A Demand-Side Alternative to Renewable Curtailment: Natural Field Experimental Evidence from Two Countries*

Daniel Lopez Garcia
Octopus Energy España

Robert D. Metcalfe
Columbia University
Centre for Net Zero & NBER

Andrew R. Schein
Centre for Net Zero

Yixin Sun
Centre for Net Zero

July 2026

Abstract

As renewable generation plays a growing role in the energy system, periods of negative wholesale electricity prices are becoming increasingly common and are expected to grow dramatically over the next ten years. Traditionally, system operators respond by paying generators, especially renewables, to curtail production, or mandating that they do so without compensation. In this study, we propose and test a novel alternative: exposing consumers to free and negative prices to stimulate demand when supply is abundant. We implemented large-scale nationwide natural field experiments simultaneously in Great Britain and Spain, with roughly 60,000 residential customers in each country randomized to receive varying financial incentives to "turn up" their electricity. We found that demand increased substantially as prices fell to zero, but paying customers to consume beyond zero price yielded little additional response. Households with electric vehicles and rooftop solar had a larger elasticity than households without such technologies, suggesting demand turn-up becomes more effective as households adopt low-carbon technologies. In Great Britain, consumption was largely shifted from adjacent hours, while in Spain the increase appeared to represent net new demand. We develop a welfare framework showing conditions under which demand turn-up improves on curtailment; we find that by inducing additional consumption in periods and locations where the marginal cost of supply is low, demand turn-up generates welfare gains that curtailment leaves uncaptured.

*We thank Joe Grainger, from Centre for Net Zero (CNZ), for his critical logistics and project management. We thank Jack Ashworth, Greg Baxter, Tessa Hillen, and Vania Milkova, from Octopus Energy Limited (OEL) in Great Britain, and Claudia Cuñat, Anna Casas Llopart, and Iñigo Vallejo, from Octopus Energy España, S.L.U. (OEE), for their support and partnership in this project. Their commitment to the implementation of the intervention was remarkable. We thank Navya Kumar for excellent research assistance. We thank Andy Hackett, from CNZ, for specific comments on the text of the paper itself. We thank Sophie Alderman, Louise Bernard, Emma Burns, Natalia Fabra, Simon Gill, Marten Ovaere, Iona Penman, Marcia Poletti, Alex Schoch, Celyn Simmonds, Phil Steele, Melita Van Steenberghe, Izzy Woolgar, and Lucy Yu for their valuable insights and comments. Conflicts of Interest: AS and YS are paid employees of Octopus Energy Group Limited (OEG), working as researchers in CNZ. DLG was a paid employee of CNZ at the time of trial implementation, and is now a paid employee of OEE. RM is a consultant to CNZ. CNZ operates as an autonomous open research institute, but CNZ, OEL, and OEE are all part of OEG. The four authors made final decisions regarding experimental design, data analysis, and writing; all mistakes herein are ours alone. The pre-registrations and pre-analysis plans for the three field experiments are: AEARCTR-0015455, AEARCTR-0016220, and AEARCTR-0016221.

1 Introduction

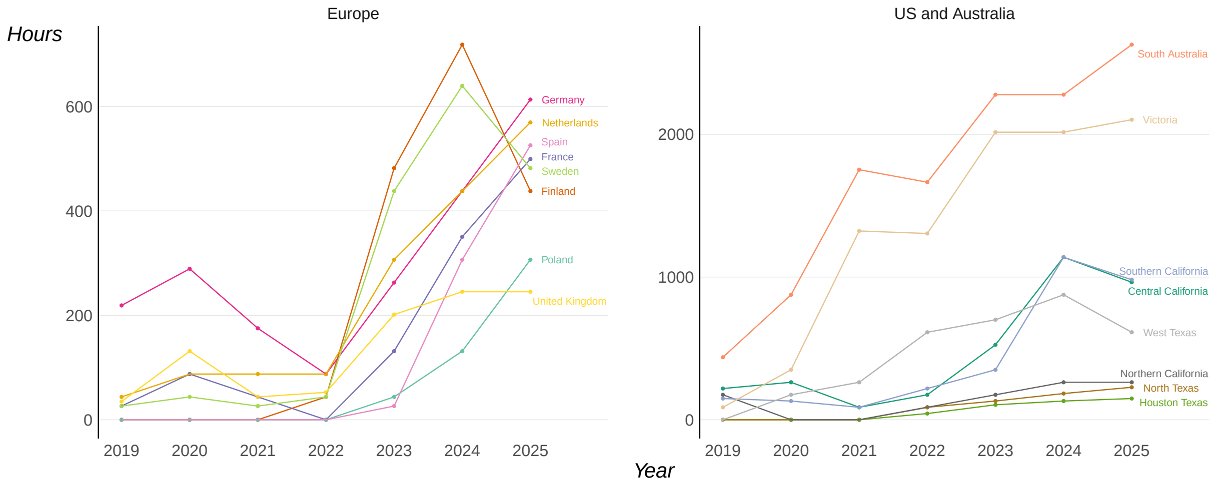
Within the coming decades, most major economies are expected to have more renewable energy generation than fossil fuel generation, and enough renewable resource that potential output could exceed total demand for many parts of the year (i.e., excess supply). The central challenge for the management of energy systems is not just whether renewable capacity can be built, but how to integrate large volumes of variable generation into power systems designed for stable, dispatchable plants. As renewable energy generation increases, power systems around the world will confront the same question: how to align demand with weather-dependent supply, and whether renewable generation be spilled (i.e., not used) in order to balance supply and demand?

One symptom of this new phase is the growing frequency of negative wholesale price episodes (Figure 1). Negative wholesale prices arise as an equilibrium outcome when supply is inflexible and demand is insufficiently responsive, amplified by non-convex generation technologies (Wolak, 2011) and by output-based support schemes that raise generators' effective marginal revenue, leading them to keep producing even when prices turn negative (Jacobsen and Schröder, 2013). The US, the UK, Spain, Australia, Germany, the Netherlands, France, and Finland – all countries with high renewable penetration – have rising hours of the year with negative prices, and both the number of hours and the number of affected markets are expected to climb steeply.

However, negative prices are only the most visible face of this mismatch. Where wholesale markets clear at a national price, as in Great Britain and Spain, transmission constraints rarely push that price below zero. Instead, the system operator manages them through redispatch, reducing generation behind the constraint and increasing generation elsewhere. In Great Britain this is achieved through payments to generators, whereas in Spain generators are instructed to curtail without compensation. Behind such a constraint, renewable energy is locally abundant and its opportunity cost is close to zero; it would otherwise be spilled. Yet the national price that all demand faces remains positive. Negative prices and constrained surplus are thus two expressions of the same underlying mismatch between abundant variable supply and inflexible demand. However, they differ in a way that matters for policy: in the first, consumers are not exposed or responsive to a price signal that already reflects system conditions; in the second, the price seen by consumers, and even retailers, fails to reflect local system conditions altogether.

In both cases, demand is rarely exposed to the true value of energy, and system op-

Figure 1: Hours with negative wholesale electricity prices
Selected regions in Europe, US, and Australia



Notes: Data from International Energy Agency on number of hours with negative wholesale prices in selected regions from 2019 to 2025.

erators have relied on curtailment to manage the surplus. Policymakers are increasingly exploring demand-side flexibility as an alternative that absorbs surplus electricity rather than discarding it. In Australia, the federal government has announced a program where retailers must offer households at least three hours of free daytime electricity to match solar generation (Australian Government Department of Climate Change, Energy, the Environment and Water, 2026). In Great Britain, the National Energy System Operator and regional Distribution System Operators have introduced dedicated "demand turn-up" (DTU) flexibility services that pay consumers to increase demand when renewable generation is abundant; the distribution-level services are locational by construction, procuring flexibility in specific constrained areas. In addition, the UK Department for Energy Security and Net Zero (2026) is launching a new trial to test removing levies (non-energy costs) during periods of high wind supply. However, without credible causal estimates of consumer responsiveness, policymakers cannot reliably evaluate whether turn-up programs are a cost-effective or welfare-enhancing substitute for curtailment. We also do not have a theoretical framework to determine the conditions under which such a policy improves economic welfare.

To address these gaps, we implemented twin nationwide natural field experiments in partnership with Octopus Energy Limited in Great Britain and Octopus Energy España, S.L.U. in Spain. We conducted ten pre-specified "turn-up" events, five in each country, between June and September 2025. Events were triggered when day-ahead wholesale electricity prices were forecast to be near zero or negative, aligning the interventions with periods of anticipated excess renewable generation.

The natural field experiments were conducted at scale, allowing precise estimation of consumption responses under real market conditions. We enrolled roughly 120,000 eligible residential customers, split across the two countries. During each event, customers were randomized into one of several treatments: discount on their electricity use, free electricity, negative prices that paid customers to consume additional electricity, a monetary prize draw (in Spain only), or a control group. Customers were informed of the event timing and their assigned incentive 24 hours in advance via email. The treatment events lasted one hour in Great Britain and two hours in Spain (the difference in length being due to differing requirements from the respective delivery partners).

We also ran a separate complementary natural field experiment, responding to localized network constraints rather than national wholesale conditions. Northern Powergrid (NPG), a Distribution Network Operator in northern England which serves 3.9 million households, faces recurring periods where local renewable generation exceeds what the distribution network can accommodate. NPG procured demand turn-up services from Octopus Energy customers in two constrained grid areas, triggering events based on real-time network conditions between January and March 2025. Unlike the two nationwide field experiments above, customers first signed up to the program and then opted into individual events. We randomized invitations among 14,477 enrolled customers across 15 event days, offering free electricity for the event duration. This field experiment provides evidence on two additional questions: whether residential demand flexibility can address localized grid constraints, and how requiring sign-up and event-level opt-in affects the volume and cost-effectiveness of turn-up achieved.

Together, these experiments separated the two over-supply issues noted above. The two nationwide trials observed behavior under system-wide abundance: because events were called when the national price signaled genuine surplus, they identified how much, and under which incentives, households shift consumption into periods when energy is plentiful. This response provides an estimate of consumer flexibility that does not hinge on the conditions of any single event. The NPG trial instead observed turn-up where the binding constraint was local network capacity rather than the national balance. Our welfare analysis looks primarily at the latter case – increasing demand where surplus is locally trapped and would otherwise be spilled, so that the energy absorbed carries an opportunity cost close to zero.

1.1 Primary findings

First, electricity demand increased substantially when prices fell from the standard rate to a 50% discount and then to free electricity, but the marginal response of moving from free to negative prices was small. In Great Britain, offering a 50% discount increased hourly electricity consumption by 0.049 kW (14.3% relative to control). Making electricity free more than doubled this effect, increasing consumption by 0.105 kW (31%). Adding a 5 pence per kWh payment on top of free electricity produced an additional increase to 0.115 kW (33.8%), but raising the payment to 15 pence per kWh did not generate further meaningful gains (0.116 kW, 34%).

Expressed as point elasticities evaluated at the control price, these effects correspond to an elasticity of -0.29 under a 50% discount, -0.31 when electricity is made free, -0.28 for free plus 5p (a 125% price reduction relative to baseline), and -0.21 for free plus 15p (175%). In summary, responsiveness was highest when customers were offered free electricity.[1]

Spanish customers exhibited the same qualitative pattern, with smaller magnitudes. A 50% discount increased hourly consumption by 0.026 kW (6%), while free electricity increased consumption by 0.053 kW (12.7%). Adding a 10 euro cent payment raised the effect only slightly to 0.056 kW (13.4%). The prize draw produced an effect comparable to the 50% discount. The corresponding point elasticities are -0.12 under a 50% discount, -0.12 when electricity is free, and -0.07 under free plus 10 cents.

The results from the localized constraints natural field experiment with NPG were closely aligned with the national British field experiment results. Among customers who signed up to the program and were notified to participate in events, in-event consumption increased by 0.411 kW. About 20.4% of customers originally invited to the program had signed up, implying an effect of approximately 0.084 kW per invited customer. This is similar in magnitude to the 0.105 kW effect observed in the free electricity group in the national British trial, where all customers were invited without any sign-up requirement. The sign-up design was more cost-effective than the non-targeted approach in the national British trial: by restricting payments to customers who actively chose to participate, it avoided compensating customers who were unaware of events or did not change

[1] We compute point elasticities by dividing the proportional change in consumption, measured as the treatment coefficient relative to mean control-group consumption, by the proportional change in price relative to the treatment-specific baseline price. Because these elasticities are evaluated over large experimental price changes, they should be interpreted as linear approximations to demand over the relevant price interval rather than local responses to infinitesimal price changes.

their behavior, reducing the payments per kWh of induced demand from £1.22 in the national British trial to £0.77 in the NPG trial.

Second, we found some meaningful and significant moderators to the price elasticity of demand. Customers on tariffs for electric vehicles and rooftop solar exhibited substantially larger treatment effects than households without these technologies in Great Britain.[2] The pattern was particularly pronounced among electric vehicle owners, who can rapidly increase charging during event windows. Similarly, customers on time-of-use tariffs responded more strongly than those on flat tariffs. These findings indicate that demand turn-up is complementary to low-carbon technology: as households adopt flexible electric technologies, the potential to align demand with periods of abundant renewable supply increases.

Third, in Great Britain we observed evidence of intertemporal substitution, whereby customers shift consumption from surrounding hours into the event window. As a result, the increase in demand during turn-up events was largely offset by reductions before or after the event, yielding no statistically significant net increase in total consumption. In contrast, in Spain we observed net demand creation, with event-period increases not offset in adjacent hours. This distinction has important economic and environmental implications. If consumption in Great Britain were reallocated away from higher-price, higher-emissions periods toward hours with abundant renewable supply, such substitution may reduce procurement costs and lower the carbon intensity of electricity use.

Fourth, we developed a partial welfare framework to evaluate when DTU improves welfare relative to curtailment in electricity systems with surplus generation and network constraints. The model combines our experimental estimates of short-run demand responsiveness with institutional features of electricity markets, including subsidy design, curtailment compensation, and system operator procurement rules. We show that DTU improves allocative efficiency by inducing consumption when the local marginal cost of electricity is low but retail prices, anchored to the national wholesale price, do not reflect local system conditions. The underlying logic is one of gains from trade: a single national price can mask locationally different marginal values, and by moving otherwise-curtailed energy to consumers who value it above its near-zero local cost, DTU realizes surplus that curtailment leaves on the table. Applying our estimates to a year of Great Britain settlement data, Scotland's 2.55 million households alone could realistically cap-

[2] Many EV and solar customers in Spain were participating in other flexibility programs and therefore were excluded from our turn-up trial.

ture approximately £18 million in annual consumer surplus this way, against a theoretical ceiling of £314 million were participation unconstrained by population.

Four findings from our welfare analysis are particularly policy-relevant. First, we find substantial mark-up in constraint-down bids in Great Britain from wind generators, leading to the possibility of fiscal savings from substituting demand turn-up for these payments to curtail.[3] Second, the welfare gain is invariant to whether curtailment is compensated: institutional design determines only the distributional incidence of surplus across agents, not its existence. Third, the gain grows with the wholesale price – the relevant wedge is between the retail price consumers face and the near-zero local cost of otherwise-curtailed generation, so it is driven by consumer surplus rather than by avoided curtailment payments, which move in the opposite direction. Fourth, because retail tariffs recover network and policy costs through volumetric charges, pass-through of low local prices is muted; lighter volumetric levies would, for the same underlying energy saving, permit larger proportional price reductions and a stronger demand response.

1.2 Relevant literature

Our analysis relates to three strands of the literature: the declining value of variable renewable generation and the role of flexibility, the market design mechanisms underlying negative prices and curtailment, and the implications of these dynamics for renewable investment and policy design.

The first strand studies how the value of renewable generation changes as penetration increases. A central result is that the market value of variable renewables declines with deployment, as generation becomes increasingly concentrated in periods of low prices (Hirth, 2013). This effect reflects both temporal correlation in output and limited demand responsiveness in the short run. In such settings, system flexibility becomes a key determinant of efficiency, as it allows consumption or storage to shift toward periods of abundant supply.

A growing literature emphasizes that demand-side flexibility can help align consumption with periods of surplus renewable generation, complementing storage and network investment as tools for managing variability (Newbery et al., 2018; Torriti, 2026). Even studies focused primarily on demand *reduction* under time-varying pricing hint at this ef-

[3] We estimate that Scottish households "turning up" would save the public sector approximately £9.5 million a year, though £8.4 million of that is a transfer from producers rather than additional welfare, with the remainder a genuine resource saving from avoided cycling costs.

fect from the other direction: Borenstein (2005*a*) finds that introducing real-time pricing modestly *increases* total annual consumption (by 0–2%), as lower off-peak prices draw in some extra demand precisely when supply is cheapest.[4] Two large studies have estimated the impact of dynamic pricing tariffs (Fabra et al., 2021) and infrequent critical peak pricing events (Fowlie et al., 2021) to reduce electricity demand at peak times. Fabra et al. (2021) estimated mean price elasticities from -0.0072 to -0.054, but noted "none of the price coefficients are significantly different from zero," in response to real-time pricing in 2016–2017. Fowlie et al. (2021) estimated a price elasticity among the most responsive subgroup (active joiners) of approximately -0.075 in response to critical peak pricing in 2012–2013. Our elasticities are substantially larger than both studies, which might be a function of the salience of our demand-response events, of it being easier for consumers to increase demand than decrease demand, and/or of consumers now having many more ways to change their household demand than 10-15 years ago because of the electrification of cooling, heating, vehicle charging, etc.[5] Moreover, most other contemporaneous existing evidence on demand responsiveness is based on observational data or stated-preference surveys, leaving uncertainty about how consumers respond to short-notice price reductions under real system conditions.

A second strand examines the market design and operational drivers of negative prices and renewable curtailment. Curtailment typically refers to available generation that is not utilized due to economic or system constraints (O'Shaughnessy et al., 2021). Empirical work shows that curtailment is often driven by localized congestion and network constraints rather than system-wide excess supply, with a small number of nodes or regions accounting for a disproportionate share of curtailed energy (Maji et al., 2025; Frysztacki and Brown, 2020). System-level modeling studies similarly highlight the interaction between transmission limits, minimum generation requirements, inflexible baseload (such as nuclear), and renewable variability in generating periods of surplus and negative prices (Frew et al., 2019; Prol and Zilberman, 2023; Novan and Wang, 2024). These

[4]See also Hinchberger et al. (2024) for a novel approach of assessing the efficiency gains of moving to dynamic electricity pricing from static pricing.

[5]Previous work has shown that greater technology in the home is important for changing energy demand (Jessoe and Rapson, 2014; Brandon et al., 2026; Bernard et al., 2026). Consistent with this pattern, several more recent studies found substantially larger demand-response elasticities than the RTP and CPP estimates discussed above, leveraging electric vehicle charging as a flexible load. Burkhardt et al. (2023) found that nighttime prices caused near-total load-shifting from households reprogramming EV charging timing – a much larger elasticity (-0.28) than their daytime CPP reduction (-0.03). La Nauze et al. (2026) found that dynamic incentives to shift EV charging toward midday solar generation were associated with an implied price elasticity of demand of -0.35. Bailey et al. (2025) found an own-price elasticity of off-peak charging of -1.59 and a cross-price elasticity of peak charging to off-peak prices of 2.10. Bernard et al. (2025) found price elasticities of -0.85 and -0.9 from a 15% and 40% reduction, respectively, in public EV charging prices. Our results, by contrast, are not confined to a single easily automated end use or to the subset of households that own an EV: the demand turn-up response we document spans the broader residential load.

findings suggest that increasing demand during periods of surplus generation may provide an alternative to curtailment, although there is limited empirical evidence on the effectiveness of such interventions in practice.

A third strand considers the implications of these dynamics for renewable investment and policy design. While curtailment can be an efficient short-run response to system constraints, persistent low or negative prices can reduce revenues and weaken investment incentives in the absence of appropriate market design (Borenstein, 2012). Policymakers have responded by adapting support mechanisms, including exposing renewable generators more directly to wholesale prices and limiting payments during negative price periods. At the same time, private contracting arrangements such as power purchase agreements increasingly incorporate provisions to manage exposure to price volatility. These developments reflect a broader shift toward integrating renewable generation into market-based systems while preserving incentives for investment (Newbery et al., 2018; Davis et al., 2023). Interestingly, we also find that a meaningful share of curtailed volume is accepted at bids exceeding even the most generous available benchmark for genuine opportunity cost (Appendix G.1), consistent with market power in constraint bidding as documented by Fabra and Imelda (2023).

Within this literature, relatively little work provides causal evidence on the effectiveness of demand-side interventions designed to absorb surplus generation. The closest related study is Yang et al. (2025), which uses survey data to elicit hypothetical willingness to shift consumption toward low-price periods. In contrast, we provide field-experimental evidence on realized consumer behavior during periods of excess supply, allowing direct comparison of demand turn-up to alternative mechanisms such as curtailment. A central uncertainty surrounding these programs is the degree to which consumers respond to short-notice requests to increase demand. In the absence of credible estimates of demand responsiveness, policymakers cannot reliably assess the cost-effectiveness of turn-up events relative to curtailment. By providing causal estimates of consumer response, our results inform both short-run operational decisions and broader market design. In the short run, they shed light on the extent to which surplus generation can be absorbed through demand turn-up rather than generation curtailment. In the medium run, they provide evidence on how consumers respond to dynamic price signals more generally, informing the potential effectiveness of real-time pricing and other mechanisms designed to improve retail price pass-through. An additional methodological advantage of our setting is that turn-up events are offered automatically to all customers rather than through voluntary enrollment. This sidesteps the selection effects that complicate welfare evaluation in opt-in dynamic pricing programs, where more price-elastic consumers are more likely to self-select into participation (Ito et al., 2023).

The rest of the paper is structured as follows. Section 2 provides policy context. Section 3 describes the experimental design and data. Section 4 presents the main results. Section 5 turns to a complementary trial responding to localized network constraints. Section 6 develops a welfare framework. Section 7 concludes.

2 Policy Background

Policymakers facing surplus generation have at their disposal several responses, operating across different time horizons. In the long run, expanding transmission and interconnection relieves the congestion that traps surplus behind constraints, and locational marginal pricing sharpens the price signals that guide where generation and demand locate (Gonzales et al., 2023; Doshi, 2026). In the medium run, reforming output-based support and widening retail price pass-through – through real-time or dynamic tariffs – makes both supply and demand more responsive to wholesale conditions. In the short run, system operators have traditionally relied on curtailment to restore balance, reducing generation when supply exceeds what the network can carry.[6] Demand turn-up is the short-run, demand-side alternative to curtailment: rather than discarding surplus, it pays consumers to absorb it. It is this short-run choice – and how it is shaped by each country's institutions – that our experiments and welfare analysis address.

The two settings: Great Britain: Great Britain's market makes the mismatch between abundant local supply and inflexible demand especially visible. Wholesale markets clear at a single national price, while transmission constraints – particularly between generation-rich Scotland and demand in England – are managed separately through the Balancing Mechanism, which *both* sets the national imbalance price and procures the constraint actions that resolve congestion. Because there is no granular locational pricing, the national price does not reflect local network conditions, so surplus behind a Scottish constraint is managed by paying generators to reduce output rather than by a local price falling toward its true near-zero value. Renewable generators are supported mainly through output-

[6]In 2024, Great Britain curtailed 8.3 TWh of wind generation and Spain curtailed 1.4 TWh (Beyond Fossil Fuels 2024).

based schemes – Contracts for Difference (CfDs) and the legacy Renewables Obligation – which historically insulated them from wholesale price risk and weakened their incentive to stop generating when prices turned negative. Recent CfD terms, introduced in Allocation Round 4, suspend subsidy when the market reference price falls below zero, though the optimal number of consecutive negative-price hours before suspension remains an open question (Van Steenberghe and Ovaere, 2026).

The two settings: Spain: Spain differs in the one respect that matters most for our welfare analysis: generators do not hold firm access rights to the transmission grid, so the system operator can mandate curtailment without compensation. Specifically, renewables curtailed in the first (Phase 1) stage of day-ahead redispatch – which accounts for the large majority of curtailed volume – are simply unwound at the day-ahead price and capture no constraint rent, in contrast to the payment a firm-access generator receives in the British Balancing Mechanism. As in Great Britain, wholesale prices clear over a wide area rather than reflecting local constraints – here through the Iberian market (MIBEL), operated by OMIE, which sets a uniform price across the coupled Spanish–Portuguese zone.[7] Spain has, however, gone further than most countries on the demand side: since 2014 households on the regulated default tariff (PVPC) have faced an hourly wholesale-linked price, giving a larger share of demand direct exposure to price than in Great Britain.[8] The contrast – Britain compensating curtailed generators, Spain mandating curtailment without payment – is the institutional difference our welfare framework turns on.

3 Experimental design and data

The design of this field experiment was to test the effect of price reductions on encouraging customers to increase electricity consumption. To understand this, we ran five events in Great Britain, and five events in Spain. Our primary outcome of interest was electricity consumption levels during “turn-up” event windows.

[7] The clearing price is uniform across the Iberian bidding zone. Newer Spanish auction-based remuneration schemes expose renewable generators more directly to market prices than earlier feed-in regimes, though legacy output-linked arrangements persist.

[8] The PVPC is linked to wholesale prices but also includes regulated charges and, since 2024, an increasing futures-market component, which dampens pass-through of spot-price fluctuations.

3.1 Sampling and eligibility into the natural field experiment

We implemented our natural field experiment using eligible customers of Octopus in Great Britain and Spain.[9] We used two slightly different sampling procedures across settings due to cross-country differences in operational constraints, available covariates, and customer populations.

Great Britain: In Great Britain, we had access to a large pool of eligible customer accounts and a rich set of covariates. Our implementing partner provided a sampling frame of 484,683 customer accounts deemed eligible for the trial. Customers were eligible if they had opted in to marketing communications, were not enrolled in similar campaigns, had a smart meter, and were not on a real-time tariff that was already responding to negative prices.

From this sampling frame, we drew a nationally representative sample of 60,000 customers. We used stratified sampling based on: (i) property value terciles interacted with urban–rural classification, (ii) property type (house versus non-house), and (iii) terciles of annual electricity consumption. Crossing these dimensions yields 36 strata. We computed each stratum's population share and used proportional allocation to select 60,000 customers. Between sampling and trial launch, 1,606 customers exited Octopus Energy or otherwise became ineligible, leaving a final sample of 58,394 customers.

Spain: In Spain, we began with the full population of domestic Octopus Energy customers and applied eligibility filters specified by our implementing partner, yielding a pool of 60,438 customers. Given the smaller customer base, we did not employ stratified sampling and instead included all eligible customers. Additional details on eligibility criteria are provided in Appendix D.1.

3.2 Treatment groups

In each country, for each event customers were randomized into six groups, as shown in Figure 2. Four groups were common across countries: (1) a pure control group, (2) a

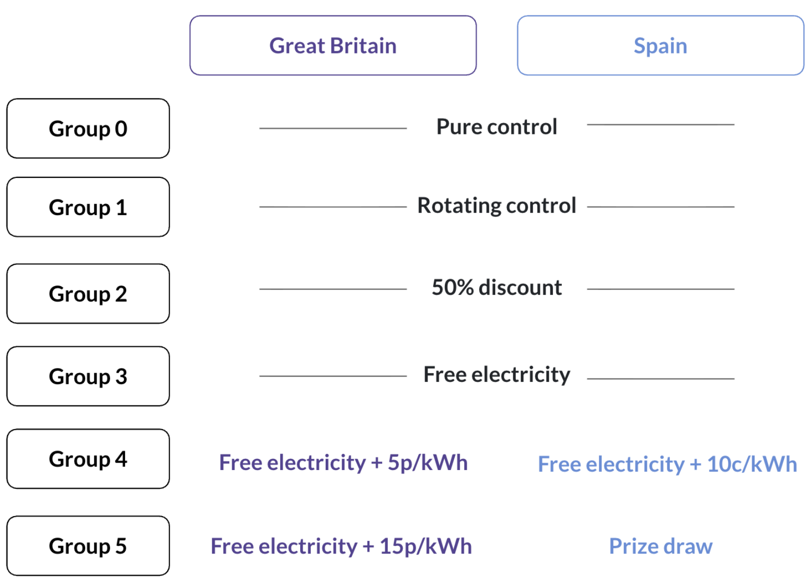
[9] As of January 2nd, 2025, Octopus Energy was the largest domestic electricity supplier in Great Britain, serving 7.8 million customers. In Spain, Octopus entered the market through the acquisition of the local supplier Umeme. As of the start of our trial, it served 400,000 domestic customers.

rotating control group, (3) a 50% discount on electricity consumption during the event, and (4) free electricity consumption during the event.

The design of the negative pricing treatment arms differed slightly between the two countries. In Great Britain, there were two negative-price groups. In the first, customers received free electricity plus an additional payment of 5 pence per kWh consumed during the event window, equivalent to approximately 25% of the typical unit rate. In the second, customers received free electricity plus 15 pence per kWh, equivalent to approximately 75% of the typical unit rate. In Spain, the negative-price treatment consisted of free electricity plus a payment of 10 euro cents per kWh consumed during the event window, approximately 80% of the typical unit rate.

Additionally, Spain included a treatment arm where customers were offered free electricity for any consumption above their baseline level. In addition to receiving free electricity above baseline, ten participants were randomly selected to each receive a €150 prize. Eligibility for the prize draw required participation in the event, defined as increasing electricity consumption relative to baseline. The baseline was defined as consumption during the same hour over the previous ten weekdays for weekdays and the previous three weekend days for weekends.

Figure 2: Treatment design for Great Britain and Spain



Notes: This figure shows the six treatment groups in each country. Groups 0–3 are identical across Great Britain and Spain: Pure Control, Rotating Control, 50% discount, and Free electricity. Groups 4 and 5 differ by country. In Great Britain, these are Free electricity + 5p/kWh and Free electricity + 15p/kWh; in Spain, they are Free electricity + 10c/kWh and a Prize draw.

3.3 Crossover randomization

In each country, we conducted five events and assigned customers to rotate through five treatment groups. In addition, a subset of customers was assigned to a pure control group and never received any treatment. To implement this rotation across events, we used two Latin squares, which generated 20 unique treatment sequences. Each sequence specified a distinct order in which a customer received the five conditions (four treatments and one rotating control), ensuring that (i) each customer experienced each condition exactly once over the five events and (ii) each condition appeared in each event position exactly once per Latin square, thereby controlling for order effects. We additionally included sequences in which customers were never treated, forming the pure control group. Examples of the crossover randomization sequences are shown in Figure 3.

Figure 3: Example of sequences in Great Britain crossover randomization

Randomisation	Event 1	Event 2	Event 3	Event 4	Event 5
Customer 1	Free + 5p	50% Discount	Free	Control	Free + 15p
Customer 2	Free	Free + 5p	Free + 15p	50% Discount	Control
Customer 3	Control	Free + 15p	Free + 5p	Free	50% Discount
			⋮		
Customer 21	Pure Control	Pure Control	Pure Control	Pure Control	Pure Control

Notes: This figure shows example treatment sequences from the Great Britain sample. We used two Latin squares to generate 20 unique rotating sequences, plus 4 sequences of customers who are never treated in any event (Pure Control). Customers were assigned to conditions across the five events using these 24 sequences. In Spain, we instead used 6 Pure Control sequences (for 26 total), reflecting the number of customers our partner wanted in the study.

We then employed matched pairs randomization to assign sampled households to these sequences, with the aim of ensuring pairwise treatment balance over time. The steps for how we conducted the matched pairs randomization are in Appendix D.2. In summary, though, for the Great Britain trial, we created blocks of 24 customers, randomly allocating 20 to a rotating-treatment sequence, and 4 to pure control "sequences". And, we created blocks of 26 customers for Spain – where, again 20 were randomly allocated to each rotating-treatment sequence, but 6 were allocated to pure control sequences. This randomization resulted in 9,710 customers in the pure control group and 48,684 customers assigned to treatment sequences in Great Britain. In Spain, 13,942 customers were randomized to the pure control group and 46,496 customers were assigned to treatment sequences.

Due to an implementation error, the fifth event in Spain inadvertently used the same treatment assignments as the fourth event. As a result, not all customers in Spain were exposed to all four treatments. Every ordered pair of distinct treatments still appears among the earlier events (coverage is retained), but carryover exposure is no longer balanced because of the duplication of the fifth event. Reassuringly, estimates that exclude the fifth event are similar to our main results, indicating that our conclusions do not depend on the duplicated event (Section 4).

Validity of ITT estimate: Our design thus results in a within-customer (within-subject) randomized field experiment. Within-subject designs increase statistical power, but they require two assumptions additional to what is typically seen in a between-subjects setting to support causal interpretation. We go through these below.

First, there are three standard identification assumptions that would also be necessary with between-subject experimental designs (List, 2026). First, the Stable Unit Treatment Value Assumption (SUTVA) holds: a customer's potential outcomes do not depend on the treatment assignments of other customers. Second, observability is ensured through a balanced panel design: treatment was assigned only to customers for whom we reliably received smart meter data. While post-assignment attrition was possible, we tested for differential attrition and found no evidence that treatment affected outcome observability. Third, statistical independence holds, as treatment assignment was randomized and therefore independent of potential outcomes. Note that our analysis focused on intent-to-treat (ITT) effects rather than average treatment effects. While we observe email delivery, we cannot verify whether customers opened or read the treatment emails; full compliance is therefore not required for identification of the ITT estimate.

Adopting a within-subject design introduces an explicit time dimension, which necessitates two additional assumptions. The fourth assumption is temporal stability, which requires that time-varying factors affecting outcomes are not confounded with treatment assignment. We address this by randomizing the order of treatment conditions using a Latin square design, ensuring that all treatment conditions appear in each event. We further include event fixed effects to net out common time shocks from the estimated treatment effects. Finally, the design requires causal transience: treatment effects must not depend on prior or future treatments. This assumption rules out both anticipation and carryover effects. Anticipation is unlikely in our setting, as customers do not know when events will occur or what treatment they will receive. However, carryover effects may arise if treatment impacts persist over time.

We distinguish between two forms of carryover. The first is habit formation, where exposure to turn-up events could affect electricity consumption even during periods with no events. We test for habit formation by comparing customers assigned to the rotating control group to those in the pure control group. If the estimated coefficients for these two groups are statistically indistinguishable, this implies that any treatment effects are short-lived and do not persist beyond the event window. Conversely, statistically significant differences between these coefficients would indicate the presence of habit formation. As discussed in Section 4, we find that the coefficient estimates for the rotating and pure control groups are statistically indistinguishable, providing evidence that treatment effects do not persist beyond the event window.

The second form of carryover is learning or fatigue, where responsiveness to treatment depends on prior exposure. Importantly, we view such learning as policy-relevant: our implementation reflects how turn-up events would plausibly be deployed in practice. Given the Latin square design, every treatment appears in every event position, and the estimated effect of each treatment averages over all levels of prior exposure. As discussed in Section 4.3, the first event a customer experienced elicited the largest response in the Free and Negative Pricing groups in Great Britain. Effects were stable across the 50% Discount groups in both countries, and were stable across Free and Negative Pricing groups in Spain.

Reporting estimates by number of prior treatments, we therefore advise that, if this policy is implemented with frequent turn-up events, the average effect of the third and fourth exposures better satisfies the conditions for an internally valid ITT in steady state. With infrequent events, the novelty effect might instead be renewed between exposures, though we cannot say whether this occurs or how large a gap would be required.

With these conditions satisfied, randomization identifies an internally valid and unbiased ITT estimate of the effect of turn-up events on household electricity consumption, absorbing fatigue effects into this ITT. Moreover, because customers were not required to opt in and therefore did not self-select into the experiment, we classify the study as a natural field experiment with unbiased treatment assignment.

3.4 Implementation of turn-up events

Turn-up events were conducted between June and September 2025. In Great Britain, each event lasted one hour, while in Spain each event lasted two hours. Events were not scheduled on predetermined dates. Instead, in coordination with the implementing partners, events were triggered when day-ahead wholesale electricity prices were close to zero or negative.

Three days prior to the first event, all sampled customers, except those assigned to the pure control group, received an informational email describing the nature of the turn-up events program and notifying them that events would occur during the study period, although the exact timing was not disclosed. Subsequently, 24 hours before each turn-up event, treated customers received a treatment-specific email informing them of the event timing and the associated financial incentive. Examples of the emails sent in Great Britain and Spain are shown in Figure A29 and Figure A30, respectively.

Remuneration for participation differed across countries. In Great Britain, remuneration was provided after all turn-up events had concluded. In Spain, customers were compensated on a rolling basis, typically one to two weeks after each event.

3.5 Experimental integrity

Balance: Table A1 reports baseline balance between treatment and control groups in Great Britain and Spain. Overall, observable characteristics were well balanced across groups in both countries. These results are consistent with successful randomization. The crossover design provides an additional safeguard against baseline imbalance. Because households cycle between treatment and control status across events, time-invariant characteristics are orthogonal to treatment assignment in specifications that use the rotating control as the reference group.

Attrition: Attrition is low in both samples and does not differ by treatment status. In Great Britain, 7% of households attrit, while in Spain the attrition rate is 8%. Attrition occurs only if a household leaves Octopus during the study period, which we believe is unrelated to our treatment. Regressions of attrition on treatment assignment yield small and statistically insignificant coefficients in both countries, indicating no evidence of differential attrition (Table A2).

Multiple hypothesis testing: Because our main analysis estimates four treatment-versus-control effects within each country, we adjust inference for this family of comparisons using Romano-Wolf stepdown p-values (Romano and Wolf, 2005). We apply the same adjustment within each heterogeneity exercise to the family of treatment-by-heterogeneity interaction terms.

Pre-analysis plan: This study was pre-registered through the AEA RCT Registry. The Great Britain trial was registered as AEARCTR-0016220, the Spain trial was registered as AEARCTR-0016221, and the NPG trial was registered as AEARCTR-0015455. The pre-analysis plan for each experiment is included in each registration. We record deviations from our pre-analysis plans in Appendix C.

3.6 Empirical strategy

To estimate the intent-to-treat (ITT) effect of turn-up events on hourly electricity consumption (kW), we use the following specification:

$$Y_{it} = \alpha + \sum_{k=1}^{K} \beta_k E_{it}^k + \mu_b + \pi_1 C_{it} + \pi_2 K_{it} + \delta_t + \gamma X_i + \epsilon_{it} \tag{1}$$

where Y_{it} represents import electricity consumption (kW) for consumer i during hour t.[10] E_{it}^k is an indicator variable equal to 1 if household i is assigned to treatment group k at time t, and 0 otherwise. μ_b are randomization-block fixed effects. These blocks were formed for our matched-pairs randomization, as explained in Appendix D.2. C_{it} is customer i's average hourly consumption in the same hour over the last 5 working days before hour t. Analogously, K_{it} is for the last 2 weekend days. δ_t are hourly fixed effects, capturing common shocks or trends affecting all customers in a specific hour. We cluster standard errors at the level of consumer i.

We additionally control for X_i, a vector of time invariant baseline customer characteristics, based on what is available in each country. In Great Britain, this vector includes (1) baseline estimated annual consumption (split into quartiles), (2) type of smart meter installed in the property, (3) Energy Performance Certificate (EPC) rating (categorical variable with 8 levels), and (4) tenure with Octopus Energy (in categories: 0-1 years, 1-2 years, 2-3 years, 3-4 years, 4-5 years, more than 5 years). In Spain, this vector includes (1) baseline afternoon consumption, defined as the mean hourly consumption between

[10]Our analysis used high-frequency administrative data on consumption from Octopus Energy Limited (Great Britain) and Octopus Energy España, S.L.U. (Spain). We pre-specified our main outcome variable as hourly electricity consumption during turn-up events (import only).

13:00 and 17:00 in May 2025 and (2) tenure with Octopus Energy (Octopus Energy Spain started in 2023, and thus the categories here are only 0-1 years, 1-2 years, 2-3 years).[11]

The β_k values are thus our coefficients of interest, as they measure the treatment effect of treatment group k, relative to the pooled control group.

4 Results

4.1 Main results: consumption during events

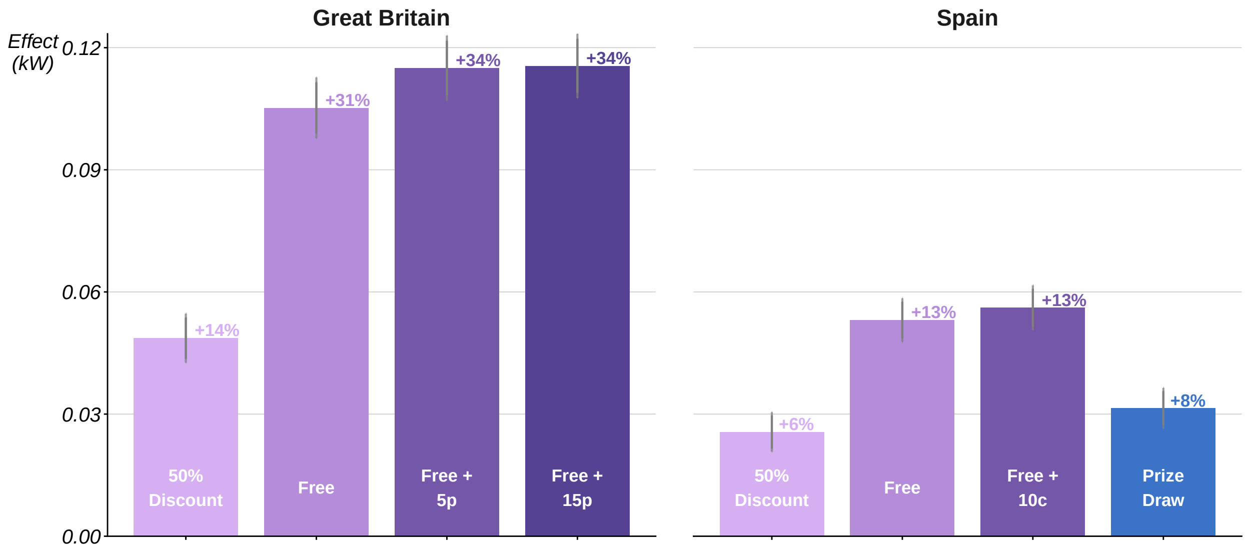
Across both countries, treatment effects were monotonically increasing and economically intuitive along the margin from no incentive to a 50% discount to free electricity, before exhibiting sharply diminishing returns to additional incentives (i.e., negative pricing). Figure 4 presents estimates from our main pre-registered specification in Equation (1). In Great Britain, customers offered a 50% discount increased hourly electricity consumption by 0.049 kW, corresponding to a 14.3% increase relative to the control group. Offering free electricity more than doubled this effect, to 0.105 kW (31%). Adding a 5p/kWh payment on top of free electricity led to only a modest additional increase, to 0.115 kW (33.8%), and increasing the payment to 15p/kWh did not yield a further meaningful increase (0.116 kW, 34%).

Spain exhibited a similar qualitative pattern, though with smaller effect sizes. Customers offered a 50% discount increased hourly consumption by 0.026 kW (6% relative to control), while offering free electricity again approximately doubled the effect to 0.053 kW (12.7%). Adding a 10c/kWh payment on top of free electricity resulted in a small additional increase, to 0.056 kW (13.4%). The prize draw produced an effect comparable in magnitude to the 50% discount, increasing consumption by 0.031 kW (7.5%).

All treatment effects remain statistically significant at the 1% level after applying Romano-Wolf adjustments that test the four treatment effects jointly within each country (Romano and Wolf, 2005). We report these adjusted p-values in brackets in Column (1) of Table A3 and Table A4.

[11] We pre-specified that we would control for estimated annual consumption in Spain. However, this measure is not well defined for many customers because it is a modeled estimate from Octopus and can be quite noisy for new customers. Octopus Energy Spain has grown quickly over the past few years, so many customers are relatively new (average tenure 0.65 years in Spain compared to 4.03 years in Great Britain). This makes the estimated annual consumption measure much less reliable in Spain. Instead, we control for observed average baseline consumption between 13:00 and 17:00 in May 2025, which is directly measured rather than modeled. This is the variable we used for our matched sampling, as documented in our pre-analysis plan, and was chosen because this is the period during which turn-up events were most likely to occur.

Figure 4: Treatment effects on hourly consumption (kW)



Notes: This figure shows treatment effects from Table A3 and Table A4, our main pre-registered specification. Lines depict 90% (dark) and 95% (light) confidence intervals. Standard errors are clustered by customer. Our outcome measure is hourly consumption during the turn-up event. In Great Britain, events were one hour long, while in Spain, events were two hours long. Percentages represent average treatment effects as a share of the consumption in the control group during the event period. All treatment-effect estimates remain statistically significant at the 1% level after Romano-Wolf adjustment for multiple hypothesis testing across treatment arms.

Table A3 and Table A4 present alternative specifications used as robustness checks. In both tables, column (2) varies the baseline control set: for Great Britain, it replaces EAC quartile with average baseline import during the event-relevant afternoon window; for Spain, it reports the pre-analysis-plan control set. Column (3) excludes both fixed effects and baseline controls. Column (4) includes customer fixed effects, absorbing time-invariant customer heterogeneity and identifying effects from within-customer changes in treatment assignment across events. Column (5) separates the rotating control group from the pure control group, using the pure control group as the omitted comparison group. In Table A4, column (6) reweights the Spanish sample to match the Great Britain sample on baseline time-of-use tariff status and baseline afternoon consumption, column (7) restricts the outcome to consumption in the first hour of the event window, and column (8) drops the fifth Spanish event, for which treatment assignments repeated the fourth event's assignments due to an implementation error. Across these specifications, the estimated treatment effects are stable in magnitude and statistical significance.

Column (5) also serves as a test for habit formation, examining whether exposure to turn-up events affects electricity consumption even during periods when customers are not actively participating. The coefficient estimates on the rotating control group are small and statistically insignificant, suggesting that treatment effects are short-lived and do not persist beyond the event window.

The payment per induced kWh of turn-up is lowest when electricity is discounted or free (see Table A13). In contrast, paying customers to consume beyond zero price substantially increases total payments while generating only minimal additional demand. Negative pricing for electricity use may nonetheless play a role in specific circumstances. If the objective is to procure a larger or more precisely targeted volume of demand, then higher incentives could be justified.

4.2 Heterogeneity

We examined treatment effect heterogeneity by interacting treatment assignment with a number of baseline customer characteristics. Because these characteristics were not randomly assigned, we interpreted the estimates as descriptive evidence about which groups delivered larger responses, not as causal effects of the characteristics themselves. Readers interested in the construction of each heterogeneity variable should see Appendix E. Table 1 summarizes the full set of heterogeneity analyses; in the text below, we focused on the dimensions where responses differed most clearly. [12] For the corresponding regression tables, we report Romano-Wolf adjusted p-values for the treatment coefficients and treatment-by-heterogeneity interaction coefficients.

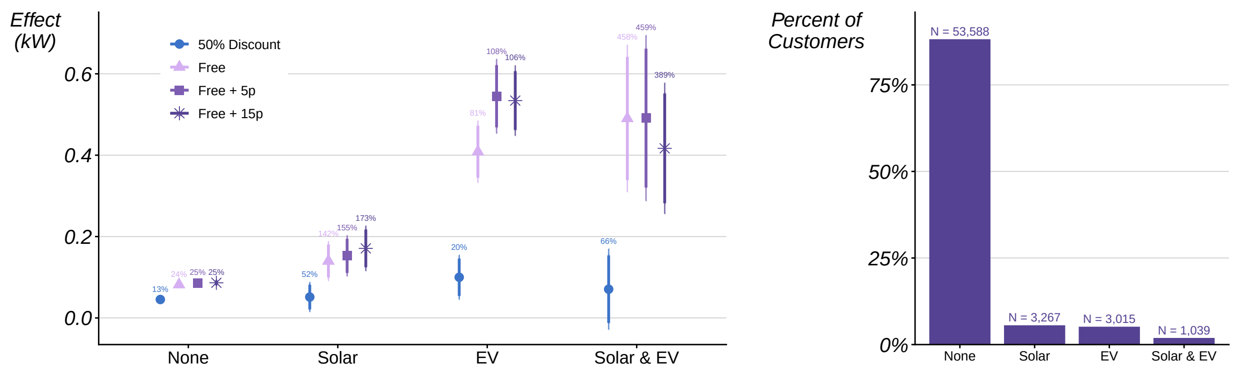
The clearest pattern was that responsiveness was largest among customers with access to flexible electrified load. In Great Britain, customers on EV and solar tariffs increased consumption substantially more than customers without low-carbon technology tariffs, especially under the free and negative-price treatments (Figure 5).[13] We did not estimate the same LCT heterogeneity in Spain because solar and most EV customers were excluded from the Spanish trial due to concurrent flexibility campaigns. Customers on time-of-use tariffs also responded more strongly in both countries, although in Great Britain this pattern partly reflected the concentration of EV and solar customers on technology-

[12] Interactions with baseline tariff types were pre-specified. Interactions with other customer characteristics were not pre-specified.

[13] As a complementary analysis, we examined whether turn-up responses among EV-linked customers displaced public charging. We were able to access charging data for Electroverse, Octopus Energy's public EV charging platform. Before the trial, 872 British turn-up customers were observed using Electroverse in 2025; interestingly, 820 were not on Intelligent Octopus Go or Octopus Go, suggesting that many did not have access to home smart charging. Table A6 provides little evidence that turn-up incentives reduced public charging before, during, or after events; if anything, we found noisy evidence that public charging increased during event windows.

specific time-of-use products. These results suggested that the largest turn-up responses came from households with both flexible load and experience responding to time-varying prices.

Figure 5: Heterogeneity by LCT ownership, Great Britain



Notes: This figure shows how treatment effects vary by EV or solar ownership. Ownership is inferred from customers' baseline tariffs. Customers on an exporting tariff are classified as having solar, while customers on one of Octopus' EV tariffs (Octopus Go or Intelligent Octopus) are classified as having an EV. Lines depict 90% (dark) and 95% (light) confidence intervals. Standard errors are clustered at the customer level. The outcome measure is hourly electricity consumption during the event. The right hand side shows the share of customers with each ownership type. The corresponding regression table is reported in Table A7.

Households with higher baseline consumption delivered larger absolute increases during events, but proportional responses were smaller among customers with higher baseline consumption. Similarly, in Great Britain, higher property-value quintiles showed larger absolute increases, while the proportional gradient was less distinct. In Spain, treatment effects did not show a clear monotonic income gradient. Together, these patterns pointed to the importance of having flexible load available at baseline.

4.3 Fatigue and anchoring effects

We found that in the Free and Negative Pricing groups, customers' response in the first event they experienced was larger than in subsequent events (Figure A4). In Great Britain, the second-event treatment effect level dropped by roughly 25% relative to the first in the Free, Free + 5p, and Free + 15p arms. In Spain, the drop between the first and second event was less pronounced, although the third event saw a 20–30% decline in the Free and Free + 10c arms. In both countries, responses appeared to stabilize across the final two events.

Part of this decline in absolute treatment effects was because baseline consumption was declining during this period. Thus fatigue is less evident when looking at treatment

Table 1: Summary of treatment effect heterogeneity by household characteristic

Characteristic	Results	Treatment-effect heterogeneity
Low-carbon technology tariff	Figure 5 and Table A7	EV and solar tariff customers responded more strongly, especially under free and negative-price treatments. EV tariff customers showed the largest effects.
Baseline tariff type	Figure A6 and Table A8	Time-of-use customers responded more strongly in both countries.
Baseline consumption	Figure A8 and Table A11	Higher-consumption households delivered larger absolute increases, while proportional responses were smaller among customers with higher baseline consumption.
Socioeconomic status	Figures A7 and A10 and Tables A9 and A10	In Great Britain, higher property-value quintiles showed larger absolute responses, but the proportional gradient was less distinct. Customers eligible for the Warm Home Discount (WHD), a UK government scheme providing support to households in fuel poverty, exhibited smaller treatment effects than customers who were not eligible. In Spain, effect sizes were quite consistent across income quintiles.
Geographic area	Figures A11 and A12	Responses were positive across regions but varied geographically.
Email engagement	Figure A9 and Table A12	Customers who opened at least one event email responded much more strongly than customers who never opened an event email.

Notes: Each row summarizes estimates from a regression that interacts treatment assignment with the listed household or customer characteristic. Treatment effects are measured in hourly electricity consumption during the event window. The Results column lists the corresponding figures and regression tables. The figures report point estimates with 90% and 95% confidence intervals; standard errors are clustered at the customer level. Proportional comparisons are computed relative to control-group consumption in the same subgroup. Appendix E describes how each heterogeneity variable is constructed. Regression results tables report Romano-Wolf adjusted p-values. The adjustment is applied within each heterogeneity analysis to the family of treatment-by-heterogeneity interaction coefficients; main-effect coefficients are not included because they are not interpreted causally.

effects as a percentage of baseline consumption. In Spain, there was no sign of fatigue on this percentage measure. We also investigated whether longer intervals between events led to a rebound in treatment effects but found no evidence of this within the range of spacings observed in our trial, although it is possible that substantially longer gaps between events could restore effect sizes closer to those observed on first exposure.[14]

Notably, the 50% Discount groups had stable estimates across both countries. This points to a novelty effect specific to being advertised something that is "Free". We thus conclude that there is evidence of fatigue from repeated exposure. One interpretation is that the ITT for the Free and Negative Pricing groups is more accurately captured by the average effect of events three and four, while the ITT for the 50% Discount groups is less affected by novelty/fatigue effects.

We also did not find anchoring effects when we looked at heterogeneity in current treatment effects by the customer's first treatment assignment (Figure A5). The response to each incentive was broadly similar regardless of which treatment the customer initially received. Customers thus seem to have evaluated each event on its own terms rather than anchoring against a reference point established by their initial exposure.

4.4 Difference between Great Britain and Spain

Consumers responded more strongly to incentives in Great Britain than in Spain, despite similar average customer rewards. In fact, in all cases of comparable incentives, the response was twice as strong in Great Britain as in Spain. We do not have a definitive explanation for this cross-country difference, but we explored several contributing factors.

First, differential engagement did not appear to be the driver. Email open rates were similar across countries: the average open rate was 66.3% in Spain and 62.4% in Great Britain, suggesting comparable exposure to the intervention.

Second, sample composition explains part of the gap. Spain's trial excluded customers on a special EV tariff and most solar exporters, at the delivery partner's request, as they were enrolled in other flexibility campaigns. Applying a similar exclusion in Great Britain narrowed the gap: under the Free treatment, the effect fell from 0.105 kW (all

[14] Remuneration timing also differed between countries: British customers were paid only after all events concluded, while Spanish customers were paid after each event. It is unclear how this affects the fatigue comparison. Faster feedback could sustain engagement, but could also disincentivize customers who were dissatisfied with their reward size.

customers) to 0.082 kW (non-LCT customers), versus 0.053 kW in Spain – so differences in flexible electrified load explain part, but not all, of the difference.

Finally, events lasted two hours in Spain versus one in Great Britain, which may mechanically dilute Spain's hourly response. For example, a one-hour task like running a washing machine could occur in either hour. Consistent with this, Spain's total two-hour consumption increase is comparable to Great Britain's one-hour increase. Furthermore, Spain's first-hour effect is similar to its full-event average, indicating the response is sustained across both hours rather than concentrated in one (Table A4).

4.5 Temporal demand substitution

While the primary objective of turn-up and negative pricing events is to absorb excess renewable supply, an additional question is whether these interventions generated net new demand or instead shifted consumption across time. If customers increased usage during the event window by reducing consumption in other periods, particularly during hours when wholesale prices were higher, this has two important implications. First, from the supplier's perspective, temporal substitution may lower procurement costs if demand is shifted away from high-price hours and into low- or negative-price hours; customers may also benefit if they are on a time-of-use tariff. Second, high-price periods typically coincide with when marginal generation is more carbon-intensive. Shifting consumption away from these hours could therefore reduce the emissions associated with electricity use.

To assess intertemporal substitution, we first estimated a two-way fixed effects event-study specification, examining treatment effects on hourly electricity consumption from 24 hours before to 72 hours after each event. As shown in Figure 6a and Figure 7a, we found some evidence of sustained elevated consumption in the hours immediately following events, but no notable reductions in consumption in the pre- or post-event windows more broadly.

However, any displacement may have been too diffuse across the surrounding hours to detect reliably in an event-study framework. We therefore conducted a complementary displacement analysis, aggregating consumption across four windows and regressing each on treatment assignment: (1) cumulative consumption in the 24 hours before the event, (2) consumption during the event window itself, (3) cumulative consumption in the 72 hours following the event, and (4) total consumption across the full −24 to +72 hour window. Event notification emails were typically sent approximately 24 hours before the event; we excluded any hours occurring before the notification was sent to avoid contaminating the pre-event window with periods during which customers could not yet have responded.

To improve statistical power, we pooled treatment groups with similar response profiles for this analysis. In Great Britain, we pooled the Free, Free + 5p, and Free + 15p groups, whose treatment effects were statistically indistinguishable in Table A3. In Spain, we pooled the Free and Free + 10c groups on the same basis (Table A4). We also report results from unpooled treatment groups.

The results reveal a contrast between the two countries. In Great Britain, we found statistically significant evidence of intertemporal substitution: the pooled Free + Negative group showed a significant reduction in pre-event consumption, and the post-event window exhibited a cumulative downward pattern consistent with load shifting, together implying complete displacement of the during-event effect (Figure 6b and Table A14).

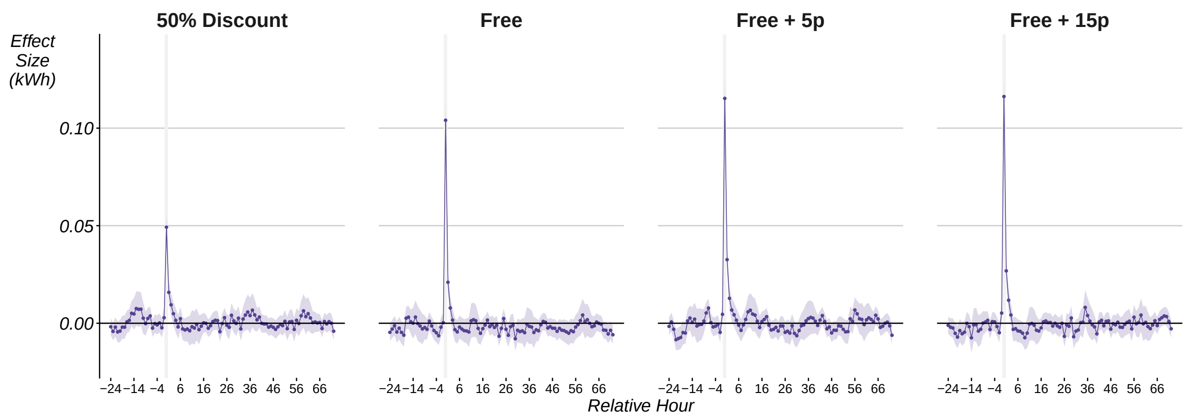
In Spain, by contrast, cumulative effects remained close to zero in both the pre- and post-event windows, suggesting that the treatment effects in Spain reflected net increases in electricity consumption rather than demand shifted from adjacent periods. Looking at overall consumption from −24 to +72 hours, consumption increased by 0.094 kWh for the 50% discount group, 0.114 kWh for the Free group, and 0.113 kWh for the Free + 10c group, consistent with net demand creation (Figure 7b and Table A14).

We urge caution in interpreting displacement estimates for Great Britain. The 50% discount group in Great Britain and all groups in Spain showed no significant displacement, which we would have expected to see if displacement were a systematic feature of customer behavior. The heterogeneity analyses also yielded no consistent evidence of displacement across subgroups (Figure A14, Figure A16, Figure A18). More fundamentally, intertemporal substitution is inherently difficult to quantify precisely: any displacement is likely to be diffuse across a long post-event window. We therefore interpret the evidence as consistent with some displacement in Great Britain, while remaining cautious about its precise magnitude.

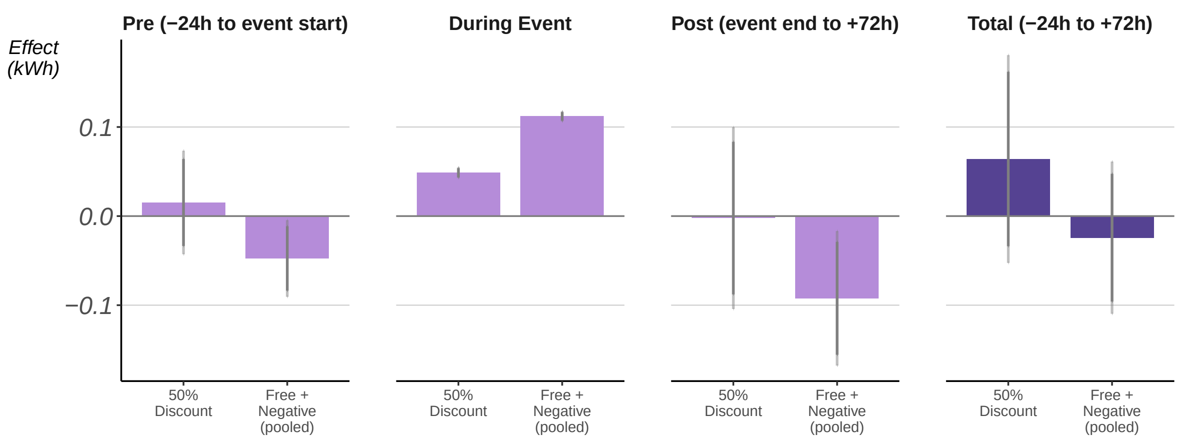
The source of this cross-country difference is not fully clear. One factor may be communication design: in Great Britain, event emails explicitly encouraged customers to shift consumption into the turn-up window, whereas the Spanish communications did not emphasize demand reallocation. A second, non-mutually exclusive explanation is

Figure 6: Effect of treatment, before and after event window - Great Britain

(a) Event Study



(b) Aggregate displacement

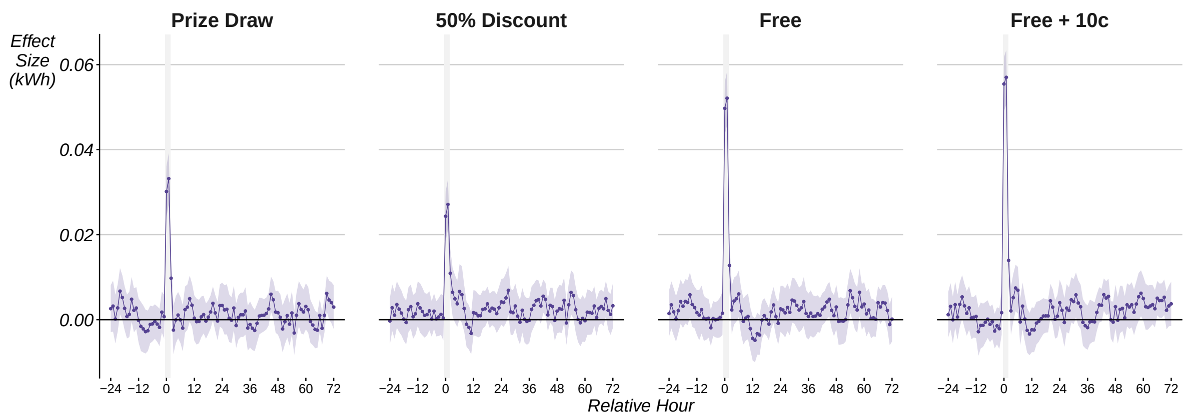


Notes: Panel a shows estimates from an event study analysis where we interact treatment with relative time since the event. Shaded regions are the 95% confidence interval. Panel b plots the impact of treatment on cumulative electricity consumption in different windows. The first pane is from −24 hours to the event start. The second pane is the hour of the event. The third pane is the total consumption from the end of the event to 72 hours afterwards. The fourth pane is total consumption from −24 to +72 hours around the event. Shaded areas indicate 95% confidence intervals; standard errors are clustered at the customer level.

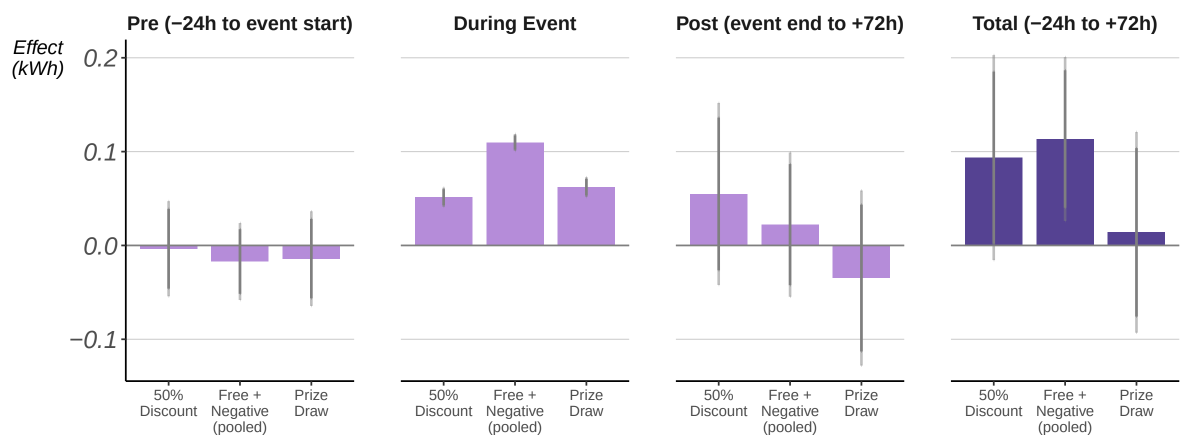
that baseline consumption patterns and flexibility differed across the two contexts. If British households had more discretionary or shiftable loads, such as electric heating, hot water systems, or electric vehicles, they may have been better positioned to move demand across hours in response to price signals. By contrast, if a larger share of consumption in Spain was less temporally flexible, responses may have taken the form of incremental usage during the event window rather than substitution away from other periods. A third

Figure 7: Effect of treatment, before and after event window - Spain

(a) Event Study



(b) Aggregate displacement



Notes: Panel a shows estimates from an event study analysis where we interact treatment with relative time since the event. Shaded regions are the 95% confidence interval. Panel b plots the impact of treatment on cumulative electricity consumption in different windows. The first pane is from −24 hours to −1 hour before the event. The second pane is the two-hour event window. The third pane is the total consumption during hours 1 to 72. The fourth pane is total consumption from −24 to +72 hours around the event. These coefficients represent treatment effects on cumulative consumption relative to the control group. Shaded areas indicate 95% confidence intervals; standard errors are clustered at the customer level.

explanation concerns air conditioning, where the higher prevalence of air conditioning in Spain may have led to more net demand creation rather than displacement.

4.6 Impact of turn-ups on climate sentiment

By passing through the benefits of abundant renewable generation to households, demand turn-up could make the energy transition more tangible. Customers may experience renewables not only as an abstract climate-policy objective, but as something that can lower their own electricity costs when supply is plentiful. To test whether this changed broader climate and energy-policy sentiment, we surveyed customers in Great Britain and Spain after the five national turn-up events, including both customers assigned to receive turn-up communications and customers in the pure control group who did not receive communications. For a full discussion of the methodology and results, see Appendix F.

We find no evidence that assignment to the turn-up treatments changed climate or energy-policy sentiment. Across both countries, treatment effects are small and statistically indistinguishable from zero for all survey questions. This suggests that, in our setting, demand turn-up affected behavior through the immediate financial incentive rather than by measurably shifting broader attitudes toward renewable energy, flexibility, or climate policy. This null result should be interpreted in light of the intervention's scale: the events were episodic, and monetary savings for most households were modest. Our estimates therefore do not rule out the possibility that more sustained exposure to cheap renewable electricity, for example through a tariff that more continuously reflects wholesale market conditions, could affect how households perceive the private benefits of decarbonization.

5 Demand flexibility for locational network constraints: evidence from Northern Powergrid

The main turn-up experiments were triggered by national wholesale price conditions; events occurred when day-ahead prices were forecast to be near zero or negative across Great Britain and Spain. But the operational case for demand-side flexibility extends beyond national surplus management to localized network constraints, where local renewable generation exceeds distribution-network capacity, creating binding bottlenecks. In Great Britain, these constraints are managed by Distribution Network Operators (DNOs), who must either curtail generation, invest in network reinforcement, or procure flexibility from consumers in the affected area.

Northern Powergrid (NPG), which serves around 3.9 million homes and businesses across the North East of England and Yorkshire, faces this challenge. Rapid growth in wind and solar in its network area has created recurring periods where local generation exceeds what the distribution network can accommodate. To test the extent to which domestic demand side response could help them solve these issues, NPG procured demand turn-up services from Octopus Energy Limited, the same energy retailer with whom we partnered on our larger British turn-up experiment. The retailer offered residential customers in two specific constrained GSP areas, Spennymoor and Ferrybridge B, free electricity during periods when NPG required increased demand to relieve local network pressure. Unlike the main experiment, events here were not triggered by wholesale prices, but by NPG's real-time assessment of distribution network constraints. This makes the NPG trial a direct test of whether residential demand flexibility could substitute for curtailment at the distribution level.

5.1 Study design and methodology

This study used a crossover randomized controlled design to evaluate the impact of power-up events in the NPG region. We accessed a pool of 14,477 customers who had signed up to the program by the start of the trial on January 29th 2025, drawn from two network areas: 9,000 in Spennymoor and 5,477 in Ferrybridge B.[15] Participation required customers to take action at two distinct stages. This opt-in structure distinguished the NPG trial from the main turn-up experiment in an important way: whereas turn-up participants were simply sent an email informing them of an upcoming event and their assigned incentive, power-up customers first actively chose to enroll in the program before any events took place, and then separately opted in to each individual event when notified. The power-ups sample is therefore, by construction, more selected than the turn-up sample, since participants had already demonstrated willingness to engage with demand flexibility before any event occurred.

NPG called power-up events between January and March 2025 based on its operational assessment of network conditions, with neither the timing nor duration determined in advance. Over the trial period, there were 19 events across 15 event days, with some days featuring multiple events at different times of day. Because events on the same

[15] A copy of the email invitation to sign up can be found in Figure A31.

day share a common demand environment and because pre- and post-event windows would overlap for same-day events, we treat the event day as the unit of observation for our analysis.

For each event day, customers were allocated to treatment (invited to participate) or control, with the allocation prioritizing customers who had received the fewest prior invitations, ensuring balance in treatment exposure over the trial. In Spennymoor, 2,400 customers were invited per event; in Ferrybridge B, 1,857. These invitation volumes were set based on expected turn-up and NPG's flexibility requirements in each area. Invited customers received an email notification, typically 24 hours in advance, and were required to confirm their participation by opting in to that specific event. The majority of customers received two invitations (Figure A20) across all the events. The treatment was free electricity for the duration of the event window.

5.2 Empirical Specification

We estimate two specifications that reflect the double opt-in structure of the trial. Let Y_{it} denote electricity consumption for customer i at half-hour t. Our first specification estimates the reduced-form effect of being invited to a power-up event among customers already signed up to the program:

$$Y_{it} = \alpha + \beta^{SU}\text{EventInvite}_i + \gamma' X_i + \delta_t + \zeta_z + \varepsilon_{it} \tag{2}$$

where EventInvite_i is an indicator equal to one if customer i is signed up and received an invitation to the event, X_i is a vector of baseline customer characteristics, δ_t are half-hour fixed effects, and ζ_z are zone fixed effects for Spennymoor and Ferrybridge B. The coefficient β^{SU} captures the intent-to-treat effect of invitation, conditional on prior program enrollment.

Our second specification estimates the effect of actual event participation. Because opt-in to a specific event is potentially endogenous, we instrument for it using the invitation assignment, yielding a two-stage least squares estimator:

$$\text{OptIn}_{it} = \pi_0 + \pi_1 \text{EventInvite}_{it} + \gamma' X_i + \delta_t + \zeta_z + \nu_{it} \qquad \text{(First Stage)}$$

$$Y_{it} = \alpha + \beta^{EO}\widehat{\text{OptIn}}_{it} + \gamma' X_i + \delta_t + \zeta_z + \varepsilon_{it} \qquad \text{(Second Stage)}$$

where OptIn_{it} is an indicator equal to one if customer i opted in to the specific event. The coefficient β^{EO} captures the local average treatment effect of event participation for customers who opted in to an event when invited. Because participation requires action at two stages (program enrollment and event-specific opt-in), both layers of selection are embedded in this estimate. The gap between β^{EO} and β^{SU} therefore reflects event-level opt-in behavior conditional on prior program enrollment.

The baseline controls X_i match those used in the main turn-up regressions: average weekday and weekend half-hourly demand, meter type, EPC rating, and tenure with Octopus Energy. The outcome is measured at the half-hourly level, though for presentational clarity all regression tables report results aggregated to the hourly level. This specification differs from our pre-registered analysis plan; however, we adopt it here to ensure comparability across the turn-up and NPG trials. We report the pre-analysis plan specification as a robustness check.

5.3 NPG Results

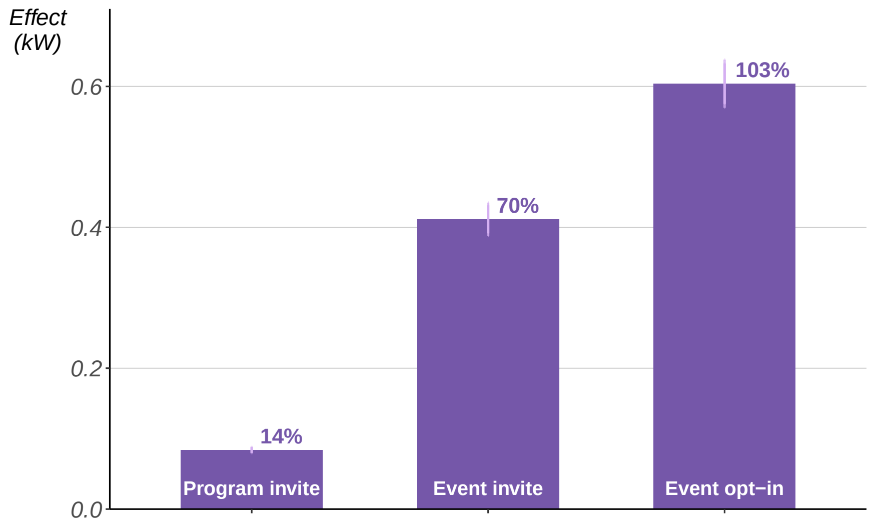
5.3.1 Main results: impact of NPG "power-up" events on event consumption

Consistent with the results from the Great Britain trial, free electricity substantially increased consumption during the event. Among those signed up to a power-up event, providing free electricity increased consumption by 0.411 kW on average, rising to 0.604 kW for customers who opted into the event (Figure 8 and Table A15). Both estimates were statistically significant and robust across specifications, including models without covariates and two-way fixed effects models (Table A16). These translate to point elasticities of -0.704 for the signed-up group, and -1.033 for the event opt-in group.

To make these estimates comparable to the Great Britain trial, we also present an implied impact of being invited to participate. In the NPG trial, roughly 20.4% of contacted customers signed up, and those who signed up increased consumption by 0.411 kW when invited. Scaling this estimate by the enrollment rate implies an effect of 0.084 kW per contacted customer. In level terms, this estimate is close to the 0.105 kW effect in the British Free Electricity group, where customers faced no sign-up requirement and were simply emailed about the event.

The comparison differs in proportional terms. Because the NPG trial was conducted in winter, baseline consumption during event periods was higher than in the summer

Figure 8: Effect of invitation, sign-up, and opt-in to NPG "power-up" events (kW)



Notes: This figure reports estimates of the effect of power-up events on half-hourly electricity consumption in the NPG region, reported at the hourly level for consistency with Table A3 and Table A4, the turn-up results. The first bar presents the implied impact of being invited to participate, calculated by scaling the effect among signed-up customers (bar 2) by the sign-up rate (20.4%). The second bar reports the impact of being invited to power-up, among customers who signed up (Equation (2); Column (1) of Table A15). The third bar reports the impact of opting into the event, instrumented for by the encouragement (Equation (Second Stage); Column (4) of Table A15). Baseline controls include average consumption at the same hour over the last 10 weekdays, average consumption at the same time over the last 2 weekends, meter type, EPC rating, estimated annual consumption, and tenure with Octopus Energy. The control mean was 0.584 kW and refers to mean hourly consumption among uninvited customers during event half-hours. Standard errors are clustered at the customer level.

British trial (0.584 kW versus 0.34 kW). The same absolute increase therefore translated into a smaller percentage increase in NPG, about half of the percentage increase of the British trial. We therefore interpret the NPG and British estimates as similar in operational magnitude, but not necessarily as evidence of equal behavioral responsiveness. The difference in percentage effects is consistent with both higher winter baseline load and the additional friction introduced by program enrollment. We do not interpret the cross-trial differences causally, given differences in implementation, season, targeting, and customer selection.

The opt-in design of the NPG trials has a clear advantage when looking at cost per kWh of turn-up procured. Whereas the free electricity events in the main Great Britain turn-up experiment cost £1.22 per kWh of additional demand elicited (Table A13), the NPG power-up events cost just £0.77 per kWh (Table A17). The higher cost in the Great Britain trial reflects the fact that, without an opt-in requirement, payments go to all invited customers including those who were unaware of the event or made no active change to their behavior. The active opt-in design of the NPG trial thus avoided these payments to non-responding customers, delivering an average of 1.4 MWh of additional demand per event at substantially lower cost.

Heterogeneity in treatment effects by tariff type, shown in Figure A22, mirrors the pattern found in the main Great Britain turn-up experiment: customers on time-of-use tariffs and those with low-carbon technologies such as EVs and solar panels exhibit substantially larger responses, consistent with these customers having greater flexible load available to shift during event windows.

5.3.2 Dynamic consumption response: event study analysis

To examine how the consumption response evolved around events and how consumption evolved over differing treatment lengths, we estimated a TWFE specification. For each customer and each half-hour relative to the start of a given event, we estimated the interaction between treatment assignment and an indicator for each relative time period, including customer and half-hour fixed effects and clustering standard errors at the customer level. We focus on the 12 hours before and 12 hours after an event, and restricted this analysis to events where neither the 12 hours before nor the 12 hours after overlapped with another event; this left 11 events for analysis.

The event-study results confirm that the consumption response is sharply concentrated within event windows, as shown in Figure A23. Demand responses in the first two half-hours are similarly large across 30-minute, 1-hour, and 1.5-hour events. During 1.5-hour events, the response falls slightly in the third half-hour. For the longest event (4 hours), the per half-hour increase in consumption is smaller in magnitude but sustained throughout the event window, consistent with customers having a finite capacity to increase demand that becomes spread across a longer duration.

Across all event lengths, there is no evidence of anticipatory reductions in consumption in the hours before events. Displacement effects are also not apparent from the post-event coefficients. We suspect displacement is likely to be diffuse across the long pre- and post-event windows and therefore difficult to detect in an event-study framework. We return to this in the temporal demand substitution analysis in Section 5.3.3.

5.3.3 Temporal demand substitution

To further assess whether the consumption increases during events represented net demand creation or intertemporal substitution, we aggregated consumption across three windows: the 24 hours before an event, the event window itself, and the 72 hours following the event. Given the overlapping windows created by multiple events per day, we restricted this analysis to event days with no overlap, leaving 6 event days.

We found suggestive evidence of displacement, shown in Table A18. The resulting coefficients before and after the event are noisy, leading to a noisy overall estimate of total consumption in the −24 to +72 hours that encompass the event. However, we suspect there may be displacement that was too diffuse to reliably detect. Taking the point estimates at face value, the net effect over the full window was positive, but about a fourth of the size of the effect during the event. This would be consistent with some intertemporal demand substitution, but overall net demand creation. The presence of some temporal substitution was also consistent with what we found in turn-up events in Great Britain (Section 4.5).

6 Welfare analysis

This section develops a partial-equilibrium framework to characterize when demand turn-up (DTU) constitutes a welfare-improving substitute for curtailment. Our goal is not to provide a full structural model, but to isolate the key forces determining (i) cost-effectiveness relative to curtailment and (ii) allocative efficiency in the presence of network constraints and imperfect retail price pass-through.

We focus our welfare analysis on the case of curtailment in a transmission-constrained region with surplus generation. A growing body of evidence suggests that a substantial share of renewable curtailment is driven by local network congestion rather than economy-wide low wholesale prices: congestion-related constraints account for the majority of curtailed energy in several high-renewables systems, concentrated in specific nodes or transmission corridors rather than uniformly across time (Frysztacki and Brown, 2020; Maji et al., 2025), and recent evidence for California traces a large fraction of curtailment to transmission bottlenecks rather than system-wide oversupply (Perry et al., 2026). National over-supply, signaled by negative wholesale prices, is a distinct case; we treat it as a subordinate extension at the end of this section (Section 6.6).

A key input into our framework is the estimated demand response from our natural field experiment. Across the 50% discount and free tariff treatments, we estimated a price elasticity of demand of approximately -0.3. This provides direct empirical evidence that households do respond to short-run price incentives, even in the absence of self-selection into participation in the program. In other words, observed DTU was not driven solely by compositional effects, but reflected a measurable average behavioral response at the population level. Heterogeneity analysis further suggests that this responsiveness is likely to increase with the diffusion of low-carbon technologies, as EV and heat pump owners exhibited significantly higher elasticities (see Figure 5, and Figure A22).

How this documented demand response compares to curtailment costs is shaped by the prevalence of output-based support. The generator would only agree to curtailment if the payment covers everything it would have earned by staying on, including output-based subsidies from schemes such as Contracts for Difference. Thus curtailment is not free for the system operator, and demand turn-up's central appeal is that it avoids this compensation while letting generation proceed. But output-based subsidies also complicate what counts as a saving: when DTU keeps a generator running, the subsidy that curtailment would have avoided is now paid out, partly offsetting the avoided compensation. Correctly netting these flows is central to the cost-effectiveness comparison, which we develop in Section 6.5.

In brief, DTU events are welfare-improving across a wide range of procurement prices. The underlying logic is one of gains from trade: a single national price can mask locationally different marginal values, and by moving otherwise-curtailed energy to consumers who value it above its near-zero local cost, DTU realizes surplus that curtailment leaves on the table. The social cost of serving one more unit via DTU rather than curtailing is just the resource cost incurred to make that unit available, so DTU is welfare-optimal when demand is procured at the real resource cost of curtailment – a threshold we estimate at around £3/MWh. In this way, DTU temporarily replicates the outcome that locational marginal pricing would deliver at the constrained node.

The remainder of the section proceeds as follows. Section 6.1 sets out the model, calibrates its parameters, and proves the central result: a welfare-maximizing DTU price equal to the cycling cost of curtailment, invariant to whether curtailment is compensated. Section 6.2 decomposes the net welfare gain at that price and shows it to be small, positive, and identical across the two countries. Section 6.3 draws out the distributional and institutional implications. Section 6.4 quantifies how much constrained curtailment

British demand could realistically absorb. Section 6.5 turns to cost-effectiveness, showing why the avoided fiscal outlay is an unreliable guide to the welfare gain. Section 6.6 extends the framework to the case of national oversupply.

6.1 Welfare model

We evaluate the welfare effects of DTU events using the marginal value of public funds (MVPF) framework (Hendren and Sprung-Keyser, 2020; Hahn et al., 2026). We treat the system operator and subsidy administrator collectively as the "public sector," and model the retailer as a competitive pass-through agent that transmits procurement cost changes one-for-one to consumers.

A consumer in the constrained region faces retail price $R = w + N$, where w is the national wholesale price and N captures non-energy costs invariant to the DTU mechanism. Let p denote the DTU procurement price: the system operator's payment to the retailer per MWh of demand procured, with $p > 0$ a payment from the system operator to the retailer and $p < 0$ the reverse. With full pass-through, the consumer price falls from R to $R - (w + p)$, so consumers save $(w + p)$ relative to baseline; DTU lowers the consumer price whenever $p > -w$. With baseline consumption Q_0 and elasticity $\varepsilon < 0$, this induces additional demand $\Delta Q(p) = -\varepsilon Q_0(w+p)/R > 0$. Because the retailer is a pass-through conduit, this $(w + p)$ reduction simply reflects its own procurement cost falling by the same amount: w because the constrained energy is local surplus with near-zero opportunity cost, and p because the system operator funds it.[16]

Producers (generators): Under DTU the generator sells the electricity, receiving the wholesale price w and the per-unit renewable support payment σ; under curtailment it receives compensation ϕ but bears cycling and maintenance cost κ.[17] The per-unit

[16] No agent finances an unmatched w – the retailer nets to zero, and the wholesale payment it makes to the generator is the mirror of the wholesale cost the consumer no longer bears. The redispatch needed to balance the wider system is identical whether the constrained region is cleared by curtailment or by DTU, so DTU calls forth no additional generation: the w is a transfer between consumer and generator, not a net resource cost.

[17] κ reflects genuine mechanical wear from active curtailment – chiefly cycling and stress in wind turbines – and has no clean analogue for solar PV, which is curtailed via inverter power limiting with no moving parts. We hold κ fixed across the Great Britain and Spain illustrations for comparability; the empirical calibration in Appendix G.1 is drawn from, and applies to, wind curtailment specifically. To the extent Spain's curtailed volume includes a larger solar share than Great Britain's, its true κ is likely smaller.

producer surplus gain from DTU relative to curtailment is therefore:

$$\Delta\text{PS} = \underbrace{(w+\sigma)}_{\substack{\text{wholesale revenue}\\ \text{+ support payment}}} - \underbrace{(\phi-\kappa)}_{\substack{\text{curtailment comp}\\ \text{- cycling/maintenance}}} = w+\sigma+\kappa-\phi. \tag{3}$$

Government: The net government cost of DTU relative to curtailment reflects the DTU payment and support payment on each additional MWh generated, minus avoided curtailment compensation:

$$\Delta\text{GovtCost} = p+\sigma-\phi. \tag{4}$$

Institutional design: The curtailment compensation ϕ depends on how or whether system operators remunerate constrained-off generators. This market design feature differs across our two settings.

In Great Britain, a constrained-off generator is paid its own accepted constraint-down bid. Under competitive bidding, that bid would just recover what the generator genuinely forgoes – its output-based support payment (σ), together with the real cycling and maintenance cost that curtailment imposes (κ), giving a competitive benchmark of $\phi = \sigma + \kappa$. We therefore write $\phi = \mu\sigma + \kappa$, where $\mu = 1$ is that competitive benchmark and $\mu > 1$ is a strategic mark-up over the forgone subsidy – bidding above cost that a fully competitive constraint market would compete away. In Spain, generators hold no firm access right and are curtailed by instruction rather than by bid, so no such payment arises and $\phi = 0$. We show the resulting welfare accounting under these two institutional designs in Table 2.

We calibrate the two cost parameters from British constraint-down bids over the year to 18 February 2026. Benchmarked against the CfD top-up a generator forgoes when curtailed, bids exceed this fundamental by a volume-weighted average of about £11/MWh (26% of the CfD top-up) – this is the gross wedge reported in Table A21 in Appendix G.1. We estimate that of this £11, roughly £3/MWh (7%) is genuine cycling and maintenance cost, which enters the model as κ, while about £8/MWh (19%) is strategic mark-up, corresponding to $\mu \approx 1.19$, a wedge we read as market power in constraint bidding, consistent with Fabra and Imelda (2023).[18]

Combining (3) and (4) and aggregating across inframarginal and induced demand

[18] Discussions with Octopus Energy Generation suggest cycling and maintenance costs are on the order of £3/MWh. The full regime-by-regime analysis, including the treatment of Renewables Obligation and merchant generators, is set out in Appendix G.1.

Table 2: Welfare accounting under compensated and uncompensated curtailment

Regime	Curtailment comp., ϕ	ΔPS	ΔGovtCost
Great Britain	$\mu\sigma + \kappa$	$w + \sigma(1-\mu)$	$p + \sigma - \mu\sigma - \kappa$
Spain	0	$w + \sigma + \kappa$	$p + \sigma$

yields the MVPF in the form of Hendren and Sprung-Keyser (2020); Hahn et al. (2026). Of the $(w + p)$ price reduction on consumers' baseline consumption, the w component is the consumer-generator transfer discussed above – funded by no one and net-zero in social welfare – while the public-sector-mediated component $Q_0 \cdot p$ enters consumers' willingness to pay and the public sector's cost identically, and so cancels in the ratio:

$$\text{MVPF} = \frac{1 + \dfrac{\Delta\text{CS} + \Delta\text{PS}}{R}(-\varepsilon)}{1 + \dfrac{\Delta\text{GovtCost}}{R}(-\varepsilon)}, \tag{5}$$

where R is the retail price, $-\varepsilon > 0$ is the absolute price elasticity of demand, and the two "1" terms are the canceling government-mediated inframarginal transfer $Q_0 \cdot p$. The full derivation is given in Appendix G.2.

Proposition 6.1.1. *The marginal MVPF is strictly decreasing in p. For any value of ϕ, marginal MVPF equals 1 at $p = \kappa$, the welfare-maximizing DTU price.*

Proof. See Appendix G.2.

At $p = \kappa$, the system operator pays retailers exactly the cycling cost of curtailment – the only real resource cost at stake. The producer surplus gain – net of the wholesale-price transfer w, which washes against consumer surplus (see Appendix G.2), leaving $\sigma + \kappa - \phi$ – exactly offsets the remaining government cost, regardless of how ϕ allocates the cycling cost. For $p < \kappa$ the marginal MVPF exceeds one; for $p > \kappa$ it falls below one. The *overall* MVPF at $p = \kappa$ – total willingness to pay divided by total government cost – is by contrast a little above one, approximately 1.05 in our British calibration, with the welfare gain accruing as consumer surplus rather than as a fiscal saving; we decompose it in Section 6.2.

The result has an instructive connection to locational marginal pricing: with dispatch-based subsidies intact, self-curtailment would drive constrained-region prices to $-(\kappa + \sigma)$, overshooting the optimum, because the extra σ of consumer saving is a government transfer valued at less than face value at the margin. The efficient outcome would require suspending output-based subsidies when prices turn negative, so the effective floor remains at $-\kappa$ – precisely what DTU at $p = \kappa$ delivers.

6.2 Net benefits at the optimal price

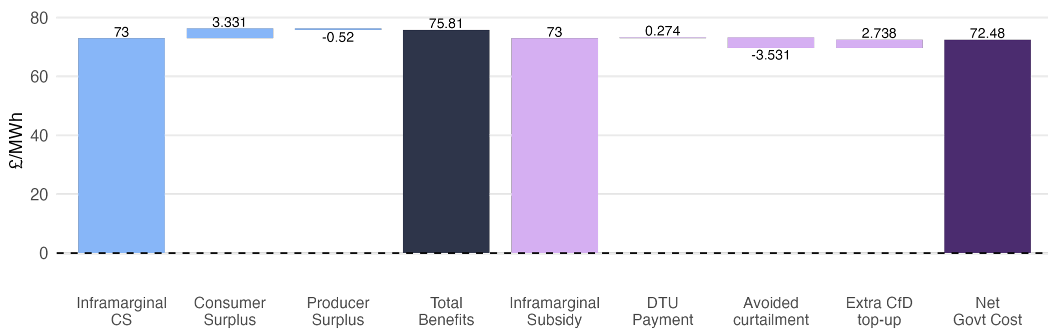
Figure 9 decomposes this net gain at the optimal price for each country. In Great Britain, almost all of the gain is consumer surplus – households consuming electricity they value above its near-zero local cost. Producer surplus is actually slightly negative here, a side-effect of producers losing their mark-up in constraint-down bids. On the fiscal side, the DTU payment and the extra subsidy paid on newly generated output are largely offset by the curtailment payments avoided. What remains is a small positive net benefit – an overall MVPF of about 1.05. The gain is real but modest, and shows up as a benefit to consumers rather than a saving to the public purse. Spain reaches the same total by a different route: with no curtailment payment to avoid, the offsetting entries instead show up as producer surplus, as the generator recovers the subsidy and cycling cost it would otherwise have lost. That the two totals match despite this different composition is Proposition 6.1.1 made visible.

Welfare is highest when national wholesale prices are highest: For illustration, the waterfall fixes the wholesale price at £70/MWh (a commonly cited average wholesale price), but this is an input that varies by half-hour, and the welfare gain rises with it. It is tempting to judge demand turn-up by how much it saves on curtailment payments. However, welfare gains come not from the constraint payments avoided, but from the cheap electricity that reaches consumers when otherwise-wasted generation is put to use. That gain is largest exactly when the payment saving is smallest. When wholesale prices are high, the gap between what households pay and the near-zero local value of constrained power is at its widest, so letting that power through delivers the most benefit to consumers – yet high wholesale prices also shrink the subsidy top-up, and with it the curtailment payment the system operator avoids. Thus welfare gains and fiscal savings move in opposite directions with respect to wholesale price.[19]

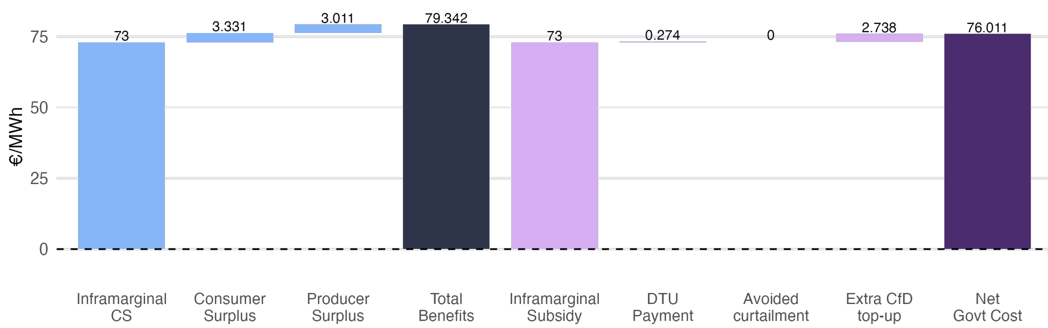
[19] Formally, at a given p the consumer-surplus gain rises with w (induced demand is proportional to $w + p$), while the avoided curtailment payment $\phi = \mu\sigma + \kappa$ falls, since a higher w compresses the subsidy top-up σ. See Section 6.1.

Figure 9: Net benefits waterfall at optimal DTU price ($p = \kappa$)

(a) Great Britain



(b) Spain



Notes: This figure decomposes the net welfare benefits of demand turn-up at the optimal DTU price $p = \kappa =$ £3/MWh (€3/MWh) for Great Britain (top panel) and Spain (bottom panel). Net benefits are computed as total willingness to pay minus net government cost, with inframarginal transfers – which appear symmetrically in both – netted out. The remaining components are consumer surplus from induced demand, producer surplus (reflecting generator cycling cost savings and, in Spain, recovered output-based subsidy payments), DTU payment, avoided curtailment costs, and extra output-based subsidy top-up on new generation. Net benefits are identical across the two panels, illustrating Proposition 6.1.1: the welfare gain from DTU at the optimal price is independent of whether curtailment is compensated (Great Britain) or mandated without payment (Spain). Assumes a wholesale price of £70/MWh (€70/MWh), an output-based subsidy of £30/MWh (€30/MWh), a retail price of £240/MWh (€240/MWh), and a price elasticity of demand of −0.3. We hold κ fixed across both panels for comparability; as noted in Section 6.1, κ is chiefly a wind-turbine mechanical-wear cost with no clean solar analogue, so this figure should be read as illustrative rather than as a claim that Great Britain and Spain share an identical true κ.

6.3 Welfare considerations and policy implications

Institutional design shifts the distribution of surplus, but not net benefits: The distinction between Great Britain and Spain is one of property rights over transmission access: British generators hold a firm right and must be compensated when curtailed (Newbery et al., 2018), while Spanish generators hold no firm access right and can be curtailed without payment.

Net benefits are nevertheless identical, because every item that constitutes a government saving in Great Britain has a mirror image in Spain on the other side of the welfare accounts: the cycling cost compensated in a British bid appears as a fiscal saving, while in Spain the same cost, borne by the generator, appears as a producer surplus gain; the forgone subsidy saved under British DTU is, in Spain, simply recovered by the generator. Even the strategic mark-up is a transfer from government to producer that washes out in aggregate welfare. The ledger entry differs; the total does not (Figure 9).

Welfare-neutrality, however, does not imply the transfers are harmless, as illustrated by the strategic mark-up. In aggregate, the mark-up is a transfer from government to producer that washes out. But it remains a real fiscal expenditure, paid above fundamental cost whenever curtailment is chosen over DTU, and transfers that a welfare analysis treats as a wash may matter a great deal to policymakers.

The British and Spanish cases are better understood as two points on a continuous design spectrum, with the compensation rule ϕ free to take a range of values. This is because ϕ is a contract choice. A related point is that, as Van Steenberghe and Ovaere (2026) have shown, CfD-type contract design shapes curtailment incentives, where a contract that withholds top-up during constrained periods would shrink ϕ, and with it the compensation for the lost subsidy.

Stripped of the accounting, the underlying logic is simple gains from trade: whenever generation would otherwise be curtailed, value sits on the table – avoided equipment wear, electricity reaching consumers nearer its true momentary cost, and revenue the system operator could earn from retailers seeking cheap constrained electricity – and DTU realizes it. The institutional design governs the distribution of these gains, not their existence.[20]

Figure A26 plots the marginal MVPF across the full range of p. It is strictly decreasing and crosses one at $p = \kappa$, so pricing DTU above the cycling cost destroys value at the margin, while pricing below it leaves achievable welfare uncaptured. The two country curves are nearly coincident despite their different ϕ. The institutional difference between Great Britain and Spain affects how the resulting surplus is divided at the optimum, not whether it exists there (Proposition 6.1.1). In practice, though, even the divided gains at the optimum are largely unrealized, for reasons we turn to now.

[20] Relatedly, DTU does not unravel the standard variable tariff in the way one-way price exposure can in the presence of adverse selection (Borenstein, 2005*b*), because participants capture a genuine system efficiency gain rather than a cross-subsidy withdrawn from other consumers; we discuss this in Appendix G.3.

Barriers to realization: In practice, these gains are largely unrealized: no mechanism in either country lets a retailer literally purchase constrained energy at locational prices.

In Great Britain, this is technically feasible through the Balancing Mechanism, but DTU faces significant frictions there: the short gate-closure window leaves little time to coordinate a consumer response at scale, and registration and settlement complexity weighs more heavily on distributed demand than on large generators.[21] One partial exception is the Demand Flexibility Service, whose day-ahead notification eases the coordination problem. Its reach is limited, however, by being seasonal, centrally procured, and subject to annual decisions on whether to run it. The priority in Great Britain is therefore to reduce frictions within the existing Balancing Mechanism – longer notification windows, lighter-touch metering for aggregated demand, clearer commercial routes for retailers to access constrained-region pricing – with locational signals in the Demand Flexibility Service serving a similar end.

Table 3: Comparison of curtailment and DTU frameworks: Great Britain and Spain

	Great Britain	**Spain**
Generator access rights:	Firm – generators must be compensated when curtailed	Non-firm – system operator can mandate curtailment without payment
Curtailment cost to government:	High – constraint-down payments cover forgone CfD top-up, cycling cost, and strategic mark-up	Low – direct fiscal cost is near zero
Generator locational siting incentives:	Weaker – curtailment risk is socialized, reducing incentive to site away from constrained areas or co-locate storage	Stronger – generators bear curtailment risk directly, creating incentives to co-locate storage and site in less constrained locations
DTU market infrastructure:	Balancing Mechanism exists in principle but access barriers limit demand-side participation in practice	Balancing markets exist but mandated curtailment creates no commercial pathway for retailers to access cheap constrained electricity
Primary distributional consequence:	Reduced fiscal burden on the system operator (avoided curtailment compensation)	Improved generator utilization and profitability (recovered subsidy and cycling costs)
Aggregate welfare:	Identical at optimal DTU price $p = \kappa$ (Proposition 6.1.1)	

Notes: This table summarizes the key institutional differences between Great Britain and Spain relevant to the cost-effectiveness and welfare analysis of demand turn-up. Despite these differences, the welfare gain from DTU at the optimal price is identical across the two settings, as shown formally in Section 6.1.

In Spain, the constraint is more fundamental: administrative curtailment means no

[21] Gate closure occurs one hour ahead of delivery, with acceptances arriving at any point within that window, potentially leaving only minutes to mobilize demand. The Virtual Lead Party framework introduced under BSC modification P344 has eased aggregator access but has not eliminated this asymmetry.

market exists through which demand can bid to absorb excess supply, so the same institutional choice that allows curtailment without payment also removes the price signal that would attract demand-side turn-up. Thus in Spain and similar non-firm access regimes, the more foundational step is to create a mechanism through which demand can participate during curtailment at all.

Our experimental evidence (Section 4) suggests meaningful demand response is achievable under both settings given salient price signals; the binding constraint is market access, not consumer behavior. Full locational marginal pricing would deliver the outcome endogenously, transmitting surplus signals directly through prices so that demand responds without administrative intervention – though even then, dispatch-based subsidies may push prices below the welfare optimum, leaving a residual role for a DTU-like mechanism. The British and Spanish cases, and this locational-pricing limit, sit at different points on a single spectrum of market designs (Table 3).

6.4 The size of the prize in Great Britain

The welfare framework characterized the per-unit gain from demand turn-up; we now quantify how much constrained curtailment British demand could absorb, and the consumer surplus this would generate. We focus on Great Britain simply because its granular balancing and settlement data are more transparent and publicly accessible than Spain's.

Methods: We used Great Britain's National Energy System Operator's bid-acceptance data for the period 19 February 2025 to 18 February 2026, identifying curtailment from local network constraints as system-operator-accepted bid volume in half-hours with a positive system sell price, and distinguishing it from national over-supply, which we associate with negative system prices.[22]

We restricted our attention to wind generation, the source of essentially all constrained-off renewable volume in the sample. In periods of positive system prices – consistent with local rather than national constraints – we identify approximately 8.4 million MWh of system-operator-actioned wind curtailment. This is the curtailment our mechanism aims to reduce.

We then applied the demand response estimated in our natural field experiment to

[22]The sign of the system price is a proxy: a positive price alongside curtailment indicates the system as a whole was not long, so the binding constraint was local rather than national.

each curtailment half-hour, to compute the induced demand and resulting consumer surplus. We assumed the retail price fell by the prevailing system price – the wedge between the national price and the near-zero local value of constrained electricity – and applied our estimated elasticity of -0.3 to a baseline half-hourly household consumption of 0.17 kWh. We set the demand turn-up price $p = 0$ throughout, rather than the welfare-optimal $p = \kappa$ derived in Proposition 6.1.1.[23] Since the marginal MVPF exceeds one for all $p < \kappa$, this choice leaves some feasible welfare gain uncaptured, so the prize we estimate below is conservative relative to a fully optimized mechanism. Consumer surplus on the induced demand was the standard triangle $\frac{1}{2}\Delta p\,\Delta Q$; consistent with Section 6.1, inframarginal savings are transfers and net out of the welfare measure.

Two points about how to interpret this exercise are worth making explicit. First, the curtailment payment ϕ is sized to cover everything the producer forgoes by being curtailed, including the cycling cost κ; producers are consequently already fully compensated for that cost, so when demand turn-up removes the need for the payment, the saving accrues entirely to the public sector[24] rather than to the producer. Second, the public sector saving is a net figure: the curtailment payment the system operator avoids, minus the subsidy top-up now paid out on the newly generated output (via the Low Carbon Contracts Company for CfD units, or suppliers directly for Renewables Obligation supported generators) – not the gross avoided payment alone.

Results: Figure 10 decomposes the realistic Scotland-capped prize into genuine welfare and pure transfers, applying the $\phi = \mu\sigma + \kappa$ accounting of Section 6.1 to the actual bid and price data underlying Section 6.4, rather than to the illustrative calibration used in Figure 9.

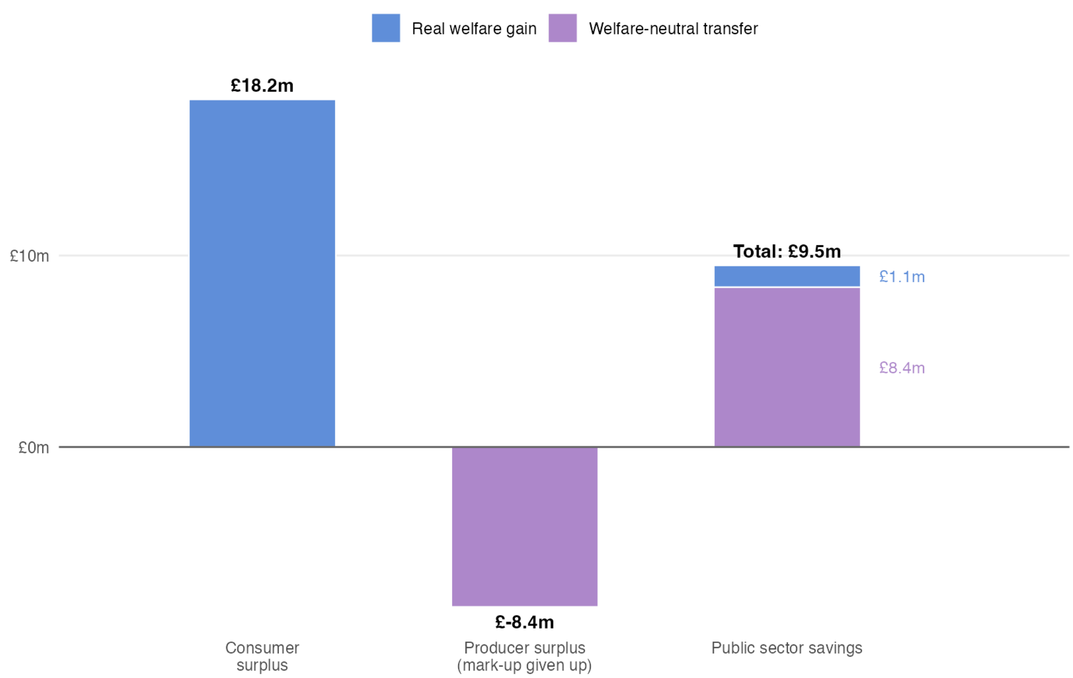
Consumer surplus from induced demand is £18.2 million. Against this, producers give up £8.4 million in mark-up – the strategic component of ϕ identified in Appendix G.1 ($\mu \approx 1.19$), reversed because demand turn-up lets the constrained generator run rather than be paid to switch off. This £8.4 million reappears, in full, as public sector saving:

[23] This choice means that no separate payment from the system operator to the retailer, or vice versa, enters the accounting. Averaged across curtailment half-hours (volume-weighted), the system price was approximately £75/MWh (7.5p/kWh); relative to a typical retail price of around £240/MWh (24p/kWh), the assumed price cut takes the unit rate to roughly £165/MWh (16.5p/kWh) on average – though the actual price cut, and hence the induced demand, varies half-hour by half-hour with the prevailing system price rather than sitting fixed at this average.

[24] We use "public sector" loosely, to mean the billpayer collectively, regardless of which body administers a given payment. Curtailment payments, CfD top-ups, and Renewables Obligation costs are administered by different bodies – the system operator, the Low Carbon Contracts Company, and suppliers, respectively – but all are recovered through volumetric charges on final electricity consumption. Whichever body disburses or avoids a given payment, the pound ultimately comes from, or returns to, the same place.

the mark-up is a transfer from the public sector to producers under curtailment, and demand turn-up simply stops that transfer, so what the producer loses the public sector exactly recovers.

Figure 10: Genuine welfare vs. transfers, Scotland-capped demand turn-up



Notes: This figure decomposes the Scotland-capped prize (2.55m households, DTU price $p = 0$) into consumer surplus, producer surplus given up, and public sector savings. The public sector savings bar splits into the strategic mark-up reversed (purple, a transfer that exactly cancels the producer's loss) and the avoided cycling/maintenance cost (blue, a genuine resource saving realized through the fiscal channel, and therefore counted again in total genuine welfare of £19.3m). This analysis follows standard practice in the MVPF literature (Hendren and Sprung-Keyser, 2020) in evaluating net public sector cost at face value, without closing the budget constraint or adjusting for the marginal cost of the revenue used to finance curtailment compensation. In Great Britain this revenue is raised via Balancing Services Use of System (BSUoS) charges, recovered from electricity billpayers. To the extent BSUoS is levied volumetrically, financing it imposes a deadweight loss on electricity consumption at the margin, which we do not net out in the main analysis. Because demand turn-up requires financing less of this levy than curtailment does, incorporating it would strengthen demand turn-up's advantage over curtailment.

The public sector saving also captures the genuine cycling and maintenance cost, $\kappa \approx$ £3/MWh, that curtailment would otherwise have compensated the generator for incurring. Unlike the mark-up, κ has no offsetting loser: producers are financially indifferent to it either way – compensated for a real cost under curtailment, no compensation but no cost under demand turn-up – so it is a genuine resource saving rather than a transfer, and belongs in the welfare total (£19.3 million) as well as the public sector ledger (£9.5 million). Of that £9.5 million, only £8.4 million – the reversed mark-up – is a transfer that cancels against producer surplus; the remaining £1.1 million is additional to, not separate from, the £18.2 million of consumer surplus.

This exercise is Scotland-capped: it takes the same population constraint developed in Figure A27 and Figure A28 – the maximum number of households that a given half-hour's curtailed volume would actually require, but capped at the estimated 2.55 million households in Scotland – and applies it to every curtailment half-hour in the sample rather than to a single illustrative participation level. The £18.2 million consumer-surplus figure is consequently smaller than the £314 million theoretical ceiling reported in Figure A27, which assumes participation unconstrained by population; the gap between the two is the subject of Figure A28 and the discussion below. For the underlying participation-sweep relationship, the shape of the ceiling, and the household requirement that drives the gap between the realistic and theoretical figures, see Figure A27 and Figure A28 in the appendix.

Why the realistic prize is modest, and why it will grow: That the realistic figure is a small fraction of the £314 million ceiling reflects the limited scale of residential demand response, so demand turn-up through residential consumers is best understood as a complement to storage, transmission, and indeed curtailment, not a full substitute.

However, two present-day conditions, if they changed, would raise it. First, the exercise is a static snapshot of a quantity rising rapidly with renewable deployment, while the relevant flexibility is increasingly that of high-consumption, price-elastic low-carbon technologies (Figure 5) and of flexible loads – electrolyzers, data centers, charging hubs, and storage – that need not be residential. As an illustration, if every Scottish household had an EV and exhibited the elasticity and baseline consumption we observed among EV-tariff customers, the same Scotland-capped consumer surplus would rise from £18.2 million to approximately £64 million – still short of the £314 million ceiling, but a substantial share of the gap closed by technology adoption alone. Second, retail tariffs recover network and policy costs through substantial volumetric charges, so even a fall in the energy price to its local marginal cost produces only a modest proportional reduction in the price consumers face (Borenstein and Bushnell, 2022). The demand-side signal is throttled by tariff structure rather than by the underlying value of constrained electricity.[25] If network and policy costs were recovered less heavily through volumetric charges, the same energy wedge would mean a larger proportional price reduction, a larger per-household response, and a greater share of curtailment within reach.

[25] Partly with this in mind, the UK Department for Energy Security and Net Zero (2026) recently announced a levy exemption on DTU on a trial basis.

6.5 Cost-effectiveness versus welfare, and the case for targeting

The consumer surplus above is the welfare gain; it is not the same as the fiscal flows the mechanism mobilizes, and the gap between them carries a practical lesson. As with carbon abatement, an intervention can carry a high headline cost per unit yet still be good value, because much of the spending is a transfer that returns to recipients rather than a resource cost. Demand turn-up separates the two sharply. Measured as the payment per unit of *induced* consumption – pounds paid per additional kWh actually turned up – it is expensive, ranging from approximately £1.22 to £1.76/kWh in our experiment (Table A13), against an energy value of a few pence.

Yet this does not make it poor value: almost all of that payment is a transfer back to consumers, not a resource cost, which is why the MVPF remains near one. The operational lesson is that the gain may be best captured by procuring the induced response directly – paying for the demand actually turned up – rather than by cutting the price faced by all consumption in the constrained region, which would move a transfer an order of magnitude larger than the gain it delivers. In practice this favors an explicit opt-in or enrollment stage: a welfare analysis is indifferent to this, but where cost-effectiveness or the optics of paying for non-response matter, targeting dominates a market-wide price cut.

Demand displacement: One additional consideration strengthens the cost-effectiveness case. As documented in our nationwide Great Britain trial (Table A14) and suggested by the NPG trial (Table A18), DTU induces demand shifting across hours: increased consumption during low-cost constrained periods partly substitutes for consumption that would otherwise have occurred in higher-cost hours. The relevant cost of DTU to the retailer is therefore not the payment for additional demand in the event hour alone, but the net change in procurement costs across all hours – which will generally be lower, and may be negative even when the contemporaneous DTU price is positive. Retailers should in principle be willing to participate at prices that appear costly in the event hour, since the cross-hour procurement savings offset part of the event-hour cost – though how much of this saving the retailer retains, rather than passing through to customers, depends on the tariff structure. We treat this cost saving as an important possibility to keep in mind. However, to be conservative, our welfare framework and estimates have ignored cross-hour displacement, taking each event hour in isolation.

6.6 How national over-supply fits in

The analysis above concerns regional constraint payments, arising when transmission bottlenecks prevent surplus generation in one area from reaching deficit areas elsewhere – the Scottish wind problem being the leading example. Here the market failure is clear: the wholesale price transmits the wrong signal to local consumers, and DTU corrects a genuine allocative distortion. A distinct case is negative national wholesale prices, when total system supply exceeds demand; here the wholesale market is functioning correctly, and the welfare rationale is weaker. Even then, however, negative national prices in Spain or Great Britain often arise because the interconnectors linking them to the broader European grid are themselves constrained. The regional constraint problem does not disappear at the national border; it reappears at the level of cross-border transmission capacity. The practical implication is simply that the same welfare case for DTU applies at each level: the gains from trade between a surplus region and a deficit region are real whether the binding bottleneck is a Scottish or Galician transmission line or a full Anglo-Irish or Spanish–French interconnector.

7 Conclusion

This study provides the first large-scale, cross-country field experimental evidence on consumer responsiveness to time-bound electricity price *reductions* and negative pricing in the context of increasing renewable penetration. We randomized 120,000 households across Great Britain and Spain into varying financial incentives, finding that residential demand was elastic when prices fell toward zero. However, we identified a "zero-price" threshold: 50% and 100% price decreases were associated with a price elasticity of demand of -0.3, while the marginal gains from further payments ("negative" pricing) were minimal. While we found evidence of intertemporal shifting in Great Britain, in Spain our results suggested net demand creation.

The results highlight the close relationship between demand-side flexibility and the broader electrification of the economy. Households equipped with low-carbon technologies, particularly electric vehicles, exhibited elasticities four to five times greater than traditional households. This suggests that technical potential for demand turn-up to balance intermittent renewable supply may grow as the global vehicle fleet and heating systems transition to electricity. Our findings indicate that demand turn-up is possible without any sign-up or opt-in selection on the part of customers, as demonstrated by our main trials; however, responsiveness and cost-effectiveness both increase when a sign-up stage is introduced to select for more engaged customers.

From a system operation perspective, demand turn-up offers a partial alternative to curtailment as a mechanism for managing surplus renewable generation. Rather than relying solely on supply-side reductions, system operators can leverage short-run demand responses to help manage periods of congestion and excess generation. Our results also indicate that the effectiveness of this tool depends on the composition of the demand side and the presence of flexible electrified loads, implying that its value will evolve alongside ongoing changes in system design and electrification.

From a welfare perspective, demand turn-up offers an improvement over traditional curtailment by addressing the allocative inefficiency associated with retail prices remaining high due to fixed levies and limited pass-through even during periods of low marginal costs of generation. Aligning consumption with these periods of surplus lets consumers access electricity near its true local cost, generating a real efficiency gain rather than a transfer. This gain is identical in size regardless of whether curtailment is compensated, though it surfaces differently depending on institutional design. In systems where generators are remunerated for curtailment, the gain partly shows up as a reduced fiscal burden on the system operator. In systems where generators must curtail without compensation, demand turn-up instead improves their utilization and profitability, since generation that would otherwise have been wasted is now sold. In both cases, consumers capture direct value from access to cheaper electricity; the institutional difference determines only how the remaining surplus is distributed between government and generators, not whether the underlying welfare gain exists.

References

Australian Government Department of Climate Change, Energy, the Environment and Water (2026), 'Solar Sharer Offer'.

Bailey, M. R., Brown, D. P., Shaffer, B. and Wolak, F. A. (2025), 'Show me the money! a field experiment on electric vehicle charge timing', *American Economic Journal: Economic Policy* **17**(2), 259–284.

Bernard, L., Hackett, A., Metcalfe, R. D., Panzone, L. A. and Schein, A. R. (2025), The

Impact of Dynamic Prices on Electric Vehicle Public Charging Demand: Evidence from a Nationwide Natural Field Experiment, Technical report, Centre for Net Zero.

Bernard, L., Metcalfe, R. and Schein, A. R. (2026), Critical peak rewards: Evidence from a nationwide demand response program, Working paper, Centre for Net Zero. AEA RCT Registry: AEARCTR-0012451.

Borenstein, S. (2005*a*), 'The long-run efficiency of real-time electricity pricing', *Energy Journal* **26**(3).

Borenstein, S. (2005*b*), 'Time-varying retail electricity prices: Theory and practice', *Electricity deregulation: choices and challenges* **4**, 317–356.

Borenstein, S. (2012), 'The private and public economics of renewable electricity generation', *Journal of Economic Perspectives* **26**(1), 67–92.

Borenstein, S. and Bushnell, J. B. (2022), 'Headwinds and Tailwinds: Implications of Inefficient Retail Energy Pricing for Energy Substitution', *Environmental and Energy Policy and the Economy* **3**(1), 37–70.

Brandon, A., Ferraro, P. J., List, J. A., Metcalfe, R. D., Price, M. K. and Rundhammer, F. (2026), 'Do the effects of nudges persist? Theory and evidence from 38 natural field experiments', *Review of Economic Studies* p. rdag012.

Burkhardt, J., Gillingham, K. T. and Kopalle, P. K. (2023), 'Field experimental evidence on the effect of pricing on residential electricity conservation', *Management Science* **69**(12), 7784–7798.

CEPA (2026), Impact assessment of P462: The removal of subsidies from bid prices in the balancing mechanism, Final report to BSC panel, Cambridge Economic Policy Associates (CEPA), prepared for Elexon. Accompanying databook available from Elexon.

Davis, L. W., Hausman, C. and Rose, N. L. (2023), 'Transmission impossible? Prospects for decarbonizing the US grid', *Journal of Economic Perspectives* **37**(4), 155–180.

Dechezleprêtre, A., Fabre, A., Kruse, T., Planterose, B., Sanchez Chico, A. and Stantcheva, S. (2025), 'Fighting climate change: International attitudes toward climate policies', *American Economic Review* **115**(4), 1258–1300.

Department for Energy Security and Net Zero (2025), Public attitudes tracker: Technical overview, spring 2025, Technical report, UK Government.

Department for Energy Security and Net Zero (2026), 'Government to make "plug-in solar" available within months'.

Doshi, G. (2026), 'Chicken and egg: Electricity transmission and investment in renewable energy', *Journal of Public Economics* **256**, 105598.

Fabra, N. and Imelda (2023), 'Market power and price exposure: learning from changes in renewable energy regulation', *American Economic Journal: Economic Policy* **15**(4), 323–358.

Fabra, N., Rapson, D., Reguant, M. and Wang, J. (2021), Estimating the elasticity to real-time pricing: evidence from the Spanish electricity market, *in* 'AEA Papers and Proceedings', Vol. 111, American Economic Association 2014 Broadway, Suite 305, Nashville, TN 37203, pp. 425–429.

Fowlie, M., Wolfram, C., Baylis, P., Spurlock, C. A., Todd-Blick, A. and Cappers, P. (2021), 'Default effects and follow-on behaviour: Evidence from an electricity pricing program', *The Review of Economic Studies* **88**(6), 2886–2934.

Frew, B., Cole, W., Denholm, P., Frazier, A. W., Vincent, N. and Margolis, R. (2019), 'Sunny with a chance of curtailment: operating the US grid with very high levels of solar photovoltaics', *iScience* **21**, 436–447.

Frysztacki, M. and Brown, T. (2020), 'Modeling Curtailment in Germany: How Spatial Resolution Impacts Line Congestion', *arXiv preprint arXiv:2009.03801* .

Gonzales, L. E., Ito, K. and Reguant, M. (2023), 'The investment effects of market integration: Evidence from renewable energy expansion in Chile', *Econometrica* **91**(5), 1659–1693.

Hahn, R. W., Hendren, N., Metcalfe, R. D. and Sprung-Keyser, B. (2026), 'A Welfare Analysis of Policies Impacting Climate Change', *American Economic Review* **116**(7), 2368–2421.

Hendren, N. and Sprung-Keyser, B. (2020), 'A unified welfare analysis of government policies', *The Quarterly Journal of Economics* **135**(3), 1209–1318.

Hinchberger, A. J., Jacobsen, M. R., Knittel, C. R., Sallee, J. M. and van Benthem, A. A. (2024), The efficiency of dynamic electricity prices, Technical report, National Bureau of Economic Research.

Hirth, L. (2013), 'The market value of variable renewables', *Energy Economics* **38**, 218–236.

Integrated National Energy and Climate Plan (PNIEC 2023-2030) (2024), Technical report, Ministerio para la Transición Ecológica y el Reto Demográfico.

Ito, K., Ida, T. and Tanaka, M. (2023), 'Selection on welfare gains: Experimental evidence from electricity plan choice', *American Economic Review* **113**(11), 2937–2973.

Jacobsen, H. K. and Schröder, S. (2013), 'Curtailment of renewable generation: Economic optimality and incentives', *Energy Policy* **49**, 663–675.

Jessoe, K. and Rapson, D. (2014), 'Knowledge is (less) power: Experimental evidence from residential energy use', *American Economic Review* **104**(4), 1417–1438.

La Nauze, A., Friesen, L., Lim, K. L., Menezes, F., Page, L., Philip, T. and Whitehead, J. (2026), 'Can electric vehicles aid the renewable transition? Evidence from a field experiment incentivizing midday charging', *Journal of Public Economics* **260**, 105696.

List, J. A. (2026), *Experimental economics: Theory and practice*, University of Chicago Press.

Maji, D., Irwin, D., Shenoy, P. and Sitaraman, R. K. (2025), A First Look at Node-Level Curtailment of Renewable Energy and its Implications, *in* 'Proceedings of the 16th ACM International Conference on Future and Sustainable Energy Systems', pp. 293–304.

Newbery, D., Pollitt, M. G., Ritz, R. A. and Strielkowski, W. (2018), 'Market design for a high-renewables European electricity system', *Renewable and Sustainable Energy Reviews* **91**, 695–707.

Novan, K. and Wang, Y. (2024), 'Estimates of the marginal curtailment rates for solar and wind generation', *Journal of Environmental Economics and Management* **124**, 102930.

O'Shaughnessy, E., Cruce, J. and Xu, K. (2021), Solar PV curtailment in changing grid and technological contexts, Technical report, National Renewable Energy Laboratory (NREL), Golden, CO (United States).

Perry, F. N., Perry, M. E., Leonard, S., Ochs, G., Yu, L., Xing, J., Vadlamani, A. and Lee, B. C. (2026), Curtail to Compute: How Data Centers Can Use Surplus Renewable Energy, Technical report, Next 10, California, USA.

Prol, J. L. and Zilberman, D. (2023), 'No alarms and no surprises: Dynamics of renewable energy curtailment in California', *Energy Economics* **126**, 106974.

Romano, J. P. and Wolf, M. (2005), 'Stepwise multiple testing as formalized data snooping', *Econometrica* **73**(4), 1237–1282.

Torriti, J. (2026), 'Demand turn-up in electricity markets: Economic lessons from non-energy sectors', *Energy Research & Social Science* **135**, 104664.

Van Steenberghe, M. and Ovaere, M. (2026), Market Inefficiencies in Renewable Support Policies: Evidence from Offshore Wind Contracts for Difference in Great Britain, Working Paper 2025/1124, Ghent University, Department of Economics.

Wolak, F. (2011), Designing a competitive wholesale electricity market, *in* 'Handbook of Regulatory Economics', Edward Elgar Publishing.

Yang, Y., Raman, G. and Peng, J. C.-H. (2025), 'Shaping residential electricity demand with negative pricing', *Nature Energy* **11**(1), 58–65.

Appendix

Table of Contents

A Tables

Table A1: Balance Table

Variable	Control Mean	Treatment	N
A. Great Britain			
Octopus tenure	4.06	-0.00	58,320
	[1.78]	(0.02)	
Estimated annual consumption (kWh)	3901.33	-1.81	58,320
	[3190.80]	(8.97)	
Energy rating	4.22	-0.01	58,320
	[0.78]	(0.01)	
TOU tariff	0.18	-0.00	58,320
	[0.38]	(0.00)	
Joint F-Test		0.64	
B. Spain			
Octopus tenure	0.65	0.01*	60,427
	[0.59]	(0.01)	
Baseline afternoon consumption (kW)	0.67	-0.00	60,438
	[0.49]	(0.00)	
TOU tariff	0.28	-0.00	60,423
	[0.45]	(0.00)	
Joint F-Test		0.31	

Notes: This table reports baseline balance between households assigned to treatment and pure-control households, separately for Great Britain (Panel A) and Spain (Panel B). For each covariate, the *Control Mean* column reports the mean in the pure-control group; standard deviations are in brackets. The *Treatment* column reports the coefficient on an indicator for treatment assignment from a separate OLS regression of the covariate on treatment status with randomization-block fixed effects; standard errors are in parentheses. *N* is the number of non-missing observations used in each row-specific regression. The *Joint F-Test* row reports the p-value from a test that all listed covariates are jointly unrelated to treatment assignment within panel. Estimated Annual Consumption is measured in kWh. Current energy rating is based on the official A–G efficiency scale and is mapped to a numerical value for analysis. TOU tariff is an indicator for enrollment in a time-of-use tariff. Baseline afternoon consumption is the mean hourly consumption between 13:00 and 17:00 in May 2025, before the start of the trial.

Table A2: Attrition

Sample	Mean Attrition (Treated)	Mean Attrition (Pure Control)	Difference	p-value
Great Britain	0.068	0.069	-0.002	0.540
Spain	0.071	0.070	0.001	0.627

Notes: This table reports attrition rates for treated and pure control customers in Great Britain and Spain. Attrition is defined as leaving Octopus Energy during the study period. The difference column reports the coefficient from a regression of attrition on treatment assignment with randomization-block fixed effects. P-values are from a test of the null hypothesis that attrition rates are equal across groups.

Table A3: Great Britain Robustness, impact of treatment on hourly consumption (kW)

Model:	Main (1)	kWh Control (2)	No Controls (3)	Customer FE (4)	Pure Control (5)
Variables					
50% Discount	0.049^{***}	0.049^{***}	0.050^{***}	0.046^{***}	0.051^{***}
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
	[0.00]				
Free	0.105^{***}	0.105^{***}	0.105^{***}	0.101^{***}	0.107^{***}
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
	[0.00]				
Free + 5p	0.115^{***}	0.116^{***}	0.116^{***}	0.112^{***}	0.117^{***}
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
	[0.00]				
Free + 15p	0.116^{***}	0.116^{***}	0.117^{***}	0.113^{***}	0.117^{***}
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
	[0.00]				
Rotating Control					0.004
					(0.003)
Fixed-effects					
Covariates	Yes	Yes	No	No	Yes
Event	Yes	Yes		Yes	Yes
Block	Yes	Yes		Yes	Yes
Customer				Yes	
Fit statistics					
Control Mean	0.340	0.340	0.340	0.340	0.337
Observations	271,040	271,040	271,349	271,225	271,040

Clustered (Customer) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Notes: This table reports ITT estimates of the effect of one-hour turn-up events on electricity consumption during the event in Great Britain. The main specification, Column (1), includes event and randomization-block fixed effects and controls for baseline weekday consumption, baseline weekend consumption, customer tenure category, estimated annual consumption (EAC) quartile, current energy rating, and meter type. Column (2) replaces EAC quartile with average baseline import during the event-relevant afternoon window. Column (3) omits controls and fixed effects. Column (4) adds customer fixed effects. Column (5) separates the rotating control group from the pure control group, using the pure control group as the omitted category. The reported control mean is mean consumption in the omitted control group for each specification. Standard errors are clustered at the customer level. Romano-Wolf multiple-hypothesis-adjusted p-values are reported in square brackets below the standard errors for the treatment coefficients in Column (1). The correction treats the four treatment-versus-control effects in the main specification as a single family.

Table A4: Spain Robustness, impact of treatment on hourly consumption (kW)

Model:	Main (1)	PAP Controls (2)	No Controls (3)	Customer FE (4)	Pure Control (5)	Reweighted (6)	First Hour (7)	Drop Event 5 (8)
Variables								
50% Discount	0.026***	0.026***	0.027***	0.025***	0.025***	0.019***	0.024***	0.024***
	(0.002)	(0.003)	(0.003)	(0.003)	(0.003)	(0.002)	(0.003)	(0.003)
	[0.00]							
Free	0.053***	0.053***	0.054***	0.053***	0.053***	0.042***	0.052***	0.052***
	(0.003)	(0.003)	(0.004)	(0.004)	(0.003)	(0.003)	(0.003)	(0.003)
	[0.00]							
Free + 10c	0.056***	0.056***	0.058***	0.059***	0.056***	0.042***	0.056***	0.057***
	(0.003)	(0.003)	(0.004)	(0.004)	(0.003)	(0.003)	(0.003)	(0.003)
	[0.00]							
Prize Draw	0.031***	0.031***	0.032***	0.033***	0.031***	0.023***	0.030***	0.029***
	(0.003)	(0.003)	(0.003)	(0.004)	(0.003)	(0.002)	(0.003)	(0.003)
	[0.00]							
Rotating Control					-0.0005			
					(0.003)			
Fixed-effects								
Covariates	Yes	Yes	No	No	Yes	Yes	Yes	Yes
Event	Yes	Yes		Yes	Yes	Yes	Yes	Yes
Block	Yes	Yes		Yes	Yes	Yes	Yes	Yes
Customer				Yes				
Fit statistics								
Control Mean	0.418	0.418	0.418	0.418	0.418	0.259	0.412	0.437
Observations	600,686	542,831	600,740	600,640	600,686	600,672	300,409	483,487

Clustered (Customer) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Notes: This table reports ITT estimates of the effect of two-hour turn-up events on hourly electricity consumption during the event in Spain. The main specification, Column (1), includes event and randomization-block fixed effects and controls for baseline weekday consumption, baseline weekend consumption, customer tenure category, and average baseline import during the event-relevant afternoon window. Column (2) uses the pre-analysis-plan control set. Column (3) omits controls and fixed effects. Column (4) uses customer fixed effects. Column (5) separates the rotating control group from the pure control group, using the pure control group as the omitted category. Column (6) reweights Spanish customers to match the Great Britain sample on baseline time-of-use tariff status and baseline afternoon consumption. Column (7) restricts the sample to the first hour of each Spain event. Column (8) drops Spain Event 5, for which treatment assignments repeated Event 4 assignments due to an implementation error. The reported control mean is mean consumption in the omitted control group for each specification. Standard errors are clustered at the customer level. Romano-Wolf multiple-hypothesis-adjusted p-values are reported in square brackets below the standard errors for the treatment coefficients in Column (1). The correction treats the four treatment-versus-control effects in the main specification as a single family.

Table A5: Impact of treatment on hourly consumption (kW), no solar / EV

	GB	GB, No solar/EV	Spain
Model:	(1)	(2)	(3)
Variables			
50% Discount	0.05***	0.04***	0.03***
	(0.003)	(0.003)	(0.003)
Free	0.11***	0.08***	0.05***
	(0.004)	(0.003)	(0.003)
Free + 5p	0.12***	0.09***	
	(0.004)	(0.003)	
Free + 15p	0.12***	0.09***	
	(0.004)	(0.003)	
Free + 10c			0.06***
			(0.003)
Prize Draw			0.03***
			(0.003)
Fit statistics			
Control Mean	0.34	0.34	0.42
Observations	271,040	245,889	542,831

*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Notes: This table reports regression estimates of the impact of treatment on hourly electricity consumption (kW) during turn-up events. Column (1) presents results for the full Great Britain sample, column (2) restricts the Great Britain sample to households without electric vehicles or solar, and column (3) reports results for Spain. All specifications follow the main empirical model and include the same set of controls and fixed effects as in the baseline results.

Table A6: Impact of treatment on public charging

	I(Charged)			kWh charged		
	Pre	During	Post	Pre	During	Post
Model:	(1)	(2)	(3)	(4)	(5)	(6)
Variables						
50% Discount	-0.0002	0.004	-0.001	-0.013	0.037	-0.011
	(0.001)	(0.004)	(0.001)	(0.010)	(0.047)	(0.010)
Free	0.0009	0.003	-0.0003	0.008	0.015	-0.002
	(0.001)	(0.004)	(0.001)	(0.012)	(0.039)	(0.010)
Free + 5p	0.001	0.004	8.94×10^{-6}	0.007	0.065	-0.003
	(0.001)	(0.004)	(0.001)	(0.012)	(0.051)	(0.010)
Free + 15p	0.0003	0.004	-0.0005	-0.0007	0.033	0.001
	(0.001)	(0.004)	(0.001)	(0.010)	(0.039)	(0.011)
Fixed-effects						
Event	Yes	Yes	Yes	Yes	Yes	Yes
Fit statistics						
Control Mean	0.003	0.005	0.004	0.031	0.042	0.03
Observations	102,791	4,310	304,832	102,791	4,310	304,832

Clustered (Customer) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Notes: This table tests whether British turn-up incentives displaced public EV charging among customers observed using Electroverse. The dependent variables are an indicator for any Electroverse charging and total Electroverse kWh in a given hour, measured separately over the 24 hours before the event, the event hour, and the 72 hours after the event. All specifications include event fixed effects and baseline controls; standard errors are clustered by customer.

Table A7: Great Britain heterogeneity by low-carbon technology tariff

Model:	(1)
Variables	
50% Discount	0.045***
	(0.003)
	[0.00]
Free	0.082***
	(0.003)
	[0.00]
Free + 5p	0.086***
	(0.003)
	[0.00]
Free + 15p	0.086***
	(0.003)
	[0.00]
50% Discount × Baseline LCT: Solar	0.006
	(0.019)
	[0.85]
Free × Baseline LCT: Solar	0.058**
	(0.025)
	[0.07]
Free + 5p × Baseline LCT: Solar	0.067***
	(0.026)
	[0.05]
Free + 15p × Baseline LCT: Solar	0.085***
	(0.028)
	[0.02]
50% Discount × Baseline LCT: EV	0.055*
	(0.028)
	[0.14]
Free × Baseline LCT: EV	0.327***
	(0.039)
	[0.00]
Free + 5p × Baseline LCT: EV	0.459***
	(0.047)
	[0.00]
Free + 15p × Baseline LCT: EV	0.448***
	(0.044)
	[0.00]
50% Discount × Baseline LCT: Solar & EV	0.025
	(0.051)
	[0.85]
Free × Baseline LCT: Solar & EV	0.408***
	(0.092)
	[0.00]
Free + 5p × Baseline LCT: Solar & EV	0.405***
	(0.104)
	[0.00]
Free + 15p × Baseline LCT: Solar & EV	0.330***
	(0.082)
	[0.00]
Baseline LCT: Solar	-0.107***
	(0.011)
Baseline LCT: EV	0.048*
	(0.026)
Baseline LCT: Solar & EV	-0.095***
	(0.028)
Fixed-effects	
Event	Yes
Block	Yes
Fit statistics	
Control Mean	0.341
Observations	284,334

Clustered (Customer) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Notes: This table reports the regression underlying Figure 5. The omitted groups are the control group and customers without EV or solar tariff markers. Treatment coefficients report effects for the omitted LCT group, LCT coefficients report baseline differences in the control group, and interactions report differential treatment effects relative to the omitted LCT group. The outcome is hourly event-window consumption. The specification uses the main Great Britain controls and event and randomization-block fixed effects; standard errors are clustered by customer. Romano-Wolf adjusted p-values are reported in square brackets below treatment and interaction coefficients. The adjustment treats these coefficients as a single family and excludes LCT main effects because they are not interpreted causally.

Table A8: Heterogeneity by baseline tariff type

Great Britain

Model:	(1)
Variables	
50% Discount	0.045***
	(0.003)
	[0.00]
Free	0.081***
	(0.004)
	[0.00]
Free + 5p	0.084***
	(0.004)
	[0.00]
Free + 15p	0.087***
	(0.004)
	[0.00]
50% Discount × Baseline TOU: Classic ToU	0.009
	(0.010)
	[0.37]
Free × Baseline TOU: Classic ToU	0.034***
	(0.012)
	[0.03]
Free + 5p × Baseline TOU: Classic ToU	0.031**
	(0.012)
	[0.05]
Free + 15p × Baseline TOU: Classic ToU	0.026**
	(0.012)
	[0.10]
50% Discount × Baseline TOU: New ToU	0.028
	(0.020)
	[0.30]
Free × Baseline TOU: New ToU	0.268***
	(0.029)
	[0.00]
Free + 5p × Baseline TOU: New ToU	0.357***
	(0.035)
	[0.00]
Free + 15p × Baseline TOU: New ToU	0.328***
	(0.031)
	[0.00]
Baseline TOU: Classic ToU	-0.046***
	(0.005)
Baseline TOU: New ToU	0.043
	(0.065)
Fixed-effects	
Event	Yes
Block	Yes
Fit statistics	
Control Mean	0.332
Observations	284,240

Clustered (Customer) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Spain

Model:	(1)
Variables	
50% Discount	0.023***
	(0.003)
	[0.00]
Free	0.049***
	(0.003)
	[0.00]
Free + 10c	0.048***
	(0.003)
	[0.00]
Prize Draw	0.028***
	(0.003)
	[0.00]
50% Discount × Baseline TOU: TOU	0.012*
	(0.007)
	[0.46]
Free × Baseline TOU: TOU	0.019**
	(0.008)
	[0.11]
Free + 10c × Baseline TOU: TOU	0.040***
	(0.008)
	[0.00]
Prize Draw × Baseline TOU: TOU	0.012*
	(0.007)
	[0.46]
50% Discount × Baseline TOU: Dynamic	0.010
	(0.010)
	[0.81]
Free × Baseline TOU: Dynamic	0.004
	(0.012)
	[0.89]
Free + 10c × Baseline TOU: Dynamic	0.005
	(0.012)
	[0.89]
Prize Draw × Baseline TOU: Dynamic	0.010
	(0.011)
	[0.81]
Baseline TOU: TOU	0.020***
	(0.004)
Baseline TOU: Dynamic	0.010*
	(0.005)
Fixed-effects	
Event	Yes
Block	Yes
Fit statistics	
Control Mean	0.422
Observations	600,672

Clustered (Customer) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Notes: This table reports the regressions underlying Figure A6. The omitted groups are the control group and fixed-tariff customers. Treatment coefficients report effects for fixed-tariff customers, tariff coefficients report baseline differences in the control group, and interactions report differential treatment effects relative to fixed-tariff customers. Each country uses its main-specification controls and fixed effects; standard errors are clustered by customer. Romano-Wolf adjusted p-values are reported in square brackets below treatment and interaction coefficients. Within each country, the adjustment treats these coefficients as a single family and excludes tariff main effects because they are not interpreted causally.

Table A9: Heterogeneity by socioeconomic status

Great Britain: property value

Model:	(1)
Variables	
50% Discount	0.029***
	(0.006)
	[0.00]
Free	0.068***
	(0.007)
	[0.00]
Free + 5p	0.067***
	(0.007)
	[0.00]
Free + 15p	0.073***
	(0.007)
	[0.00]
50% Discount × Property Value: Q2	0.017**
	(0.009)
	[0.18]
Free × Property Value: Q2	0.025**
	(0.011)
	[0.13]
Free + 5p × Property Value: Q2	0.018*
	(0.010)
	[0.18]
Free + 15p × Property Value: Q2	0.018
	(0.011)
	[0.18]
50% Discount × Property Value: Q3	0.023**
	(0.010)
	[0.13]
Free × Property Value: Q3	0.037***
	(0.011)
	[0.01]
Free + 5p × Property Value: Q3	0.056***
	(0.012)
	[0.00]
Free + 15p × Property Value: Q3	0.040***
	(0.012)
	[0.01]
50% Discount × Property Value: Q4	0.020**
	(0.010)
	[0.18]
Free × Property Value: Q4	0.049***
	(0.012)
	[0.01]
Free + 5p × Property Value: Q4	0.074***
	(0.014)
	[0.00]
Free + 15p × Property Value: Q4	0.073***
	(0.013)
	[0.00]
50% Discount × Property Value: Q5	0.032***
	(0.011)
	[0.03]
Free × Property Value: Q5	0.082***
	(0.014)
	[0.00]
Free + 5p × Property Value: Q5	0.110***
	(0.016)
	[0.00]
Free + 15p × Property Value: Q5	0.100***
	(0.015)
	[0.00]
Property Value: Q2	-0.011**
	(0.005)
Property Value: Q3	-0.018**
	(0.008)
Property Value: Q4	-0.034***
	(0.010)
Property Value: Q5	-0.048***
	(0.012)
Fixed-effects	
Event	Yes
Block	Yes
Fit statistics	
Control Mean	0.278
Observations	265,664

Clustered (Customer) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Spain: household income

Model:	(1)
Variables	
50% Discount	0.025***
	(0.007)
	[0.00]
Free	0.046***
	(0.007)
	[0.00]
Free + 10c	0.042***
	(0.007)
	[0.00]
Prize Draw	0.018***
	(0.006)
	[0.05]
50% Discount × HH Income: Q2	-0.005
	(0.009)
	[0.99]
Free × HH Income: Q2	0.005
	(0.009)
	[0.99]
Free + 10c × HH Income: Q2	0.002
	(0.009)
	[0.99]
Prize Draw × HH Income: Q2	0.018**
	(0.009)
	[0.34]
50% Discount × HH Income: Q3	-0.009
	(0.008)
	[0.94]
Free × HH Income: Q3	-0.004
	(0.009)
	[0.99]
Free + 10c × HH Income: Q3	0.007
	(0.009)
	[0.99]
Prize Draw × HH Income: Q3	-0.004
	(0.008)
	[0.99]
50% Discount × HH Income: Q4	0.005
	(0.008)
	[0.99]
Free × HH Income: Q4	0.007
	(0.009)
	[0.98]
Free + 10c × HH Income: Q4	0.009
	(0.009)
	[0.94]
Prize Draw × HH Income: Q4	0.015*
	(0.008)
	[0.51]
50% Discount × HH Income: Q5	0.007
	(0.009)
	[0.99]
Free × HH Income: Q5	0.017*
	(0.009)
	[0.45]
Free + 10c × HH Income: Q5	0.037***
	(0.009)
	[0.00]
Prize Draw × HH Income: Q5	0.026***
	(0.009)
	[0.03]
HH Income: Q2	-0.008*
	(0.004)
HH Income: Q3	0.0008
	(0.004)
HH Income: Q4	-0.0002
	(0.004)
HH Income: Q5	0.011**
	(0.004)
Fixed-effects	
Event	Yes
Block	Yes
Fit statistics	
Control Mean	0.413
Observations	590,439

Clustered (Customer) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Notes: This table reports the regressions underlying Figure A7. The omitted groups are the control group and the first socioeconomic-status quintile: property-value quintile 1 in Great Britain and household-income quintile 1 in Spain. Treatment coefficients report effects for the first quintile, quintile coefficients report baseline differences in the control group, and interactions report differential treatment effects. Specifications use the corresponding main controls and fixed effects; standard errors are clustered by customer. Romano-Wolf adjusted p-values are reported in square brackets below treatment and interaction coefficients. Within each country, the adjustment treats these coefficients as a single family and excludes socioeconomic-status main effects because they are not interpreted causally.

Table A10: Great Britain heterogeneity by Warm Home Discount eligibility

Model:	(1)
Variables	
50% Discount	0.051***
	(0.003)
	[0.00]
Free	0.109***
	(0.004)
	[0.00]
Free + 5p	0.122***
	(0.004)
	[0.00]
Free + 15p	0.119***
	(0.004)
	[0.00]
50% Discount × WHD Eligible	-0.025**
	(0.010)
	[0.01]
Free × WHD Eligible	-0.032***
	(0.012)
	[0.01]
Free + 5p × WHD Eligible	-0.056***
	(0.012)
	[0.00]
Free + 15p × WHD Eligible	-0.043***
	(0.013)
	[0.00]
WHD Eligible	0.020***
	(0.006)
Fixed-effects	
Event	Yes
Block	Yes
Fit statistics	
Control Mean	0.339
Observations	264,139

Clustered (Customer) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Notes: This table reports the regression underlying Figure A10. The omitted groups are the control group and customers who were not Warm Home Discount eligible. Treatment coefficients report effects for ineligible customers, the Warm Home Discount coefficient reports the baseline difference in the control group, and interactions report differential treatment effects for eligible customers. The specification uses the main Great Britain controls and fixed effects; standard errors are clustered by customer. Romano-Wolf adjusted p-values are reported in square brackets below treatment and interaction coefficients. The adjustment treats these coefficients as a single family and excludes the Warm Home Discount main effect because it is not interpreted causally.

Table A11: Heterogeneity by baseline consumption

Great Britain

Model:	(1)
Variables	
50% Discount	0.027***
	(0.003)
	[0.00]
Free	0.066***
	(0.005)
	[0.00]
Free + 5p	0.064***
	(0.004)
	[0.00]
Free + 15p	0.057***
	(0.005)
	[0.00]
50% Discount × EAC: Quartile 2	0.015**
	(0.006)
	[0.04]
Free × EAC: Quartile 2	0.015**
	(0.008)
	[0.05]
Free + 5p × EAC: Quartile 2	0.030***
	(0.008)
	[0.00]
Free + 15p × EAC: Quartile 2	0.037***
	(0.008)
	[0.00]
50% Discount × EAC: Quartile 3	0.023***
	(0.007)
	[0.01]
Free × EAC: Quartile 3	0.048***
	(0.009)
	[0.00]
Free + 5p × EAC: Quartile 3	0.052***
	(0.009)
	[0.00]
Free + 15p × EAC: Quartile 3	0.072***
	(0.009)
	[0.00]
50% Discount × EAC: Quartile 4	0.048***
	(0.010)
	[0.00]
Free × EAC: Quartile 4	0.105***
	(0.013)
	[0.00]
Free + 5p × EAC: Quartile 4	0.141***
	(0.014)
	[0.00]
Free + 15p × EAC: Quartile 4	0.137***
	(0.013)
	[0.00]
EAC: Quartile 2	-0.047
	(0.059)
EAC: Quartile 3	-0.124
	(0.101)
EAC: Quartile 4	-0.132
	(0.084)
Fixed-effects	
Event	Yes
Block	Yes
Fit statistics	
Control Mean	0.178
Observations	284,334

Clustered (Customer) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Spain

Model:	(1)
Variables	
50% Discount	0.018***
	(0.003)
	[0.00]
Free	0.037***
	(0.004)
	[0.00]
Free + 10c	0.039***
	(0.004)
	[0.00]
Prize Draw	0.022***
	(0.004)
	[0.00]
50% Discount × EAC: 2	0.004
	(0.005)
	[0.78]
Free × EAC: 2	0.010
	(0.006)
	[0.36]
Free + 10c × EAC: 2	0.015**
	(0.006)
	[0.11]
Prize Draw × EAC: 2	0.0006
	(0.006)
	[0.90]
50% Discount × EAC: 3	0.015**
	(0.006)
	[0.11]
Free × EAC: 3	0.030***
	(0.007)
	[0.00]
Free + 10c × EAC: 3	0.023***
	(0.007)
	[0.01]
Prize Draw × EAC: 3	0.010
	(0.006)
	[0.36]
50% Discount × EAC: 4	0.013*
	(0.007)
	[0.33]
Free × EAC: 4	0.022***
	(0.008)
	[0.03]
Free + 10c × EAC: 4	0.030***
	(0.008)
	[0.01]
Prize Draw × EAC: 4	0.025***
	(0.008)
	[0.01]
EAC: 2	-0.002
	(0.008)
EAC: 3	0.012
	(0.012)
EAC: 4	0.030
	(0.023)
Fixed-effects	
Event	Yes
Block	Yes
Fit statistics	
Control Mean	0.187
Observations	600,686

Clustered (Customer) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Notes: This table reports the regressions underlying Figure A8. The omitted groups are the control group and the first baseline-consumption quartile, defined separately within each country. Treatment coefficients report effects for first-quartile customers, baseline-consumption coefficients report baseline differences in the control group, and interactions report differential treatment effects. Each country uses its main-specification controls and fixed effects; standard errors are clustered by customer. Romano-Wolf adjusted p-values are reported in square brackets below treatment and interaction coefficients. Within each country, the adjustment treats these coefficients as a single family and excludes baseline-consumption main effects because they are not interpreted causally.

Table A12: Great Britain heterogeneity by email engagement

Model:	(1)
Variables	
50% Discount	0.018**
	(0.007)
	[0.01]
Free	0.035***
	(0.008)
	[0.00]
Free + 5p	0.051***
	(0.008)
	[0.00]
Free + 15p	0.038***
	(0.008)
	[0.00]
50% Discount × Opened Email: Opened	0.036***
	(0.008)
	[0.00]
Free × Opened Email: Opened	0.091***
	(0.009)
	[0.00]
Free + 5p × Opened Email: Opened	0.086***
	(0.010)
	[0.00]
Free + 15p × Opened Email: Opened	0.102***
	(0.010)
	[0.00]
Opened Email: Opened	-0.011**
	(0.005)
Fixed-effects	
Event	Yes
Block	Yes
Fit statistics	
Control Mean	0.334
Observations	237,071

Clustered (Customer) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Notes: This table reports the regression underlying Figure A9. The omitted groups are the control group and customers who did not open an event email. Treatment coefficients report effects for non-openers, the email-engagement coefficient reports the baseline difference in the control group, and interactions report differential treatment effects for customers who opened at least one event email. The specification uses the main Great Britain controls and fixed effects; standard errors are clustered by customer. Romano-Wolf adjusted p-values are reported in square brackets below treatment and interaction coefficients. The adjustment treats these coefficients as a single family and excludes the email-engagement main effect because it is not interpreted causally.

Table A13: Payment for each kWh of turn-up, by treatment group

Great Britain

Treatment	Total kWh (per event)	Total Payment (£)	£ per kWh	Demand Elasticity
50% Discount	440.56	571.28	1.30	-0.29
Free	952.19	1,156.97	1.22	-0.31
Free + 5p	1,040.45	1,398.61	1.34	-0.28
Free + 15p	1,045.96	1,842.03	1.76	-0.21

Spain

Treatment	Total kWh (per event)	Total Payment (€)	€ per kWh	Demand Elasticity
50% Discount	475.37	508.69	1.07	-0.12
Free	980.25	1,056.89	1.08	-0.13
Free + 10c	1,035.34	1,958.54	1.89	-0.07
Prize Draw	580.50	1,837.68	3.17	

Notes: The tables report the average total turn-up (in kWh) achieved by each treatment group per event, separately for each country. The payment per kWh is calculated as the total financial incentives paid to each treatment group divided by the total turn-up delivered by that group. For example, where customers would typically pay £0.24/kWh but were offered "free" electricity and consumed 0.5 kWh in the hour, the payment would be considered to be £0.12. The number of customers listed here is the number for the fifth event.

Table A14: Impact of treatment on aggregate consumption, −24 to +72 hours

Country	Great Britain		Spain	
Model	Separate Treatments	Pooled Treatment	Separate Treatments	Pooled Treatment
Model:	(1)	(2)	(3)	(4)
Variables				
50% Discount	0.064	0.064	0.094*	0.094*
	(0.059)	(0.059)	(0.056)	(0.056)
Free	-0.112*		0.114**	
	(0.059)		(0.056)	
Free + 5p	-0.002			
	(0.060)			
Free + 15p	0.041			
	(0.058)			
Free + Negative (pooled)		-0.024		0.113**
		(0.044)		(0.044)
Free + 10c			0.113**	
			(0.055)	
Prize Draw			0.014	0.014
			(0.055)	(0.055)
Fixed-effects				
Event	Yes	Yes	Yes	Yes
Block	Yes	Yes	Yes	Yes
Fit statistics				
Observations	272,589	272,589	276,669	276,669

*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Notes: This table reports regression estimates of the impact of turn-up treatments on aggregate consumption within a window of −24 to +72 hours around treatment exposure. "Separate treatments" specifications estimate each intervention individually, while "pooled treatment" specifications combine selected treatments as indicated. All models include event and block fixed effects. The estimates in (2) and (4) correspond to the average treatment effects visualized in Figure 6 and Figure 7 in the main text.

Table A15: Effect of invite and opt-in to NPG "power-up" events (kW)

	Event invite			Event opt-in		
Region	Pooled	Ferrybridge B	Spennymoor GSP	Pooled	Ferrybridge B	Spennymoor GSP
Model:	(1)	(2)	(3)	(4)	(5)	(6)
Variables						
Encouraged	0.411***	0.658***	0.247***			
	(0.012)	(0.025)	(0.011)			
Opted In (IV)				0.604***	0.884***	0.387***
				(0.018)	(0.033)	(0.016)
Controls						
Covariates	Yes	Yes	Yes	Yes	Yes	Yes
Half-Hour	Yes	Yes	Yes	Yes	Yes	Yes
Zone	Yes	Yes	Yes	Yes	Yes	Yes
Fit statistics						
Control Mean	0.584	0.605	0.574	0.584	0.605	0.574
Observations	325,861	117,857	208,004	325,861	117,857	208,004
Wald (1st stage), Opted In				30,389.3	17,040.7	15,008.3

Clustered (Customer) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Notes: This table reports estimates of the effect of power-up events on half-hourly electricity consumption in the NPG region, reported at the hourly level for consistency with Table A3 and Table A4, the turn-up results. Column (1) reports the impact of being invited to a power-up event across both network areas, among customers who signed up. Column (2) restricts the estimate to Ferrybridge B. Column (3) restricts the estimate to Spennymoor. Column (4) reports the impact of opting into the event, instrumented for by the encouragement, across both network areas. Column (5) restricts the estimate to Ferrybridge B. Column (6) restricts the estimate to Spennymoor. All specifications include baseline controls and half-hour and zone fixed effects. Baseline controls include average consumption at the same hour over the last 10 weekdays, average consumption at the same time over the last 2 weekends, meter type, EPC rating, estimated annual consumption, and tenure with Octopus Energy. The control mean refers to mean hourly consumption among uninvited customers during event half-hours. Standard errors are clustered at the customer level.

Table A16: Robustness: Effect of invite and opt-in to NPG "power-up" events (kW)

	Event invite			Event opt-in		
Specification	PAP	No Controls	TWFE	PAP	No Controls	TWFE
Model:	(1)	(2)	(3)	(4)	(5)	(6)
Variables						
Encouraged	0.408^{***}	0.417^{***}	0.435^{***}			
	(0.013)	(0.013)	(0.012)			
Opted In (IV)				0.599^{***}	0.610^{***}	0.640^{***}
				(0.018)	(0.018)	(0.017)
Controls						
Covariates	Yes	No	No	Yes	No	No
Half-Hour	Yes	Yes	Yes	Yes	Yes	Yes
Zone	Yes	Yes	Yes	Yes	Yes	Yes
Customer			Yes			Yes
Fit statistics						
Control Mean	0.584	0.584	0.584	0.584	0.584	0.584
Observations	326,073	338,295	338,294	326,073	338,295	338,294
Wald (1st stage), Opted In				30,424.5	31,924.8	35,419.9

Clustered (Customer) standard-errors in parentheses

*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Notes: This table reports robustness checks for the estimates in Table A15. Column (1) reports the impact of being invited to a power-up event using the pre-analysis plan specification. Column (2) reports the estimate excluding baseline controls. Column (3) reports the estimate using a two-way fixed effects specification, additionally including customer fixed effects. Column (4) reports the impact of opting into the event, instrumented for by the encouragement, using the pre-analysis plan specification. Column (5) reports the estimate excluding baseline controls. Column (6) reports the estimate using a two-way fixed effects specification, additionally including customer fixed effects. All specifications include half-hour and zone fixed effects. Baseline controls in columns (1) and (4) include average consumption at the same hour over the last 10 weekdays, average consumption at the same time over the last 2 weekends, meter type, EPC rating, estimated annual consumption, and tenure with Octopus Energy. The control mean refers to mean hourly consumption among uninvited customers during event half-hours. Standard errors are clustered at the customer level.

Table A17: Cost of turn-up, NPG

Event Date	Event Length (hours)	Total Payment (£)	Estimated Turn-up (kWh)	Number of opt-ins	Payment per kWh of Turn-up (£)
A. Total					
Total		15,900.93	20,772.77	19,836	0.77
B. Ferrybridge					
01 Feb, 2025	5	3,476.68	4,605.62	1,367	0.75
08 Feb, 2025	1	842.86	1,204.86	1,298	0.70
15 Feb, 2025	1	904.83	1,423.64	1,185	0.64
19 Feb, 2025	1	705.29	1,024.40	1,165	0.69
26 Feb, 2025	1.5	1,097.10	1,594.19	1,153	0.69
27 Feb, 2025	1.5	1,175.45	2,126.57	1,069	0.55
28 Feb, 2025	1.5	1,107.94	1,685.92	1,164	0.66
C. Spennymoor					
30 Jan, 2025	0.5	471.03	651.33	1,705	0.72
31 Jan, 2025	0.5	489.05	686.74	1,678	0.71
05 Feb, 2025	4	1,430.43	2,016.61	1,437	0.71
12 Feb, 2025	1.5	439.49	312.23	1,259	1.41
18 Feb, 2025	3	1,527.61	963.19	1,334	1.59
25 Feb, 2025	2	525.07	274.30	1,057	1.91
04 Mar, 2025	1	875.20	1,058.49	1,475	0.83
11 Mar, 2025	1	832.88	1,144.67	1,490	0.73

Notes: This table reports the total payment and turn-up achieved by customers who opted into each event in the NPG trial. The cost per kWh is calculated as the total financial incentives paid divided by the total turn-up delivered.

Table A18: Temporal displacement, NPG trial

	Event Invite				Event opt-in			
Model	Pre (-24h to event start)	During event	Post (event end to +72h)	Total (-24h to +72h)	Pre (-24h to event start)	During event	Post (event end to +72h)	Total (-24h to +72h)
Model:	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Variables								
Encouraged	-0.152	0.562***	-0.379	0.140				
	(0.122)	(0.064)	(0.299)	(0.405)				
Opted In (IV)					-0.215	0.798***	-0.538	0.199
					(0.173)	(0.090)	(0.424)	(0.574)
Fixed-effects								
Event	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Zone	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Fit statistics								
Control Mean	12.45	5.97	36.78	50.56	12.45	5.97	36.78	50.56
Observations	37,298	37,314	37,305	37,318	37,298	37,314	37,305	37,318
Wald (1st stage), Opted In					21,874.2	21,863.8	21,854.3	21,891.2

Clustered (Customer) standard-errors in parentheses

*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Notes: This table reports the impact of treatment on cumulative electricity consumption in different windows. The first four columns are for customers who signed up. The first column is from −24 hours to −1 hour before the event. The second column is total consumption during the event. The third column is the total consumption during hours 1 to 72. The fourth column is total consumption from −24 to +72 hours around the event. Columns 5-8 are analogous to the first four columns, but use the IV method to estimate the impacts among opted-in customers. Standard errors are clustered at the customer level.

Table A19: Great Britain, impact of treatment on climate sentiment

	Index	Flex	Renewables	Climate Change	Net Zero
Model:	(1)	(2)	(3)	(4)	(5)
Variables					
Treated	-0.007	0.008	-0.016	0.014	-0.033
	(0.025)	(0.042)	(0.040)	(0.041)	(0.040)
Octopus Sentiment	0.429***	0.298***	0.489***	0.474***	0.455***
	(0.012)	(0.021)	(0.019)	(0.020)	(0.020)
Fixed-effects					
Block ID	Yes	Yes	Yes	Yes	Yes
Fit statistics					
Observations	3,503	3,503	3,503	3,503	3,503

IID standard-errors in parentheses

*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Notes: This table reports estimates of the relationship between exposure to turn-up events and post-experiment climate-related sentiment among Great Britain survey respondents. The dependent variables are standardized survey outcomes: an index equal to the average of the four standardized climate-sentiment measures, and separate measures for views on energy flexibility, renewable energy, climate change, and Net Zero. "Treated" is an indicator equal to one if the customer received any turn-up communications, and zero if assigned to the pure control group. "Octopus Sentiment" is the response to the question measuring alignment with Octopus Energy's mission. All specifications include randomization block fixed effects. IID standard errors are reported in parentheses.

Table A20: Spain, impact of treatment on climate sentiment

	Index	Flex	Renewables	Climate Change	Net Zero
Model:	(1)	(2)	(3)	(4)	(5)
Variables					
Treated	-0.009	0.003	-0.013	0.004	-0.028
	(0.020)	(0.035)	(0.033)	(0.033)	(0.034)
Octopus Sentiment	0.425***	0.390***	0.476***	0.434***	0.400***
	(0.010)	(0.016)	(0.015)	(0.016)	(0.016)
Fixed-effects					
Block ID	Yes	Yes	Yes	Yes	Yes
Fit statistics					
Observations	4,876	4,878	4,878	4,878	4,876

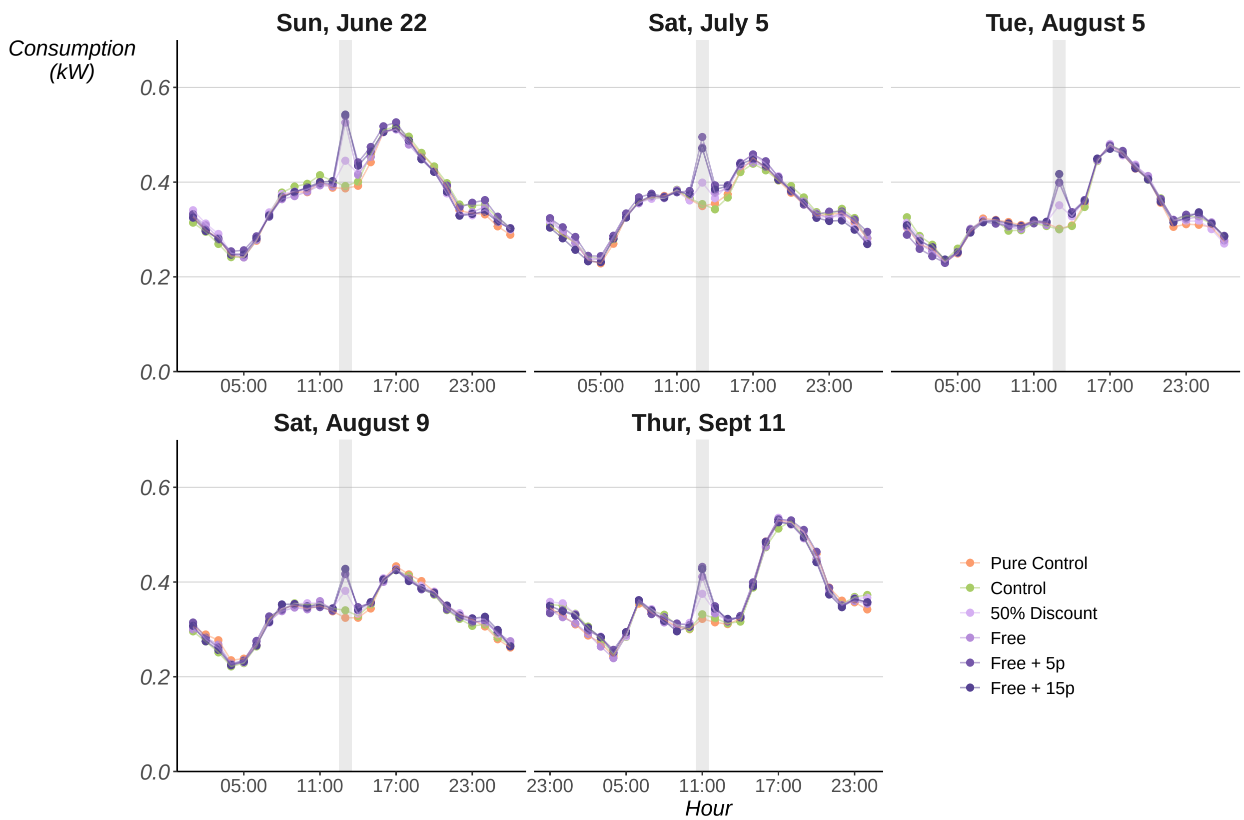
IID standard-errors in parentheses

*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Notes: This table reports estimates of the relationship between exposure to turn-up events and post-experiment climate-related sentiment among Spanish survey respondents. The dependent variables are standardized survey outcomes: an index equal to the average of the four standardized climate-sentiment measures, and separate measures for views on energy flexibility, renewable energy, climate change, and Net Zero. "Treated" is an indicator equal to one if the customer received any turn-up communications, and zero if assigned to the pure control group. "Octopus Sentiment" is the standardized response to the question measuring alignment with Octopus Energy's mission. All specifications include randomization block fixed effects. IID standard errors are reported in parentheses.

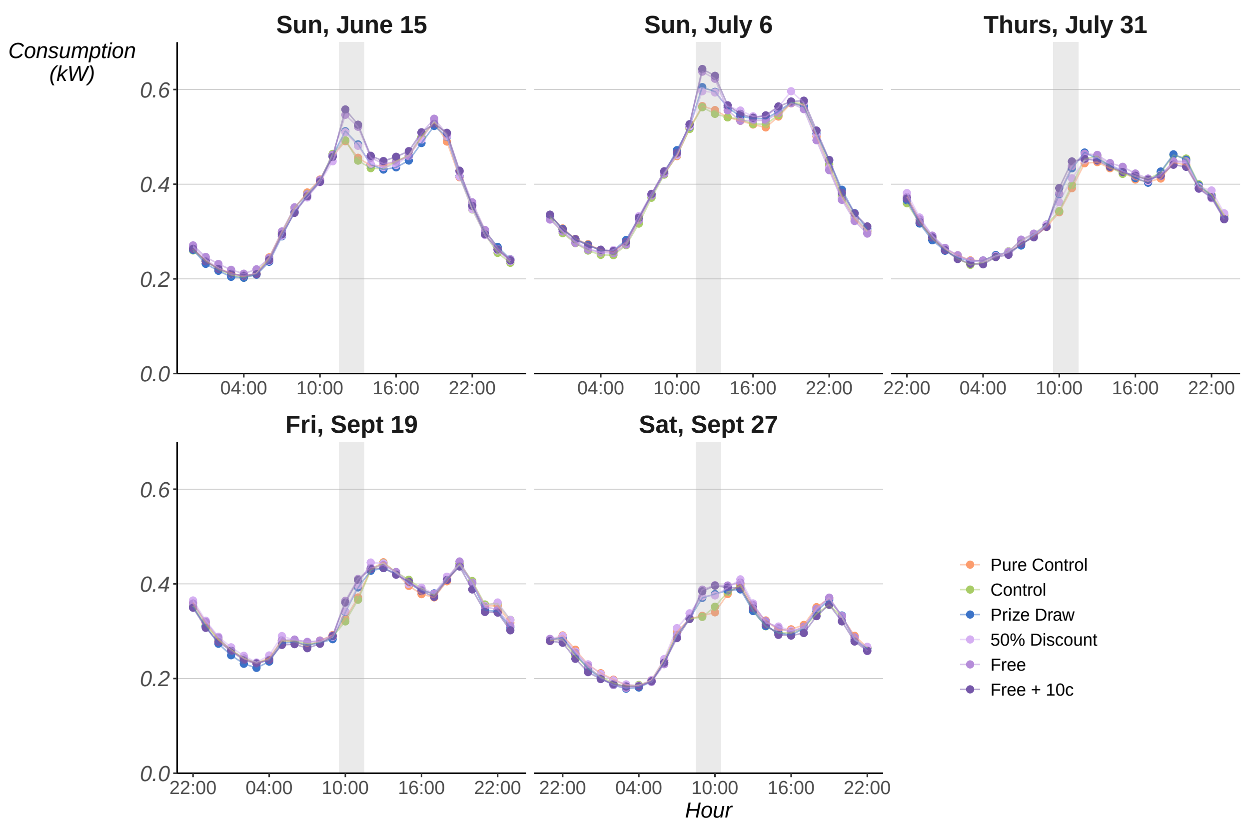
B Figures

Figure A1: Hourly electricity consumption (kW), Great Britain



Notes: This figure plots mean hourly electricity consumption (kW) by treatment group for each of the five turn-up events in Great Britain. Each panel corresponds to one event day. The shaded gray band indicates the one-hour event window.

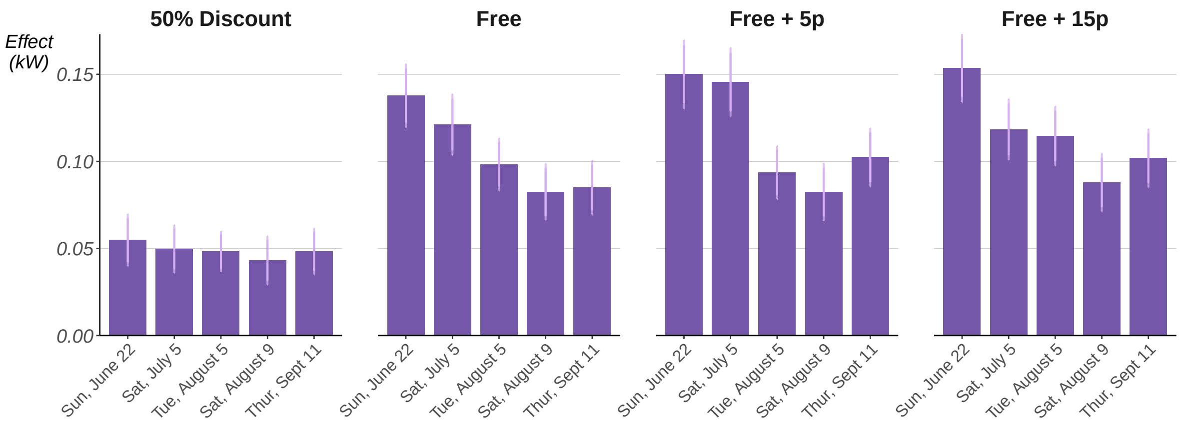
Figure A2: Hourly electricity consumption (kW), Spain



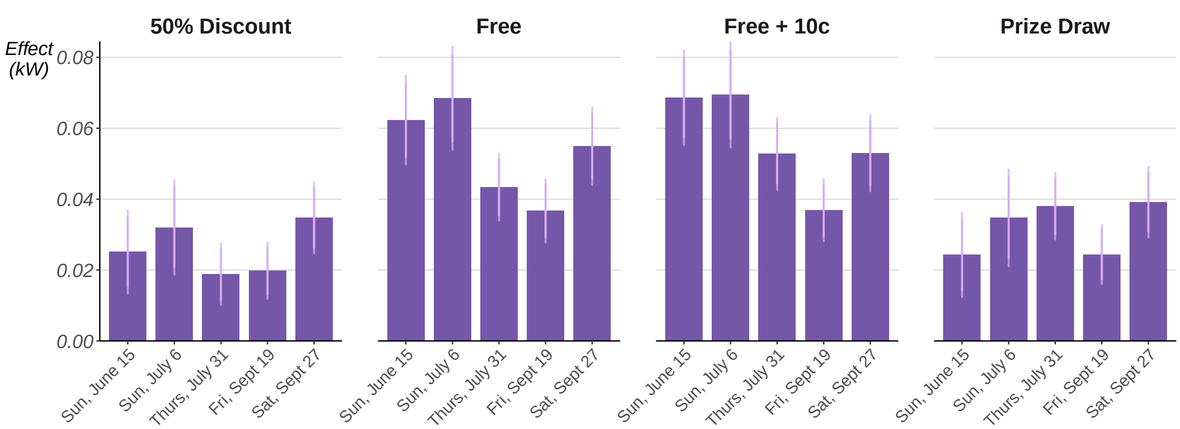
Notes: This figure plots raw mean hourly electricity consumption (kW) by treatment group for each of the five turn-up events in Spain. Each panel corresponds to one event day. The shaded gray band indicates the two-hour event window.

Figure A3: Event level treatment effects

(a) Great Britain



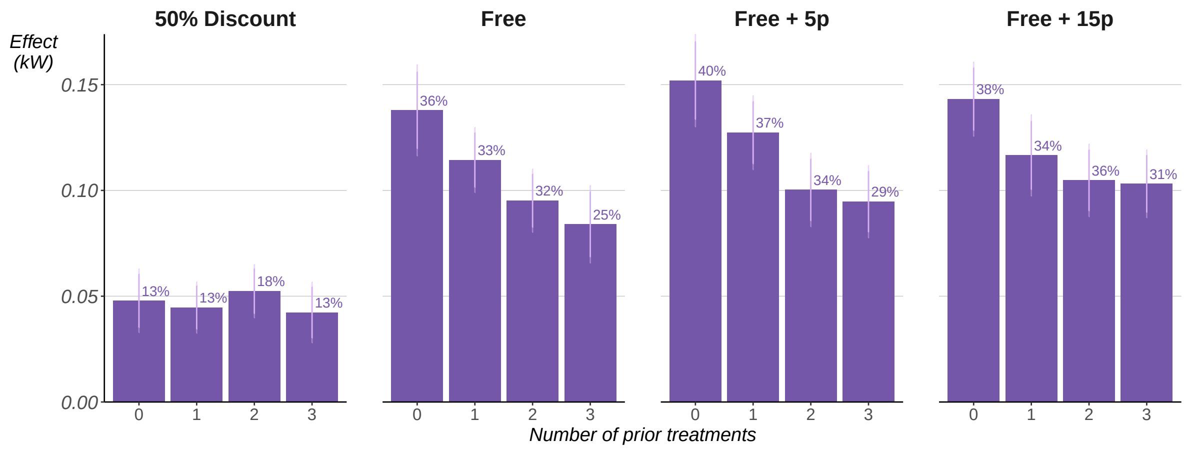
(b) Spain



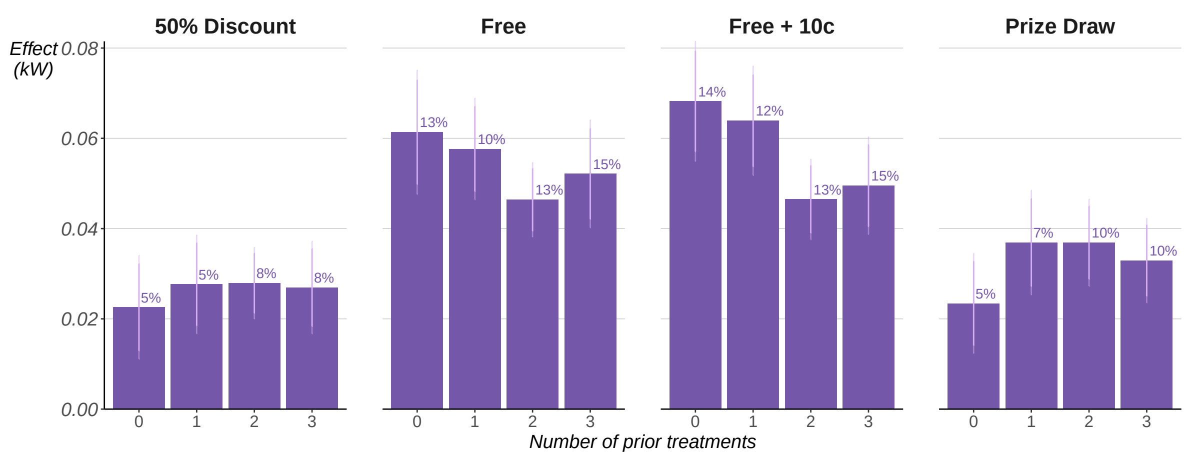
Notes: This figure reports event-specific treatment effects on hourly electricity consumption during turn-up events. Panel (a) reports Great Britain and panel (b) reports Spain. Estimates are obtained by estimating the main specification separately for each event, with treatment effects measured relative to the event control group. Bars show point estimates for each treatment arm and event date. Lines depict 90% and 95% confidence intervals. Standard errors are clustered at the customer level. In Great Britain, events were one hour long; in Spain, events were two hours long.

Figure A4: Fatigue / novelty effects

(a) Great Britain



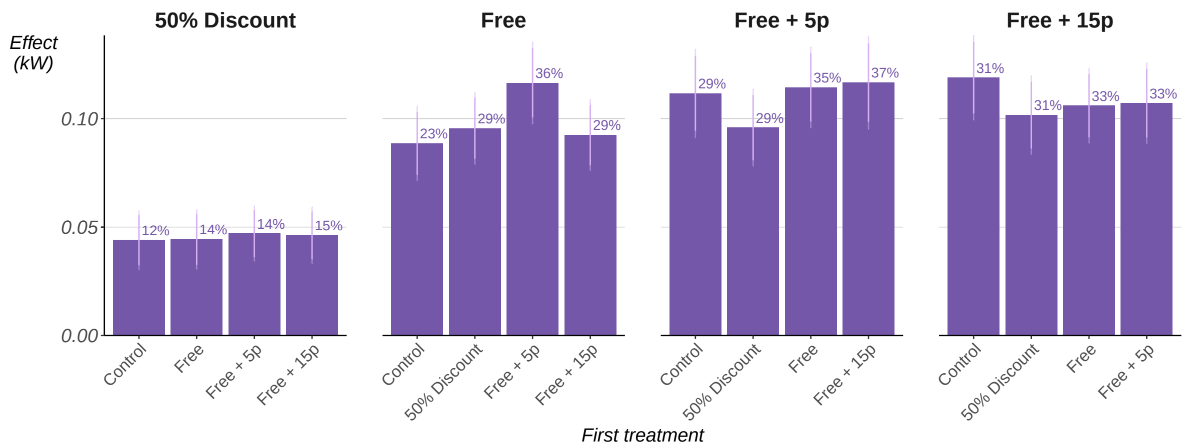
(b) Spain



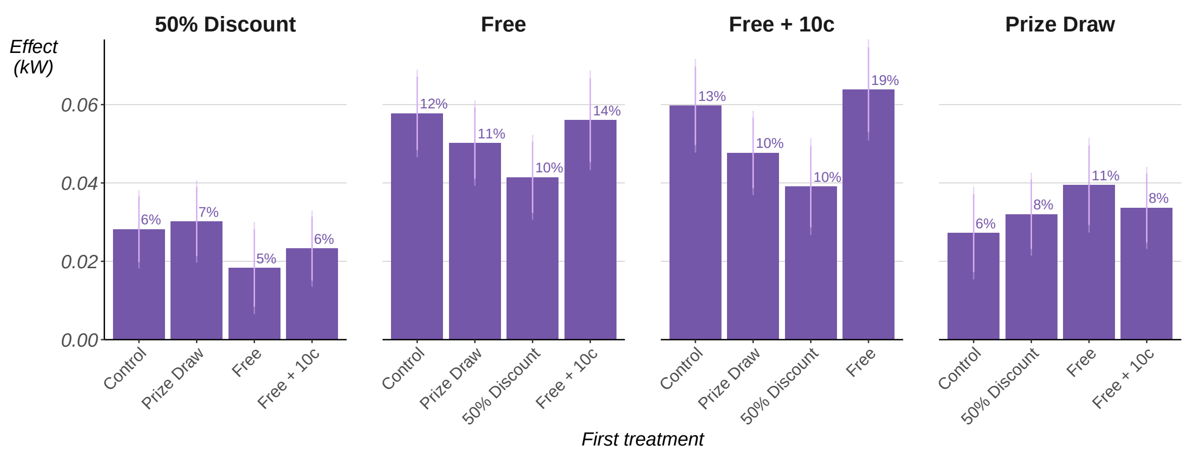
Notes: This figure shows how treatment effects vary with prior exposure to turn-up events. Panel (a) reports Great Britain and panel (b) reports Spain. Estimates are obtained by interacting current treatment assignment with the number of prior treatments received, using the main specification for each country. Bars show point estimates for each treatment arm by number of prior treatments. Lines depict 90% and 95% confidence intervals. Standard errors are clustered at the customer level. The outcome measure is hourly electricity consumption during the event.

Figure A5: Anchoring on first event

(a) Great Britain



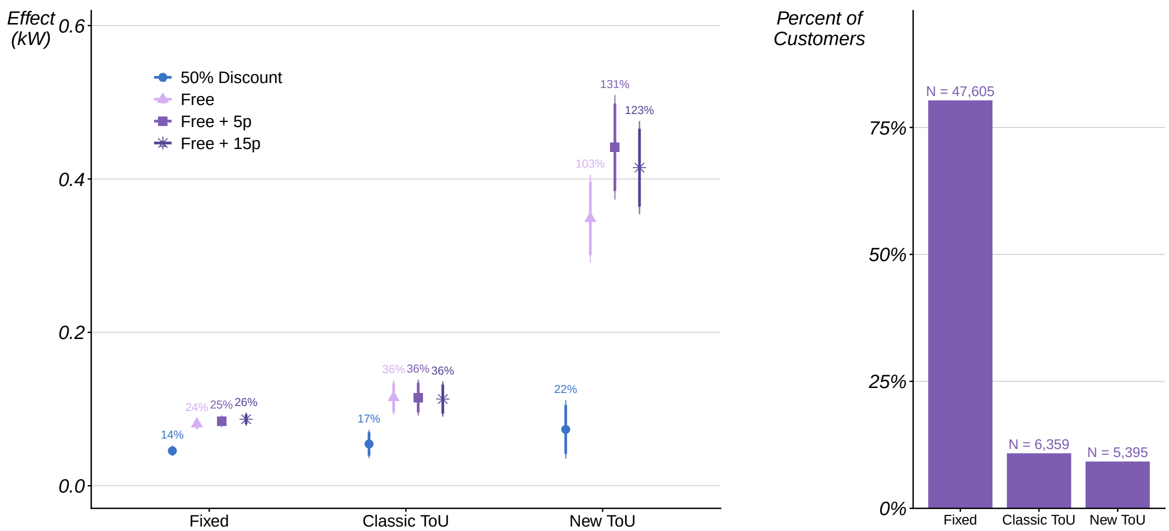
(b) Spain



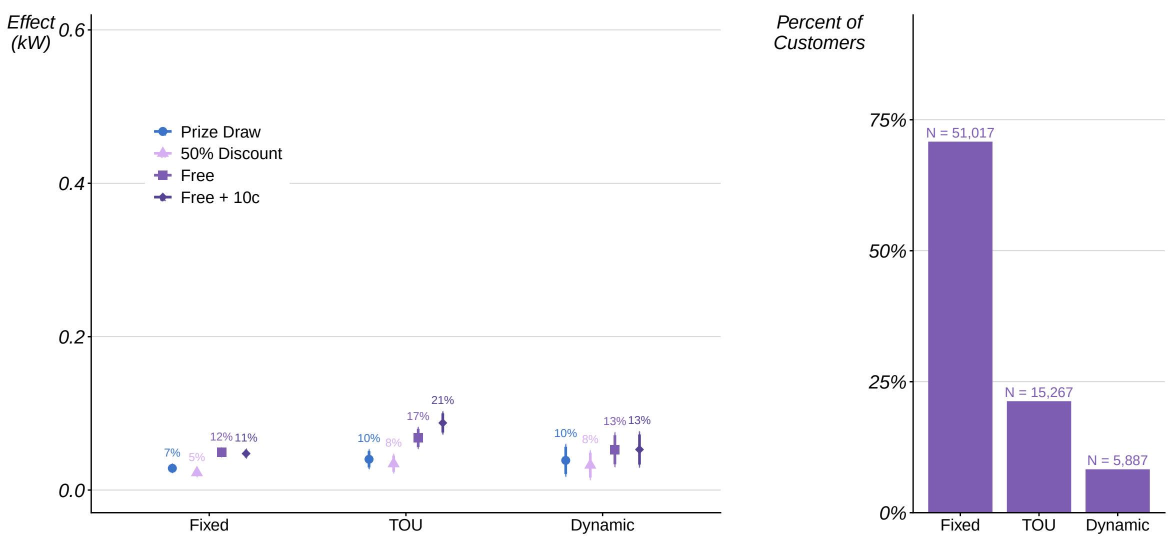
Notes: This figure shows how treatment effects vary by the first treatment a customer received. Panel (a) reports Great Britain and panel (b) reports Spain. Estimates are obtained by interacting current treatment assignment with the customer's first treatment assignment, using the main specification for each country. The sample is restricted to customer-event observations after the first event. Bars show point estimates for each current treatment arm by first treatment assignment. Lines depict 90% and 95% confidence intervals. Standard errors are clustered at the customer level. The outcome measure is hourly electricity consumption during the event.

Figure A6: Heterogeneity, by baseline tariff

(a) Great Britain



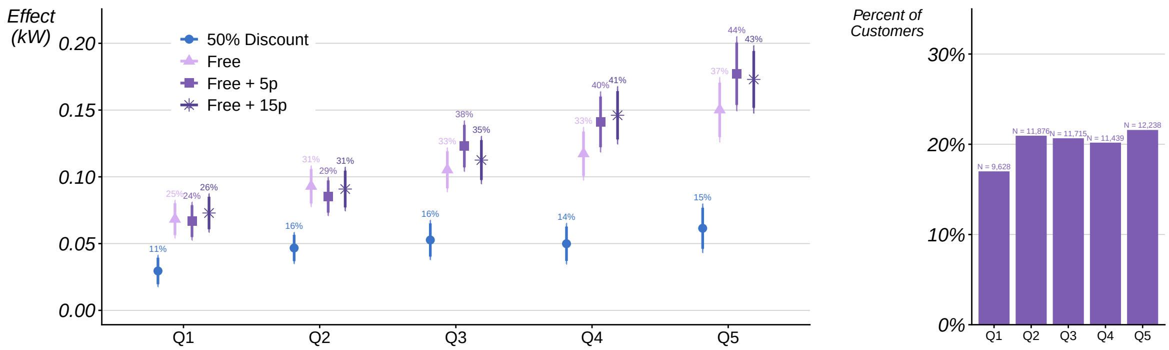
(b) Spain



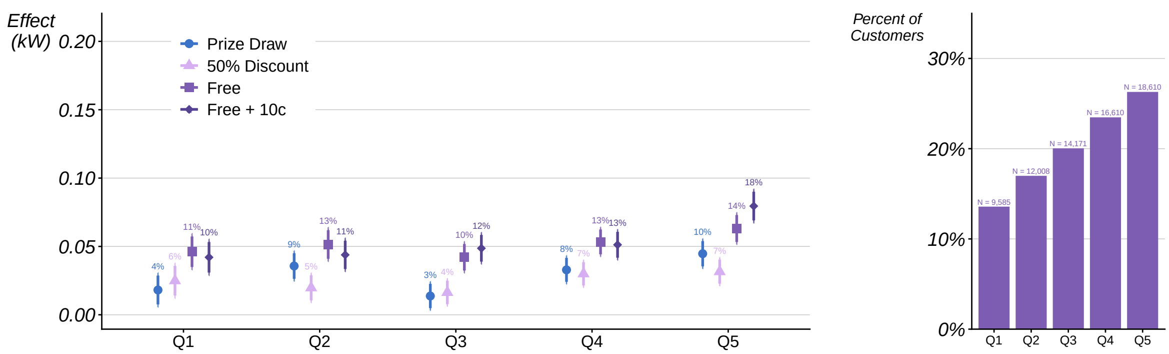
Notes: This figure shows how treatment effects vary by baseline tariff. Figure A6a presents results for Great Britain. The left-hand side plots treatment effect estimates by baseline tariff category: Fixed, Classic TOUs, which consist of traditional overnight tariffs, and New TOU, which include fixed multi-band tariffs with clearly defined peak and off-peak windows as well as real-time pricing tariffs that track wholesale market conditions. The right-hand side shows the share of customers in each baseline tariff category. Figure A6b presents results for Spain. We differentiate between Fixed, TOU tariffs where prices do not change day to day, and Dynamic TOU tariffs which are indexed to wholesale market prices. Lines depict 90% (dark) and 95% (light) confidence intervals. Standard errors are clustered at the customer level. The outcome measure is hourly electricity consumption during the event. The corresponding regression tables are reported in Table A8.

Figure A7: Heterogeneity by Socioeconomic Status

(a) Great Britain - Property Value



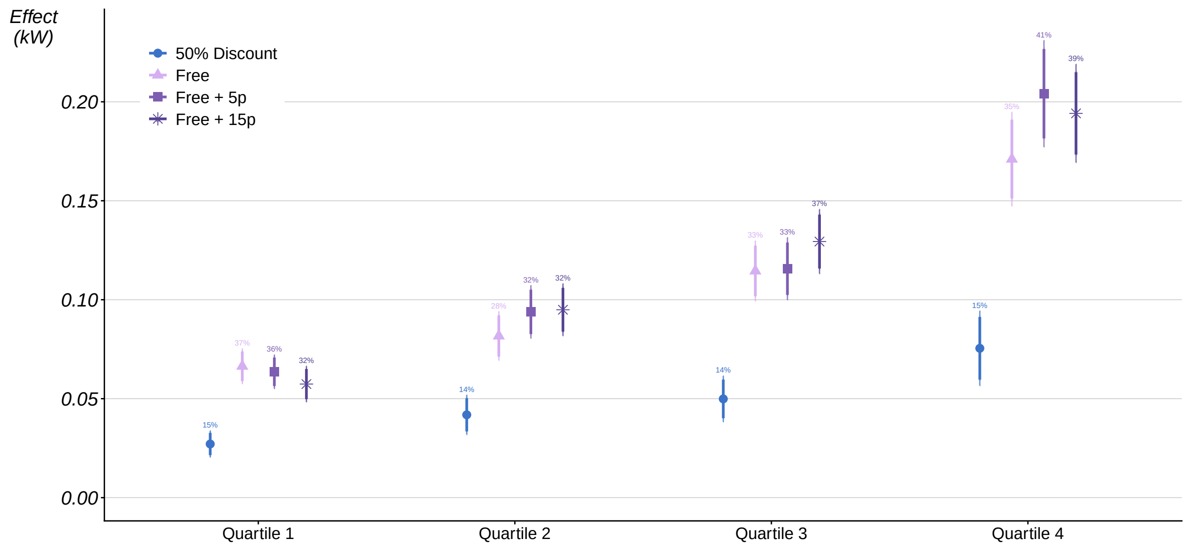
(b) Spain - Income



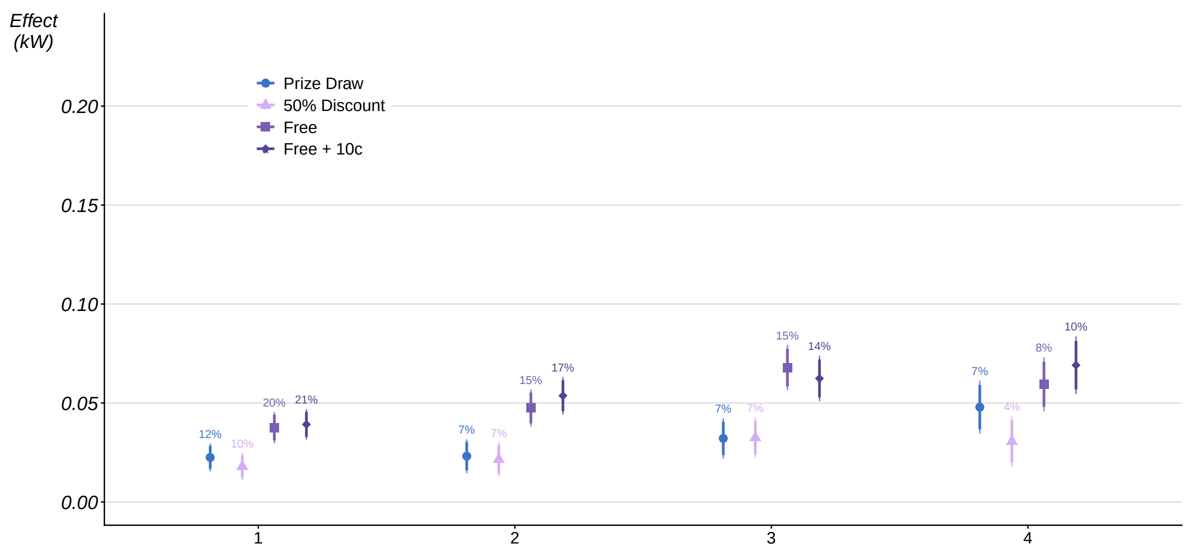
Notes: This figure shows treatment effect heterogeneity by proxies for socioeconomic status. Panel (a) shows Great Britain, where socioeconomic status is proxied by the estimated property value of the customer's residence from the WhenFresh property data. Customers are assigned to quintiles using national property-value cutoffs. Panel (b) shows Spain, where socioeconomic status is proxied by gross household income from the 2023 INE Atlas de Distribución de Renta de los Hogares. Income is measured at the census-section level and linked to customers using geocoded customer census-section codes; census sections are then divided into quintiles of gross household income. Points report estimated treatment effects on hourly electricity consumption during the event window, from regressions that interact treatment assignment with the relevant quintile indicators. Dark and light vertical bars show 90% and 95% confidence intervals, respectively. Standard errors are clustered at the customer level. The right-hand panels show the share and number of customers in each quintile. The corresponding regression tables are reported in Table A9.

Figure A8: Heterogeneity by estimated annual consumption quartiles

(a) Great Britain

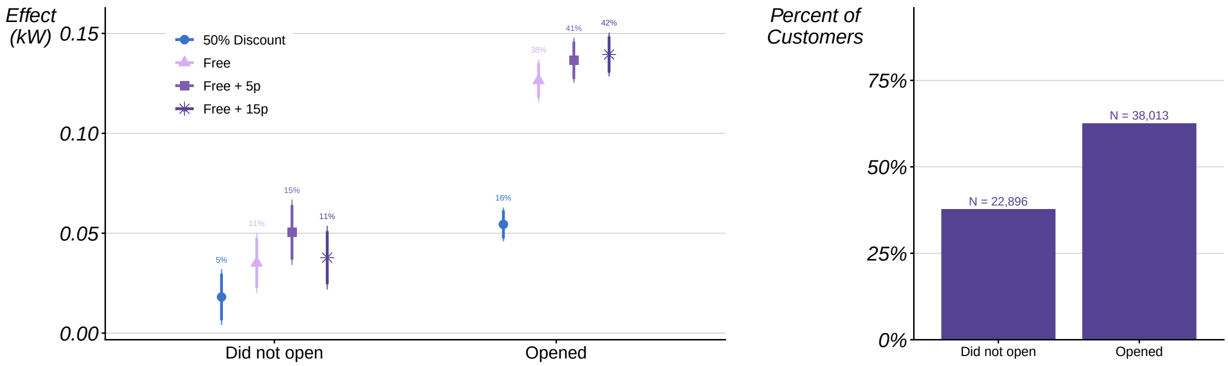


(b) Spain



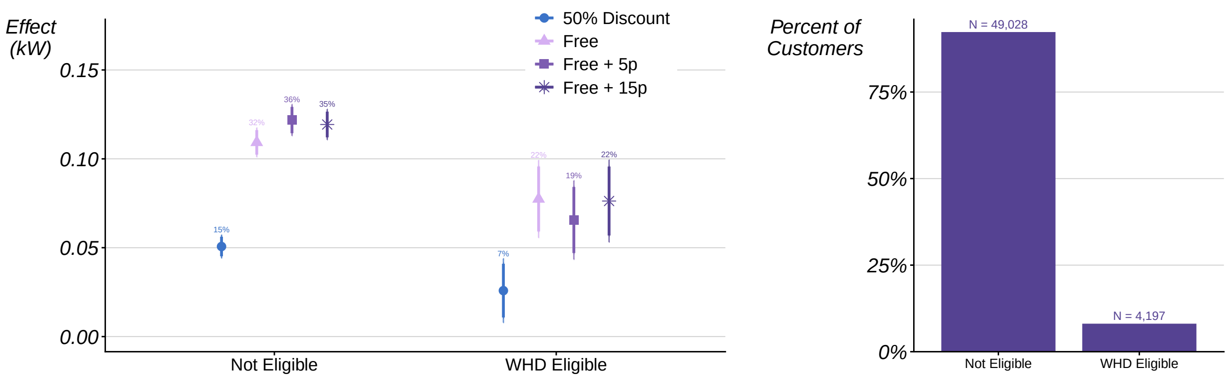
Notes: This figure shows how treatment effects vary by baseline estimated annual consumption, split into quartiles. Figure A8a shows results for Great Britain and Figure A8b shows results for Spain, with quartiles defined separately within each country. Lines depict 90% (dark) and 95% (light) confidence intervals. Standard errors are clustered at the customer level. The outcome measure is hourly electricity consumption during the event. The corresponding regression tables are reported in Table A11.

Figure A9: Heterogeneity by whether customer opened email - Great Britain



Notes: This figure shows heterogeneity in Great Britain treatment effects by whether the customer opened at least one event notification email. Estimates are obtained by interacting treatment assignment with an indicator for ever having opened an event email in the main specification. Lines depict 90% and 95% confidence intervals. Standard errors are clustered at the customer level. Percentage labels report treatment effects as a share of mean hourly consumption in the control group. The bar chart on the right reports the share of customer-event observations in each email-open category. The corresponding regression table is reported in Table A12.

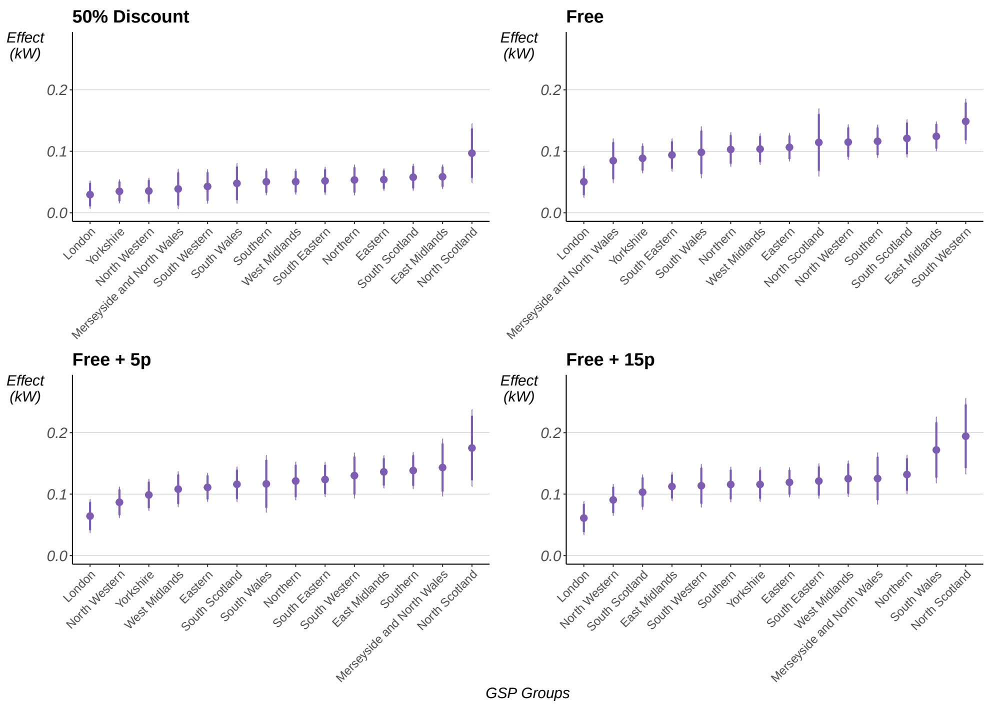
Figure A10: Heterogeneity by Warm Home Discount eligibility - Great Britain



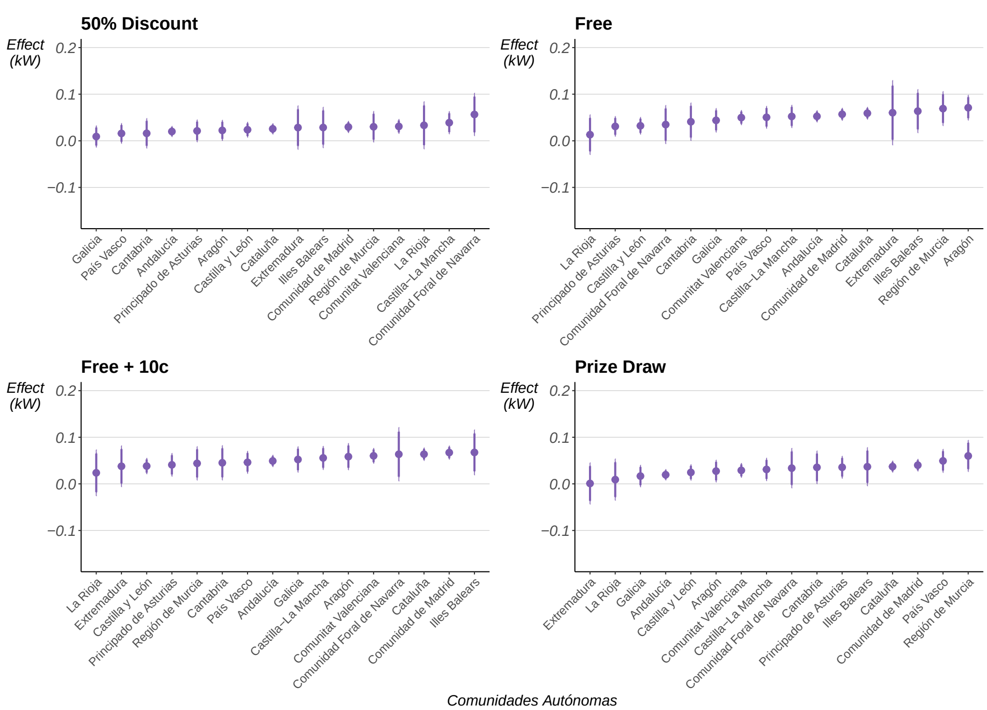
Notes: This figure shows heterogeneity in Great Britain treatment effects by whether the customer was eligible for the Warm Home Discount. Estimates are obtained by interacting treatment assignment with a Warm Home Discount eligibility indicator in the main specification. Lines depict 90% and 95% confidence intervals. Standard errors are clustered at the customer level. Percentage labels report treatment effects as a share of mean hourly consumption in the control group. The bar chart on the right reports the share of customer-event observations by Warm Home Discount eligibility. The corresponding regression table is reported in Table A10.

Figure A11: Heterogeneity by geographic region

(a) Great Britain



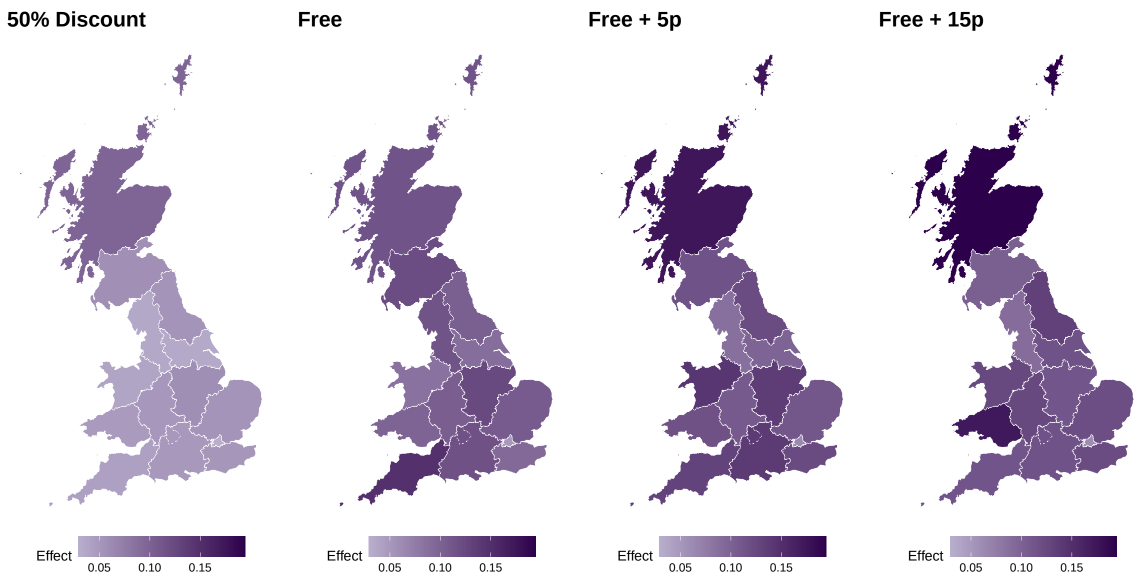
(b) Spain



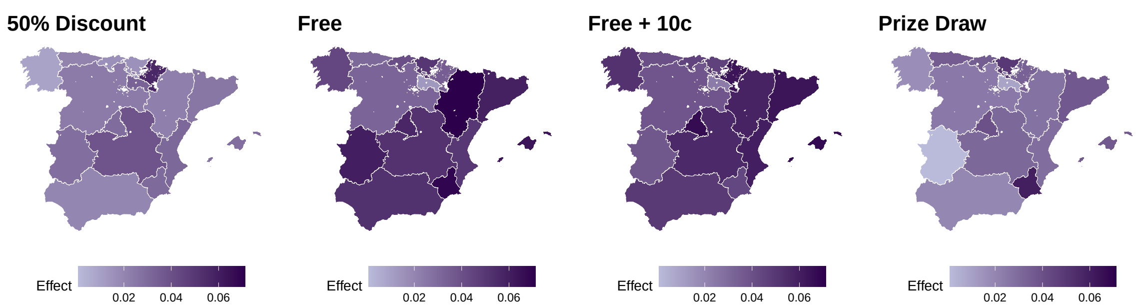
Notes: This figure shows treatment effects on hourly electricity consumption by geographic region, separately for each treatment group. Panel (a) reports estimates by Grid Supply Point (GSP) Group in Great Britain; panel (b) reports estimates by Comunidad Autónoma in Spain. Estimates are obtained by interacting treatment assignment with region indicators in the main specification. Regions are ordered by effect size within each treatment. Lines depict 90% (dark) and 95% (light) confidence intervals. Standard errors are clustered at the customer level.

Figure A12: Heterogeneity by geographic region

(a) Great Britain



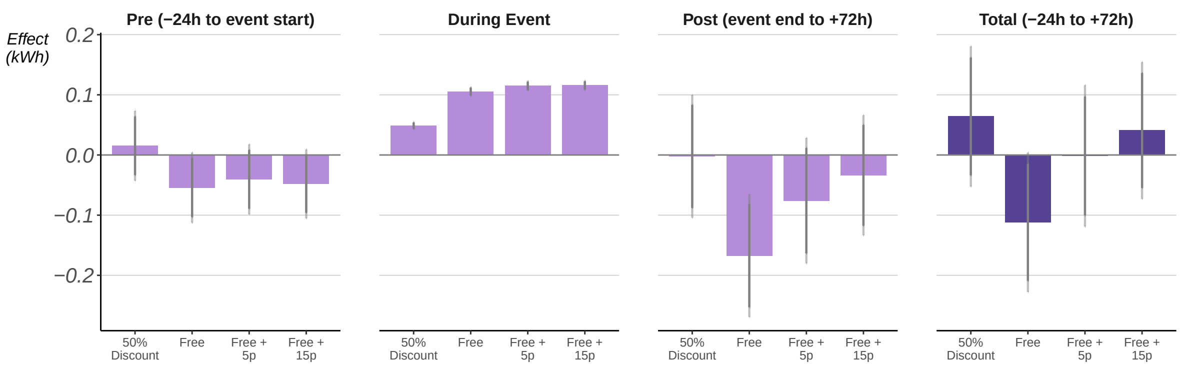
(b) Spain



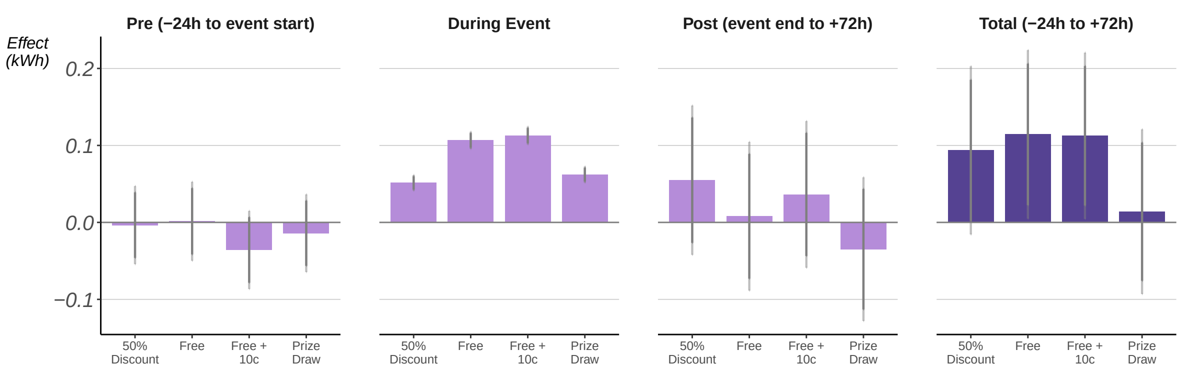
Notes: This figure maps estimated treatment effects on hourly electricity consumption by geographic region, separately for each treatment group. Panel (a) shows Great Britain, with regions defined by Grid Supply Point (GSP) Group. Panel (b) shows Spain, with regions defined by Comunidad Autónoma. Darker shading indicates larger treatment effects. Note that the color scales differ across treatments within each panel to reflect the range of effects for each incentive level.

Figure A13: Robustness: Displacement analysis with unpooled treatment variables

(a) Great Britain

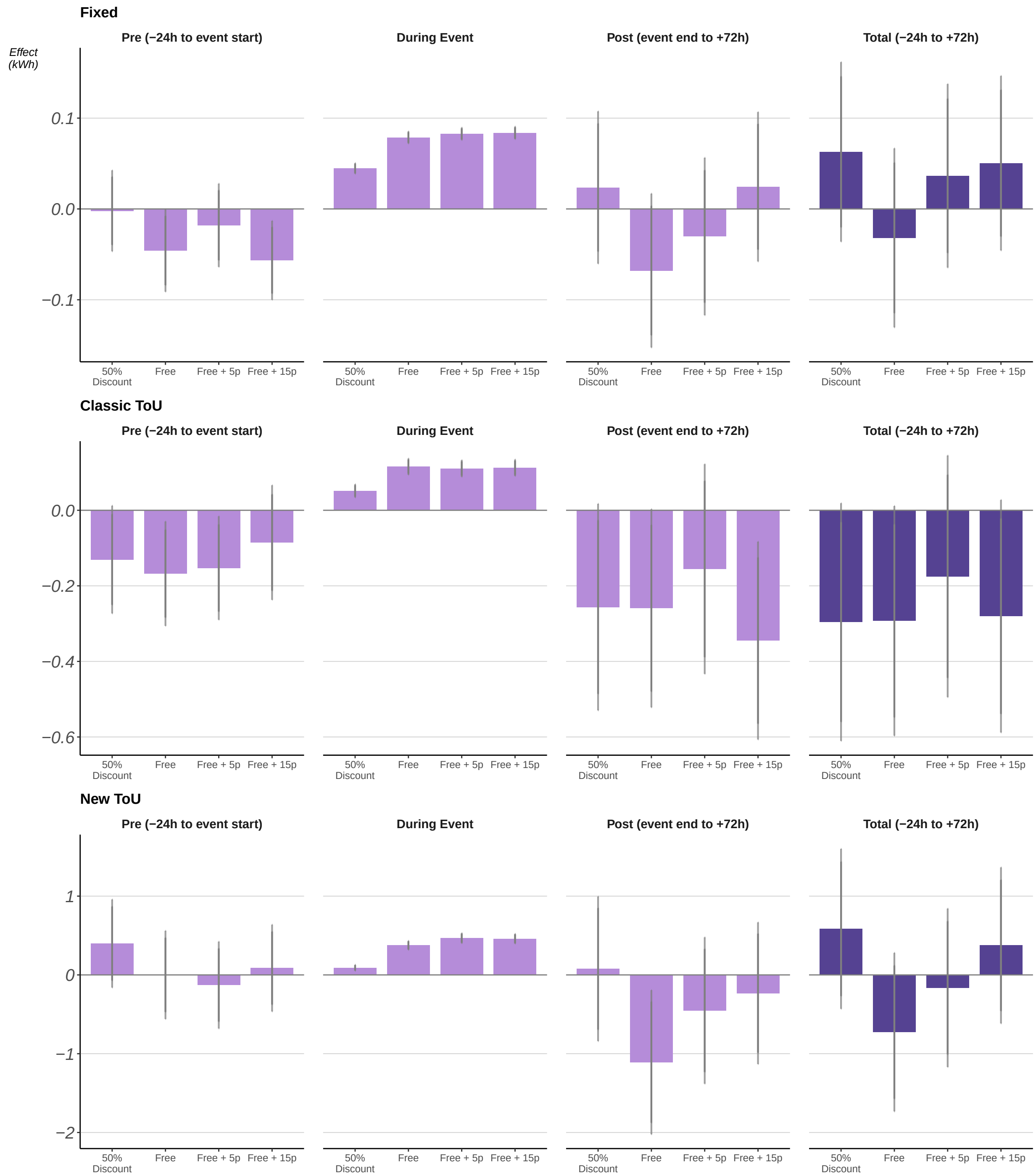


(b) Spain



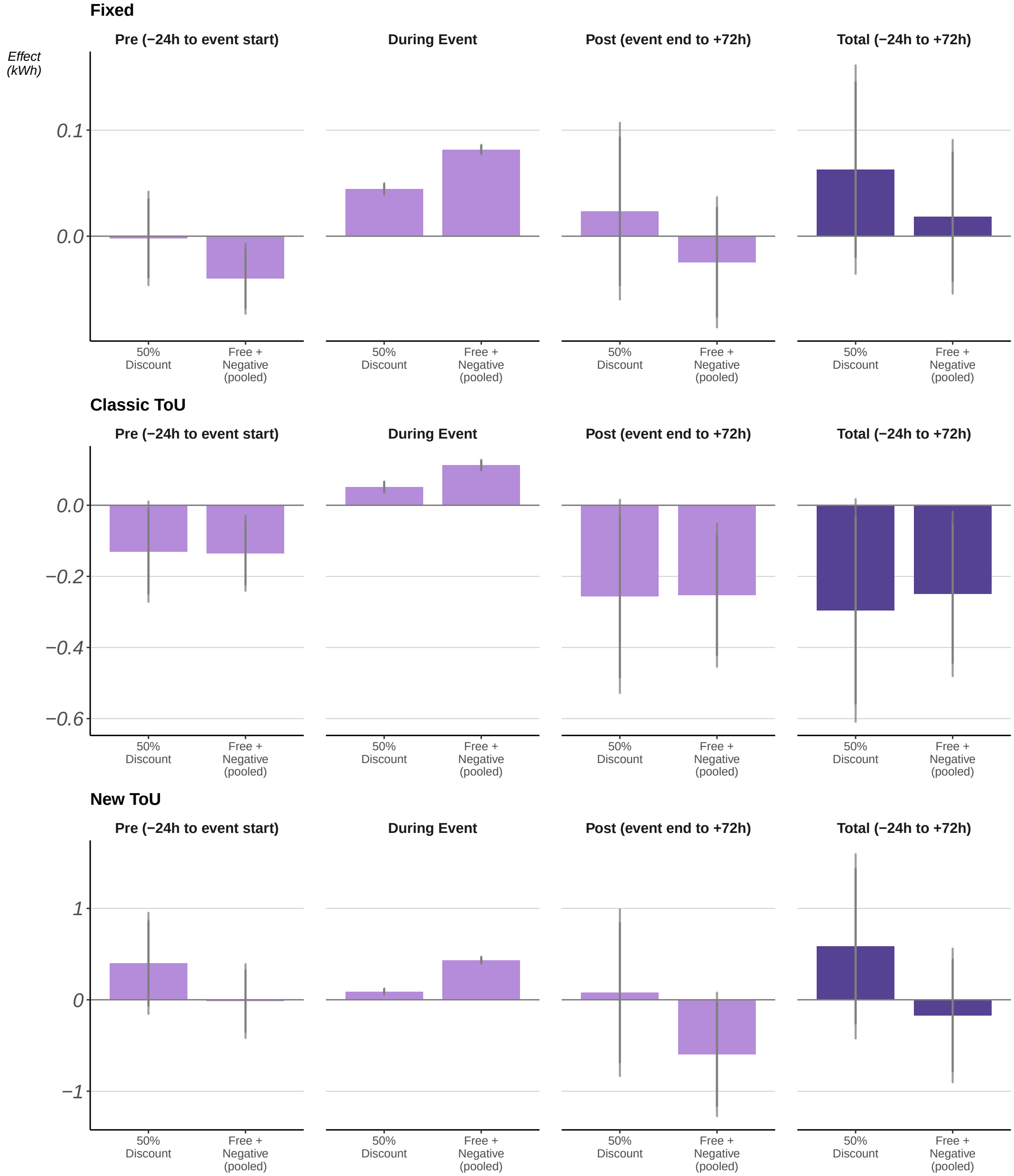
Notes: These figures plot the impact of treatment on cumulative electricity consumption in different windows. The first pane is from −24 hours to −1 hour before the event. The second pane is the event window. The third pane is the total consumption during hours 1 to 72. The fourth pane is total consumption from −24 to +72 hours around the event. Shaded areas indicate 95% confidence intervals; standard errors are clustered at the customer level.

Figure A14: Heterogeneity in displacement by baseline TOU tariff - Great Britain



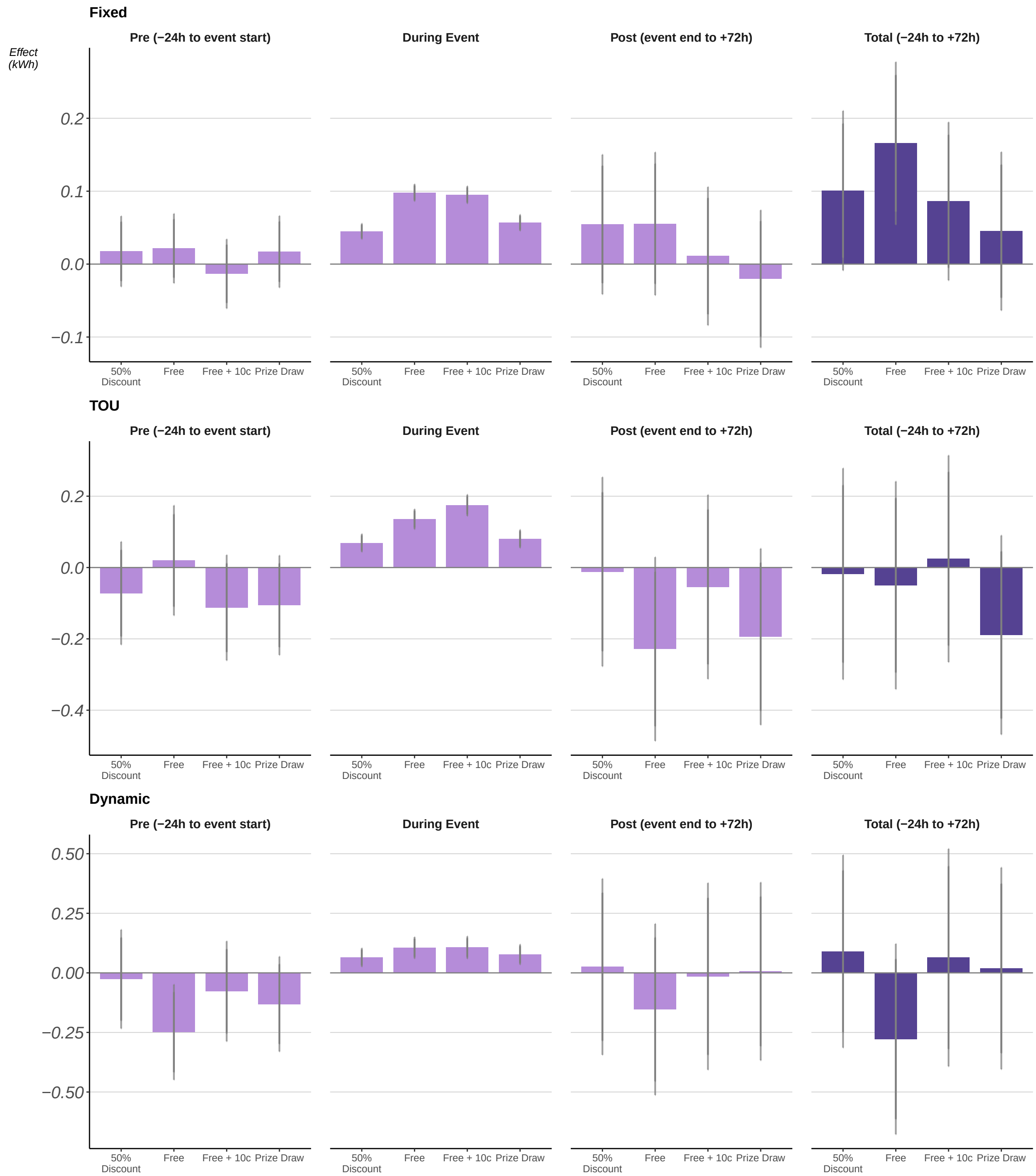
Notes: This figure shows impact of treatment on cumulative electricity consumption in different windows. Estimates come from interaction specifications that interact treatment status with tariff type and are expressed relative to the control group within each tariff. Shaded areas denote 95% confidence intervals; standard errors are clustered at the customer level.

Figure A15: Heterogeneity in displacement by baseline TOU tariff, pooled treatments - Great Britain



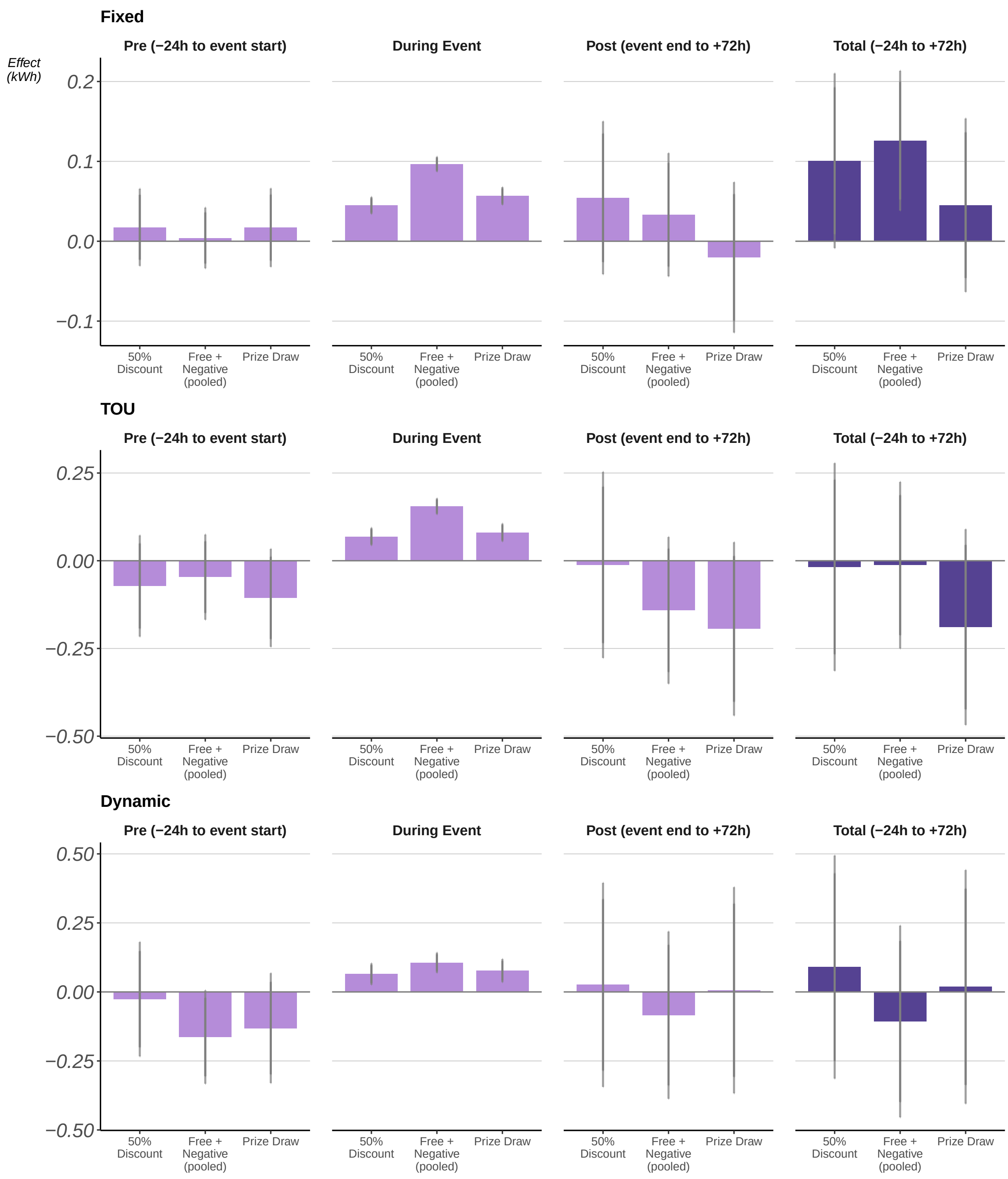
Notes: The figure shows impact of treatment on cumulative electricity consumption in different windows. Estimates come from interaction specifications that interact treatment status with tariff type and are expressed relative to the control group within each tariff. This figure pools Free, Free + 5p, and Free + 15p into one group. Shaded areas denote 95% confidence intervals; standard errors are clustered at the customer level.

Figure A16: Heterogeneity in displacement by TOU tariff - Spain



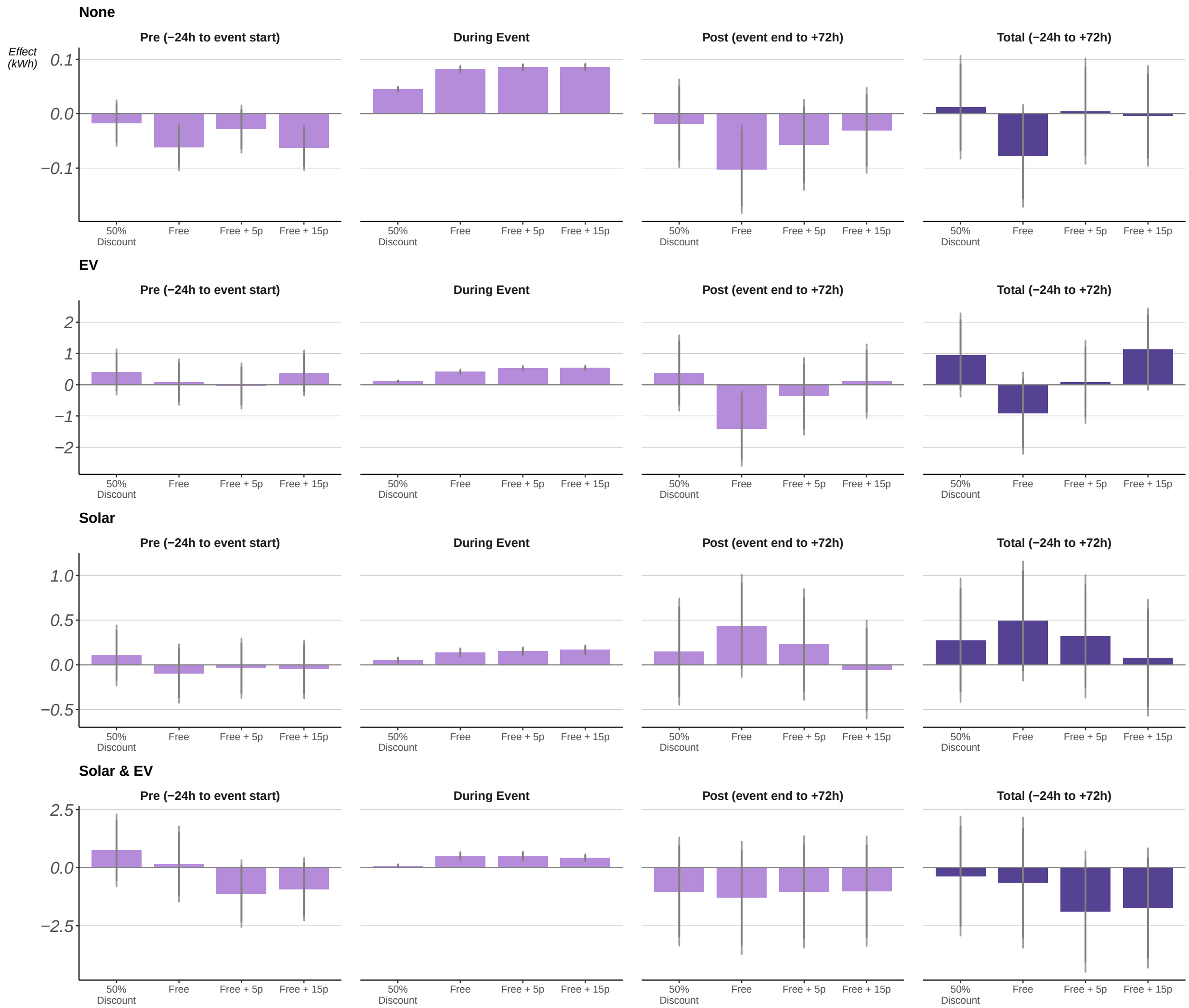
Notes: This figure shows impact of treatment on cumulative electricity consumption in different windows. Estimates come from interaction specifications that interact treatment status with tariff type and are expressed relative to the control group within each tariff. Shaded areas denote 95% confidence intervals; standard errors are clustered at the customer level.

Figure A17: Heterogeneity in Displacement by TOU tariff, pooled treatments - Spain



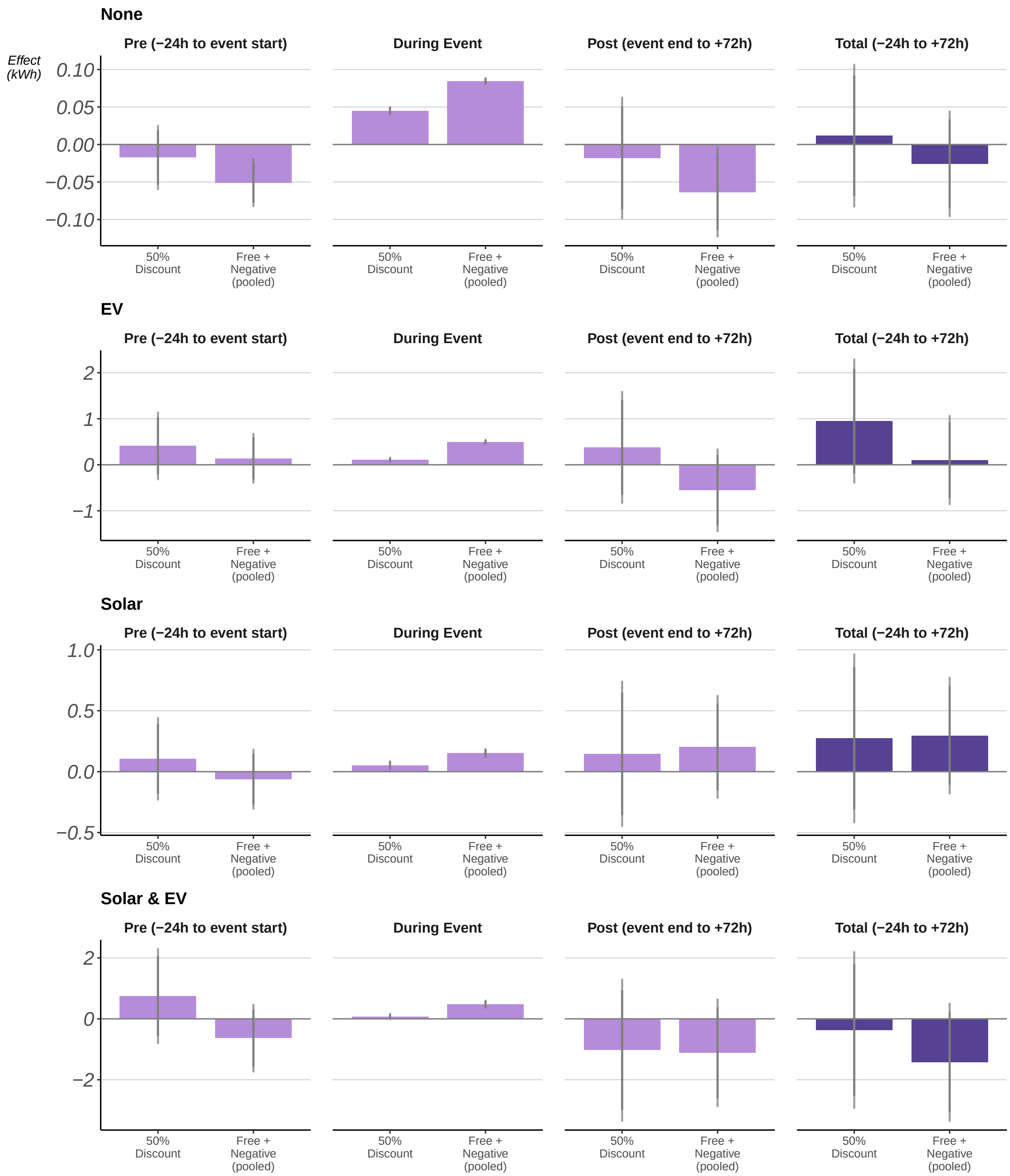
Notes: This figure shows the impact of treatment on cumulative electricity consumption in different windows. Estimates come from interaction specifications that interact treatment status with tariff type and are expressed relative to the control group within each tariff. This figure pools Free and Free + 10c into one group. Shaded areas denote 95% confidence intervals; standard errors are clustered at the customer level.

Figure A18: Heterogeneity in displacement by low-carbon technology tariff - Great Britain



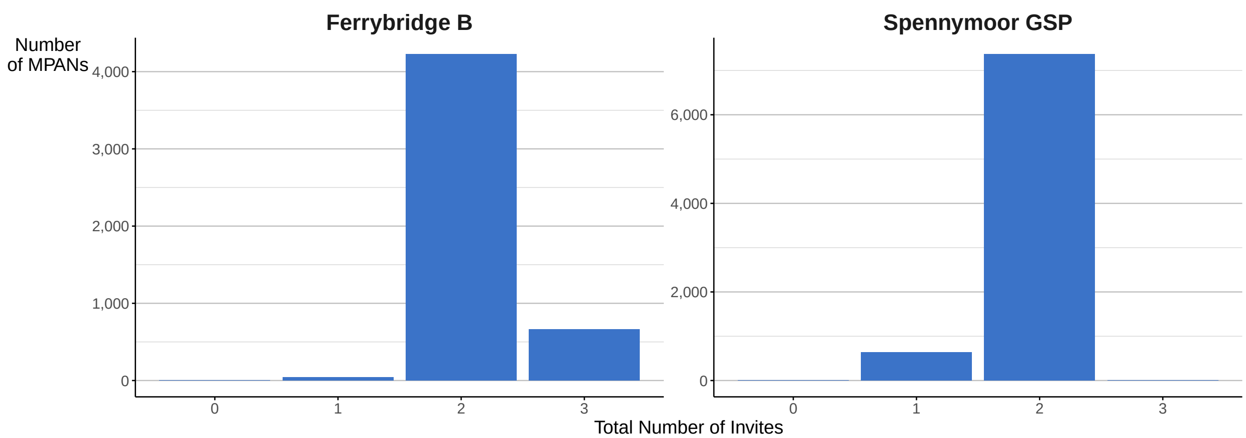
Notes: This figure shows impact of treatment on cumulative electricity consumption in different windows. Estimates come from interaction specifications that interact treatment status with LCT tariff type and are expressed relative to the control group within each LCT group. Shaded areas denote 95% confidence intervals; standard errors are clustered at the customer level.

Figure A19: Heterogeneity in displacement by low-carbon technology tariff, pooled treatments - Great Britain



Notes: This figure shows impact of treatment on cumulative electricity consumption in different windows. Estimates come from interaction specifications that interact treatment status with LCT tariff type and are expressed relative to the control group within each LCT group. This figure pools Free, Free + 5p, and Free + 15p into one group. Shaded areas denote 95% confidence intervals; standard errors are clustered at the customer level.

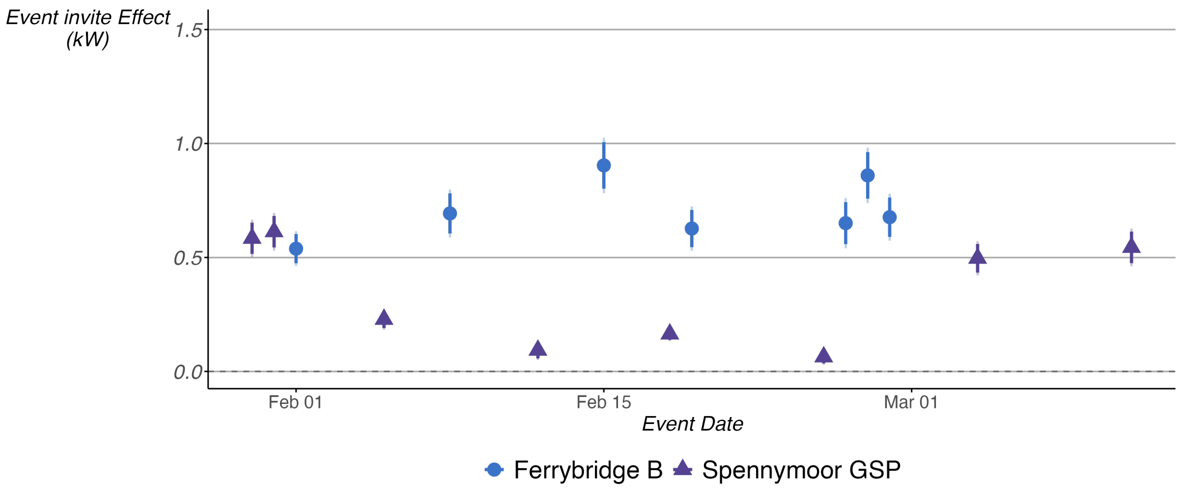
Figure A20: Number of customers invited in NPG "power-ups"



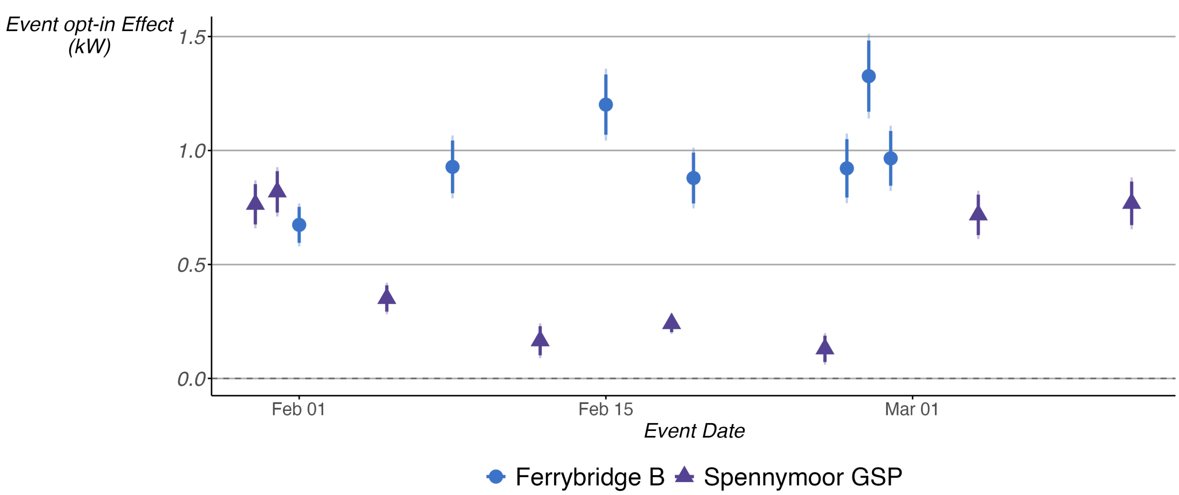
Notes: This figure plots the number of invites customers receive throughout the duration of the NPG "power-up" trials.

Figure A21: Effects of NPG “power-ups”, by event day

(a) Signed up



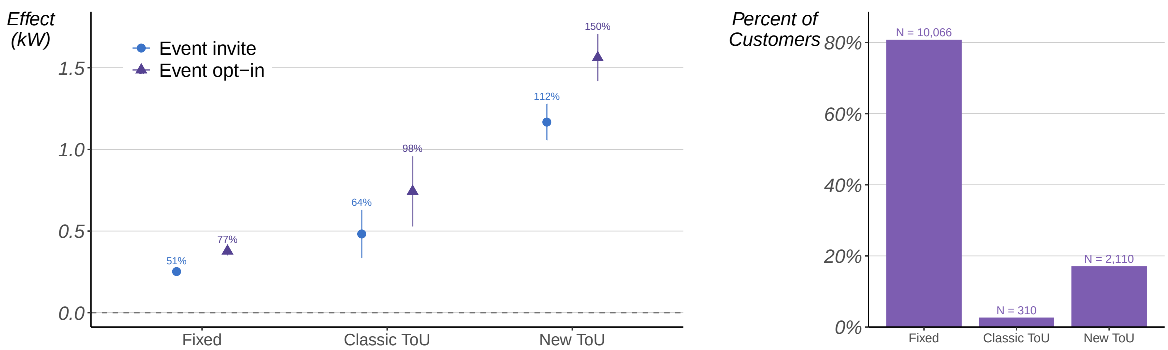
(b) Event opt-in



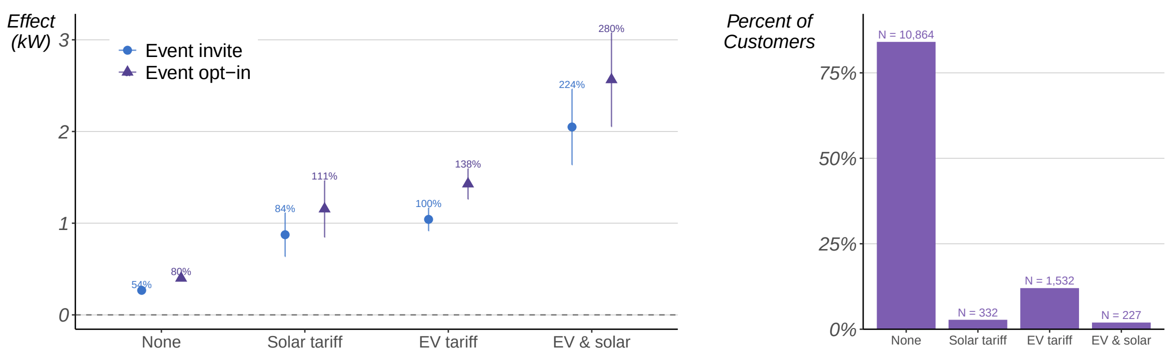
Notes: This figure reports estimates of effect of being invited (panel a) and event opt-in (panel b) separately for each event day. Each point estimate represents the effect of being invited to, or opting into, a power-up event on hourly electricity consumption on that event day. Estimates come from the main specification in Equation (2) and Equation (Second Stage), estimated separately for each event day. Lines depict 95% confidence intervals. Standard errors are clustered at the customer level.

Figure A22: Heterogeneity of NPG effects by tariff type

(a) TOU tariff type

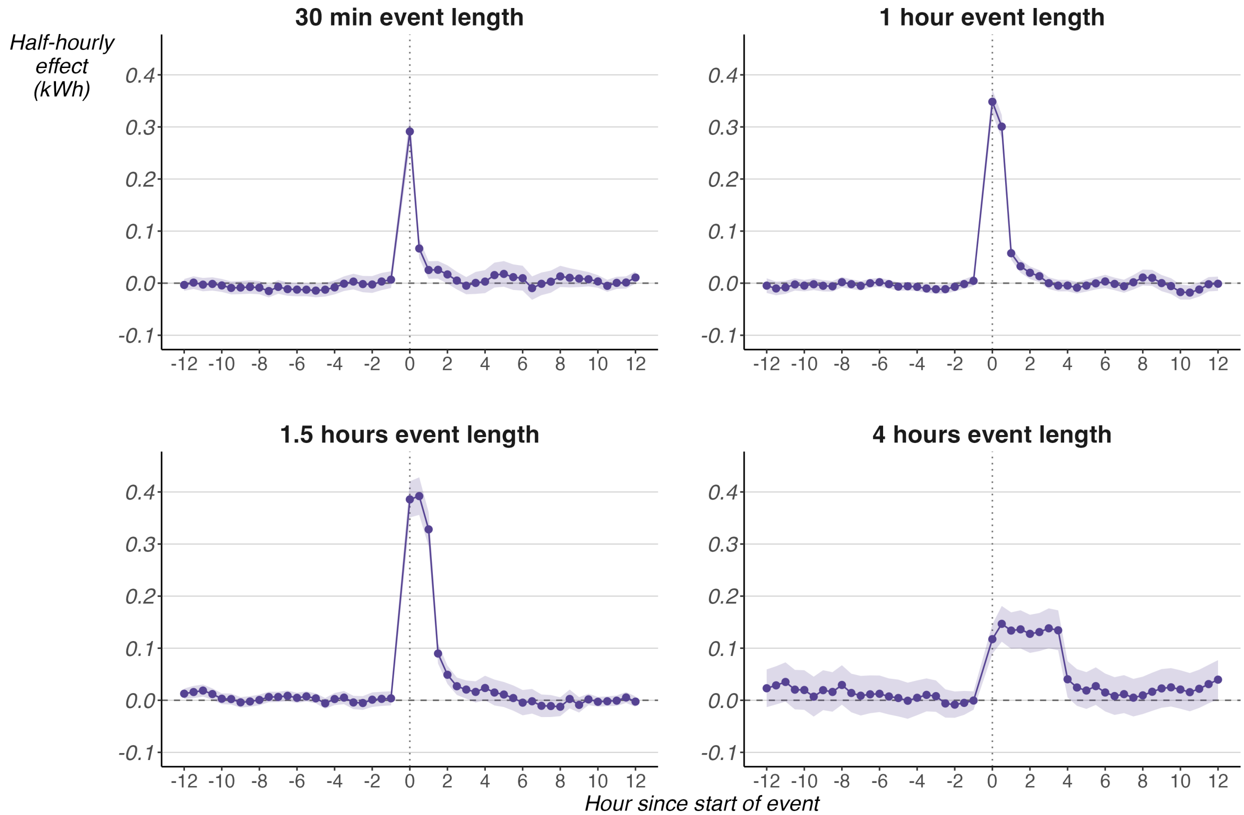


(b) LCT tariff type



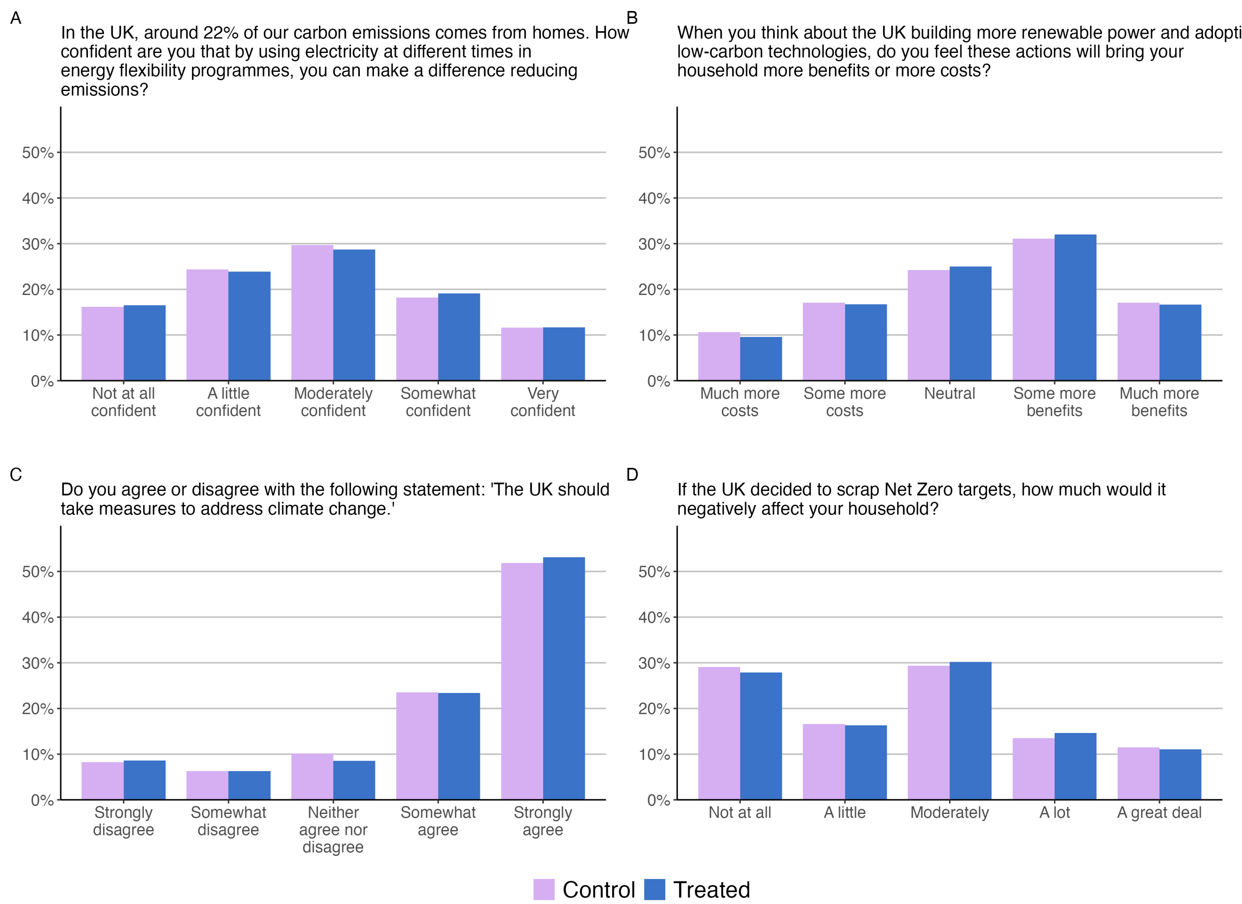
Notes: This figure reports heterogeneity in treatment effects by tariff type. Each point estimate represents the signed-up (circle) and event opt-in (triangle) treatment effects for customers on each tariff, obtained by interacting the treatment indicator with tariff type indicators in the main specification. Panel (a) shows heterogeneity by time-of-use (TOU) tariff type, distinguishing between customers on fixed tariffs, Classic TOU tariffs, and New TOU tariffs. Panel (b) shows heterogeneity by low-carbon technology (LCT) tariff type, distinguishing between customers with no LCT tariff, a solar tariff, an EV tariff, or both an EV and solar tariff. The percentage labels report the effect as a share of the control mean for that subgroup. The bar charts on the right of each panel report the share of customers in each tariff category. Lines depict 95% confidence intervals. Standard errors are clustered at the customer level.

Figure A23: NPG power-up event study analysis



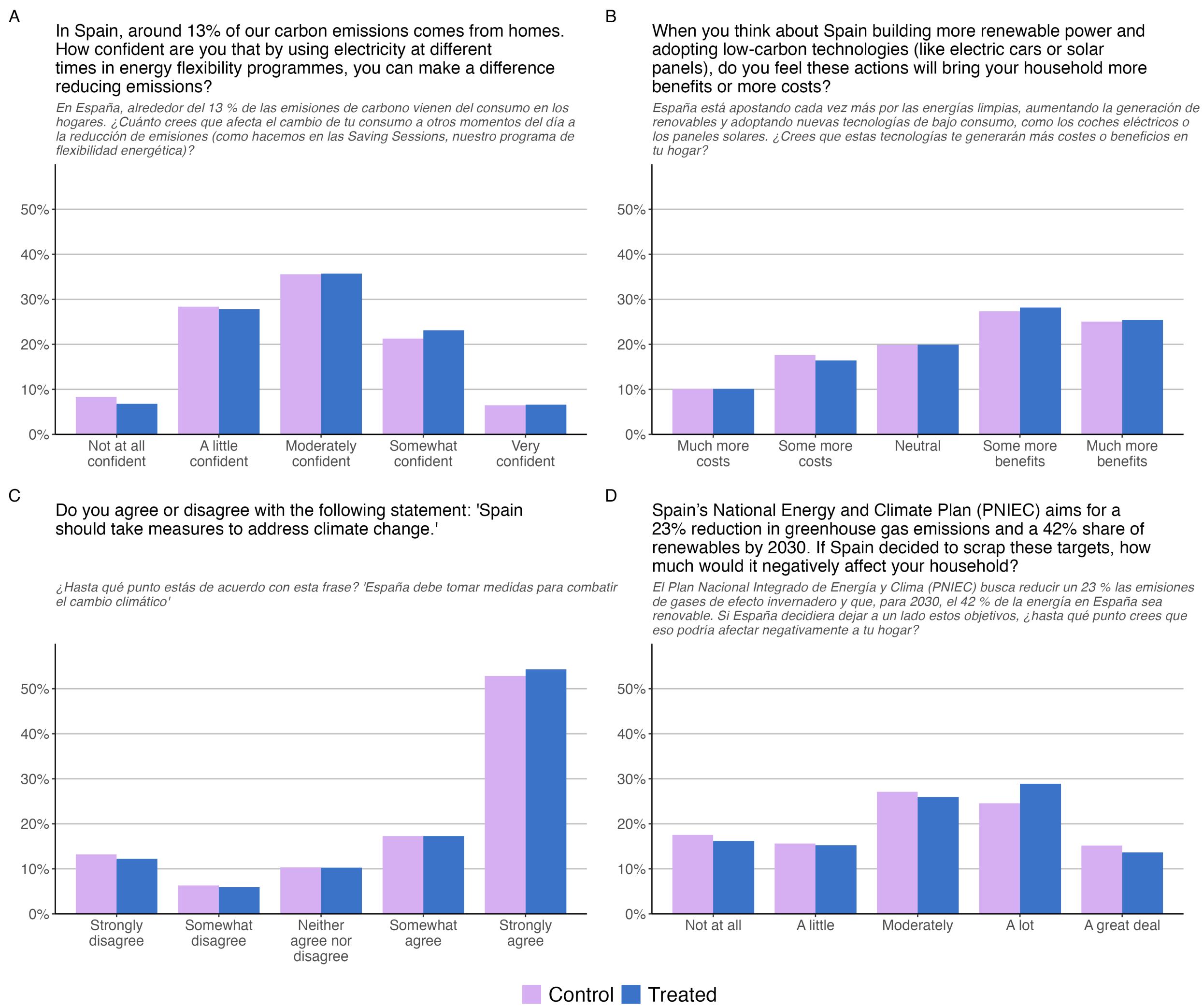
Notes: This figure reports estimates from the event-study specification, estimated separately for each event length. Each panel plots coefficients representing the half-hourly consumption response of invited customers relative to uninvited customers, at each half-hour relative to the start of the event. The omitted period is the half-hour immediately preceding the event. The sample is restricted to events where neither the 12 hours before nor the 12 hours after overlap with another event, leaving 11 events. Of the 11 events in this sample, 2 had an event length of 30 minutes, 5 had an event length of 1 hour, 3 had an event length of 1.5 hours, and 1 had an event length of 4 hours. The vertical dashed line indicates the start of the event. Shaded areas indicate 95% confidence intervals. Standard errors are clustered at the customer level.

Figure A24: Climate sentiment survey results: Great Britain



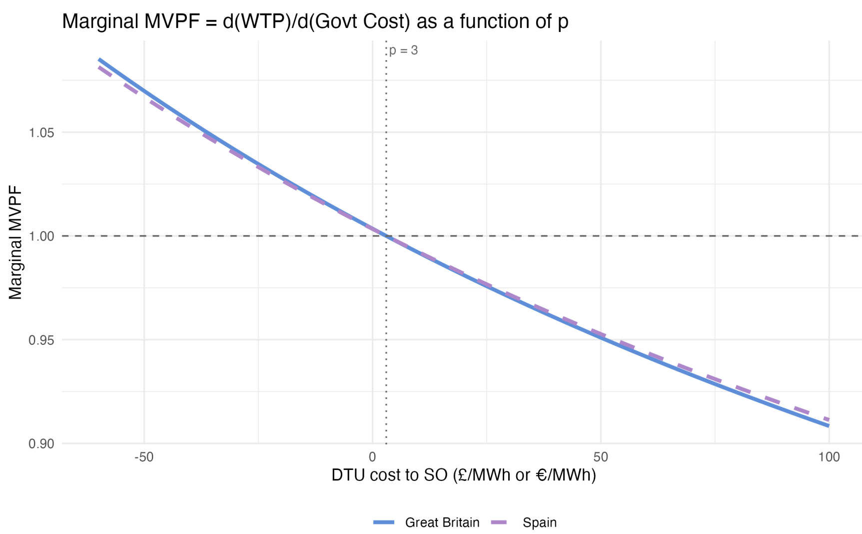
Notes: This figure shows the distribution of responses to four climate sentiment survey questions among Great Britain respondents, separately for treated customers (those who received any turn-up emails) and pure control customers. Panel A measures confidence in energy flexibility reducing emissions. Panel B captures perceived household costs versus benefits of renewable energy expansion. Panel C elicits support for government action on climate change. Panel D measures perceived household impact of abandoning Net Zero targets. Responses are measured on a five-point Likert scale. The survey was sent to 36,412 customers, with 4,260 responses received.

Figure A25: Climate sentiment survey results: Spain



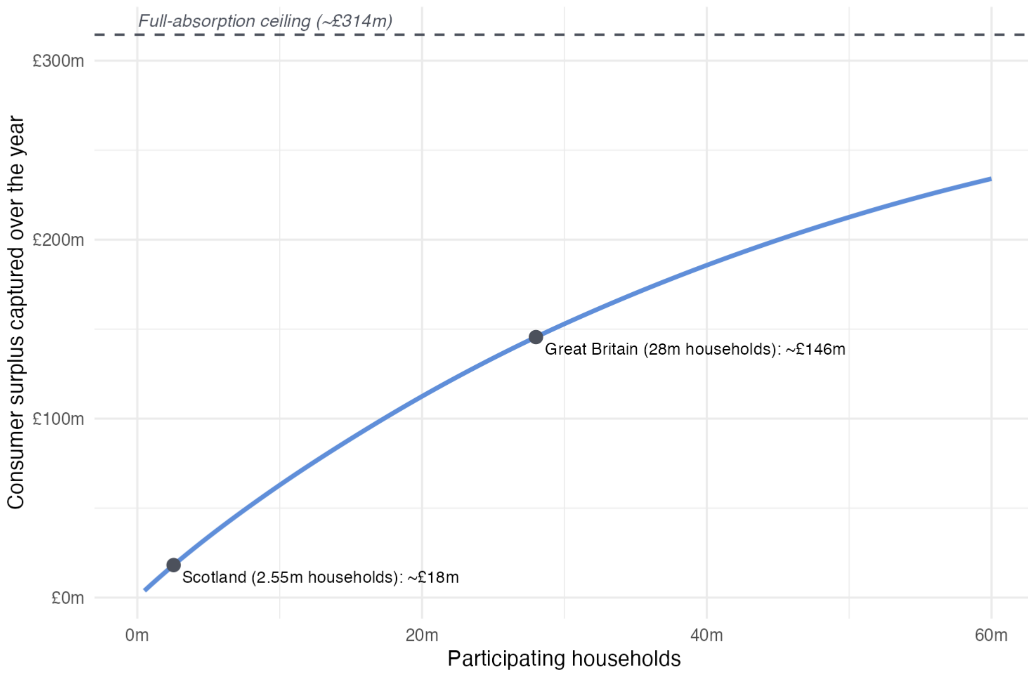
Notes: This figure shows the distribution of responses to four climate sentiment survey questions among Spain respondents, separately for treated customers (those who received any turn-up emails) and pure control customers. Panel A measures confidence in energy flexibility reducing emissions. Panel B captures perceived household costs versus benefits of renewable energy and low-carbon technology adoption. Panel C elicits support for government action on climate change. Panel D measures perceived household impact of abandoning Spain's PNIEC targets. Questions are shown in both English and Spanish. Responses are measured on a five-point Likert scale. The survey was sent to 25,999 customers, with 5,533 responses received.

Figure A26: Marginal MVPF as a function of DTU price



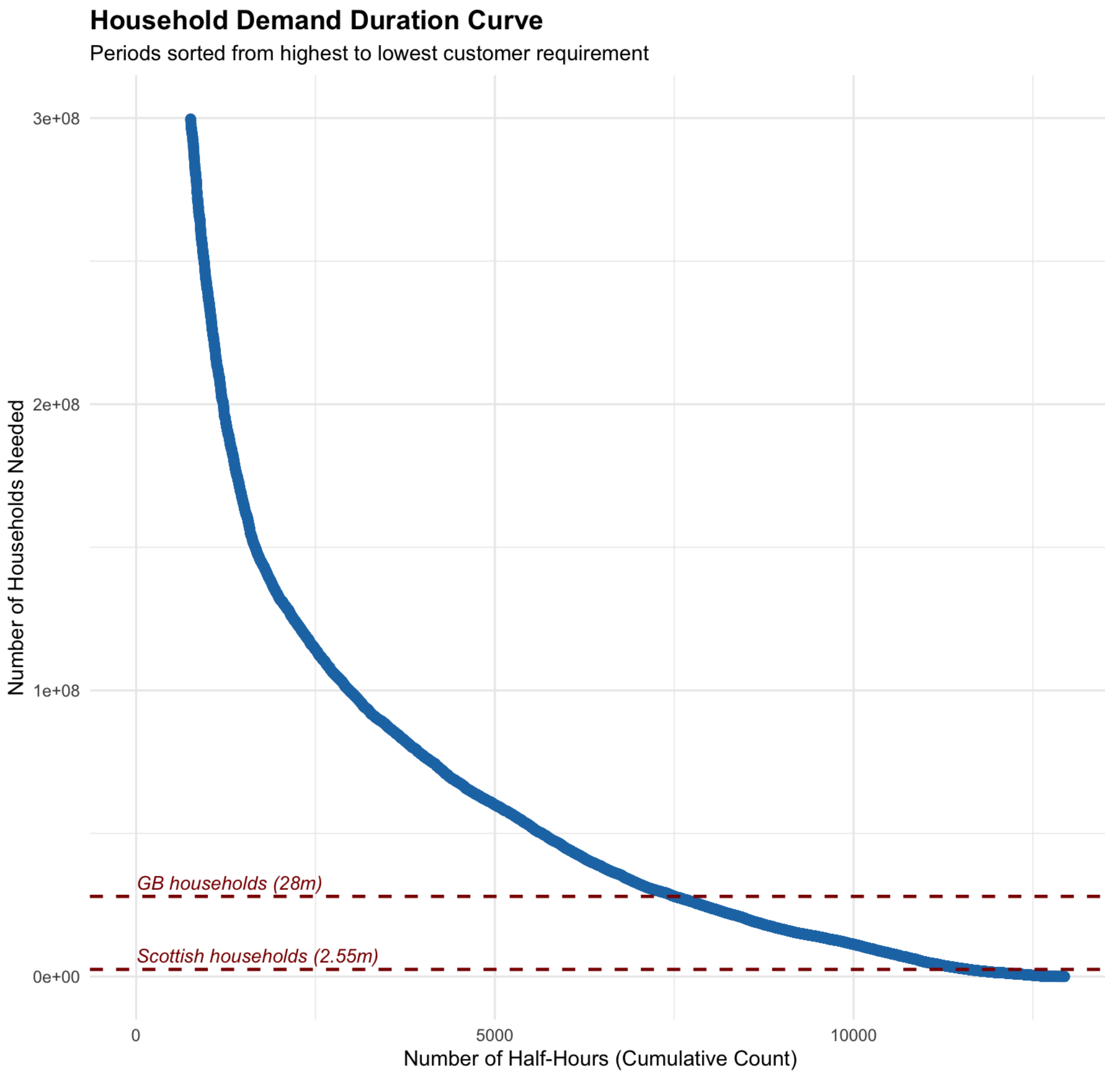
Notes: This figure plots the marginal value of public funds (MVPF) of demand turn-up as a function of the DTU price p, defined as the system operator's payment to the retailer per MWh of demand procured (negative values indicate the retailer pays the system operator). The marginal MVPF is computed as $d(\text{WTP})/d(\text{Govt Cost})$ evaluated at each value of p. The dashed horizontal line indicates MVPF = 1; the dotted vertical line indicates $p = \kappa =$ £3/MWh, the estimated cycling cost of curtailment. Results are shown separately for Great Britain and Spain. The two curves are nearly coincident, illustrating that the welfare-maximizing DTU price is independent of the institutional treatment of curtailment. Assumes a wholesale price of £70/MWh (€70/MWh), an output-based subsidy of £30/MWh (€30/MWh), a retail price of £240/MWh (€240/MWh), and a price elasticity of demand of −0.3.

Figure A27: Consumer surplus captured, by participation



Notes: This figure traces the total consumer surplus that residential demand turn-up would capture over the sample period as the number of participating households rises. For each half-hourly settlement period in which generation was curtailed due to local network constraints – identified as system-operator-accepted bid volume with a positive system sell price – we compute the induced demand from a given number of participating households by applying the price elasticity of demand estimated in our natural field experiment (-0.3) to an average half-hourly household consumption of 0.17 kWh, under an assumed retail price reduction equal to the prevailing system sell price. Consumer surplus on the induced demand is the standard triangle $\frac{1}{2}\Delta p\,\Delta Q$; inframarginal savings are transfers and net out of the welfare measure (Section 6.1). Each participation level absorbs the curtailed volume in every period up to that period's household requirement, so the curve is concave: the most easily absorbed periods are captured first, and surplus rises toward the full-absorption ceiling (dashed line) with diminishing returns. The orange markers indicate the approximate household stock of Scotland (2.55 million) and Great Britain (28 million); the horizontal axis is expressed in household-equivalents of participation and extends beyond the residential stock to represent additional non-residential flexible load. Values are aggregated over all settlement periods in the sample (19 February 2025 through 18 February 2026).

Figure A28: Household demand duration curve



Notes: Each point represents a half-hourly settlement period in which generation was curtailed due to local network constraints, identified as system-operator-accepted bid volume with a positive system sell price. The y-axis shows the number of households that would have to participate in demand turn-up to absorb that period's curtailed volume in full, computed by applying the price elasticity of demand estimated in our natural field experiment (−0.3) to an average half-hourly household consumption of 0.17 kWh, under an assumed retail price reduction equal to the prevailing system sell price. Periods are sorted from highest to lowest household requirement and plotted against their cumulative count, so the curve traces the share of curtailment periods absorbable at each level of participation. The dashed horizontal lines indicate the approximate household stock of Scotland (2.55 million) and Great Britain (28 million); periods lying above a line could not be fully absorbed even if every household at that scale participated. Values are aggregated over all settlement periods in the sample (19 February 2025 through 18 February 2026).

C Deviations from the Pre-Analysis Plans

This appendix records deviations from the pre-analysis plans for the empirical analyses in the Great Britain, Spain, and Northern Powergrid trials. The welfare framework in Section 6 was not pre-specified and is therefore not discussed here. Where possible, we report the pre-specified specification alongside the final specification.

Baseline controls: In Great Britain, the final control set follows the pre-analysis plan after adjusting the baseline consumption controls to match the hourly outcome. In Spain, we pre-specified controls for estimated annual consumption, meter type, EPC rating, and tenure. In the final analysis, we instead control for observed baseline afternoon consumption and tenure. Estimated annual consumption was not well defined for many Spanish customers because it is a modeled measure based on past consumption; but many customers were relatively new to Octopus Energy Spain and lacked historical consumption data. We therefore used the directly observed baseline consumption measure used in the matched randomization, defined as average hourly consumption between 13:00 and 17:00 in May 2025.

For NPG, the pre-analysis plan specified average half-hourly consumption during January 2025, meter type, EPC rating, and tenure. The final main specification instead uses average consumption at the same half-hour over the last 10 weekdays and the last 2 weekend days, together with meter type, EPC rating, estimated annual consumption, and tenure. We made this change to align the NPG specification with the national turn-up specifications and to control more directly for same-time baseline consumption patterns. We report the pre-analysis plan version as a robustness check in Table A16.

Sample restrictions: The main sample-construction deviations occurred in Spain. The pre-analysis plan stated that customers with solar panels, customers enrolled in Intelligent Octopus tariffs, and customers participating in related Octopus programs would be excluded. During implementation, we identified customers who had been inadvertently included despite these restrictions, including customers on export tariffs and Intelligent Octopus tariffs. These customers were excluded from the final analysis. We also excluded customers with multiple accounts when our implementing partner determined that mixed account eligibility created customer confusion and operational difficulties. Finally, customers who left Octopus Energy Spain or became ineligible before a given event were excluded for that event. These restrictions are described in Appendix D.1.

Our pre-analysis plan specified that each customer would rotate through all five conditions over the five events, following the balanced Latin square sequences in Section 3.3. Due to an implementation error, the fifth event in Spain repeated the fourth event's assignments across all sequences, so Spanish customers received their fourth-event condition twice and were never exposed to their scheduled fifth condition. Events one through four were implemented as planned. This means every treatment was still tested, but we no longer have perfect balance on treatment order. Estimates excluding the fifth event are similar to our main results (Table A4), so our conclusions do not depend on the duplicated event. This deviation applies only to Spain.

Robustness specifications: The Great Britain and Spain pre-analysis plans specified robustness checks excluding covariates and analyses separating rotating-control households from pure-control households to test for carryover effects. In the final paper, we also report customer fixed-effect specifications in Table A3 and Table A4. These specifications were not pre-specified for the national turn-up trials and are included as robustness checks, not as the main estimates. The main estimates remain the pooled-control ITT specifications described in the pre-analysis plans.

Heterogeneity and dynamics: The pre-analysis plans specified heterogeneity by baseline tariff type, smart tariff status, and anchoring to the first treatment assignment. The final paper reports these analyses, but also reports additional descriptive heterogeneity by low-carbon technology tariff status, baseline consumption, property value or income, geographic area, and email engagement. These additional heterogeneity analyses were not pre-specified and should be interpreted as descriptive subgroup comparisons.

NPG event structure and empirical specification: The NPG pre-analysis plan anticipated approximately 24 power-up events between January and March 2025. In practice, NPG called 19 events across 15 event days, with some days containing multiple events. Because same-day events share a demand environment and create overlapping pre- and post-event windows, the final analysis treats the event day as the relevant unit for several analyses, especially displacement.

The NPG pre-analysis plan stated that Spennymoor and Ferrybridge B would be analyzed separately. In the final paper, the main NPG estimates pool the two areas and include zone fixed effects, while reporting area-specific estimates in Table A15.

D Implementation Details

D.1 Eligibility

In Great Britain, our implementing partner had the following required exclusions:

- Customers who have enrolled in or been contacted about other turn-up campaigns, including NPG power-ups, LCM power-ups, and UKPN power-ups
- Customers who are enrolled in the Free Electricity Octoplus campaign, which is similar to the events in this experiment in that it gives Octopus customers free electricity at certain times
- Customers on Agile Octopus tariffs, a real-time tariff that tracks wholesale electricity prices on a half-hourly basis. These customers already respond to negative wholesale prices.

In Spain, our implementing partner had the following required exclusions:

- Self-employed customers
- Customers with an ATR code ("Acceso de Terceros a la Red") different from 2.0TD. This excluded higher voltage connections
- Customers who opted out of marketing communications
- Vulnerable customers
- Customers with open complaints
- Customers who have not received a bill in more than 60 days
- Customers whose agreement with Octopus Energy is scheduled to be renewed between 30 and 60 days after the first session begins
- Customers whose primary address is in the Canary Islands
- Customers in distribution zones unable to send smart meter data within 5 days

We additionally excluded customers participating in existing Octopus programs (Solar Family, Sun Club, and "Comunidad Octoamigos") and those enrolled in electric vehicle tariffs, since our partner planned to run concurrent flexibility campaigns with these groups. Customers exporting solar generation were also excluded due to billing complexity.

In our pre-analysis plan, we stated that these exclusions yielded a sample of 66,441 customers. After the first event, however, we identified an implementation error: 2,486 customers who were participating in other flexibility programs or enrolled in an electric vehicle tariff had been inadvertently included. After the second event, our implementing partner also determined that customers holding multiple accounts should be excluded. Some of these accounts were eligible while others were not, which created customer confusion and imposed additional operational burdens. We therefore excluded 2,853 such customers. Finally, 664 customers either left Octopus Energy Spain or became ineligible between sampling and the first event.

D.2 Matched pair randomization

In Great Britain, to assign customers to treatment conditions across the five events, we constructed two Latin squares, generating 20 unique treatment sequences. Each sequence specifies a distinct ordering in which a customer is exposed to each of the five conditions (four treatments and one rotating control). We additionally defined four sequences in which customers were never treated in any of the five events, forming a pure control group. In total, this yielded 24 distinct assignment sequences.

Within our sample, we then implemented a matched-pair randomization to assign sampled households to experimental conditions across five events, with the aim of ensuring pairwise treatment balance over time. We did this through the following steps:

1. Strata for Randomization: We created randomization strata based on three dimensions: (1) property value tercile, interacted with urban/rural classification, (2) property type (house vs. non-house), and (3) baseline tariff type

2. Blocking: Within each stratum, we ranked accounts by electricity consumption and grouped them into blocks of 24.

3. Sequence Assignment: For each block, we randomly assigned one of 24 pre-generated balanced Latin square sequences of treatments (5 treatments across 5 events) and

pure control sequences. These sequences ensure that each treatment precedes and follows every other treatment equally often across blocks.

4. Misfits: Accounts that did not fit into full blocks were pooled and randomly assigned to treatment sequences from the same set of 24 sequences.

In Spain, we followed a similar randomization strategy. We again used two Latin squares to generate 20 treatment sequences and included an additional six sequences in which customers were never treated, yielding 26 total sequences. In the absence of covariates for stratified randomization, we relied on a straightforward matched-pair randomization:

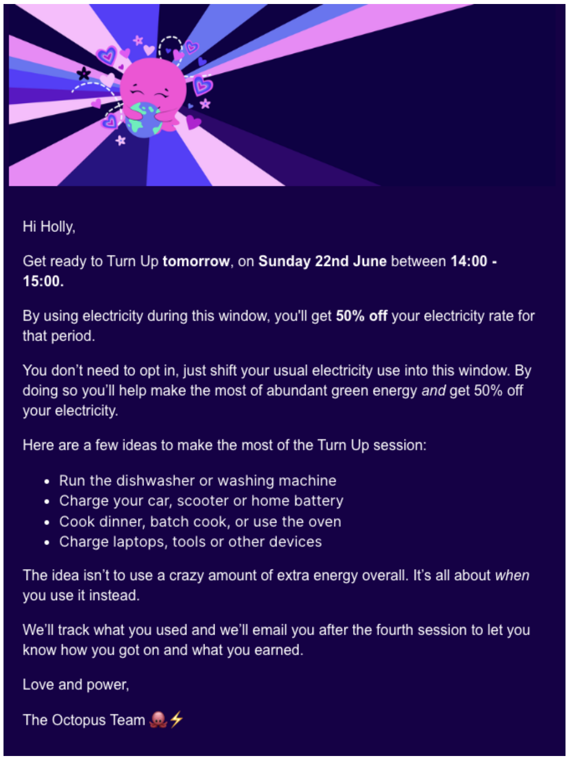
1. Blocking: We ranked customer accounts by average electricity consumption between 13:00 and 17:00 in May 2025, a period during which turn-up events were most likely to occur, and grouped them into blocks of 26.

2. Sequence Assignment: For each block, we randomly assigned one of 26 pre-generated balanced Latin square sequences of treatments (5 treatments across 5 events) and pure control sequences. These sequences ensure that each treatment precedes and follows every other treatment equally often across blocks.

3. Misfits: Accounts that did not fit into full blocks were randomly assigned to treatment sequences from the same set of 26 sequences.

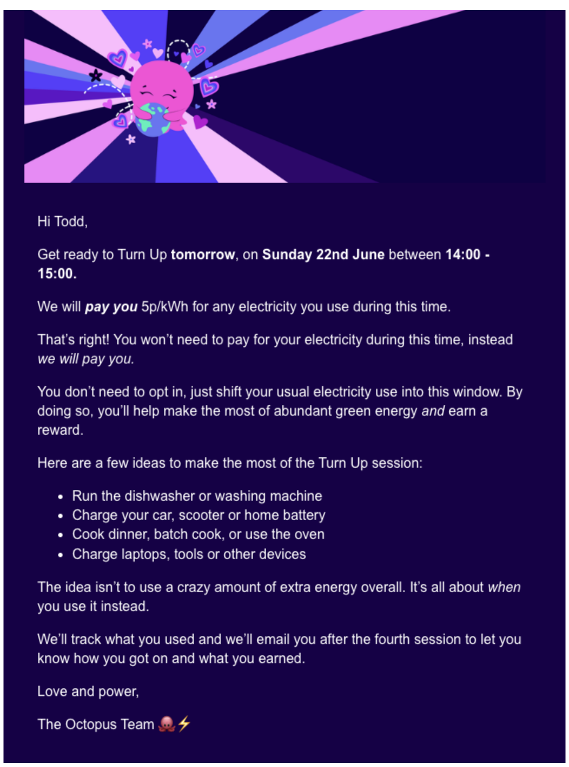
D.3 Email Delivery

Figure A29: Great Britain Treatment emails

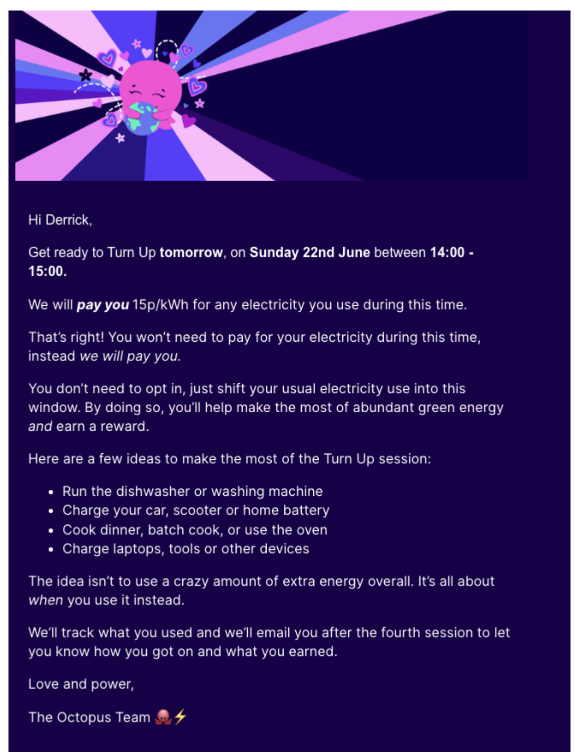


(a) Subject: Get ready to turn-up: 50% off your electricity tomorrow from 14:00 - 15:00

(b) Subject: Get ready to turn-up: free electricity tomorrow from 14:00 - 15:00

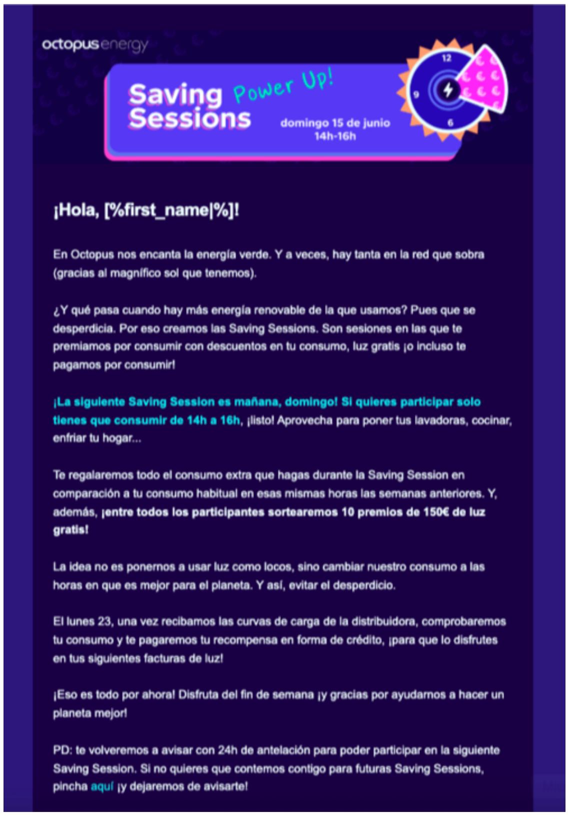


(c) Subject: Get ready to turn-up: free electricity + earn 5p/kWh tomorrow from 14:00 - 15:00

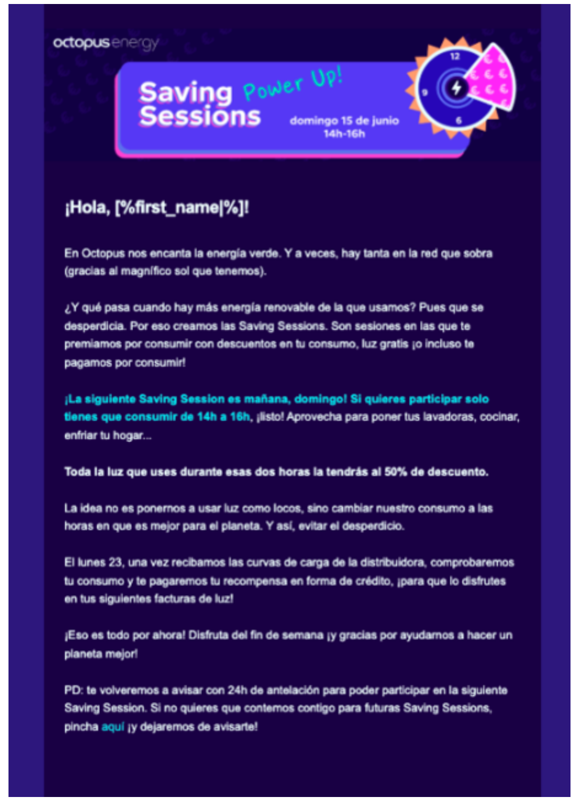


(d) Subject: Get ready to turn-up: free electricity + earn 15p/kWh tomorrow from 14:00 - 15:00

Figure A30: Spain Treatment emails



(a) Prize Draw

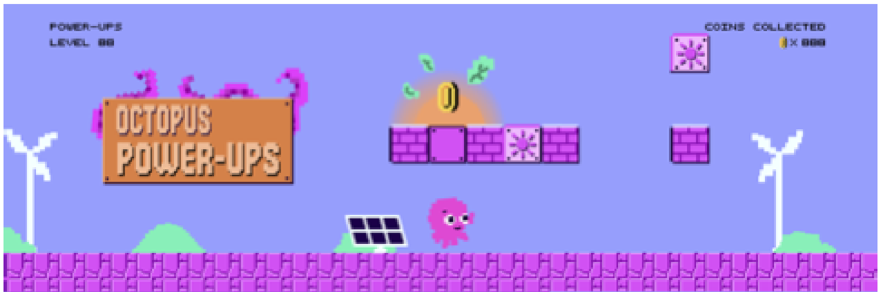


(b) 50% Discount

(c) Free electricity

(d) Free + 10c/kWh

Figure A31: NPG Email Invitation to Participate in Power-up Event



Hi Vania,

We'll make your electricity free at OCT 123, OCT 456 between **1:00 PM - 3:00 PM** on Friday 14/03/25. Get ready to Power up!

Opt me in

What do I need to do during a Power-up?

During the Power-up, there'll be plenty of green electrons in your local wires. We want you to make the most of that lovely sun and wind power while it's abundant.

Plan to use as much electricity as you need during that time: it's a perfect time to do the hoovering, run the dishwasher or washing machine, run your heat pump or immersion heater, charge up, make dinner, you name it. It's your responsibility to use electricity safely at home.

How will I get my free electricity?

We'll use your smart meter readings to work out how much electricity you used during each Power-up. Every few weeks, we'll credit your account to cover the cost of the electricity you used in Power-ups.

Your actual bills and charges will all come out as normal.

On Prepay? Make sure you've topped up enough so you don't run out of credit during a Power-up. (We'll credit you the value of your free electricity within 2 weeks of each Power-ups.)

How will I know how I did?

We'll email you every few weeks to let you know how much money you've saved thanks to the free electricity – and how much local green power you've all collectively used up – in Power-ups!

E Heterogeneity Variable Definitions

This appendix describes the pre-existing household and customer characteristics used in the heterogeneity analyses summarized in Table 1. In each case, the analysis interacted treatment assignment with indicators for the relevant characteristic and included the same baseline controls and fixed effects as the corresponding main specification. Because these characteristics were not randomly assigned, the estimates are descriptive comparisons across customer groups.

Low-carbon technology tariff. In Great Britain, we do not directly observe ownership of electric vehicles or rooftop solar. We therefore proxy for low-carbon technology (LCT) ownership using the customer's baseline tariff. Customers on an exporting tariff are classified as having solar, while customers on EV tariffs, including Octopus Go and Intelligent Octopus, are classified as having an EV. Customers with both an EV import tariff and a solar export tariff are classified separately. Customers without these tariff markers form the no-LCT group. We do not estimate LCT heterogeneity in Spain because solar and most EV customers were excluded from the Spanish trial due to concurrent flexibility campaigns.

Baseline tariff type. Baseline tariff type is measured before the first event. In Great Britain, we classify tariffs into fixed tariffs, classic time-of-use tariffs, and newer time-of-use tariffs. Classic time-of-use tariffs are traditional overnight tariffs, primarily Economy 7 and Economy 10. Newer time-of-use tariffs include fixed multi-band tariffs, dynamic pricing tariffs that track wholesale market conditions, and technology-specific tariffs linked to electric vehicles, heat pumps, solar, or batteries. In Spain, we classify tariffs into fixed tariffs, standard time-of-use tariffs with stable day-to-day price schedules, and dynamic time-of-use tariffs indexed to wholesale market prices. Note that because dynamic time-of-use tariffs already reflect low wholesale prices, control-group customers were likely responding to those prices already, which may have dampened the treatment effect of additional price decreases.

Baseline consumption. Baseline consumption is measured using observed pre-treatment electricity imports during the afternoon hours in May 2025, the period when turn-up events were most likely to occur. Specifically, we use the average hourly import between 13:00 and 17:00 local time and divide customers into country-specific quartiles. This directly measured consumption variable is distinct from modeled estimated annual consumption and is available in both countries.

Property value and income. In Great Britain, socioeconomic status is proxied by the estimated property value of the customer's residence from the WhenFresh property data. Customers are assigned to quintiles using national property-value cutoffs. In Spain, socioeconomic status is proxied by gross household income from the 2023 INE Atlas de Distribución de Renta de los Hogares, table 30824. Income is measured at the census-section level, linked to customers using geocoded census-section codes, and divided into quintiles of gross household income. In 2023, there were 36,460 census sections, with an average of 529 households per section.

Geographic area. In Great Britain, geographic heterogeneity is defined by Grid Supply Point (GSP) Group, which aggregates areas of the electricity distribution network. In Spain, geographic heterogeneity is defined by Comunidad Autónoma. These regional variables are used to describe spatial variation in treatment effects and to assess whether stronger responses are concentrated in areas that may be operationally relevant for absorbing surplus renewable generation.

Email engagement. For Great Britain, we observe whether customers opened event notification emails. We classify customers as having opened an event email if they opened at least one notification email during the trial. This variable is a proxy for awareness of the intervention, not a randomly assigned treatment, and is therefore interpreted descriptively.

Warm Home Discount eligibility. Warm Home Discount eligibility is defined using Great Britain customer records indicating whether a customer was eligible for the Warm Home Discount, a government energy-bill support scheme targeted to low-income and financially vulnerable households. We use this indicator as a proxy for household financial vulnerability. The variable is available only for Great Britain and, like the other heterogeneity measures, is interpreted as a descriptive customer characteristic rather than as a randomly assigned source of variation.

F Impact of turn-up participation on climate policy sentiment

One possible implication of turn-up events is that exposure to periods of abundant renewable generation (and potential financial gains) may help customers better connect renewable energy availability to personal financial benefits. This, in turn, could shape broader beliefs about the feasibility and desirability of climate and renewable energy policies. To examine this possibility, we conducted a post-experiment survey following the completion of the five events.

The survey consisted of four questions with responses measured on a five-point Likert scale, designed to capture customers' climate-related sentiments. We adapted questions from prior survey work on climate-policy beliefs and support, especially Dechezleprêtre et al. (2025), and from official policy-attitudes and target-setting materials in Great Britain and Spain, including the DESNZ Public Attitudes Tracker (Department for Energy Security and Net Zero, 2025) and Spain's National Energy and Climate Plan (*Integrated National Energy and Climate Plan (PNIEC 2023-2030)*, 2024).

The questions were broadly comparable across the two countries, with the Spanish version translated and slightly adapted for the local context. The first question measures customer sentiment toward energy flexibility. The second captures perceived household costs versus benefits of renewable energy expansion and low-carbon technologies. The third elicits overall support for government action on climate change, and the fourth measures perceived household impacts of abandoning Net Zero targets. The exact questions and translations are shown in Figure A24 and Figure A25. We sent the survey to 36,412 customers in Great Britain and 25,999 customers in Spain. We received 4,260 responses in Great Britain and 5,533 responses in Spain.

To test if participation in turn-up had any effect, we ran the following regression:

$$Y_i = \alpha + \beta Treated_i + \gamma OctopusSentiment_i + \mu_b + \varepsilon_i \tag{6}$$

where Y_i denotes the response to a given survey question, coded from 1 (most negative toward climate policy and renewables) to 5 (most positive). $Treated_i$ is an indicator equal to one if customer i received any turn-up emails, and zero if the customer was assigned to the pure control group. $OctopusSentiment_i$ captures responses to the survey question, "We're driving the green energy revolution – building more renewable generation, electrifying homes with heat pumps and other clean technologies, and switching to EVs. How well does our mission align with your own priorities?" This variable is included as a control to account for potential response bias or changes in sentiment toward Octopus Energy induced by participation in the turn-up events, which could otherwise be conflated with broader climate sentiment. Finally, μ_b denotes matched-pair block fixed effects from the original randomization.

Across all survey questions, we find no differences between responses from treated customers and those in the pure control group in either country, indicating that participation in turn-up events did not affect overall climate sentiment (Table A19, Table A20). Notably, the raw response distributions for the four climate-sentiment questions are strikingly similar across Great Britain and Spain (Figure A24, Figure A25), suggesting broadly comparable baseline attitudes toward climate policy in the two contexts.

G Welfare analysis: further details

G.1 Constraint-down bid mark-ups by support regime

This section sets out the bid-level analysis underlying the mark-up discussed in Section 6. The question is how far the prices generators bid to be constrained off exceed the costs a constraint-down action genuinely imposes on them. Because the relevant fundamental cost differs by support regime, we construct a separate benchmark for each and report the wedge between the accepted bid and that benchmark.

Data and sample: We used balancing-mechanism settlement data from Great Britain's National Energy System Operator for the period 19 February 2025 to 18 February 2026, restricting attention to accepted constraint-down bids – bids the system operator actioned to reduce output – in half-hours with a positive system sell price, which we associate with local network constraints rather than national over-supply (see Section 6.4). We restricted the sample to wind generation, which accounts for essentially all constrained-off renewable volume in the period. Non-wind units were removed using the OSUKED generator fuel-type register; secondary balancing-mechanism units and a small number of settlement-period artifacts in the raw data were also dropped. Bids are volume-weighted throughout, with volume measured as curtailed MWh.

Two benchmarks: For each bid we constructed two reference levels. The first, which we label the *minimum bid*, is the fundamental floor: the cost a generator genuinely avoids or forgoes by being constrained off, described regime by regime below. The second, the *maximum bid*, adds the prevailing system sell price to this floor, capturing the value of the real-time energy the generator forgoes when constrained off. The accepted bid should in principle sit between these two: above the fundamental floor, because the generator forgoes at least its support payment, but below the ceiling, because the energy itself has little local value when the network is constrained. We report wedges in levels (£/MWh), rather than as percentages, wherever the benchmark can lie near or below zero, since a percentage mark-up is uninformative when its denominator approaches zero.

Which of these two benchmarks is the "right" one against which to judge a bid depends on what the generator genuinely forgoes when constrained off, and this is less obvious than it first appears. Van Steenberghe and Ovaere (2026) decompose a CfD generator's revenue into a support payment, a day-ahead energy sale, and a real-time settlement:

$$\pi(q_{RT}, q_{DA}) = \underbrace{(P_{\text{strike}} - P_{DA})\, q_{RT}}_{\text{CfD top-up}} + \underbrace{q_{DA}\, P_{DA}}_{\text{day-ahead sale}} + \underbrace{P_{RT}\,(q_{RT} - q_{DA})}_{\text{real-time settlement}}, \tag{7}$$

where q_{DA} and q_{RT} are day-ahead and real-time volumes and P_{DA}, P_{RT} the corresponding prices. A constraint-down action reduces the generator's metered output, so on each curtailed MWh it forgoes the CfD top-up ($P_{\text{strike}} - P_{DA}$) – our minimum-bid floor, since the CfD settles on actual output – and the real-time price P_{RT} on that output as well. The day-ahead sale $q_{DA}P_{DA}$ is a settled forward position the generator retains whether or not it is constrained, and curtailment can eat into that position rather than being shielded by it, so the full P_{RT} is forgone on every curtailed MWh. Our maximum bid adds this P_{RT} to the fundamental floor; our minimum bid credits only the forgone top-up. The two benchmarks bracket the generator's private opportunity cost.[26]

For the welfare and cost-effectiveness questions of this paper, however, the minimum bid is the relevant benchmark, and the reason is the same distortion the paper is about. The real-time term in Equation (7) is settled at P_{RT}, the national system price – the price a constrained generator faces only because the market does not price location. A generator constrained off in a surplus region is thereby credited the national value of energy that, by definition, cannot reach demand elsewhere. Under locational marginal pricing the relevant price in that region would be near zero or negative, and the real-time term – and with it the gap between the maximum and minimum bids – would largely vanish. The maximum bid thus values constrained energy at a price that exists only as an artifact of national pricing, not as genuine social worth. From the system operator's and the public

[26]Whether a given curtailed MWh forgoes the real-time price depends on whether it had already been sold day-ahead: on volume the generator has pre-sold, the day-ahead revenue is retained and only the CfD top-up is forgone (our minimum bid); on volume not yet sold, the generator additionally forgoes the system price it would have earned (our maximum bid). Because we do not observe the day-ahead position q_{DA} at the unit level, we cannot tell where in this range a given bid's true opportunity cost falls, and we therefore report both benchmarks rather than a point estimate.

purse's perspective, what is truly forgone when curtailment is avoided is the subsidy alone – the minimum-bid floor – which is why we anchor on it. The maximum bid is the right lens for what a profit-maximizing generator would privately bid in today's market; the minimum bid is the right lens for what society loses.

Regime-specific floors: We assigned each wind unit to one of three support regimes, and within the Contracts for Difference (CfD) regime separated one heavily curtailed unit, discussed below.

CfD: For units holding a Contract for Difference, the fundamental floor is the top-up the generator forgoes when constrained off: the strike price less the intermittent market reference price.[27] This is the cleanest benchmark in our data, since both the strike price and the reference price are observed rather than assumed. Strike prices and CfD-to-unit mappings are taken from the Low Carbon Contracts Company registers.

Renewables Obligation (RO): For units accredited under the Renewables Obligation, the floor is the value of the Renewables Obligation Certificates the generator forgoes by not generating: the unit's certificate band (in certificates per MWh, ranging from 0.9 for onshore to 2.0 for older offshore) multiplied by an assumed certificate value of £65/MWh. We set this value to the Renewables Obligation buy-out price, which was £64.73/ROC in 2024–25 and £67.06/ROC in 2025–26, spanning our sample window.[28] Crediting this certificate value is what distinguishes a genuine mark-up from a recovery of forgone subsidy: absent the credit, RO units appear to bid far above cost, but most of that apparent excess is the certificate value they forgo. The certificate value is a blanket assumption applied across all RO units, and the resulting floor is correspondingly less precise than the CfD floor.

Merchant: Units matching neither a CfD strike nor our hand-constructed RO accreditation list are treated as merchant, with a fundamental floor of zero: with no output-based support to forgo, the whole accepted bid is a wedge above fundamental cost. This

[27] Our bid-netting approach parallels that used in CEPA's cost-benefit analysis of BSC modification P462 (CEPA, 2026), which likewise isolates a residual turndown cost by removing forgone low-carbon support from accepted bids.

[28] Source: Ofgem, Renewables Obligation buy-out price determinations. The buy-out price is the per-certificate sum suppliers must pay in lieu of presenting a certificate, and it is RPI-indexed annually. It is the regulatory floor on a certificate's value; the full market value is modestly higher once the buy-out recycle payment is included, so using the buy-out price alone is, if anything, a slight under-statement of the forgone subsidy – which would marginally overstate the Renewables Obligation wedge we report.

is a residual category, and the floor of zero is an assumption rather than a verified fact: some units classified as merchant may hold Renewables Obligation accreditation we did not capture, or Feed-in Tariff support we did not check. The merchant wedge should therefore be read as an upper bound on true merchant mark-up, with any uncaptured subsidy implying a lower true figure.

The Moray East unit: Within the CfD regime, we separate Moray East (MOWEO). On 7 April 2025, Ofgem opened an investigation into the compliance of Moray Offshore Windfarm (East) Limited with the Transmission Constraint Licence Condition (condition 20A of the Electricity Generation Standard Licence Conditions), assessing whether the unit's bid prices were excessive relative to the expected marginal cost of reducing generation during periods of constraint.[29] The opening of an investigation does not imply a finding of non-compliance. In our sample window the unit's strike price net of the reference price was close to zero, so its fundamental floor is near zero and its percentage mark-up is correspondingly extreme and uninformative. Moray East is also the largest single CfD unit by curtailed volume in our sample, so its exclusion is material; we separate it both because it is the subject of an active regulatory investigation and because its near-zero benchmark would otherwise dominate the CfD figures. Excluding it is conservative: its apparent mark-up is the largest in the sample, so excluding it understates rather than inflates the headline result. The mark-up we report in the main text, a volume-weighted wedge of approximately £11/MWh (26% above the fundamental floor), is therefore the CfD figure excluding Moray East.

Results: Figure A32 shows the bid-level evidence directly. Each point is an accepted constraint-down bid, plotted against its fundamental floor (left panel) and against the floor-plus-forgone-energy ceiling (right panel), with the dashed 45-degree line marking equality. The two panels show the same bids: the right panel is the left panel with each bid displaced horizontally by the prevailing system sell price, since the maximum bid simply adds that price to the floor; the rightward smear therefore traces the distribution of forgone energy value. The pattern that motivates the analysis is visible without any aggregation: against the floor, the mass of bids sits *above* the line – generators bid more than the support they forgo – while against the ceiling, the mass sits *below* it, confirming that they do not extract the full forgone energy value.

[29] Ofgem opened the investigation on 7 April 2025; see the Ofgem publication (accessed 26 June 2026).

Nonetheless, 14.1% of curtailed wind volume over the sample year – 1.18 million MWh – was accepted at a bid exceeding even this generous ceiling, for which no legitimate cost-recovery story remains: the maximum bid already credits the generator with the full forgone system-price value under the most favorable assumption about its day-ahead position, so a bid clearing it cannot be explained by that uncertainty alone. This is unevenly distributed across regimes: 18.0% of Moray East's volume and 13.7 and 13.8% of Merchant and RO volume exceeds the ceiling, against just 3.6% for CfD (excluding Moray East) – the regime our headline κ/μ calibration is drawn from. The calibration therefore rests on the regime least prone to this pattern; the phenomenon is markedly more pronounced elsewhere in the market.

Figure A32: Accepted constraint-down bids against fundamental benchmarks

Notes: Each point is an accepted constraint-down bid in a positive-system-price half-hour, wind units only, 19 February 2025 to 18 February 2026. The vertical axis is the accepted bid (£/MWh); the horizontal axis is the fundamental floor (left: the regime-specific minimum bid) or the floor plus the prevailing system sell price (right: the maximum bid). The dashed line is the 45-degree line of equality. Points above the line in the left panel bid above their fundamental floor; points below the line in the right panel bid below the floor-plus-forgone-energy ceiling. The right panel is the left panel with each bid shifted rightward by the system sell price, so the horizontal spread there traces the distribution of forgone energy value. Axes are clipped to $[-200, 300]$ for readability. Colors distinguish the highlighted large-volume units (Moray East, Seagreen) from other CfD, Renewables Obligation, and merchant units.

Table A21 reports the full set of results: for each regime, the volume-weighted accepted bid, fundamental benchmark, and the resulting wedge, under both the minimum-bid and maximum-bid benchmarks, and split by the sign of the benchmark. We emphasize that the wedge reported in this table is *gross* of any real curtailment cost: it is simply

Table A21: Constraint-down bid wedges by support regime

Regime	Benchmark (sign split)	Volume (MWh)	Bid (£/MWh)	Benchmark (£/MWh)	Wedge (£/MWh)	Wedge (%)
Minimum-bid benchmark (fundamental floor)						
Merchant	all	5,382,233	23.7	0.0	23.7	–
RO	all	1,085,132	75.6	64.2	11.4	17.7
CfD (ex-MOWEO)	all	348,934	54.7	43.4	11.3	26.1
	benchmark ≥ 0	316,712	59.9	49.8	10.1	20.3
	benchmark < 0	32,222	2.8	-20.0	22.9	–
CfD (MOWEO)	all	1,563,034	46.4	0.9	45.6	–
	benchmark ≥ 0	785,446	66.9	21.7	45.1	207.4
	benchmark < 0	777,589	25.8	-20.2	46.0	–
Maximum-bid benchmark (floor + forgone energy value)						
Merchant	all	5,382,233	23.7	75.4	-51.7	–
RO	all	1,085,132	75.6	130.7	-55.1	–
CfD (ex-MOWEO)	all	348,934	54.7	123.0	-68.4	–
CfD (MOWEO)	all	1,563,034	46.4	79.5	-33.1	–
	benchmark ≥ 0	1,554,008	46.6	80.0	-33.4	–
	benchmark < 0	9,026	11.1	-9.1	20.2	–

Notes: All figures are volume-weighted by curtailed MWh. The wedge is the volume-weighted mean of the accepted bid less the benchmark. This wedge is *gross*: it has not been netted of any genuine cycling or maintenance cost, and should not be read as the strategic mark-up μ discussed in Section 6.1. Section 6.1 takes the CfD (ex-Moray East) wedge reported here and nets out an assumed cycling/maintenance cost $\kappa \approx 3$/MWh (based on separate qualitative input from Octopus Energy Generation, not derived from the bid data) to isolate the strategic component μ; the wedge and μ are therefore related but numerically distinct quantities. The wedge-as-%-of-benchmark column is omitted where the benchmark is at or near zero (merchant units, by construction, and Moray East, whose net strike is near zero) or negative, since the percentage is then uninformative. Sign splits divide each regime by whether the benchmark is non-negative or negative; rows without a meaningful split (zero volume in one sign) are shown only as "all". The minimum-bid benchmark is the regime-specific fundamental floor; the maximum-bid benchmark adds the prevailing system sell price. RO floors use an assumed certificate value of £65/MWh.

accepted bid less subsidy-based floor, with nothing subtracted for genuine cycling or maintenance cost. It is therefore a different quantity from the strategic mark-up μ used in the welfare calibration of Section 6.1, which nets a further κ out of this same wedge.

The headline finding is the CfD (excluding Moray East) row: against the minimum-bid floor, the accepted bid of £54.7/MWh exceeds the fundamental of £43.4/MWh by a gross wedge of £11.3/MWh (26.1%). Of this gross wedge, Section 6.1 attributes approximately £3/MWh to genuine cycling and maintenance cost (κ), leaving approximately £8/MWh (19% of the fundamental) as the strategic component, $\mu \approx 1.19$. The same pattern of bids above fundamentals appears in the RO regime (an £11.4/MWh gross wedge over the certificate-value floor) and, by construction, in the merchant regime (the entire £23.7/MWh bid). We caution that the RO, merchant, and Moray East figures each rest on assumptions discussed above – the certificate value, the residual zero floor, and the near-zero net strike – and we accordingly anchor the main-text κ/μ calibration on the CfD (excluding Moray East) result, where the benchmark is observed rather than assumed.

One unit in the merchant group merits a brief comment, given the public attention it has received. Seagreen, Scotland's largest wind farm, has been widely cited for the scale of its curtailment payments. Our analysis is consistent with that in one respect – it is among the most-curtailed units by volume – but points to a more benign reading of its *per-unit* bidding. As Figure A32 shows, Seagreen bids at a low and remarkably stable level of roughly £15/MWh, below the merchant group's average and far below Moray East; on a per-MWh basis it is among the most restrained units in the sample, not the least. The distinction matters for interpretation. A large total curtailment payment is the product of a large curtailed volume and a per-MWh price, and a unit can accrue substantial totals while bidding at an unremarkable level per MWh, simply because it is large and frequently constrained. The figures that attract public attention are totals; the per-MWh mark-up is the quantity relevant to whether a generator is over-bidding, and on that measure Seagreen does not stand out. We make this point cautiously; it rests on a single unit and an imperfect benchmark, and Seagreen's bid, like all merchant bids, is still a positive wedge over an assumed zero floor. However, it illustrates that volume-driven payment totals and per-unit bidding behavior are distinct and should not be conflated.

G.2 The welfare model: derivation and proof

This appendix gives the full derivation of the MVPF expression (5) and proves Proposition 6.1.1. We retain the notation of Section 6.1: retail price $R = w + N$, DTU procurement price p, baseline consumption Q_0, elasticity $\varepsilon < 0$, support payment σ, cycling and maintenance cost κ, and curtailment compensation ϕ.

Induced demand and inframarginal transfer: With full pass-through the consumer price falls by $w + p$, inducing additional demand

$$\Delta Q(p) = -\varepsilon\, Q_0 \frac{w+p}{R} > 0, \qquad \frac{d\Delta Q}{dp} = -\varepsilon \frac{Q_0}{R} > 0. \tag{8}$$

A marginal increase in p raises the surplus of inframarginal consumers by Q_0 per unit – the full retail-price reduction applied to their existing consumption – and generates a triangle of surplus on the newly induced marginal demand. Of the $(w + p)$ reduction on baseline consumption, the w component is a consumer–generator transfer – the consumer no longer pays the wholesale price on energy whose local opportunity cost is near zero, and the retailer, as a pass-through conduit, remits exactly that w to the generator – so it is net-zero in social welfare and never enters the public sector's ledger. The remaining government-mediated component $Q_0 \cdot p$ is a transfer that enters consumers' willingness to pay and the public sector's cost identically, since the public sector funds exactly that part of the price reduction. This is what gives (5) its two "1" terms at the aggregate level. It is tempting to conclude that this term can simply be dropped when constructing the *marginal* WTP and government-cost expressions below, on the grounds that it cancels out of net benefits – but canceling out of a *difference* (net benefits) and canceling out of a *ratio* (the marginal MVPF plotted in Figure A26) are not the same thing. The level $Q_0(w+p)$ does drop out of WTP – GovtCost identically. Its *derivative* with respect to p, however, is the nonzero constant Q_0, and a constant added identically to the numerator and denominator of a ratio does not cancel; it must therefore be carried through into both marginal expressions below.

Marginal welfare components: The remaining components involve the induced demand $\Delta Q(p)$ from (8) and its derivative $d\Delta Q/dp$, together with the constant Q_0 just described.

Willingness to pay (marginal). The Q_0 units of baseline consumption contribute Q_0 to marginal WTP, from the inframarginal term above. Consumers who were already induced to consume at the prevailing DTU price save one unit of retail expenditure per MWh, contributing $\Delta Q(p)$ to WTP. New consumers induced by the marginal increase in p contribute zero surplus at the margin (envelope theorem). Generators experience a net producer surplus change of $\sigma + \kappa - \phi$ per MWh of DTU relative to curtailment: under curtailment their net revenue is $\phi - \kappa$ (compensation received less cycling cost incurred); under DTU it is $w + \sigma$ (market revenue plus support payment, no cycling cost), of which w is financed by consumers through the retail market and therefore does not constitute a net social gain beyond what is already reflected in consumer welfare.[30] Marginal WTP is therefore

$$\left.\frac{d\,\mathrm{WTP}}{dp}\right|_{\mathrm{marg}} = Q_0 + \Delta Q(p) + (\sigma + \kappa - \phi)\frac{d\Delta Q}{dp}. \tag{9}$$

Government cost (marginal). The Q_0 units of baseline consumption also see their retail price fall by one further penny for each additional penny of p, contributing Q_0 to

[30] Formally, the $w \cdot \Delta Q$ component of generator revenue is offset by an equal reduction in consumer surplus, leaving a net contribution of $(\sigma + \kappa - \phi) \cdot d\Delta Q/dp$ to the social welfare accounts.

marginal government cost symmetrically with WTP above. The DTU payment on existing induced demand contributes $\Delta Q(p)$; the additional DTU payment on newly induced demand contributes $p \cdot d\Delta Q/dp$; the extra support payment on new generation contributes $\sigma \cdot d\Delta Q/dp$; and the avoided curtailment payment contributes $-\phi \cdot d\Delta Q/dp$. Marginal government cost is therefore

$$\left.\frac{d\,\text{GovtCost}}{dp}\right|_{\text{marg}} = Q_0 + \Delta Q(p) + (p + \sigma - \phi)\,\frac{d\Delta Q}{dp}. \tag{10}$$

Proof of Proposition 6.1.1: We show the marginal MVPF equals one if and only if $p = \kappa$, for any $\phi \geq 0$. The marginal MVPF equals one if and only if (9) and (10) are equal. Since $Q_0 + \Delta Q(p) > 0$ appears in both, this reduces to

$$(\sigma + \kappa - \phi)\,\frac{d\Delta Q}{dp} = (p + \sigma - \phi)\,\frac{d\Delta Q}{dp}. \tag{11}$$

Since $d\Delta Q/dp = -\varepsilon\, Q_0/R > 0$, we divide both sides by $d\Delta Q/dp$ to obtain

$$\sigma + \kappa - \phi = p + \sigma - \phi, \tag{12}$$

which simplifies immediately to $p = \kappa$, independently of ϕ, σ, w, Q_0, and all demand parameters.

Monotonicity: Write $\Delta Q(p) = \Delta Q'(w + p)$ with $\Delta Q' \equiv d\Delta Q/dp = -\varepsilon Q_0/R > 0$ constant, and let $k \equiv \Delta Q'/Q_0 = -\varepsilon/R$. Substituting into (9) and (10) and factoring out Q_0,

$$\text{MVPF}^{\text{marg}}(p) = \frac{1 + k(w + p + \sigma + \kappa - \phi)}{1 + k(w + 2p + \sigma - \phi)}, \tag{13}$$

a ratio of affine functions of p, independent of Q_0. Writing $N(p)$ and $D(p)$ for the numerator and denominator, $N'(p) = k$ and $D'(p) = 2k$, so

$$\frac{d\,\text{MVPF}^{\text{marg}}}{dp} = \frac{N'D - ND'}{D^2} = \frac{k\big(D(p) - 2N(p)\big)}{D(p)^2}. \tag{14}$$

A short calculation gives $D(p) - 2N(p) = -\big[1 + k(w + \sigma + 2\kappa - \phi)\big]$ for all p – independent of p itself. Denote this bracketed quantity $M \equiv 1 + k(w + \sigma + 2\kappa - \phi)$; note M is exactly $N(\kappa) = D(\kappa)$, the value both expressions take at the crossing point. For any empirically relevant calibration $M > 0$: at our values ($w = 70, \sigma = 30, \kappa = 3, R = 240, \varepsilon = -0.3$, so $k = 1/800$), $M > 0$ holds for any $\phi < w + \sigma + 2\kappa + 1/k = 906$, comfortably satisfied by both Spain ($\phi = 0$) and Great Britain ($\phi = \mu\sigma + \kappa \approx 38.7$). Given $M > 0$, the derivative is strictly negative wherever $D(p) \neq 0$: $\text{MVPF}^{\text{marg}}(p)$ is strictly decreasing in p over any interval not containing the pole $p^\star = \big(-1/k - w - \sigma + \phi\big)/2$. At our calibration this pole lies at $p^\star \approx -431$ for Great Britain and $p^\star = -450$ for Spain – far outside the range of p plotted in Figure A26 – so the function is strictly decreasing throughout the plotted domain for both countries. Evaluating (13) at $p = \kappa$ gives $\text{MVPF}^{\text{marg}}(\kappa) = 1$ exactly (both N and D equal M there), so the marginal MVPF exceeds one for $p < \kappa$ and falls below one for $p > \kappa$.

G.3 No unraveling of the standard tariff

A natural concern, in the spirit of the adverse selection that Borenstein (2005*b*) identifies under real-time pricing, is that one-way retail price exposure could be destabilizing: flexible customers defect to DTU, the implicit cross-subsidy that sustains the standard variable tariff (SVT) unravels, and the standard tariff rises for those who remain. We do not think DTU has this structure, for three reasons.

First, DTU is not a sorting device that separates the flexible from the inflexible in the way Borenstein's mechanism requires. That mechanism relies on an irreversible tariff choice: customers permanently switch from the SVT onto real-time pricing, mechanically raising the average cost-to-serve of the shrinking SVT pool that remains, which raises the SVT price and induces further defection. DTU has no analogous structure. Standard variable tariff customers remain on the SVT and can themselves participate in any given event; there is no permanent switch, only temporary and reversible participation in individual events. There is consequently no shrinking residual pool whose average cost rises because the flexible customers left it.

Second, the energy-price component of the saving is cost-reflective rather than redistributive. Recall $R = w + N$, where N – the volumetric charge recovering network and policy costs – is unaffected by a DTU event; only the wholesale-energy component w is cut, and cut because the local energy genuinely is worth close to zero when it would otherwise be curtailed. A price reduction that tracks a real fall in marginal cost is not a transfer from non-participants: nobody pays more so that participants pay less on this component, since the forgone wholesale revenue being priced away was never revenue the system could otherwise have captured.

Third, the DTU payment p is not financed at the expense of non-participating consumers. In isolation, p is a transfer from the system operator to participating customers, and one might worry it is drawn from a common pool funded by all billpayers with nothing to show for it elsewhere. Proposition 6.1.1 rules this out at the welfare-optimal price: at $p = \kappa$, the public sector's net outlay on DTU is matched pound-for-pound by an equally sized reduction elsewhere in the ledger – avoided curtailment compensation in Great Britain, or the cycling cost the generator itself no longer bears in Spain. The payment is therefore backed by a real, contemporaneous resource saving generated by the same mechanism that funds it, not drawn down from other consumers' surplus. This is what distinguishes it from a classic cross-subsidy, in which one group's discount is financed by a mark-up on another group with no corresponding saving anywhere in the system.

Taken together: Because no cross-subsidy is being created or withdrawn, the standard tariff need not rise to restore balance, and the unraveling logic does not apply.