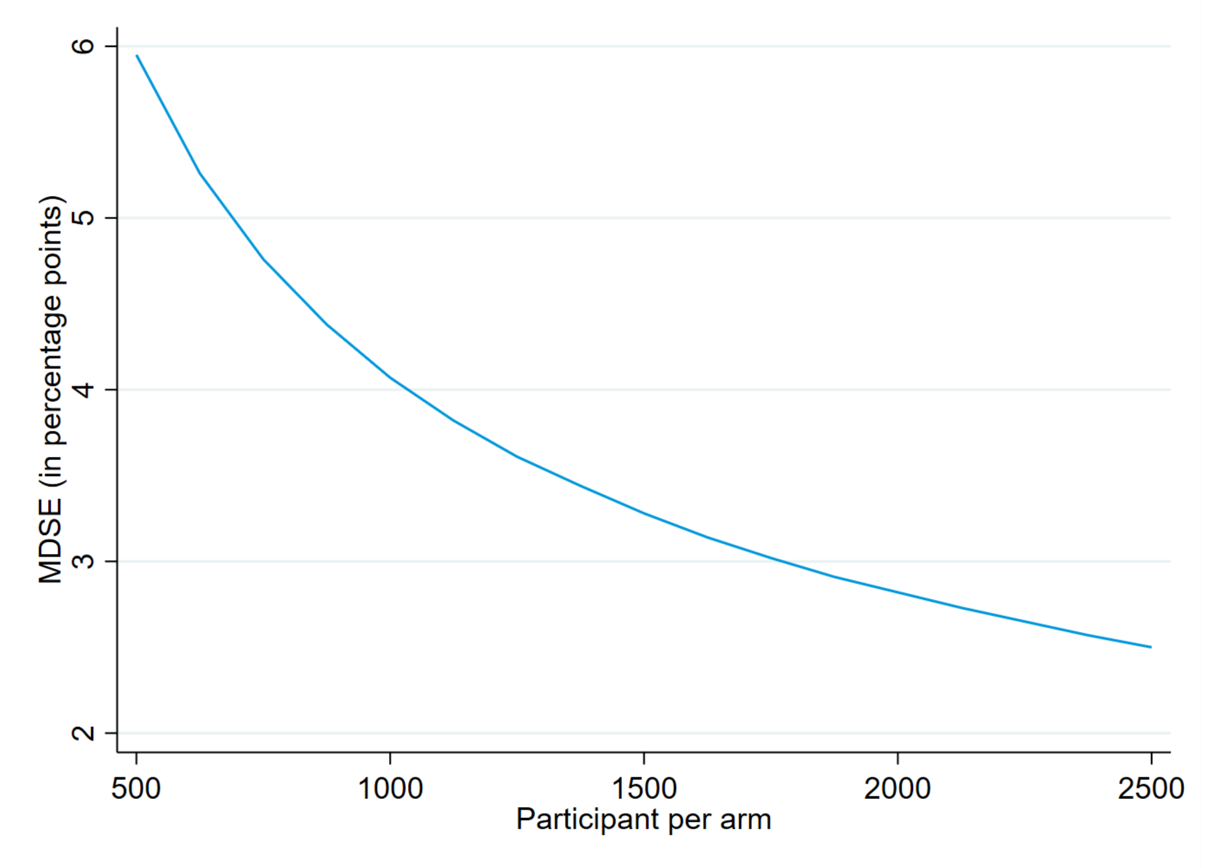
Table 18: Power for availability payment trial

# participants per arm	Total number of participants	MDES (in percentage points)
500	2,000	5.95
625	2,500	5.26
750	3,000	4.76
875	3,500	4.38
1,000	4,000	4.07
1,125	4,500	3.82
1,250	5,000	3.61
1,375	5,500	3.44
1,500	6,000	3.28
1,625	6,500	3.14
1,750	7,000	3.02
1,875	7,500	2.91
2,000	8,000	2.82
2,125	8,500	2.73
2,250	9,000	2.65
2,375	9,500	2.57
2,500	10,000	2.50

[33] We also intend to have a totally holdback group who we do not send the email invitation to. However, since this group is not essential for the analysis, we will only include this group if the number of eligible customers is sufficient.

[34] This simplification assumes each customer has one asset.

Figure 3: Power for availability payment trial



13.1.2. Power for utilisation payment trials

Table 19: Utilisation payments trials power calculation assumptions and inputs

Alpha (significance level)	5%
Power	80%
Total planned sample size	The Big Dirty Turn Down (also called the 'Domestic Scarcity Reserve Trial') had 105,320 participants. However, CrowdFlex's sample size may be higher or lower than this figure, and we show MDES for various sample sizes.
Clustered trial?	Yes, but in a slightly unusual way – observations are clustered on the individual household level (individuals participate in multiple events).
Number of trial arms	6, including the control group
Baseline outcome	0.27 kWh / half hour – this is the average consumption among Octopus Energy smart-meter customers across all half hours during 2021
Intracluster correlation coefficient (ICC)	Did not calculate - we conducted power calculation using actual (historical) data, a placebo treatment group, a semi-arbitrary event schedule, and the rule of 2.8.
Proportion of outcome variation explained by pre-intervention outcome	

What is the planned MDES for this trial?	See ‘substantive’ effect immediately below.
Anticipated substantive effect of the intervention?	Octopus Energy and ESO measured reductions of 60-64% compared to baseline, in their analyses of CrowdFlex NIA and the Big Dirty Turn Down (also called the ‘Domestic Scarcity Reserve Trial’). However, this was among customers who proactively opted in to participate in events – these participants comprised 40-50% of overall programme opt-ins. In a framework where all programme opt-ins are assumed to participate in a given event (unless they proactively opt-out), as we have suggested, the relevant reduction would then be 24-32%. However, as discussed below, this may still be an overestimate, given the novelty effect and limited number of events in CrowdFlex NIA and the Big Dirty Turn Down. As discussed below, we believe that our trial would be powered to detect much smaller effect sizes than 20% differences in consumption between trial groups.
Is the planned MDES the same as or smaller than the anticipated effect of the intervention?	Yes, assuming reductions per event that are close (i.e., not orders of magnitude lower) as those that were measured in CrowdFlex NIA and the Big Dirty Turn Down.
Have you corrected for multiple comparisons?	No – CNZ policy is to not correct for multiple comparisons if the comparisons are pre-registered in the trial protocol.

The following table describes the minimum detectable effect size as the percentage of the mean, varying the number of participants and number of treatment arms (including the control arm).

We assume:

- We are not correcting for multiple comparisons, and we are powering for treatment to control comparisons only, and not adjusting for multiple comparison is CNZ’s policy when the comparisons are pre-registered in the trial protocol.
- Individual-level randomisation.
- We account for attrition during data extraction. For example, when we requested 10k customers over a year, we only retrieved 8k as 2k of them do not have a complete year of energy consumption due to reasons such as moving.
- Each event is a two hour block.
- The unit of analysis is individual * half-hour.
- We cluster standard errors at the individual household level.
- One mini-trial with 20 ***or*** 60 events.
- We remove outliers based on the following rule: kWh/half hour > 15 or kWh/half hour < 0.01.
- We control for:
    - Baseline consumption - average energy consumption 6 months prior to the first event (average half-hourly energy consumption between).
    - GSP group (location) dummies.
    - Settlement period (dates) dummies.

We conducted power calculations using the rule of 2.8. Specifically, we created 20-event and 60-event schedules of fake events that occurred at different times of day (lasting two hours long, usually at some time in the afternoon or evening) between 01/06/2021 and 28/10/2021. (See Appendix B for the ‘fake’ events schedules.) We randomly sampled smart-meter customers (excluding customers who had participated in the Big Dirty Turn Down), increasing the sample size over successive calculations. We created a 'placebo' treatment variable, randomising by customer meter point administration number (MPAN).

We then conducted a linear regression, where we regressed energy consumption (at the customer-half hourly level) in kWh on the placebo treatment variable and the following control variables:

- 'Baseline consumption' – average energy consumption 6 months prior to the first event (average half-hourly energy consumption between 06/2021-12/2021).
- GSP dummies (for the 14 GSP groups – a high-level categorisation of customer location).

We clustered the standard errors at the MPAN level since the treatment was assigned at this level. We then multiplied the standard error on the coefficient for the placebo treatment variable by 2.8 to obtain the MDES (in kWh). To obtain MDES in % of mean, we divided MDES in kWh by the mean kWh per half hour consumption (0.27 kWh / half hour) in our sample.

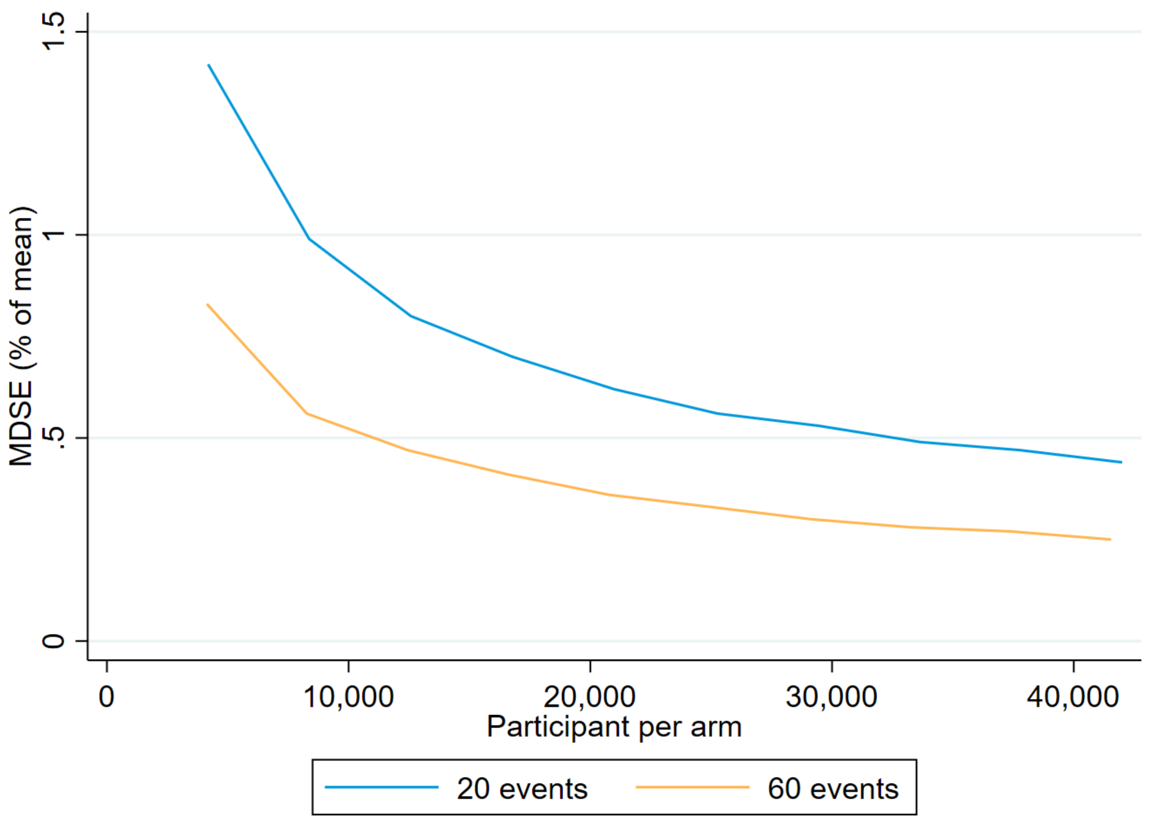
Table 20: Power for utilisation payment mini-trial of 20 events

Participants per arm	Total sample excluding control group	MDES (in kWh)	MDES (in % of mean)
4,189	20,945	0.003805	1.42
8,370	41,850	0.002652	0.99
12,578	62,890	0.002156	0.8
16,773	83,865	0.00189	0.7
20,990	104,950	0.001686	0.62
25,272	126,360	0.001523	0.56
29,445	147,225	0.001422	0.53
33,641	168,205	0.00133	0.49
37,789	188,945	0.001254	0.47
42,001	210,005	0.001179	0.44

Table 21: Power for utilisation payment mini-trial of 60 events

Participants per arm	Total sample excluding control group	MDES (in kWh)	MDES (in % of mean)
4,134	20,670	0.002201	0.83
8,272	41,360	0.001481	0.56
12,446	62,230	0.001238	0.47
16,600	83,000	0.001086	0.41
20,775	103,875	0.000963	0.36
25,009	125,045	0.000879	0.33
29,135	145,675	0.000806	0.3
33,280	166,400	0.000762	0.28
37,377	186,885	0.000722	0.27
41,550	207,750	0.000675	0.25

Figure 4: Power for utilisation payment trials



Despite being well powered for only 20 events, we believe there are various reasons to prefer 60 'turn-down' events, at least in mini-trial 1. (Note that we suggest only 20 turn-up events.) These include:

1. A key research question is whether customers' turn-down will sustain over time, across many events (as opposed to decaying) – this question would be easier to answer with 60, than just 20, events.
2. Furthermore, as specified in the exploratory analysis, we are interested in understanding the effect of different event parameters (e.g., duration of the event) on customers' energy usage. This question would also be hard to answer with only 20 events.

As discussed in Table 20 and Table 21 above, if we expect effect sizes as large as 20% – e.g., treatment group consumption during events being 20% lower than control group consumption during the same half-hours – we are very well-powered even at sample sizes as low as 1,000 participants per arm. However, we urge caution in having too much confidence in this result, for a few reasons:

1. The effect sizes we see over many events may be lower than the very large effect sizes obtained in the smaller, more novel CrowdFlex NIA, Big Dirty Turn Down (Domestic Scarcity Reserve Trial), and Demand Flexibility Service events. Moreover, these effects were actually among the subset of customers who opted in to each event, whereas we suggest that CrowdFlex work such that customers opt in to CrowdFlex overall and then are *assumed* to participate in an event unless they proactively opt out.
2. The effect sizes from CrowdFlex NIA, Big Dirty Turn Down (Domestic Scarcity Reserve Trial), and Demand Flexibility Service events are based on a set of 'baselining' methodologies (comparing actual consumption to forecast consumption) that are still being refined and may have led to systematic overestimates of flexibility.
3. Treatment versus treatment comparisons are also of interest to the consortium – and these comparisons may feature much smaller differences than treatment versus control group differences.
4. CrowdFlex consortium members are interested in various analyses we have classified as exploratory, in particular the examination of treatment effects for particular subgroups (such as examining flexibility by LCT ownership, tariff type, location, etc.). These involve splitting up the treatment groups into yet smaller cells, and it would be advantageous to maintain sufficient power for these comparisons too.

13.2. Balancing power and budget concerns

We need to balance the desire for high precision against the costs of large sample sizes. In this section, we present various sample size considerations.

13.2.1. Balancing power and budget for the availability payment trial

We are well-powered to detect meaningful differences between groups at relatively low sample sizes for the availability payment trial. In the below table, a key input is the number of participants *excluding* control group(s) – control group participants receive no incentive and thus have no impact on the overall budget for incentive costs.

Table 22: Availability payment trial MDES and incentive costs depending on various sample sizes

Parameters						Outputs	
Number of trial arms in each mini-trial	**# Partici-pants excluding control group**	**# Partici-pants per arm**	**% receiving incentive**	**Average incentive/ asset/ month**	**Trial duration (months)**	**MDES (in percent-age points)**	**Cost**
4	1,500	500	10	£7.50 (this is a rounded average of the three availability payment levels)	12	5.95	£13,500
			20			5.95	£27,000
			50			5.95	£67,500
			50			5.95	£67,500
4	3,000	1,000	10			4.07	£27,000
			20			4.07	£54,000
			40			4.07	£108,000
			50			4.07	£135,000
4	4,500	1,500	10			3.28	£40,500
			20			3.28	£81,000
			40			3.28	£162,000
			50			3.28	£202,500
4	**7,500**	**2,500**	10			2.5	£67,500
			20			2.5	£135,000
			40			**2.5**	**£270,000**
			50			2.5	£337,500

13.2.2. Balancing power and budget for the utilisation payment trials

In both Octopus Energy's recent 1-hour Saving Sessions (part of the Demand Flexibility Service) and 2-hour events during Octopus Energy's Big Dirty Turn Down (which followed CrowdFlex NIA), Octopus Energy measured an average turn-down of 0.5 kWh per 'participating household'. (This indicates 0.5 kW turn-down during Saving Sessions events, versus 0.25 kW turn-down during Big Dirty Turn Down events.) Note that these averages came from the 40-50% of customers who proactively opted in to participate in each event, meaning the turn-down among *all* Big Dirty Turn Down customers was likely closer to 0.25 kWh (over the two hours).

In both Octopus Energy's events during Octopus Energy's Big Turn Up, Octopus Energy measured an average turn-up of 1 kWh per 'participating household'. This means turn-up came from the ~23% of customers who proactively opted in to participate in each event, meaning the turn-up among all Big Turn Up customers was likely closer to 0.8 kWh.

As discussed in 9.5.1 (opt-in and opt-out), we suggest that suppliers participating in CrowdFlex do not require proactive opt-in to events by participants. Thus, we believe it would be reasonable to assume an average turn-down of 0.25 kWh per participant in the CrowdFlex programme, per 2-hour event (or equivalently, mean per-participant-hour 0.125 kW turn-down).

With these considerations in mind, we show below how different sample sizes and other parameters translate into MDES and overall trial costs associated with incentive payments to customers.

For each trial specification, we have calculated[35]:

1. A higher cost: with an average turn-down of 0.4 kWh/participant/event, and a turn-up of 0.9 kWh/participant/event. This serves as the upper-bound cost.
2. A realistic cost: with an average turn-down of 0.2 kWh/participant/event, and a turn-up of 0.6 kWh/participant/event.

Table 23: Utilisation payment trials (turn-down) MDES and incentive costs depending on various sample sizes and number of events

Parameters						**Outputs**		
Number of trial arms in mini-trial	**# Partici-pants exclud-ing control group**	**# Partici-pants per arm**	**Average incen-tive level / kWh**	**Average turn-down per partici-pant (kWh)**	**# events per mini-trial**	**MDES (% of mean)**	**MDES (kWh)**	**Cost per mini-trial**
6	21,000	4,200	£0.56 (this is a rounded average of the three payment levels)	0.2	60	0.83	0.002201	£141,120
					20	1.42	0.003805	£47,040
				0.4	60	0.83	0.002201	£282,240
					20	1.42	0.003805	£94,080
6	**41,000**	**8,200**		0.2	60	0.56	0.001481	£275,520
				0.4		0.56	0.001481	£551,040
6	62,000	12,400		0.2	60	0.47	0.001238	£416,640
				0.4		0.47	0.001238	£833,280
6	62,000	12,400		0.2	60	0.47	0.001238	£416,640
				0.4		0.47	0.001238	£833,280
7	74,400	12,400		0.2	60	0.47	0.001238	£499,968
				0.4		0.47	0.001238	£999,936

[35] We have calculated costs conservatively by assuming different customers participate in turn-down events from the ones who participate in turn-up events. In reality, some customers may serve ‘double-duty’ by sometimes being asked to turn-up and other times being asked to turn-down. However, for simplicity, these calculations ignore that potential budget saving.

Another important simplification that these calculations involve is that we ignore customers who ‘churn’ out of the trial, e.g., because they move out of their property. This may mean that these costs are slight underestimates, although there are also some potential savings associated with this churn.

Table 24: Utilisation payment trials (turn-up) MDES and incentive costs depending on various sample sizes and number of events

Parameters						Outputs		
Number of trial arms in mini-trial	**# Participants excluding control group**	**# Participants per arm**	**Average incentive level / kWh**	**Average turn-up per participant (kWh)**	**# events per mini-trial**	**MDES (% of mean)**	**MDES (kWh)**	**Cost per mini-trial**
6	**21,000**	**4,200**	£0.56 (this is a rounded average of the three payment levels)	0.6	60	0.83	0.002201	£423,360
					20	1.42	0.003805	£141,120
				0.9	60	0.83	0.002201	£635,040
					20	1.42	0.003805	£211,680
6	41,000	8,200		0.6	60	0.56	0.001481	£826,560
				0.9		0.56	0.001481	£1,239,840
6	62,000	12,400		0.6	60	0.47	0.001238	£1,249,920
				0.9		0.47	0.001238	£1,874,880
6	62,000	12,400		0.6	60	0.47	0.001238	£1,249,920
				0.9		0.47	0.001238	£1,874,880
7	74,400	12,400		0.6	60	0.47	0.001238	£1,499,904
				0.9		0.47	0.001238	£2,249,856

ESO has proposed that CrowdFlex will need to keep overall incentive costs in the range of £8-10m. Some of these incentives must go to paying 'availability payments' (see Section 13.2.1) rather than per-event utilisation payments, and some budget may need to be preserved for trials or pilots on a market-based solution for load managed areas (flexibility service 2), leaving a range of £7-8m for utilisation payments. Given the ambition to do three mini-trials during the course of CrowdFlex, each with its own 'batch' of events, this means that each mini-trial should have overall incentive costs in the range of £2-3m.

The assumed costs are sensitive to average turn-down and turn-up. If the average turn-down is lower than the 0.25 kWh per event that we expect (based on analysis of customer behaviour in Big Dirty Turn Down events), the total cost of incentive payments will be lower.

With all of these considerations in mind, our initial recommendation is for 3 batches of 60 turn-down and 20 turn-up events and with 8,200 non-waitlist-control customers in turn-down events, and 4,200 non-waitlist-control customers in the turn-up events. At this sample size, the covariate-adjusted minimum detectable effect size is approximately 1.42% (turn-up) and 0.56% (turn-down) – reasonable in the context of most energy consumption analyses, and far lower than the 20% effect size we cautiously expect.

However, we note that a much larger sample size, if somehow feasible from a cost perspective, would be preferable – especially given the consortium's interest in subgroup analyses that will create smaller analysis cells within each arm. Since we are conducting interim analyses after each mini-trial, we can flexibly adapt sample sizes for subsequent mini-trials after the first mini-trial based on interim analysis findings.

To summarise, we recommend:

1. Availability payment trials
    - At least 2,500 customers in each incentive variation.
    - Two trials, each of 12 months duration.

    Under these conditions, we would be confident in our ability to detect a 2.5 percentage point difference between the lowest incentive group and customers who do not receive any incentive (against a baseline of 10%). We estimate the incentives to then cost between £135,000 and £270,000 per trial, and between £270,000 and £540,000 across the two availability trials.
2. Utilisation payment trials
    - 60 turn-down events and 20 turn-up events per mini-trial.
    - 4 batches of events: 3 mini-trials + one ‘final’ batch of events that is more open to non-consortium flexibility service providers
    - 8,200 customers per treatment arm in each turn-down event.
    - 4,200 customers/treatment arm in each turn-up event.

    Under these conditions, we would be confident in our ability to detect a 0.56% difference in consumption between treatment and control for turn-down events, and a 1.42% difference in consumption between treatment and control for turn-up events.

    Per mini-trial, for turn-down incentives, costs would be £275,520 (assuming 0.2kWh/event average turn-down) to £551,040 (assuming 0.4kWh/event average turn-down).

    Similarly, per mini-trial, for turn-up incentives, costs would be £141,120 (assuming 0.6kWh/event average turn-up) to £211,680 (assuming 0.9kWh/event average turn-up.

    Combining costs for turn-down *and* turn-up incentives in a given mini-trial, we then estimate the incentives to cost £416,640 to £762,720 and **£1,666,560 to £3,050,880 in total across four batches of events (three mini-trials + ‘final’ batch of events).**

Combining these costs (availability payments, turn-down utilisation payments across three mini-trials plus fourth batch of events, and turn-up utilisation payments across three mini-trials plus fourth batch of events), we estimate **the total cost for incentives to be £1,936,560 to £3,590,880**.

14. Analytical strategy

14.1. Availability payment trial

Primary analysis

$$Y_{it} \sim \text{bernoulli}(p_{it}),\ p_{it} = \text{beta0} + \text{beta1 Low-Incentive}_i + \text{beta2 Medium-Incentive}_i + \text{beta3 High-Incentive}_i + \text{Date } \mathbf{tau} + \text{Covars } \mathbf{gamma}$$

- Only include customers who are in the availability payment trial.

Variable description

- Y_{it}: A binary indicator that equals 1 if asset *i* is (1) opted in to allow flexibility service providers to control the asset remotely and (2) available to be controlled by the flexibility service provider on day *t.*
  - For EVs, this means household *i*'s EV charger is plugged in between 11:30pm - 5:30am on date *t.*
  - For heat pumps, that means household *i* does not override the automatic control between [timeframe] on date *t.*
- Low-Incentive_i: A binary variable that equals 1 if asset *i* is randomised into the low incentive group.
- $\text{Medium-Incentive}_i$: A binary variable that equals 1 if asset *i* is randomised into the medium incentive group.
- High-Incentive_i: A binary variable that equals 1 if asset *i* is randomised into the high incentive group.
- Date: Date fixed effects.
- Covar: a vector of covariates (some are time invariant) that includes:
  - Baseline consumption (6-month average before the first event).
  - Location - outward postcode.
  - Whether they participated in CF: NIA, the Big Dirty Turn Down, or other past flexibility trials.
  - Whether they are participating in other concurrent flexibility trials (e.g., DFS, saving sessions).
  - Whether they have a heat pump.
  - Whether they have an EV.
  - Whether they have an at-home charger.
  - Tariff type.
  - Tenure with the energy supplier in months (dummies in 6-month bands).
- Standard errors are clustered at the customer level as customers may have multiple assets.

Coefficient of interest

- beta1, beta2, beta3 are the causal effects of low, medium and high incentives on whether asset *i* is available to be automatically controlled by the flexibility service provider, compared to not being incentivized at all.

14.2. Utilisation payment trial

This section describes the analytical approach for the first mini-trial.

Primary Analysis - Energy usage during the event window

Y_{it} = beta0 + beta1 Low-Incentive$_i$ + beta2 Medium-Incentive$_i$ + beta3 High-Incentive$_i$ + beta4 Low-Incentive x Second-Half$_t$ + beta5 High-Incentive x Second-Half$_t$ +

Event **tau**+ Covars **gamma** + e_{it}

- Only use half-hours within the event windows.
- Run the regression separately for turn-up/turn-down events.

Variable description

- Y_{it}: Household *i* electricity consumption in half-hour *t.*
- Low-Incentive$_i$: A binary variable that equals 1 if household *i* is randomised into the low incentive group.
- Medium-Incentive$_i$: A binary variable that equals 1 if household *i* is randomised into the medium incentive group.
- High-Incentive$_i$: A binary variable that equals 1 if household *i* is randomised into the high incentive group.
- Second-Half: A binary variable that equals 1 if *t* is in the 2nd half of the mini-trial.
- Event$_t$: Event fixed effects (note that these are essentially time fixed effects, and therefore apply to individuals who are not participating in the event itself).
- Covars$_{it}$: a vector of covariates (some are time invariant) that includes:
    - Baseline consumption (6-month average before the first event).
    - Hour of the day fixed effects.
    - Month of the year fixed effects.
    - Location - outward postcode.
    - Whether they participated in CF: NIA, the Big Dirty Turn Down, or other past flexibility trials.
    - Whether they are participating in other concurrent flexibility trials (e.g., DFS, saving sessions).
    - Whether they have a heat pump.
    - Whether they have an EV.
    - Tariff type.
    - Tenure with the energy supplier.
- e_{it}: error term. We will use standard errors clustered at the household level.

Coefficient of interest

- beta1, beta2, beta3 are the causal effects of low, medium and high incentives on households' energy consumption within a flexibility event respectively, compared to not being incentivized (the waitlisted control).
- beta4 represents the (marginal) carryover effect of being exposed to low incentive previously on the effect of high incentive.
- beta5 represents the (marginal) carryover effect of being exposed to high incentive previously on the effect of low incentive.

Secondary analyses

Changes in consumption before/after the event

- We will perform the same analysis in the same way as the primary analysis, except:
- We will use data 24 hours prior and 24 hours after the event window.
- Change the outcome variable:
  - Y_{it}: Household *i* average hourly electricity consumption in the 24 hours before plus the 24 hours after the event.
  - Exclude events that are fewer than 48 hours apart.

Exploratory analyses

A number of exploratory analyses will be included in the pre-registration of the trial. They include:

- We will examine how flexibility differs by characteristics of households, both of their physical structure and their occupants.
  - For example: do customers on dynamic ToU tariffs such as Octopus Agile respond (less/more) to 'response' flexibility events than non-Agile customers, given that Agile customers already provide 'routine' flexibility (as shown in CF NIA)?
- We will examine proxy measures for how many customers participated in events. We have set up events as opt-out; thus we no longer observe formal opt-in as a measure of participation. Instead, exploratory analyses will examine how many customers seem to reduce their consumption by at least X% compared to their supplier-calculated 'baseline'. We will plot this out in a cumulative probability distribution (% of participants that turned down at least X% against -10% <= X <= 150%).
- We will examine the performance of the baselining methodology (or of various baselining methodologies) by comparing non-event treatment-actual-vs-baseline difference to treatment-vs-control group difference. Well-performing baselining methodologies would show no difference between these differences.
  - Though note that nonzero differences may not be due to poor performance of the baselining methodology; because of carryovers, treatment-vs-control may be nonzero even during non-event periods. However, fortunately, this question is also one we can investigate via exploratory analysis.
  - A related exploratory analysis concerns whether treatment group customers 'game' the system by trying to increase their non-event baseline.

Understanding the effect of different event parameters

As specified in an earlier section, events will differ in their specifications (e.g., time of the event, duration of event, etc).

$$Y_{it} = \text{beta0} + \text{beta1 Low-incenive}_i + \text{beta2 Medium-incentive}_i + \text{beta3 High-incentive}_i + \text{gamma Covars}_{it} + e_{it}$$

- Only use data within the event window.
- Run the regression separately for turn-up/turn-down events.
- Exclude customers in the wait-list control group.

This specification differs from the primary analysis, as we replaced the event fixed effects with specific event characteristics.

Variable description

- Y_{it}: Household *i* electricity consumption (kWh) in half-hour *t.*
- Low-Incentive$_i$: A binary variable that equals 1 if household *i* is randomised into the low incentive group.
- Medium-Incentive$_i$: A binary variable that equals 1 if household *i* is randomised into the medium incentive group.
- High-Incentive$_i$: A binary variable that equals 1 if household *i* is randomised into the high incentive group.
- Covars$_{it}$: a vector of covariates (some are time invariant) that includes:
    - Covariates that vary over households:
        - Baseline consumption (6-month average before the first event).
        - Location - outward postcode.
        - Whether they participated in CF: NIA.
        - Whether they have a heat pump.
        - Whether they have an EV.
        - Tariff type.
    - Covariates that vary over time:
        - Month of the year dummies.
        - Hour of the day dummies.
        - Day of the week dummies.
        - Duration of the event (in hours).
        - Notice period of the event (number of hours).
        - Hours of the event.
- e_{it}: error term. We will use robust standard errors clustered at the household level

gamma is the estimated vector of interest. This is a vector of coefficients that describes the correlation between a household characteristics / event specification on energy consumption. We suggest reporting the coefficient on regardless of its statistical significance:

- Baseline consumption.
- Location dummies.
- Whether the household have a heat pump.
- Whether the household have an EV.
- Tariff type.
- Month of the year dummies.
- Duration of the event.
- Notice period of the event.
- Hour of the event.

Understanding how the effect of the treatment varies over different event parameters

We will treat turn-up and turn-down separately for the following analysis.

1. Daily/seasonal timing of trial

We will conduct separate regressions of the primary and secondary analysis for events in summer / winter (for turn-up and turn-down events separately).

2. Persistence / intensity of flexibility events

We follow Ito et al. (2018)[36] to understand persistent effects as follow:

$$Y_{it} = b0 + b1\ \text{Low-Incentive}_i + b2\ \text{Medium-Incentive} + b3\ \text{High-Incentive}$$
$$+ b_{C,incentive}\ \text{CycleC}_t \times (\text{Low-Incentive}_i + \text{Medium-Incentive}_i + \text{High-Incentive}_i)$$
$$+ \text{tau Event}_t + \text{gamma Covars}_{it} + e_{it}$$

- Y_{it}: Household *i* electricity consumption (kWh) in half-hour *t.*
- Low-Incentive$_i$: A binary variable that equals 1 if household *i* is randomised into the low incentive group.
- Medium-Incentive$_i$: A binary variable that equals 1 if household *i* is randomised into the medium incentive group.
- High-Incentive$_i$: A binary variable that equals 1 if household *i* is randomised into the high incentive group.
- C: cycle = 1 to 10. We split the events into 10 equally sized cycles (6 events each).
- CycleC$_t$: A binary variable that equals 1 if time *t* falls within cycle *C.*
- The incentive-cycle specific coefficients ($b_{C,incentive}$) represent the persistent effect throughout repeated interventions. We suggest plotting these coefficients for a graphical representation similar to Figure 3 in Ito et al. (2018).

3. Firmness of supply

Our interpretation of this question is that it's not a question for randomisation or even quasi-experimental methods to attack. Instead, we assume CNZ's data science team will examine how well they can predict curtailment (or turn-up) magnitude of held-back events based on training data events (i.e., WP8).

4. Location

We will conduct separate regressions of the primary analysis for participants in different geographic GSP groups (for turn-up and turn-down events separately).[37]

5. Survey outcomes

We will conduct a pre-post analysis to understand whether customer's satisfaction with their energy supplier and their subjective comfort changes after CrowdFlex:

$$Y_{it} = b0 + b1\ \text{post}_t + \text{gamma Covars}_{it} + e_{it}$$

- Y_{it} : Household *i*'s self-reported outcome in survey *t*, including:
    - Satisfaction: self-reported satisfaction with participation in CrowdFlex, reported in a Likert scale.
    - Subjective comfort: subjective comfort measure, reported in a Likert scale.
    - Technology adoption: whether customer *i* adopts:
        - A smart thermostat.
        - A heat pump.
        - An electric vehicle.

---

[36] Equation (2) from Ito, K., Ida, T., & Tanaka, M. (2018). Moral suasion and economic incentives: Field experimental evidence from energy demand. *American Economic Journal: Economic Policy*, *10*(1), 240-67.

[37] North Scotland, Central and Southern Scotland, North East England, North West England, Yorkshire, Merseyside, Cheshire, North Wales and North Shropshire, East Midlands, West Midlands, South Wales & South West England, Eastern England, Southern England, London, South East England, Northern Ireland

- $post_t$: binary variable that indicates whether *t* happens after CrowdFlex.
- $Covars_{it}$: a vector of covariates that includes:
    - Covariates that vary over households:
        - Tenure with the energy suppliers.
        - Location - outward postcode.
        - Whether they have a heat pump.
        - Whether they have an EV.
        - Tariff type.
        - Whether they consent to automatic EV charging.
    - Covariates that vary household and time:
        - Month of the year when household *i* responds to the survey (dummies).
        - Number of CrowdFlex events that household *i* were invited to by survey *t.*
- e_{it}: error term. We will use robust standard errors clustered at the household level.

Missing Data

Since participants are customers of energy suppliers who use their smart meter data for legitimate interests, we will still be able to track their energy consumption even if they decide to opt out of the trial. We only expect missing data for the following situations:

- The participant switches energy supplier.
- The participant moves out of the property.
- The participant explicitly asks to stop sharing their energy consumption data.
- A faulty meter reading.

We will not impute for participants' consumption if they fall into reasons 1-3. However, we believe that missing data due to reason (4) will likely be balanced across treatment groups, hence this does not threaten the internal validity of the analytical strategy. We will use predictive mean matching with consumption in other hours (up to 12 months before trial start) and trial arms predictors, creating the number of imputed datasets that equals the % of missing observations, and pooling the estimates from each. We will impute missing values using sequential multiple imputation). We will impute read periods containing missing reads, negative reads, and reads that imply consumption of more than X kWh per hour. We assume that very high reads reflect misreads.

14.3. Graphical analysis

To understand how the impact of the events varies based on number of events, we run a similar analysis to Allcott and Rogers (2014)[38]. For each hour *h*, the regression is:

$$Y_{it} = beta^{h}\ Treat_{i} + tau_{t} + gamma\ Covars_{it} + e_{it},\ \forall\ t \in [h - 12, h + 12]$$

- Use all data within the experimental period and one month before the first event.
- The regression is at the hour level, that is, we are conducting (24 x Duration of CrowdFlex in days) + (24 x 31) regressions.

[38] Allcott, H., & Rogers, T. (2014). The short-run and long-run effects of behavioral interventions: Experimental evidence from energy conservation. *American Economic Review*, *104*(10), 3003-37. Equation (1), and Figure (2).

Variable description

- Y_{it}: Household *i* electricity consumption in hour *t.*
- $Treat_i$: A vector indicating which treatment group household *i* is randomised to, base group will be the control group.
- tau_t : Hour fixed effects (note that these are essentially time fixed effects, and therefore apply to individuals who are not participating in the event itself).
- $Covars_{it}$: a vector of covariates that includes:
  - Baseline consumption (6-month average before the first event).
  - Hour of the day fixed effects.
  - Location - outward postcode.
  - Whether they participated in CF: NIA.
  - Whether they have a heat pump.
  - Whether they have an EV.
  - Whether they have an at-home charger.
  - Tariff type.
  - Tenure with the energy supplier.
- e_{it}: error term. We will use robust standard errors clustered at the household level.

<u>Coefficient of interest</u>

$beta^h$ represents the average treatment effect (kWh / hour) at different hours in the trial. We will plot $beta^h$ against time.

15. Ethical issues & review

Table 25: Ethical risk level summary

This trial was self-assessed as being:	Low risk
The reason for assessment was:	The precedent has been set by previous trials, in particular CrowdFlex NIA.
Link to completed ethics review form (if medium or high level risk)	

What were the key ethical considerations for the project?

- Customers excessively ration electricity consumption during events, leading to harm.
- Customers fail to understand that "turn-up" event electricity still has a cost (it is not "free" and therefore extra to their usual consumption), leading to increased overall usage and therefore higher net bills. For LIV customers, this could therefore lead to a worsening of their situation, rather than an improvement.
- Sudden turn-on and turn-off damage heat pump or EV equipment.
- Unsocialised domestic flexibility disproportionately benefits high income households.

Did you seek informed consent from participants? Yes, customers have to opt-in to participate in the trial.

Have you considered whether harms to participants might arise and how you will deal with them?

1. Previous research found evidence of customers with health conditions engaging in self-rationing strategies such as (1) keeping one room in the house warm, (2) only boiling one cup of water at a time (3) using thermoses to keep water warm and (4) using hot water bottles to keep warm if ill – regardless of participation in energy flexibility events.[39]
    a. Insofar as demand response could lead to self-rationing, we need to ensure our recruitment materials steer customers away from engaging in demand response if doing so would put them at risk, as discussed in CrowdFlex deliverable D2.2 and a subsequent interview with a specialist in flexibility trials at Citizens Advice.
    b. We will conduct surveys to collect data on customer experience, comfort, understanding of the flexibility of event logistics, understanding of the purpose of flexibility, views on and satisfaction with event logistics, and the behaviours customers report employing to shift their demand. These surveys will be sent at various points during the trial to a random subset of customers – including at the beginning (to obtain baselines) and end of the CrowdFlex program, but also at various points during the program. Ideally, survey completion would be compensated to encourage high participation.
2. The trial will involve automation. Customers will generally never be 'out of pocket' for participating in flexibility events – they can only *earn* financial compensation, not lose money, by participating in events. The automation involved in this trial will likely concern charging times for batteries, especially electric vehicle batteries.

[39] Citizens Advice (2020): Understanding high and low electricity usage. Retrieved from: https://www.citizensadvice.org.uk/about-us/our-work/policy/policy-research-topics/energy-policy-research-and-consultation-responses/energy-policy-research/understanding-high-and-low-electricity-usage/

a. We allow customers to override the automation anytime. However, a Citizens Advice report on demand response notes that automation in any form poses questions of who manages the risk of mishaps, who is accountable, and who has control over appliances.[40] The Citizens Advice authors note: 'A key protection will be the provision of an override function that is easy to identify and use, and can be used at the consumer's discretion without incurring a financial penalty other than forgoing the incentive to use off-peak power.' Insofar as this might limit the 'effectiveness' of automated demand response, they argue that 'the essential principle is that automation should be a tool to help consumers shift their usage if they wish to do so, not a way of forcing them to'. This principle underpinned CrowdFlex:NIA and should continue to underpin the automation involved in CrowdFlex.
b. We recommend conducting accessibility testing prior to full trial launch of CrowdFlex, too.

[40] Citizens Advice (2014): Take a walk on the demand side, Retrieved from: https://www.citizensadvice.org.uk/about-us/our-work/policy/policy-research-topics/energy-policy-research-and-consultation-responses/energy-policy-research/take-a-walk-on-the-demand-side/

16. Risks

Table 26: Key risks and mitigations

Risk	Strategy to mitigate risk	Responsibility	Timeframe (if applicable)
Customers excessively ration electricity consumption during events, leading to harm.	We will conduct an interim survey to understand whether there is a backfire (e.g., severely reduced comfort and customer satisfaction).		After event X
Randomisation failure.	We will seed the energy suppliers' email addresses into each trial arm to ensure that emails are going out as expected.		Start of each mini-trial
Technology failure (e.g., Ohme charger).	Accessibility testing prior to full launch to ensure (1) opt-out is easy and (2) the potential harm of automatic charging (e.g., depreciation of the battery) is communicated appropriately.		Prior to launch
Missing data.	Conduct trial runs in the start phase.		Start phase

Appendices

17. Appendix A: Mini-trial 1 example event schedule

event_num	mini_trial	date	time start	time end	Turn up/down	Regions
1	1	01/11/2023	18:00	20:00	Down	All regions
2	1	04/11/2023	17:00	19:00	Down	All regions
3	1	06/11/2023	11:00	13:00	Down	Just 4 regions
4	1	09/11/2023	20:00	22:00	Down and Up (depending on region)	Just 2 regions
5	1	13/11/2023	13:00	15:00	Down	All regions
6	1	15/11/2023	18:00	20:00	Down	6 regions
7	1	16/11/2023	19:00	21:00	Down	All regions
8	1	19/11/2023	18:00	20:00	Down	All regions
9	1	21/11/2023	17:00	19:00	Down and Up (depending on region)	All regions
10	1	24/11/2023	11:00	13:00	Down	All regions
11	1	28/11/2023	20:00	22:00	Down	Just 4 regions
12	1	30/11/2023	13:00	15:00	Down and Up (depending on region)	Just 2 regions
13	1	01/12/2023	18:00	20:00	Down	All regions
14	1	04/12/2023	19:00	21:00	Down and Up (depending on region)	6 regions
15	1	06/12/2023	18:00	20:00	Down	All regions
16	1	09/12/2023	17:00	19:00	Down	All regions
17	1	13/12/2023	11:00	13:00	Down and Up (depending on region)	Just 4 regions
18	1	15/12/2023	20:00	22:00	Down	Just 2 regions
19	1	16/12/2023	13:00	15:00	Down	All regions
20	1	19/12/2023	18:00	20:00	Down and Up (depending on region)	6 regions
21	1	21/12/2023	19:00	21:00	Down	All regions
22	1	24/12/2023	18:00	20:00	Down	All regions
23	1	28/12/2023	17:00	19:00	Down	Just 4 regions
24	1	30/12/2023	11:00	13:00	Down	Just 2 regions
25	1	31/12/2023	20:00	22:00	Down and Up (depending on region)	All regions
26	1	03/01/2024	13:00	15:00	Down	6 regions
27	1	05/01/2024	18:00	20:00	Down	All regions
28	1	08/01/2024	19:00	21:00	Down and Up (depending on region)	All regions
29	1	12/01/2024	18:00	20:00	Down and Up (depending on region)	Just 4 regions

event_num	mini_trial	date	time start	time end	Turn up/down	Regions
30	1	14/01/2024	17:00	19:00	Down and Up (depending on region)	Just 2 regions
31	1	15/01/2024	11:00	13:00	Down	All regions
32	1	18/01/2024	20:00	22:00	Down	6 regions
33	1	20/01/2024	13:00	15:00	Down and Up (depending on region)	All regions
34	1	23/01/2024	18:00	20:00	Down	All regions
35	1	27/01/2024	19:00	21:00	Down	Just 4 regions
36	1	29/01/2024	18:00	20:00	Down	Just 2 regions
37	1	30/01/2024	17:00	19:00	Down and Up (depending on region)	All regions
38	1	02/02/2024	11:00	13:00	Down	6 regions
39	1	04/02/2024	20:00	22:00	Down	All regions
40	1	07/02/2024	13:00	15:00	Down	All regions
41	1	11/02/2024	18:00	20:00	Down	Just 4 regions
42	1	13/02/2024	19:00	21:00	Down and Up (depending on region)	Just 2 regions
43	1	14/02/2024	18:00	20:00	Down	All regions
44	1	17/02/2024	17:00	19:00	Down and Up (depending on region)	10 regions
45	1	19/02/2024	11:00	13:00	Down	All regions
46	1	22/02/2024	20:00	22:00	Down and Up (depending on region)	All regions
47	1	26/02/2024	13:00	15:00	Down	Just 4 regions
48	1	28/02/2024	18:00	20:00	Down and Up (depending on region)	Just 2 regions
49	1	02/03/2024	19:00	21:00	Down	All regions
50	1	04/03/2024	18:00	20:00	Down	6 regions
51	1	07/03/2024	17:00	19:00	Down and Up (depending on region)	All regions
52	1	11/03/2024	11:00	13:00	Down	All regions
53	1	13/03/2024	20:00	22:00	Down	Just 4 regions
54	1	14/03/2024	13:00	15:00	Down and Up (depending on region)	10 regions
55	1	17/03/2024	18:00	20:00	Down	All regions
56	1	19/03/2024	19:00	21:00	Down	6 regions
57	1	22/03/2024	18:00	20:00	Down and Up (depending on region)	All regions

event_num	mini_trial	date	time start	time end	Turn up/down	Regions
58	1	26/03/2024	17:00	19:00	Down and Up (depending on region)	All regions
59	1	28/03/2024	11:00	13:00	Down	All regions
60	1	29/03/2024	20:00	22:00	Down	All regions

18. Appendix B: Fake events schedules for power calculations

Table 27: Events schedule we used for 20-event mini-trials

event_num	mini_trial	date	time start	time end
1	1	01/06/2021	18:00	20:00
2	1	11/06/2021	17:00	19:00
3	1	16/06/2021	11:00	13:00
4	1	23/06/2021	20:00	22:00
5	1	05/07/2021	13:00	15:00
6	1	15/07/2021	18:00	20:00
7	1	20/07/2021	19:00	21:00
8	1	27/07/2021	18:00	20:00
9	1	08/08/2021	17:00	19:00
10	1	18/08/2021	11:00	13:00
11	1	23/08/2021	20:00	22:00
12	1	30/08/2021	13:00	15:00
13	1	11/09/2021	18:00	20:00
14	1	21/09/2021	19:00	21:00
15	1	26/09/2021	18:00	20:00
16	1	03/10/2021	17:00	19:00
17	1	15/10/2021	11:00	13:00
18	1	25/10/2021	20:00	22:00
19	1	30/10/2021	13:00	15:00
20	1	06/11/2021	18:00	20:00
21	2	01/01/2022	19:00	21:00
22	2	11/01/2022	18:00	20:00
23	2	16/01/2022	17:00	19:00
24	2	23/01/2022	11:00	13:00
25	2	04/02/2022	20:00	22:00
26	2	14/02/2022	13:00	15:00
27	2	19/02/2022	18:00	20:00
28	2	26/02/2022	19:00	21:00
29	2	10/03/2022	18:00	20:00
30	2	20/03/2022	17:00	19:00
31	2	25/03/2022	11:00	13:00
32	2	01/04/2022	20:00	22:00
33	2	13/04/2022	13:00	15:00
34	2	23/04/2022	18:00	20:00
35	2	28/04/2022	19:00	21:00
36	2	05/05/2022	18:00	20:00
37	2	17/05/2022	17:00	19:00
38	2	27/05/2022	11:00	13:00
39	2	28/05/2022	20:00	22:00
40	2	29/05/2022	13:00	15:00
41	3	01/06/2022	18:00	20:00
42	3	11/06/2022	19:00	21:00
43	3	16/06/2022	18:00	20:00
44	3	23/06/2022	17:00	19:00
45	3	05/07/2022	11:00	13:00
46	3	15/07/2022	20:00	22:00
47	3	20/07/2022	13:00	15:00
48	3	27/07/2022	18:00	20:00
49	3	08/08/2022	19:00	21:00
50	3	18/08/2022	18:00	20:00

51	3	23/08/2022	17:00	19:00
52	3	30/08/2022	11:00	13:00
53	3	11/09/2022	20:00	22:00
54	3	21/09/2022	13:00	15:00
55	3	26/09/2022	18:00	20:00
56	3	24/10/2022	19:00	21:00
57	3	03/11/2022	18:00	20:00
58	3	08/11/2022	17:00	19:00
59	3	15/11/2022	11:00	13:00
60	3	27/11/2022	20:00	22:00

Table 28: Events schedule we used for 60-event mini-trials

event_num	mini_trial	date	time start	time end
1	1	01/06/2021	18:00	20:00
2	1	04/06/2021	17:00	19:00
3	1	06/06/2021	11:00	13:00
4	1	09/06/2021	20:00	22:00
5	1	13/06/2021	13:00	15:00
6	1	15/06/2021	18:00	20:00
7	1	16/06/2021	19:00	21:00
8	1	19/06/2021	18:00	20:00
9	1	21/06/2021	17:00	19:00
10	1	24/06/2021	11:00	13:00
11	1	28/06/2021	20:00	22:00
12	1	30/06/2021	13:00	15:00
13	1	01/07/2021	18:00	20:00
14	1	04/07/2021	19:00	21:00
15	1	06/07/2021	18:00	20:00
16	1	09/07/2021	17:00	19:00
17	1	13/07/2021	11:00	13:00
18	1	15/07/2021	20:00	22:00
19	1	16/07/2021	13:00	15:00
20	1	19/07/2021	18:00	20:00
21	1	21/07/2021	19:00	21:00
22	1	24/07/2021	18:00	20:00
23	1	28/07/2021	17:00	19:00
24	1	30/07/2021	11:00	13:00
25	1	31/07/2021	20:00	22:00
26	1	03/08/2021	13:00	15:00
27	1	05/08/2021	18:00	20:00
28	1	08/08/2021	19:00	21:00
29	1	12/08/2021	18:00	20:00
30	1	14/08/2021	17:00	19:00
31	1	15/08/2021	11:00	13:00
32	1	18/08/2021	20:00	22:00
33	1	20/08/2021	13:00	15:00
34	1	23/08/2021	18:00	20:00
35	1	27/08/2021	19:00	21:00
36	1	29/08/2021	18:00	20:00
37	1	30/08/2021	17:00	19:00
38	1	02/09/2021	11:00	13:00
39	1	04/09/2021	20:00	22:00
40	1	07/09/2021	13:00	15:00

41	1	11/09/2021	18:00	20:00
42	1	13/09/2021	19:00	21:00
43	1	14/09/2021	18:00	20:00
44	1	17/09/2021	17:00	19:00
45	1	19/09/2021	11:00	13:00
46	1	22/09/2021	20:00	22:00
47	1	26/09/2021	13:00	15:00
48	1	28/09/2021	18:00	20:00
49	1	01/10/2021	19:00	21:00
50	1	03/10/2021	18:00	20:00
51	1	06/10/2021	17:00	19:00
52	1	10/10/2021	11:00	13:00
53	1	12/10/2021	20:00	22:00
54	1	13/10/2021	13:00	15:00
55	1	16/10/2021	18:00	20:00
56	1	18/10/2021	19:00	21:00
57	1	21/10/2021	18:00	20:00
58	1	25/10/2021	17:00	19:00
59	1	27/10/2021	11:00	13:00
60	1	28/10/2021	20:00	22:00
61	2	01/01/2022	18:00	20:00
62	2	04/01/2022	17:00	19:00
63	2	06/01/2022	11:00	13:00
64	2	09/01/2022	20:00	22:00
65	2	13/01/2022	13:00	15:00
66	2	15/01/2022	18:00	20:00
67	2	16/01/2022	19:00	21:00
68	2	19/01/2022	18:00	20:00
69	2	21/01/2022	17:00	19:00
70	2	24/01/2022	11:00	13:00
71	2	28/01/2022	20:00	22:00
72	2	30/01/2022	13:00	15:00
73	2	31/01/2022	18:00	20:00
74	2	03/02/2022	19:00	21:00
75	2	05/02/2022	18:00	20:00
76	2	08/02/2022	17:00	19:00
77	2	12/02/2022	11:00	13:00
78	2	14/02/2022	20:00	22:00
79	2	15/02/2022	13:00	15:00
80	2	18/02/2022	18:00	20:00
81	2	20/02/2022	19:00	21:00
82	2	23/02/2022	18:00	20:00
83	2	27/02/2022	17:00	19:00
84	2	01/03/2022	11:00	13:00
85	2	02/03/2022	20:00	22:00
86	2	05/03/2022	13:00	15:00
87	2	07/03/2022	18:00	20:00
88	2	10/03/2022	18:00	20:00
89	2	14/03/2022	17:00	19:00
90	2	16/03/2022	11:00	13:00
91	2	17/03/2022	20:00	22:00
92	2	20/03/2022	13:00	15:00
93	2	22/03/2022	18:00	20:00
94	2	25/03/2022	19:00	21:00

95	2	29/03/2022	18:00	20:00
96	2	31/03/2022	17:00	19:00
97	2	01/04/2022	11:00	13:00
98	2	04/04/2022	20:00	22:00
99	2	06/04/2022	13:00	15:00
100	2	09/04/2022	18:00	20:00
101	2	13/04/2022	19:00	21:00
102	2	15/04/2022	18:00	20:00
103	2	16/04/2022	17:00	19:00
104	2	19/04/2022	11:00	13:00
105	2	21/04/2022	20:00	22:00
106	2	24/04/2022	13:00	15:00
107	2	28/04/2022	18:00	20:00
108	2	30/04/2022	19:00	21:00
109	2	03/05/2022	18:00	20:00
110	2	05/05/2022	17:00	19:00
111	2	08/05/2022	11:00	13:00
112	2	12/05/2022	20:00	22:00
113	2	14/05/2022	13:00	15:00
114	2	15/05/2022	18:00	20:00
115	2	18/05/2022	18:00	20:00
116	2	20/05/2022	17:00	19:00
117	2	23/05/2022	11:00	13:00
118	2	27/05/2022	20:00	22:00
119	2	29/05/2022	13:00	15:00
120	2	30/05/2022	18:00	20:00
121	3	01/06/2022	19:00	21:00
122	3	04/06/2022	18:00	20:00
123	3	06/06/2022	17:00	19:00
124	3	09/06/2022	11:00	13:00
125	3	13/06/2022	20:00	22:00
126	3	15/06/2022	13:00	15:00
127	3	16/06/2022	18:00	20:00
128	3	19/06/2022	19:00	21:00
129	3	21/06/2022	18:00	20:00
130	3	24/06/2022	17:00	19:00
131	3	28/06/2022	11:00	13:00
132	3	30/06/2022	20:00	22:00
133	3	01/07/2022	13:00	15:00
134	3	04/07/2022	18:00	20:00
135	3	06/07/2022	19:00	21:00
136	3	09/07/2022	18:00	20:00
137	3	13/07/2022	17:00	19:00
138	3	15/07/2022	11:00	13:00
139	3	16/07/2022	20:00	22:00
140	3	19/07/2022	13:00	15:00
141	3	21/07/2022	18:00	20:00
142	3	24/07/2022	18:00	20:00
143	3	28/07/2022	17:00	19:00
144	3	30/07/2022	11:00	13:00
145	3	31/07/2022	20:00	22:00
146	3	03/08/2022	13:00	15:00
147	3	05/08/2022	18:00	20:00
148	3	08/08/2022	19:00	21:00

149	3	12/08/2022	18:00	20:00
150	3	14/08/2022	17:00	19:00
151	3	15/08/2022	11:00	13:00
152	3	18/08/2022	20:00	22:00
153	3	20/08/2022	13:00	15:00
154	3	23/08/2022	18:00	20:00
155	3	27/08/2022	19:00	21:00
156	3	29/08/2022	18:00	20:00
157	3	30/08/2022	17:00	19:00
158	3	02/09/2022	11:00	13:00
159	3	04/09/2022	20:00	22:00
160	3	07/09/2022	13:00	15:00
161	3	11/09/2022	18:00	20:00
162	3	13/09/2022	19:00	21:00
163	3	14/09/2022	18:00	20:00
164	3	17/09/2022	17:00	19:00
165	3	19/09/2022	11:00	13:00
166	3	22/09/2022	20:00	22:00
167	3	26/09/2022	13:00	15:00
168	3	28/09/2022	18:00	20:00
169	3	01/10/2022	18:00	20:00
170	3	03/10/2022	17:00	19:00
171	3	06/10/2022	11:00	13:00
172	3	10/10/2022	20:00	22:00
173	3	12/10/2022	13:00	15:00
174	3	13/10/2022	18:00	20:00
175	3	16/10/2022	19:00	21:00
176	3	18/10/2022	18:00	20:00
177	3	21/10/2022	17:00	19:00
178	3	25/10/2022	11:00	13:00
179	3	27/10/2022	20:00	22:00
180	3	28/10/2022	13:00	15:00