AI in Charge: Large-Scale Experimental Evidence on Electric Vehicle Charging Demand*

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

Andrew R. Schein
Centre for Net Zero

Cohen R. Simpson
University College London

Yixin Sun
Centre for Net Zero

January 2026

Abstract

One of the promising opportunities offered by AI to support the decarbonization of electricity grids is to align demand with low-carbon supply. We evaluated the effects of one of the world's largest AI managed EV charging tariffs (a retail electricity pricing plan) using a large-scale natural field experiment. The tariff dynamically controlled vehicle charging to follow real-time wholesale electricity prices and coordinate and optimize charging for the grid and the consumer through AI. We randomized financial incentives to encourage enrollment onto the tariff. Over more than a year, we found that the tariff led to a 42% reduction in household electricity demand during peak hours, with 100% of this demand shifted to lower-cost and lower-carbon-intensity periods. The tariff generated substantial consumer savings, while demonstrating potential to lower producer costs, energy system costs, and carbon emissions through significant load shifting. Overrides of the AI algorithm were low, suggesting that this tariff was likely more efficient than a real-time-pricing tariff without AI, given our theoretical framework. We found similar plug-in and override behavior in several markets, including the UK, US, Germany, and Spain, implying the potential for comparable demand and welfare effects. Our findings highlight the potential for scalable AI managed charging and its substantial welfare gains for the electricity system and society. We also show that experimental estimates differed meaningfully from those obtained via non-randomized difference-in-differences analysis, due to differences in the samples in the two evaluation strategies, although we can reconcile the estimates with observables.

*We thank Lucy Bourke and Joe Grainger, from Centre for Net Zero (CNZ), for their critical support with logistics and background research. We thank Jaye Cribb, from Octopus Energy Limited (OE), for devising the strategy used to detect ownership of an electric vehicle based on electricity usage, and Charlotte Avery, from CNZ, for further refining and implementing that strategy. We thank Francesca Corby, Geraldine De Boisse, and Katrin Weingartner, from OE, for their comments and collaboration regarding study design and implementation, in particular construction of the email-based treatments used for customer experimentation. We thank George Doucy, Ben Ferris, Lilla Hoddosy, Ryan Jenkinson, Will Reeves, Richard Sephton, and Henry Tagg, all from OE, for their guidance around operational aspects of the EV charging algorithms and associated data collection. We thank Andy Hackett, from CNZ, for specific comments on the text of the paper itself. We thank Navya Kumar for excellent research assistance. We thank Omar Asensio, Susanna Berkouwer, Louise Bernard, Daniel Björkegren, David Byrne, Gus Chadney, Sheng Chai, Sylvain Chassang, Daniel Lopez Garcia, Charlie Hutchinson, Alex Imas, Maria Jacob, Seung Min Kim, Justin Kirkpatrick, Alex Schoch, Ben Sprung-Keyser, Phil Steele, Chad Syverson, Arthur van Bentham, Casey Wichman, Izzy Woolgar, Lucy Yu, and seminar participants at Wharton, Manchester Workshop on the Economics of Energy Transition, and Harvard's Workshop on the Economics and Policy of Electric Transportation Charging Infrastructure for their valuable insights and comments. This work would not have been possible without financial support from the Abdul Latif Jameel Poverty Action Lab's (J-PAL) King Climate Action Initiative (K-CAI). Conflicts of Interest: AS and YS are paid employees of Centre for Net Zero (CNZ), a part of the Octopus Energy Group Limited (OEG), RM is a consultant, and CS was a paid employee of CNZ at the time of the trial but is no longer employed by CNZ. The four authors made final decisions regarding experimental design, data analysis, and writing; all mistakes herein are ours alone. AEARCTR-0013037 and USC IRB: UP-23-00733.

1 Introduction

Artificial intelligence (AI) presents both major risks and opportunities for the power grid (International Energy Agency, 2025*a*). Growing research highlights that AI's own electricity needs, driven by expanding data centers, are straining grids worldwide (De Vries, 2023; Aljbour et al., 2024; Bogmans et al., 2025; Chen, 2025), a concern now central to public and policy debates (Erdenesanaa, 2023; Kolbert, 2024). Yet AI also offers tools to improve energy efficiency and flexibility by forecasting, coordinating, and optimizing electricity demand in increasingly dynamic systems (Schweppe et al., 1980, 1981; Antonopoulos et al., 2020; Biswas et al., 2024; Boopathy et al., 2024; Sandalow et al., 2024).

AI has the potential to optimize energy demand management for both residential and industrial customers. In theory, the most efficient approach to energy demand management is real-time pricing (RTP), which directly exposes consumers to the marginal cost of electricity (Nicolson et al., 2018; Borenstein, 2005*b*; Hinchberger et al., 2024). However, a significant body of evidence suggests that consumers are reluctant to engage with day-ahead or real-time prices, due to the cognitive burden and effort required to optimize their usage, along with the price uncertainty they must constantly monitor (Harding and Sexton, 2017; Fabra et al., 2021). In this context, AI managed tariffs offer a compromise: they allow real-time responsiveness while insulating customers from volatile prices and enabling supplier risk management. AI can also coordinate millions of devices to smooth aggregate demand, stabilize wholesale prices, and ensure local grid stability.

Electric vehicles (EVs) exemplify both the challenges and promise of AI in energy systems. As flexible, mobile loads with storage capacity, EVs can absorb or supply power to help balance renewables. Yet optimal managed charging requires solving complex, real-time problems of cost minimization, congestion avoidance, and user preferences, precisely the kinds of tasks AI excels at (Rigas et al., 2014; Kaack et al., 2022; Yaghoubi et al., 2024). With EV sales now exceeding 20% of new cars globally (International Energy Agency, 2025*b*) and projected to drive 15% of future energy demand growth (Agency, 2024), unmanaged charging could worsen peaks and strain infrastructure (Li and Jenn, 2024; Bailey, Brown, Shaffer and Wolak, 2025; Bernard et al., 2025). Managed charging with AI instead aligns demand with low-cost, low-emission periods, reducing system stress.

Despite increasing policy and commercial interest in AI managed charging (Hildermeier et al., 2022), causal evidence on consumer engagement with AI-based tariffs remains scarce (Black et al., 2024). Dozens of managed-charging pilots exist (Lowell et al., 2017; Bradley et al., 2018; Seamonds and Lowell, 2018; NYSERDA, 2021; Anwar et al., 2022; Jones et al., 2022; SEPA, 2024), but most lack credible counterfactuals or sufficient scale.[1] We address this gap with what we believe is the first natural field experiment on the world's largest AI managed EV tariff.

In partnership with Octopus Energy Limited, the UK's largest electricity supplier, we conducted a large-scale randomized encouragement design involving 13,233 suspected EV owners across Great Britain.[2] The intervention promoted the Intelligent Octopus (IO) Go tariff, which combines time-of-use (ToU) pricing with AI managed charging.[3] Customers faced a lower off-peak rate (approx. 60% discount) overnight; if the AI chose to charge during the day, it still billed at that off-peak rate. In return, Octopus controls the charging, although the consumer can override the algorithm.

IO Go is now, to the best of our knowledge, the world's largest managed EV charging tariff, serving approximately 300,000 customers across the UK and expanding rapidly in the US, Germany, France, Italy, and Spain. Currently, the tariff manages roughly 2GW of power in Great Britain (Green, 2025). In our setting, the AI managed charging tariff uses real-time wholesale prices as a key input into scheduling charging, without exposing consumers to these granular prices. It combines linear programming with machine learning models to minimize expected energy costs, reduce grid congestion, and support grid balancing, all subject to customer-set constraints (a ”ready-by" deadline and target state of charge). The AI tariff we studied represents a partial and centrally coordinated implementation of RTP. It preserves key elements of allocative efficiency by shifting load away from high-cost periods but bypasses price-based incentives at the individual level.[4]

The field experiment allowed us to estimate the causal effect of financial incentives

[1]In the United States alone, over $100 million has been allocated across at least ten recent or ongoing managed EV charging pilots (SEPA, 2024), yet none constitutes a proper field experiment with a credible counterfactual. Moreover, none deploy AI-based charging management at meaningful scale, and all suffer from limited sample sizes, severely undermining statistical power and the reliability of their conclusions. These shortcomings underscore the need for rigorous empirical experiments.

[2]Ownership was modeled from load profiles consistent with home chargers.

[3]AI optimization by Kraken Technologies uses real-time data on EVs, local and national loads, prices across many markets, and grid conditions, applying machine learning and linear programming to optimize charging.

[4]While one way to assess the efficiency of this AI managed charging tariff is by estimating its correlation with wholesale prices (Hinchberger et al., 2024), this proves challenging in our case. The AI is performing a complex real-time optimization to maximize overall welfare, not just in wholesale markets, but also in ancillary services, while simultaneously optimizing consumer welfare. This includes aggregating and coordinating user preferences to ensure EVs are charged when needed. Additionally, the sufficient statistics approach to evaluating tariff efficiency overlooks total net benefits, particularly the role of tariff take-up and the own-price and cross-price elasticity (over time) of those who take up the tariff. AI has the promise of making RTP tariffs palatable (i.e., reducing the need for consumer attention and reducing high price uncertainty) in the future for consumers.

on adoption of the AI managed charging tariff, and to use this variation to instrument for the impact of managed charging on electricity consumption. Treated trial participants were emailed and randomly assigned to one of four groups: (1) email only; (2) £5/month; (3) £50/month; or (4) £50/month conditional on not overriding the AI managed charging, with all payments lasting three months. A pure control group received no contact. We pre-specified two primary outcomes: (1) tariff take-up and (2) electricity consumption, with particular focus on consumption during the evening peak demand period from 16:30-20:30, over the 12 months following the encouragement (we followed our pre-analysis plan, unless otherwise stated).

In addition to the experimental analysis, we conducted a supplementary pre-specified difference-in-differences (DiD) analysis using observational data to estimate the impact of managed charging on electricity consumption among customers who voluntarily adopted the tariff outside of the trial. We compared electricity use before and after adoption for these customers, relative to a control group of similar customers who had not yet adopted, exploiting variation in adoption timing. This approach provides complementary evidence on the effects of managed charging in a real-world, self-selected yet scaled-up setting.

The intervention we studied represents a bundled treatment that combines ToU pricing, automated scheduling, and real-time algorithmic optimization. We cannot cleanly separately identify the contribution of each component (although we provide some suggestive evidence of the contribution of the AI algorithm). However, as electricity systems grow more complex, with greater price volatility from renewable generation and weather variability, the proliferation of distributed assets, and increased participation of those assets across multiple markets, purely deterministic or rule-based scheduling may become less effective. In such settings, more advanced AI architectures for coordination and adaptive control may become increasingly valuable (Si et al., 2025; Wang et al., 2025). At the same time, understanding consumer acceptance of AI managed systems remains critical, since there is far less empirical evidence on how households respond to AI-driven automation, as opposed to more familiar ToU or deterministic scheduling systems. We show that customer acceptance, as measured by the frequency of overriding the algorithm, is a sufficient statistic for measuring the welfare of such tariffs.

1.1 Primary findings

We report seven main sets of findings. First, email-based encouragements increased adoption of the AI managed charging tariff. Across all treatment groups, assignment raised enrollment in IO Go relative to the control. Even simple email contact (no incentive) increased take-up by 3.4 percentage points; £50/month incentives roughly doubled that effect. Importantly, we suspect that these take-up rates represented a lower bound, as enrollment was constrained by technical compatibility: trial participants needed a supported charger or EV to join (although most chargers and EVs in the market were eligible). Lacking data on compatibility, we cannot identify the true eligible population. Thus, the measured effects likely understate the potential impact in a fully compatible setting.[5] We find a price elasticity of the take up of the managed tariff to be -0.143 (from a £150 to a £15 subsidy).

Second, we observed strong retention and widespread acceptance of AI managed charging among trial participants. Trial participants who enrolled in IO Go largely remained on the tariff, with post-incentive take-up rates statistically indistinguishable from those during the first three months. Among adopters, we observe high adherence to the AI managed charging schedule: over half never overrode the AI managed charging schedule, on any given day there is a 1% likelihood of overriding, and only 2.3% of total electricity consumption occurred via overrides. These patterns suggest that, once adopted, AI managed charging integrates smoothly into daily routines with minimal disruption, underscoring its potential for long-term grid flexibility. Additionally, these overrides tended to be customer-specific rather than correlated across space or time, indicating that they are unlikely to generate coincident demand spikes that would strain the grid.

Importantly, these override percentages are far below the threshold where RTP would dominate an AI-ToU tariff, according to a theoretical model we developed to quantify the conditions under which AI managed charging (with overrides) has welfare dominance over RTP. This is even true when the AI algorithm only controls 25% of the household's energy demand (roughly the amount of energy from EV charging in our sample). Moreover, the majority of overrides do not occur during the peak time. These empirical and theoretical findings suggest that AI may be the best feasible option in the presence of attention constraints and other real frictions. In our theoretical model comparing welfare under AI-managed and non-managed tariffs, the override rate is a sufficient statistic for

[5] According to informal estimates shared by Octopus Energy staff via personal communication, 60–70% of EV owners were likely compatible with IO Go in 2024, based on available EV and charger integrations.

welfare ordering: when the override rate is sufficiently low, the managed tariff dominates all alternatives, including RTP.

Third, using treatment assignment as an instrument for tariff adoption, we estimated that AI managed charging reduced household peak-period electricity consumption by 42%, with the entirety of that load shifted to overnight off-peak hours. This change in peak demand takes their consumption down to similar levels as non-EV users. There was no overall change in total electricity use, indicating that the program induced temporal load shifting rather than increased consumption. This shifting was similar across all of the randomized encouragement groups (email, low incentive, high incentive, and high incentive plus cost to override). The overall electricity demand through the charger was equivalent to around 7,500 vehicle miles traveled per year, which is in line with the UK average.

Fourth, the AI appeared to enhance responsiveness to wholesale electricity prices beyond the effects of the peak to off-peak shifting. We examined consumption patterns of trial participants who signed up to the AI tariff versus those who signed up to a similar ToU pricing regime but with no AI management of charging (often due to technical compatibility issues that meant they could not sign up to the AI tariff). Comparing the two groups, we found that participants who signed up to IO Go exhibited higher elasticity with respect to wholesale electricity prices during peak evening and off-peak overnight hours, whereas daytime responsiveness is similar across groups. While we cannot interpret these differences causally due to non-random assignment, the results suggest that the AI managed scheduling in IO Go helps shift consumption away from high system prices within periods of the day, i.e., above and beyond between-period shifting that a static ToU tariff is designed to achieve.

Fifth, DiD estimates were smaller than the experimental estimates. We believe this difference was due to differing samples. While both the experimental and DiD approaches reveal similar directional patterns of load shifting, the experimental instrumental variables estimates indicated a 42% reduction in peak-period consumption, compared to just 8% in the DiD estimates. An important contributor to this gap is that many customers in the DiD sample were already on time-of-use tariffs before adopting IO Go, leaving them with less potential for further consumption change. After reweighting the DiD sample to match the experimental group on prior tariff type, the estimated peak reduction rises to 22%: narrowing, but not eliminating, the gap. This weaker peak effect is compensated for by a decline in daytime, non-peak consumption. We interpret these findings as evidence that compositional differences between self-selected IO Go adopters and our RCT sample explain much of the discrepancy, with voluntary adopters exhibiting higher baseline flexibility and thus smaller observable peak reductions. We hypothesize that the intraday differences in impacts reflect greater daytime flexibility among the self-selected participants in the DiD sample from smart technology.

Sixth, we estimated the consumer, producer, grid, and social welfare benefits of this AI managed charging tariff. The reallocation of electricity use to periods with substantially lower rates reduced trial participants' bills by £343 per year, an estimated 18% reduction in cost per kWh. When benchmarked against the retailer's standard flat tariff, the estimated savings rise to roughly £650 per year.[6] The electricity retailer experienced similar savings in procurement costs (which include wholesale power prices and non-energy charges such as transmission and distribution fees), implying near-complete pass-through of savings to participants, at least during our study period. Using 2050 grid projections, we further estimated that if all EV households adopted this tariff, the resulting load shifting would flatten the load-duration curve, lowering peaks and raising off-peak demand, thereby improving efficiency and reducing the need to build at least five to six large nuclear power stations to meet future electricity demand.

The impact on CO2e emissions were substantial.[7] The resource cost per tonne of CO2e abated was extremely low (negative) for consumers. We estimated a resource cost of –£888 per tonne, an order of magnitude lower than that of the next-best technology (Gillingham and Stock, 2018; Gosnell et al., 2020; Hahn et al., 2024).

Lastly, we leveraged the fact that this AI managed tariff has also been deployed in several other markets, including the US, Spain, and Germany, to assess its cross-market efficiency. We used detailed data on charging behavior, household electricity consumption, overrides, and both consumer and market prices. In these markets, we found similar plug-in and override behavior in several other markets, including the UK, US, Germany, and Spain, implying the potential for comparable demand and welfare effects. Given that the override rate is a sufficient statistic for a welfare in our theoretical model, we show the mechanism-level scalability and welfare dominance of this tariff across many countries.

[6] The counterfactual for our £343 estimate is the average bill of the control group during the experimental period, which includes many tariffs. The £650 figure uses our estimated consumption treatment effects relative to the bill under the standard flat tariff, which most customers in our trial sample were on prior to the experiment.

[7] Electricity-related emissions are generally lowest overnight, when wind and nuclear generation dominate, and highest during the evening peak, when falling solar output and rising demand bring more gas generation online.

1.2 Contribution to the existing literature

Our study contributes to three strands of research: tariff switching and dynamic pricing, electric-vehicle (EV) charging, and the economics of AI.

Tariff switching and dynamic pricing. Closest to our work are randomized experiments by Fowlie et al. (2021) and Ito et al. (2023), who studied consumer adoption and demand under time-varying tariffs. They showed that incentives increased switching, that adoption correlated with price elasticity, and that opt-in vs. opt-out designs affected persistence. Yet both examined ex-ante fixed rates, without uncertainty or automation, and thus could not address the challenges of RTP or AI managed tariffs that optimize around spot prices (with the exception of critical peak pricing). In our setting, once consumers enrolled and revealed their state of charge and ready by preferences, all intraday adjustments were automated, minimizing behavioral effort and thus selection on their elasticity, which is why our study differs from Einav et al. (2013) and Ito et al. (2023).[8]

Three related other studies speak to automation in similar contexts. Blonz et al. (2025) examined the impact of automated temperature set-points through a smart thermostat in an opt-in field experiment with ToU pricing. While their system had no AI component (the automation was deterministic, and did not update over time or use machine learning), they convincingly demonstrated the benefits of the automation (raising temperature set-points in the summer during the peak price part of the day) and low override rates. Our findings of "set and forget" (via automation) having similarly low overrides support their results. Bailey et al. (2024) and Khanna et al. (2024) studied utility-managed demand response in opt-in experiments, again without AI. We extended these important contributions by leveraging a natural field experiment with AI deployed in a scaled-up market product.

Beyond these related studies, a broader literature on tariff switching and design documents significant inertia and inattention in energy markets, often attributed to high switching costs (Hortaçsu et al., 2017; Byrne et al., 2022; Gravert, 2024; Garcia-Osipenko et al., 2025). Our study is the first, to the best of our knowledge, to estimate switching costs in a natural field experiment by randomizing financial incentives for tariff switching. We found a price elasticity of switching to an AI tariff to be -0.143, suggesting the importance of switching costs.[9]

[8] Residual factors such as control loss or privacy concerns may still influence adoption (Bailey and Axsen, 2015; Moser, 2017).

[9] Prior work in health, telecommunications, and energy has typically estimated switching elasticities using structural

We also developed a theoretical model extending Joskow and Tirole (2006*b*), who analyzed optimal tariff menus under competition. We added an AI managed option with partial non-compliance ("overrides") and derived a closed-form crossover rate at which automation yields higher welfare than RTP even with small transaction costs. This embeds behavioral frictions — attention costs, uncertainty, limited compliance — into classical welfare rankings, complementing Borenstein (2005*b*), who assumed full responsiveness. Our framework linked automation quality and override behavior to welfare outcomes, bridging idealized RTP theory with friction-aware retail models.[10]

EV charging. Our work also extends empirical studies on managed and flexible EV charging (Burkhardt et al., 2023; Garg et al., 2024; La Nauze et al., 2024; Bailey, Brown, Shaffer and Wolak, 2025; Bernard et al., 2025). Prior pilots examined manual or deterministic charging responses, often in small or selected samples. Bailey, Brown, Myers, Shaffer and Wolak (2025) compared managed vs. ToU charging in a framed experiment, and Burkhardt et al. (2023) tested off-peak price incentives, both without AI and at limited scale. We studied the world's largest AI managed EV tariff within a natural field experiment, embedding automation, randomization, and market realism. This enabled the first causal evidence on large-scale AI coordination of household electricity demand.

Our findings also inform the broader EV policy literature on infrastructure, adoption, and system costs (Zhou and Li, 2018; Archsmith et al., 2022; Powell et al., 2022; Rapson and Muehlegger, 2023*b*,*a*; Heid et al., 2024; Turk et al., 2024; Asensio et al., 2025; Dorsey et al., 2025; Gillingham et al., 2025). By showing that AI scheduling lowered consumer bills, reduced grid system peaks, and shifted electricity load to low-emission periods, we demonstrated a key pathway through which managed charging can cut grid investment needs and raise welfare.

Mechanism design and AI. Finally, the paper contributes to the mechanism design literature by showing how a simple, algorithmic mechanism can implement efficient coordination with minimal information requirements. Under the Intelligent Octopus tariff, users reveal only two primitives — their desired state of charge and ready-by time — while the algorithm optimizes charging schedules subject to these constraints. This restricted-revelation structure enables efficient load shifting without the need for complex preference elicitation or strategic communication. In fact, while there is much het-

models or by experimentally varying the information provided about costs and benefits across options.

[10]See also Holland and Mansur (2006); Poletti and Wright (2020); Imelda et al. (2024) for quantitative simulations of RTP gains and Borenstein (2007); Allcott (2011); Fabra et al. (2021); Pébereau and Remmy (2023) for evidence of adoption barriers.

erogeneity in those two primitives from the samples, the effect of the tariff is similar across all types.

Conceptually, the design parallels the emphasis by Budish (2011) on simple, approximately efficient mechanisms and extends recent work in dynamic and algorithmic mechanism design (Athey and Segal, 2013; Parkes and Wellman, 2015; Bergemann and Välimäki, 2019). In doing so, the field experiment demonstrates how AI managed tariffs can serve as practical, data-driven mechanisms that align private incentives with system-level efficiency. The AI optimization substitutes for strategic communication: incentive alignment is embedded in the system's architecture, not in agents' reasoning. The experiment thus offers large-scale evidence that automated, data-driven mechanisms can deliver efficiency and incentive compatibility through simple user inputs. These results support speculation that AI agents will better satisfy the rationality assumptions underlying mechanism design frameworks than human agents (Varian, 1995; Bergemann and Välimäki, 2019).

We also contribute to the emerging literature on AI's causal effects on economic outcomes (Agarwal et al., 2023; Brynjolfsson et al., 2025; Björkegren et al., 2025).[11] Although some have highlighted the challenges of identifying AI's effects in labor markets (Brynjolfsson et al., 2019; Frank et al., 2019), these identification barriers are less pronounced in energy and environmental contexts. AI often lowers the cost of effort, making it difficult to disentangle AI from price effects and isolate its unique impact. This matters because reducing effort, through tools, automation, or process improvements, has long influenced labor markets, yet existing studies rarely separate this effect from AI itself. In our setting, the reduction in effort occurs on the retailer side in responding to wholesale prices or ancillary markets, while consumer prices and effort remain unaffected by the AI's optimization.

The paper is structured as follows: Section 2 provides an overview of the natural field experiment, including the sampling, the intervention (i.e., the AI-EV tariff), the randomized encouragement, the randomization procedure, and available data. Section 3 develops the demand estimates from the field experiment. Section 4 complements Section 3 with a difference-in-difference analysis of the intervention, and attempts to reconcile with the experimental estimates. Section 5 presents the welfare estimates of the intervention and the randomized encouragement, Section 6 presents descriptive statistics of customer behavior on an AI managed tariff in three other countries, and finally, Section 7 concludes.

[11] This complements work showing that AI data centers could impact on energy flexibility and thus emissions (Knittel et al., 2025).

2 Experimental design

This section describes the design and implementation of our natural field experiment. We begin by detailing the eligibility criteria and sampling strategy used to construct the study population. We then describe the AI managed charging tariff and its underlying automation, before outlining the randomized encouragement design used to induce uptake. We conclude by summarizing our randomization procedure and the data used in the analysis.

2.1 Eligibility into the field experiment

We implemented the experiment in partnership with Octopus Energy, the United Kingdom's largest electricity supplier. We defined the target population as residential customers satisfying three criteria: (i) they were likely to own an EV; (ii) as of December 2023, they had only ever subscribed to conventional flat-rate or variable-rate tariffs – that is, they had no prior engagement with managed tariffs; and (iii) they resided in houses (rather than apartments/flats or mobile homes), which are likely suitable for at-home charging.

As we lacked administrative records on EV ownership, we inferred it using household electricity consumption data from smart meters. Following internal technical guidance, we defined "suspected charging events" as four to twelve consecutive half-hourly intervals with half-hourly total usage exceeding 3.5 kWh, consistent with Level 2 (7 kW) home charging. We averaged the number of such events per week over a ten-week window in summer 2023 (July–August),[12] and classified as "suspected EV owners" those customers with between 0.5 and 4 events per week (these thresholds were chosen in consultation with Octopus Energy colleagues with subject matter expertise on typical EV charging frequency among customers).[13]

In summary, our sampling strategy yielded 13,233 trial participants. In this sample, 86% were on flat-price tariffs. Given that most customers in the UK are on flat tariffs,

[12]We used the summer for this approach as there was less likely to be heat pumps or any other electric heating blocking the EV charging signal.

[13]We are unable to provide formal performance statistics for the EV detection algorithm because we do not have ground-truth data on customers' low-carbon technology ownership. However, after the first round of encouragements in February 2024, we surveyed 305 customers who received an encouragement but had not signed up for Intelligent Octopus Go. We received 73 responses. Of these respondents, only four reported not owning an EV (5%). This high proportion of true EV owners among those we predicted to have an EV suggests that the detection algorithm likely had a low Type I error rate (i.e., few false positives).

our trial population is likely to resemble the broader set of prospective adopters of managed charging technologies along this dimension. This contrasts with early adopters of IO Go, who were typically more engaged, more technologically inclined, and potentially more environmentally motivated than the mainstream EV-owning population.[14] In other words, unlike early adopters already enrolled in smart tariffs, we believe our trial participants were more similar to the broader population of potential adopters of managed charging technologies.[15]

2.2 The intervention: Intelligent Octopus (IO) Go

IO Go is a residential electricity tariff that combines time-of-use pricing with AI management of automated EV charging. Through the Octopus app, customers set a target battery level (i.e., state of charge) and departure time (i.e., ready by time); Octopus then schedules charging to meet those targets (Figure 1a). In exchange, IO Go provides a favorable electricity rate of £0.07/kWh during a fixed six-hour off-peak window (23:30–05:30).[16] This rate applies not only to EV charging but to all household consumption during that period. Charging that occurs outside the scheduled off-peak window is still billed at the off-peak rate if it is initiated by the AI automation. Customers retain the ability to manually override this schedule via a mobile app ("bump charging", shown in Figure 1b, or sustained override shown in Figure 1c). The overriding is easy and costless – customers open the app, go to the tab in Figure 1a, and press the "Bump Charge" button. When customers charge outside of the schedule through overriding, it is billed at the higher rate so there is a financial cost to overriding. The applicable tariff schedule for 2024 is shown in Figure 2.

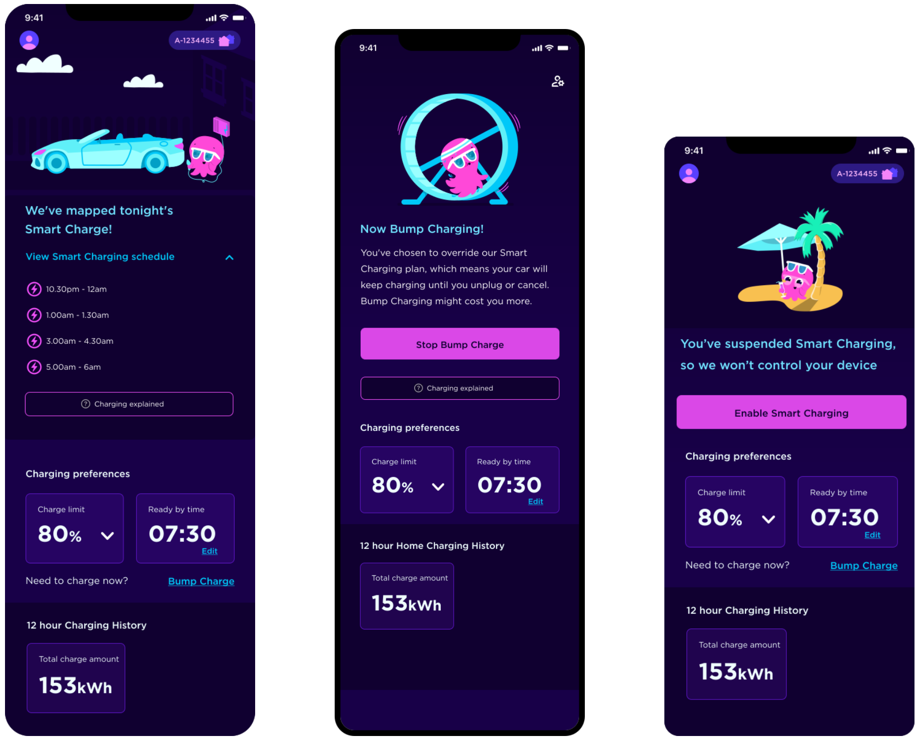
Beyond its engineering implementation, the tariff design embodies a core principle from mechanism design. It functions as a large-scale restricted-revelation mechanism. Each customer specifies two primitives — a ready-by time and a desired state of charge — which together summarize individual flexibility (Figure 1a). With these constraints, the algorithm allocates charging across time to minimize system costs while meeting user needs. No detailed reporting of preferences, utilities, or price sensitivities is required. In mechanism-design terms, it is a simple direct mechanism defined over a coarse message

[14] According to a recent report, 91% of customers are on flat-rate tariffs in the UK.

[15] Given that they had adopted their EV for a few months before we contacted them, we expect response rates would be slightly lower than if we contacted them immediately after buying their EV (Nicolson et al., 2017). However, that would raise difficult identification issues since we would not have a clear baseline time period.

[16] Octopus Energy's standard tariff Flexible Octopus charges a flat rate of approximately £0.22–£0.27 per kWh throughout the day, depending on region, as shown in Figure 2.

Figure 1: IO Go Charging Controls via Mobile App



(a) Daily Charging Schedule | (b) Temporary Override | (c) Sustained Override

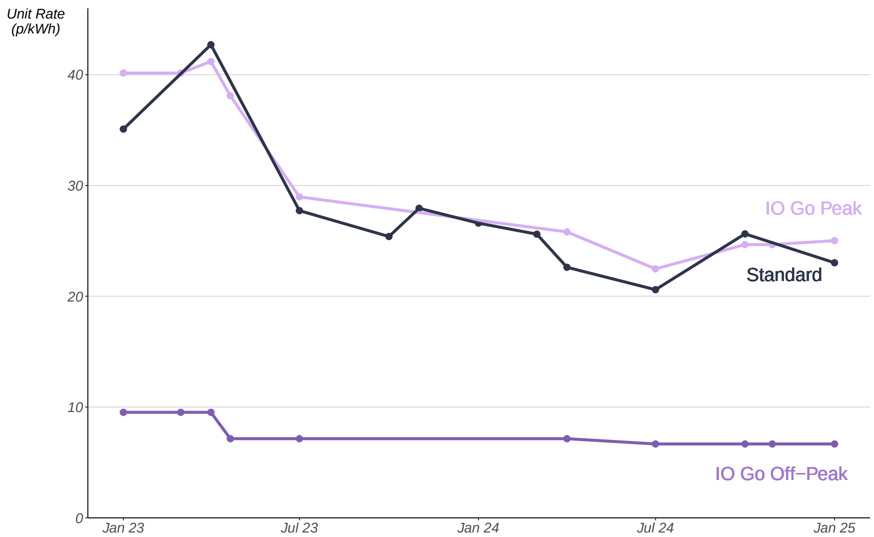
Notes: Example screens from the Intelligent Octopus Go mobile application. Panel (a) shows the AI-generated charging schedule, while panels (b) and (c) depict temporary and sustained user overrides of managed charging.

space, yet it achieves near-efficient outcomes because the AI optimizes using observed demand and wholesale price data. The design echoes the notion of approximate competitive equilibrium from Budish (2011).

The IO Go tariff uses AI to generate charging schedules based on prices in many markets. The first is wholesale electricity prices, adapted to the market structure of each region where it is offered. In Great Britain, where wholesale prices are national, schedules are optimized using a blend of day-ahead and intraday wholesale prices, with intraday data available on a rolling-hourly basis. In most European markets, the optimization is updated daily based on day-ahead wholesale prices, while in the United States, the algorithm relies on forecasts of real-time prices. The optimization window can extend up to 24 hours into the future, accommodating consumer-specific “ready-by” times. These may require intraday charging plans – such as afternoon readiness – or more commonly, overnight charging plans, particularly in Great Britain.

Following the baseline price-driven optimization, the system supports a secondary ad-

Figure 2: Tariff Rates



Notes: This figure shows the tariff rates for Intelligent Octopus Go customers during the off-peak overnight period (23:30–05:30, dark purple) and the peak daytime period (05:30–23:30, light purple). For comparison, we also include the standard non-time-varying tariff (black) from Octopus Energy, the main counterfactual tariff.

justment layer for participation in ancillary service markets and real-time grid-balancing operations. In this mode, the EV portfolio is managed as a coordinated fleet with portfolio-level volume targets, allowing for dynamic adjustment of charging schedules in response to market trades or direct requests from system operators. This enables dual operational modes: price-only optimization, in which vehicles charge during the lowest-cost periods (often aligning with high renewable output), and price-plus-volume optimization, in which charging is redistributed to meet specific aggregate energy delivery constraints while maintaining cost efficiency.

The optimization engine integrates two complementary computational approaches: classical linear programming and machine learning–based forecasting (the AI currently uses no large language models to predict or optimize). The linear programming component solves for the cost-minimizing charging schedule subject to technical and operational constraints, such as maximum charging rates, user readiness requirements, and fleet-wide volume limits. The machine learning component enhances this process by providing predictive inputs to the optimization, including forecasts of EV availability (based on historical connection patterns), expected charging requirements for vehicles not yet connected, and local load predictions, including home consumption and on-site solar generation when applicable. These forecasts are updated in real time, enabling the optimization to adapt to changing conditions. A diagram outlining the structure of the AI scheduling framework is shown in Figure 3.

From an efficiency point of view, a key advantage of IO Go's automation is to mitigate charging demand inefficiencies at the consumer level due to human error, bounded rationality, or rational inattention. In environments where consumers are either rationally inattentive or lack the sophistication to respond optimally to high-frequency price signals, automated optimization under IO Go may yield strictly higher social welfare (from a net benefits or resource cost-effectiveness point of view) than consumer-managed RTP (see Appendix G for our predictions).

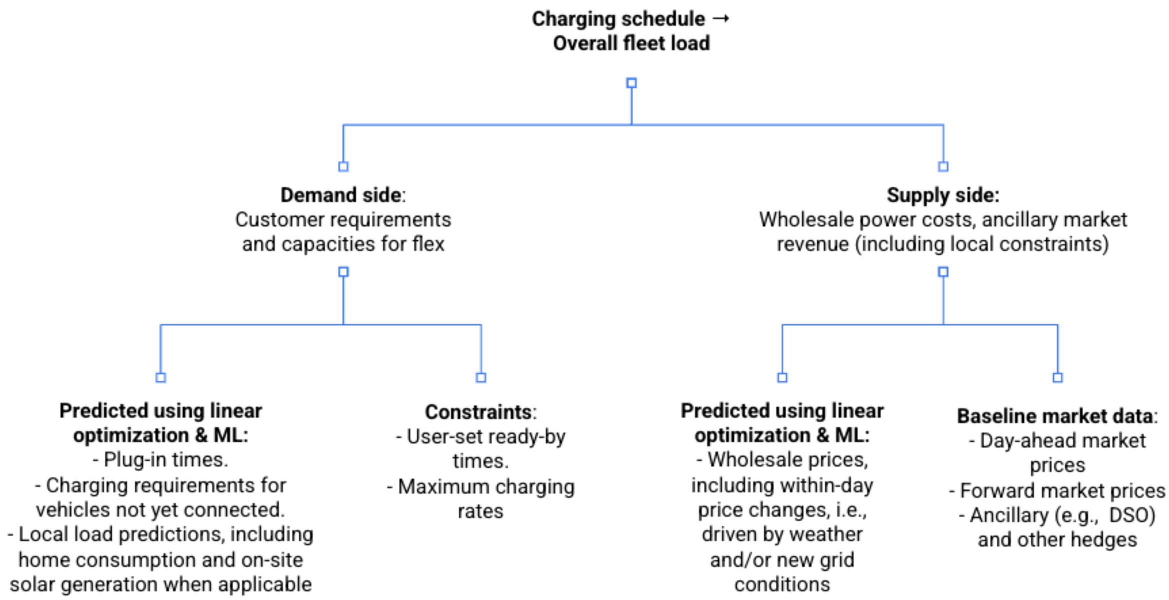
In addition, the British electricity system — like many others — includes multiple ancillary and flexibility markets, with overlapping mechanisms for valuing flexibility. Market participants respond not only to wholesale price signals but also to a range of network, balancing, and capacity markets, including the Capacity Market, Balancing Mechanism, and various frequency response and distribution network services. This complexity means that optimizing against a single real-time or day-ahead wholesale price would fail to capture value available in other markets. IO Go's automation enables the retailer to allocate charging efficiently across these interacting price signals.

2.3 The randomized encouragement design

The trial participants in our sample were randomized into one of five arms:

1. Control group (n = 2,205): no outreach
2. Email (n = 7,720): encouragement email with no financial incentives for signing up to IO Go
3. Email + £5/month incentive to sign up (n = 1,101): encouragement email with offer of £5/month for three months for signing up to IO Go
4. Email + £50/month incentive to sign up (n = 1,102): encouragement email with offer of £50/month for three months for signing up to IO Go
5. Email + £50/month incentive to sign up (n = 1,105), no bump charging (overriding): encouragement email with offer of £50/month for three months for signing up to IO Go. Additionally, they pay £2 of their incentive for each day they "bump charged" – that is, overrode the AI control at least once per day. This was designed to make

Figure 3: AI Managed Scheduling Framework



Notes: The optimization is implemented by Kraken Technologies, which ingests both demand-side inputs (customer requirements, plug-in behavior, and local load predictions) and supply-side inputs provided by Octopus Energy. The supply side inputs can be complex. Distribution Network Operators procure local flexibility to manage constraints and reliability. Sustain services address long-term capacity needs, Secure services manage predictable network peaks, Dynamic services respond to real-time constraints, and Restore services support recovery after faults. At the national level, NESO procures ancillary and balancing services to maintain system stability. Frequency response maintains 50 Hz operation through Dynamic Containment, Dynamic Moderation, and Dynamic Regulation. Reserve services provide backup capacity, split into Quick and Slow Reserve, while the Balancing Mechanism matches supply and demand in real time. Technical services include Black Start, inertia, reactive power, short-circuit level, and constraint management. The Capacity Market provides availability payments for system stress events via T-1 (one year ahead) and T-4 (four years ahead) auctions. Wholesale markets provide energy price signals across various timeframes. Day-ahead markets allow trading for the following day. Intraday markets enable near–real-time adjustments, while imbalance pricing penalizes deviations from contracted positions. Peak-related levies – including the Capacity Market levy, Distribution Use of System (DUoS), and Transmission Network Use of System (TNUoS) charges – further incentivize load shifting and peak reduction. Not all of these markets and signals are directly acted on; they are noted to illustrate the range of supply-side inputs considered in optimizing demand. Octopus Energy is responsible for forecasting wholesale and ancillary market prices and other market conditions, while Kraken integrates these inputs to generate charging schedules subject to user constraints and technical limits.

overriding even more costly and thus understand trial participants' willingness to tolerate even less control over their charging schedule.

Each treatment arm received a single encouragement email from Octopus Energy's marketing team, reproduced in Appendix D.1. These messages included a prominent call to action, a theoretical £700/year savings estimate (based on historical usage modeling), and tariff-specific details, with minor variations in content to reflect the assigned incentive level. The "no bump" group was explicitly informed that their monthly payment would be reduced by £2 for each day they initiated a manual override of managed charging.

Encouragement emails were dispatched in two waves: the first on February 15, 2024, and the second on March 20, 2024. Incentive payments were credited to trial participants' Octopus Energy account balances, accruing daily over a 90-day period conditional on maintaining an active IO Go contract. In effect, this structure functioned as a discount on the customer's electricity bill, lowering the effective cost of household energy use during the incentive window. This setup was very natural and in-line with how the retailer administers incentives in other programs.

Note that it was required by our implementing partner, Octopus Energy, that our encouragement emails for IO Go inform customers that Octopus Energy offers other tariffs that may have better met their needs. Importantly, IO Go is not compatible with all chargers or vehicles; for these circumstances, Octopus Energy offers an alternative EV tariff: Octopus Go, a time-of-use EV tariff offering a fixed off-peak rate for electricity. The major differences from Intelligent Octopus Go are: (1) it offers one fewer hour of cheap overnight rate (00:30-5:30, instead of IO Go which is 23:30-5:30) (2) its off-peak rate is higher than IO Go's (£0.085/kWh, as compared to £0.07/kWh for IO Go; for exact rates, see Figure A1b), and (3) Octopus Go does not incorporate AI managed charging and thus has no bonus off-peak windows. This creates the possibility that encouragements influenced uptake of tariffs other than IO Go, posing a potential channel for exclusion restriction violations. We test and discuss this in Section 3.1.

2.3.1 Hypotheses

The design of the field experiment allowed us to test two main pre-specified hypotheses:

1. Increasing the incentive payment for adopting IO Go will increase actual adoption.

2. Adoption of IO Go tariff will shift electricity consumption from peak (16:30–20:30) to off-peak (23:30-5:30) hours.

These two hypotheses are not from the same family and thus we did not correct for multiple hypothesis testing across them. However, we will go into some of the testable predictions from the design below.

2.4 Tying the Experiment to a Theoretical Framework

From a welfare analysis perspective, we expect the IO Go tariff to generate distinct welfare outcomes relative to real-time pricing (RTP), depending on the scope for optimization and the behavioral frictions consumers face. IO Go can outperform RTP when it leverages price signals from multiple markets, such as ancillary service opportunities that are typically excluded from the RTP signal.[17] Our theoretical model focuses on behavioral frictions rather than multi-market optimization; however, we acknowledge that both are important drivers of the extent to which automation might or might not outperform RTP.

We consider EV charging under four tariff regimes: flat (the baseline counterfactual our sample begins on), time-of-use (ToU) without automation, RTP without automation, and the AI-assisted tariff "IO Go" with possible household overrides ("bump charging"). Households require a given amount of energy by a deadline, have specific plug-in hours, and face inconvenience costs when charging at less-preferred times. The retailer's total cost depends on total load and wholesale prices, net of ancillary-service revenues. Under IO Go, charging is centrally scheduled to minimize these costs; under RTP, households self-schedule but face *attention costs* and *bill-volatility risk*. ToU households respond only to fixed peak/off-peak prices.

We present the full model and formal predictions in Appendix G, which builds upon Borenstein (2005*a*,*b*); Joskow and Tirole (2006*a*,*b*). Here, we summarize the key intuition: automation under IO Go will generally increase welfare relative to RTP when users face attention or risk costs. However, under IO Go, households may *override* the optimized schedule when their immediate value from charging exceeds the deferred benefit net of

[17] In principle, RTP could incorporate such signals, but doing so would require a richer and less transparent price vector, as well as greater computational and behavioral demands on consumers. IO Go may instead yield lower welfare than RTP when intertemporal optimization across days offers larger gains than within-day shifting, or if users systematically fail to plug in their EVs when charging would be most cost-effective. However, evidence from RTP studies suggests that cross-day shifting is rare, and our data show no systematic plug-in failures.

a hassle cost. Such overrides can increase user utility but raise system costs if they shift load into expensive or high-emission hours, thereby narrowing the welfare gap between IO and RTP.

Let W^{IO} and W^{RTP} denote per-EV welfare without overrides, and define the baseline welfare gap:

$$\Delta W_0 = W^{\mathrm{IO}} - W^{\mathrm{RTP}}. \tag{1}$$

If overrides occur at rate $\bar{\beta}$ with per-override welfare loss λ, total IO welfare becomes:

$$W^{\mathrm{IO+O}} = W^{\mathrm{IO}} - \bar{\beta}\,\lambda. \tag{2}$$

We approximate λ empirically as:

$$\hat{\lambda} \approx p_{\mathrm{peak}} \cdot \Delta\tilde{c}_{\mathrm{eff}} \cdot q^{O}, \tag{3}$$

where p_{peak} is the probability that an override lands in a peak period, $\Delta\tilde{c}_{\mathrm{eff}}$ is the peak–off-peak price spread (£/kWh), and q^{O} is the mean energy shifted per override.

Lemma. If $\Delta W_0 > 0$ and $\lambda > 0$, the override rate at which IO welfare equals RTP welfare is:

$$\bar{\beta}^{\star} = \frac{\Delta W_0}{\lambda}. \tag{4}$$

This crossover override rate $\beta^{\star}$ is a sufficient statistic for understanding the relative efficiency of IO Go and RTP. In a stylized UK calibration with RTP price elasticity of demand $\varepsilon = -0.2$ (consistent with or above values in the literature), and that EV demand represents roughly 25% of total electricity load, we obtain $\bar{\beta}^{\star} \approx 0.33$ overrides per day at moderate attention costs. Higher attention costs or lower RTP elasticities increase this threshold. Appendix G.1 provides the full calibration for the UK and the other IO Go markets.

Our model builds on strands of the energy economics and operations literature that model demand under dynamic electricity pricing. Threshold-style decision rules, in which users act only when net private benefits exceed a frictional cost, appear in online EV charging optimization under RTP (Yi et al., 2019), where a dissatisfaction penalty plays a role similar to our hassle cost ϕ_i. Menu-based contract designs for EV charging (Ghosh and Aggarwal, 2017) and joint welfare-maximizing scheduling algorithms (Huang et al., 2023) adopt multi-stage decision structures that parallel our two-stage extension, in which households first adopt a tariff and then choose whether to override automation.

Our framework also relates to the generalized Roy model of Ito et al. (2023), who study voluntary take-up of dynamic pricing plans using marginal treatment effects to characterize heterogeneity in welfare gains from adoption. In their setting, the key policy-relevant object is the optimal *adoption cutoff*; in ours, it is the *crossover override rate* at which an AI managed schedule ceases to dominate RTP in welfare. Both identify a threshold along a behavioral margin — adoption or overrides — at which the welfare ranking of competing regimes changes.[18]

Relative to this literature, our contribution is to embed such a behavioral threshold in an AI optimization framework for EV charging, explicitly linking override behavior to aggregate welfare outcomes and producing sharp, testable predictions for our experimental design. Based on our parameterization – with a price elasticity of demand under RTP of $\varepsilon = -0.2$ and moderate attention costs – our simulations imply that an average override rate below $\bar{\beta}^{\star} \approx 0.33$ per day would make the AI tariff welfare-dominant relative to RTP.

2.5 Randomization

Random assignment was implemented using a block randomization procedure to improve covariate balance across trial arms (Moore and Schnakenberg, 2023). Prior to assignment, trial participants were grouped into 1,109 blocks of twelve customers based on Mahalanobis distance calculated over a set of pre-encouragement variables predictive of EV ownership and electricity consumption. These included historical electricity usage, tenure as a customer with Octopus Energy, and past engagement with smart tariff onboarding. Within each block, two trial participants were assigned to the pure control group, seven to the £0/month encouragement group, and one each to the £5/month, £50/month, and £50/month (no bump) groups. This distribution across treatment arms reflected a balance between budget constraints and the desire to offer realistic incentive

[18] Joskow and Tirole (2006*b*) present a general model of retail electricity competition in which tariff menus are designed to elicit efficient real-time demand response, subject to transaction and metering frictions. In their framework, sufficiently low frictions imply that high-granularity pricing schemes such as RTP dominate coarser alternatives. Borenstein (2005*b*) similarly show that RTP improves allocative and investment efficiency relative to flat or ToU rates in competitive markets, assuming full compliance and ignoring behavioral frictions. Our model extends these frameworks in two key ways: (i) we introduce an AI managed regime (IO) that automates price response, removing household attention costs but allowing partial non-compliance through overrides; and (ii) we model overrides as a distinct behavioral margin with its own welfare implications. This richer friction structure implies that IO Go can outperform RTP even when RTP's transaction costs are small, provided override rates remain below the crossover threshold $\bar{\beta}^{\star}$, and conversely that high override rates can reverse the ranking — an effect absent from the original Joskow–Tirole and Borenstein formulations.

levels, on the one hand, with the goal of maximizing power (by increasing sign-up to IO Go), on the other. Randomization was performed separately within each of the United Kingdom's 14 electricity distribution regions to ensure geographic stratification. This procedure yielded excellent covariate balance across trial arms (Table A1, and Figure A2).

In our pre-analysis plan, we calculated the minimum detectable LATE when using the encouragements as instruments for take-up of IO Go, and performed power calculations separately for each encouragement-specific instrument. We found that to achieve 80% power required detecting a 30–40% reduction in peak-period consumption.

We used the randomized encouragement to estimate the intent-to-treat effect (ITT) of receiving the encouragement on take-up of AI managed charging, and then used the encouragement as an instrument to identify the local average treatment effect (LATE) of AI managed charging on outcomes. To identify the ITT effect, we satisfied the three standard requirements for a valid randomized assignment mechanism: probabilistic assignment, individualistic assignment, and unconfoundedness.

First, every participant had a positive probability of being assigned to any of the encouragement groups or control (probabilistic assignment). Second, assignment depended only on pre-encouragement covariates and is therefore independent of other participants' potential outcomes (individualism). Finally, encouragement assignment was independent of potential outcomes (unconfoundedness).

In addition to having used an assignment mechanism that ensures unbiased estimation, the study benefited from high observability. We tracked each participant's electricity consumption before, during, and after tariff adoption. Data loss only occurred when customers left Octopus or stopped using smart meters. Twelve months after encouragement, 97% of participants still provided usable smart-meter readings, with attrition rates balanced between the treatment and control groups. Finally, there was no selection into the sample and no notification of randomization or an experiment to the sample, limiting any Hawthorne or John Henry effects. Thus, we can call our experiment a natural field experiment (Harrison and List, 2004).

2.6 Data

Our analysis drew on high-frequency administrative data from Octopus Energy, Great Britain's largest electricity supplier, with personal identifiers removed. We focused on two primary outcomes for our two hypotheses: (1) take-up of the IO Go tariff, measured as a binary indicator for whether a trial participant held an active IO Go contract during week t; and (2) electricity demand, observed at half-hourly resolution and expressed in kilowatt-hours (kWh). We aggregated consumption to the week × hour-of-day level. Electricity demand includes both EV-related and non-EV household load; concretely, we summed smart-meter readings (which measure consumption from all appliances, not just the EV charger) across all meter point administration numbers linked to each trial participant account. These outcomes are observed from January 1, 2024 through March 31, 2025.[19] All of these decisions were in our pre-analysis plan (AEARCTR-0013037).

In Figure 4, we show the average hourly electricity consumption in the pre-encouragement period (i.e., January 2024), showing an expected evening peak beginning around 16:30, coinciding with the system peak in Great Britain when wholesale prices, network congestion, and emissions intensity are typically highest. Baseline electricity consumption for our trial participants was much higher than that of a random sample of 5,000 other Octopus Energy customers with smart meters, but without EVs. Notably, we also observe substantial overnight consumption, consistent with EV charging demand. We believe this is partly driven by plugging the EV after work in the evening and the default scheduling behavior in common EV chargers, which sometimes pre-set charging to off-peak times.[20] Our intervention thus tests the additional impact of managed charging in a setting where at least some users are already defaulted into off-peak charging schedules.

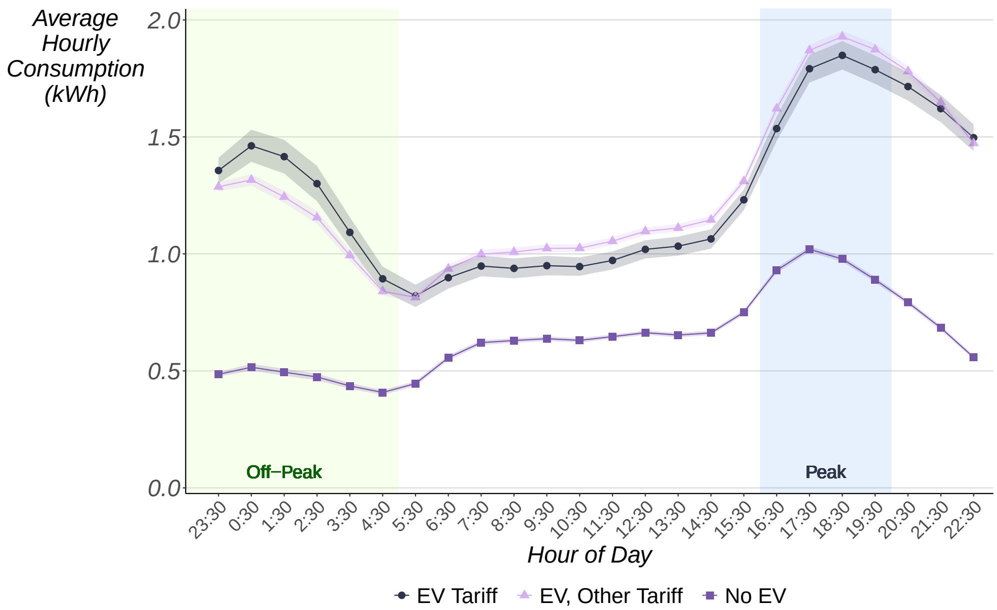
We focus our analysis on off-peak periods (23:30-05:30) and peak periods (16:30-20:30), following our pre-registration. Off-peak periods correspond to IO Go's hours of cheap overnight rates. Peak hours capture the period of highest intensive domestic electricity consumption (Few et al., 2022). That said, it is worth noting that the definition of "peak" and "off-peak" may change and themselves become more variable by day and season in the coming years.

For experiment participants who adopted the IO Go tariff, we also collected customer settings and high-frequency telemetry on charging behavior. The settings include: (1) the ready-by time, defined as the user-specified time by which the vehicle should be fully or partially charged; and (2) the desired state of charge by that time. The telemetry data include: (1) plug-in and unplug timestamps; (2) charging start and charging end

[19] A meter point administration number (MPAN) is a unique ID for an electricity supply point (e.g, a house) in relation to a specific area of the UK's national electricity grid. Tariff-contracts and half-hourly measurements of electricity use are tied to account identifiers via MPANs. A single account can have multiple MPANs with different tariff agreements that are simultaneously active.

[20] Since June 2022, the UK's *Electric Vehicles (Smart Charge Points) Regulations 2021* have required that all new private EV chargers include a default charging mode set outside of peak hours (8–11am and 4–10pm), along with a randomized delay function to reduce grid strain.

Figure 4: Pre-Trial (January 2024) Hourly Consumption



Notes: This figure shows average hourly electricity consumption across the sample in January 2024, prior to the start of the trial. Trial participants are grouped by (1) those who later spent more than 50% of the trial period on Intelligent Octopus Go or Octopus Go (2) those who remained on other tariffs, and (3) a random 20,000 sample of other Octopus Energy customers. Shaded regions show the 95% confidence interval. Octopus Go is an alternative time-of-use EV tariff designed for customers who either had hardware that was incompatible with IO Go or did not wish to enroll in automation. Green shaded box indicates IO Go off-peak hours (23:30–05:30), when electricity is charged at £0.07/kWh; all other hours are billed at the standard variable rate. Blue shaded box indicates typical system peak hours (16:30–20:30), which are highlighted to show times of heightened grid stress.

timestamps; and (3) an indicator for whether the charge was automatically dispatched by Octopus or manually overridden by the user. These data allowed us to reconstruct intended charging preferences, actual charging demand, and deviations from automated control.

We used additional administrative data at the daily level to estimate consumer and supplier benefits. First, to measure benefits to consumers, we used administrative data from Octopus Energy detailing each customer's unit rate per kWh. Second, to estimate benefits to the electricity supplier (in procurement cost savings), we used administrative data on each customer's wholesale energy costs, as well as non-energy costs, per settlement period. Non-energy costs include Transmission Network Use of System (TNUoS) and Distribution Use of System (DUoS) charges, capacity market payments, and policy costs (such as charges supporting Contracts for Difference); some of these charges vary by period of day (e.g., DUoS). This yields an imputed per-kWh cost that combines the energy and non-energy costs of supplying electricity.

For an exploration of sub-group heterogeneity, we used area-level deprivation data from the UK-wide composite Index of Multiple Deprivation (IMD), small area measures of relative deprivation across each of the United Kingdom. Areas were ranked from the most deprived area (rank 1) to the least deprived area, based on income, employment, education, health, crime, barriers to housing and services, and the living environment. This index was constructed by Parsons and mySociety (2021) using methods from Abel et al. (2016). This version harmonizes the constituent country-specific IMDs to a common England-anchored scale, enabling deprivation comparisons across the UK's statistical reporting areas.

To quantify CO_2e impacts, we integrated data from WattTime, a U.S.-based nonprofit that produces historical estimates of the marginal operating emissions rate, or the emissions associated with the marginal change in load on the grid (WattTime, 2022). WattTime data is available at five-minute intervals.

3 Experimental results

This section presents the main empirical findings from our field experiment. We begin by documenting the impact of the encouragements on trial participant enrollment into the AI managed EV charging tariff. We then examine how these changes in enrollment influenced electricity consumption patterns, using both reduced-form and instrumental variable approaches. Finally, we explore heterogeneity in treatment effects, user behavior under the managed charging regime, and the added value of automation beyond conventional time-of-use pricing. We followed our pre-analysis plan, but state where we added or deviated away from it, why we did, and their implications for interpretation in Appendix C.

3.1 Impact of encouragement on take-up of managed charging

We begin by estimating the effect of encouragement on trial participants' likelihood of adopting a managed EV-charging tariff. Specifically, we estimate the following linear probability model for whether trial participant i is enrolled in the IO Go tariff in week t:

$$D_{it}^{\mathrm{IO}} = \pi_0 + \pi_1 \mathbf{Z}_i + \pi_2 P_{it} + \pi_3 (\mathbf{Z}_i \times P_{it}) + \mu_b + \mu_t + \epsilon_{it} \tag{5}$$

Here, D_{it}^{IO} is a binary indicator equal to one if trial participant i is on an IO Go contract in week t. Z_i is a set of four binary indicators capturing the encouragement assignment, with the control group omitted as the reference category. P_{it} denotes whether participant i is in the incentive period in week t, where the incentive period is the three months after the start of the trial. We include fixed effects for randomization block μ_b and calendar week μ_t. Standard errors are clustered at the level of participant and week.

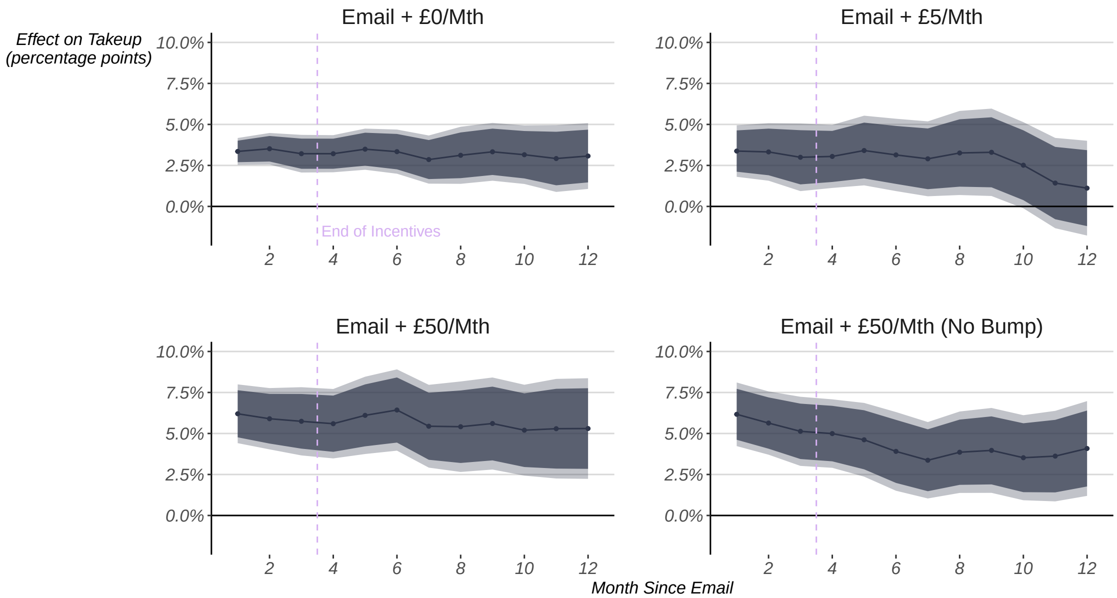
We observe an increase in take-up across all treatment groups, as shown in Table A2. During the 90-day incentive window, the £0/Month and £5/Month groups each increased the probability of take-up by approximately 3.4 percentage points, while the £50/Month and £50/Month (No Bump) groups nearly doubled that effect, reaching 5.9 and 5.7 percentage points, respectively. Take-up in the control group is also rising over time, but remains consistently lower than that observed in any encouragement arm (Figure A3). At the end of the three-month incentivization period, take-up in the control group stood at 2.7%, compared to 7.0% and 6.8% in the £0/Month and £5/Month groups, respectively. The higher incentives resulted in the greatest adoption, with the £50/Month and £50/Month (No Bump) groups reaching 9.3% and 9.2%.[21]. We calculated a price elasticity of 0.143 between the £0/Month and £50/Month groups using the arc elasticity formula, which compares the change in take-up rates relative to the midpoint of both the take-up and total incentive levels (£0 vs £150).

Across most encouragement arms, post-incentive enrollment remained stable. We formally test this in Table A2, interacting the encouragement indicators and the post-incentive period indicator in Table A2. The persistent retention after the incentive period suggests that the initial encouragements had durable effects even in the absence of continued subsidy payments. Only the £50/Month (No Bump) experienced a statistically significant drop of 1.6 percentage points after incentives ended. These results can be further visualized in Figure 5, which shows estimates of Equation (5), split by month since treatment. Retention dropped only in the group facing restrictions on manual overrides, suggesting that managed charging is more likely to succeed when framed as a convenient default rather than a rigid mandate. Encouragements that showcase consumer benefits and flexibility can foster lasting adoption.

Importantly, take-up is mechanically constrained by compatibility: trial participants need either their charger or EV to be supported by IO Go in order to enroll. For those

[21] Tariff contracts can change at the level of the day, and thus there could be some worry that take-up analysis at the weekly level could be biased. We run a robustness check of take-up at the daily level in Table A3, and find extremely similar results to Table A2

Figure 5: Impact of Encouragements on Take-up of Managed Charging Tariff (IO Go)



Notes: This figure plots the intention-to-treat effects of four email-based encouragements to adopt a AI managed charging tariff (Intelligent Octopus Go), split by month since receiving the email. The outcome is a binary indicator for weekly use of the tariff. Each panel corresponds to a different encouragement group, varying in the level of offered financial incentive. Shaded areas represent 90% (dark) and 95% (light) confidence intervals, with standard errors clustered at the participant and week level. The dashed vertical line indicates the end of the 90-day incentive period.

who do not enroll and have an EV, we do not observe their vehicle or charger details, and there is no comprehensive national data on EV and charger ownership to fill this gap. Therefore, we are unable to estimate what proportion of trial participants actually have compatible equipment.[22] While IO Go is compatible with some of the most popular home chargers (notably Ohme, MyEnergi/Zappi, and Hypervolt) this still represents only a subset of the overall charger market. Compatibility via direct vehicle integration covers several major EV brands, including Tesla, BMW, and VW, but again excludes some others. Thus, the measured take-up rate represents a conservative estimate, limited by the extent to which trial participants' vehicles or chargers were compatible with IO Go.[23,24]

[22] In February 2024, we sent a survey to 305 trial participants who we had emailed as part of a pre-trial pilot (where pilot trial participants are not in the sample of our main analyses). We received 68 responses. Among the 56 respondents who expressed interest in signing up for IO Go, 17 indicated that they had not done so due to device incompatibility. We also show in Figure A5 the completion rate for participants who started signing up for IO Go but did not complete onboarding. 23%, did not complete onboarding, and this appeared balanced across encouragement groups. These individuals were not included in our take-up rate. This does not represent the overall incompatibility rate, for two reasons. First, many customers likely checked whether their device was compatible before starting sign-up, since Octopus provides a compatibility survey on the sign-up page. Second, there may also be other reasons beyond compatibility for not completing onboarding, such as hassle factors associated with onboarding.

[23] See this link for compatibility with charge points. IO Go is also compatible with several major EV brands.

[24] To further contextualize these take-up rates, the email open rate was 78% across all treatment groups, with most opens occurring on the day following delivery and no qualitative variation across treatments (Figure A4). Noncompliance with the intended IO Go take-up can thus be a result of not receiving or opening the encouragement email.

As IO Go is not yet compatible for all customers, Octopus Energy offers an alternative EV tariff, Octopus Go, described in Section 2.3. As a result of our email-based encouragements, we observed a nontrivial increase in take-up of Octopus Go; our speculation is that this was driven by customers who owned devices incompatible with IO Go. Compared to the effects on take-up of Intelligent Octopus Go, the impacts on Octopus Go adoption were smaller – roughly 30% as large as those for the £0/Month and £5/Month groups, and 8–11% as large as those for the £50/Month and £50/Month (No Bump) groups, although the effects in the latter two groups are not statistically significant (Table A5). The results suggest that, although uptake was concentrated among adopters of the managed-charging product, the encouragements induced a more general shift toward tariffs designed for EV needs more generally. This affects how we interpret the impacts of IO Go on electricity consumption, which we will discuss more in Section 3.3.[25]

Additionally, we tested for selection on levels – that is, whether adoption of IO Go could be explained by observable baseline characteristics, and in particular whether structural winners (those with higher expected bill savings) were more likely to enroll. We estimated a logit regression of IO Go take-up on encouragement assignment, expected structural winnings, and baseline covariates. We estimated four specifications that progressively increased flexibility: (1) a baseline model including only the incentive and expected winnings; (2) an expanded model adding additional covariates; (3) a model allowing for interactions among covariates; and (4) a final specification incorporating nonparametric controls for expected savings.

We found no evidence of selection on levels: the coefficient on expected winnings was small and statistically insignificant across all specifications. The results, presented as marginal effects at the means of the covariates in Table A6, suggest that customers with greater expected financial savings were not systematically more likely to adopt. There is some evidence of selection on socioeconomic status – households in the second and third terciles of the Index of Multiple Deprivation were more likely to enroll than those in the most deprived tercile. In addition, customers who were already on a time-of-use tariff prior to the trial were more likely to take up IO Go.

The explanatory power of these models is extremely low: the squared correlation coefficients from the propensity-score estimates are near zero, and the estimated propensity scores themselves occupy a narrow range. This limited variation implies that most

[25] IO Go has another rival tariff, Agile Octopus, which is a variable-rate tariff with prices linked to the day-ahead wholesale cost of electricity. Take-up of Agile was quite low in our sample, and we found no effect of our encouragement design on take-up of Agile Octopus.

of the heterogeneity in IO Go take-up arose from unobserved factors rather than observable characteristics. Moreover, the propensity scores produced by the four models are not highly correlated with one another, underscoring the instability of the selection equations. As a result, we were unable to estimate marginal treatment effects to further understand selection on slope and heterogeneous treatment effects.

3.2 Impact of encouragements on electricity consumption

We next assess whether encouragement-induced take-up translated into changes in electricity consumption over the course of the day. Specifically, we estimated intention-to-treat (ITT) effects of each encouragement arm on hourly electricity use (kWh) using the following specification:

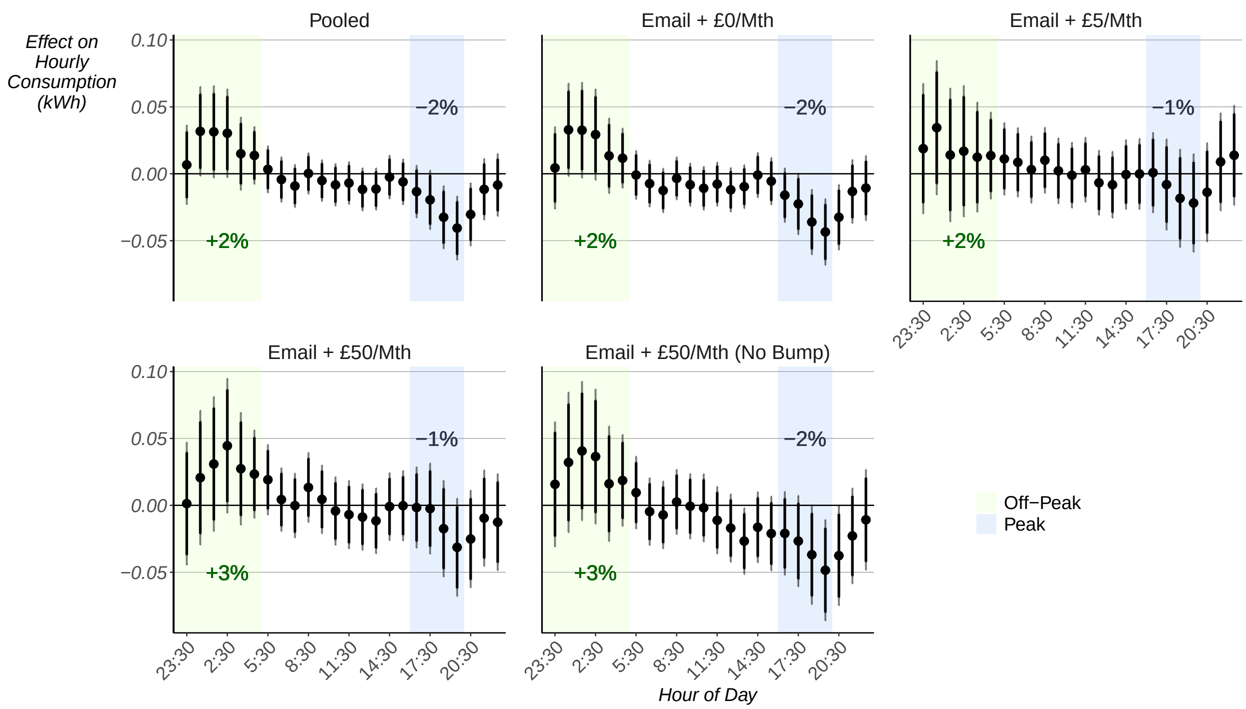
$$Y_{iht} = \alpha + \beta \mathbf{Z}_i + \gamma X_{ih} + \psi_b + \psi_t + \varepsilon_{iht} \tag{6}$$

where Y_{iht} denotes mean electricity consumption for user i in hour h during week t, $\mathbf{Z}_i$ is a vector of binary indicators for each encouragement assignment, X_{ih} is user i's average pre-encouragement January 2024 consumption during hour h,[26] and ψ_b and ψ_t are fixed effects for block and week, respectively. Standard errors were clustered by user and by week (Colin Cameron and Miller, 2015). To obtain hour-specific effects, we estimated the model separately for each hour of the day. To analyze broader periods (i.e., peak vs. off-peak), we grouped hours into the relevant period and ran the regression on those aggregates.

In Figure 6, we present the ITT estimates by encouragement arm, split by hour-of-day. All groups showed increased consumption during the overnight off-peak period (23:30–05:30) and modest reductions during peak hours (16:30–20:30), consistent with the tariff's incentive to shift usage. Patterns are broadly similar across arms, although the £0/Month and £50/Month (No Bump) treatments exhibit the largest declines in peak consumption. When pooling all treatments into a single encouragement indicator and estimating a regression over all hours within the specified peak and off-peak windows, we find that receiving any encouragement increased off-peak consumption by 2% (0.019 kWh) and reduced peak consumption by 2% (–0.026 kWh), with no overall effect on total usage (Table A4).

[26] This control variable was not specified in our pre-analysis plan, but we included it to enhance the precision of our estimates.

Figure 6: Impact of Encouragements on Hourly Electricity Consumption



Notes: This figure shows intention-to-treat effects of four email-based encouragements on hourly electricity use (kWh), using data from the 12 months after sending out the email (estimated using Equation (6), split by hour-of-day). The first panel defines encouragement as a binary indicator for whether the user received any encouragement. Each subsequent panel represents a separate encouragement arm with varying incentive levels. Lines depict 90% (dark) and 95% (light) confidence intervals. Standard errors are clustered by participant and week. Green shaded box indicates IO Go off-peak hours (23:30–05:30), when electricity is charged at £0.07/kWh; all other hours are billed at the standard variable rate. Blue shaded box indicates typical system peak hours (16:30–20:30), which are highlighted to show times of heightened grid stress. Our outcome measure is hourly consumption, with hours defined as starting on the half-hour to align with IO Go's pricing structure. Percentages represent treatment effects as a share of the control group trial participants who are not on an EV tariff, for off-peak (23:30–05:30, green) and peak (16:30–20:30, blue) periods. Estimates come from regressions pooling all hours in each window, as reported in Table A4 and defined in Equation (6).

3.3 Impact of managed charging on electricity consumption

To obtain the causal impact of adoption of the managed charging tariff on electricity consumption, we used an instrumental variables estimation. In the first stage, we instrumented tariff adoption with random assignment to any of the four email-based encouragements:

$$D_{iht} = \pi_0 + \pi_1 Z_i + \gamma X_{ih} + \mu_b + \mu_t + \epsilon_{iht} \tag{7}$$

where Z_i is a binary indicator for whether the user received any encouragement. X_{ih} denotes average baseline consumption for user i in hour h. The specification includes fixed effects for randomization block (μ_b) and calendar week (μ_t).

As shown in Section 3.1, our encouragements increased adoption of both the managed-charging tariff (IO Go) and the EV-oriented time-of-use alternative (Octopus Go). To preserve the exclusion restriction, given that our instrument affects both products, we define D_{it} as an indicator for take-up of either IO Go or Octopus Go.[27]

In the second stage, we regressed hourly electricity use on predicted tariff status:

$$Y_{iht} = \alpha + \beta \hat{D}_{iht} + \gamma X_{ih} + \psi_b + \psi_t + \varepsilon_{iht} \tag{8}$$

where Y_{iht} is the mean daily consumption for user i in hour h of week t (i.e., the mean daily consumption at each hour, averaged across the week). Standard errors are clustered at both the participant and week levels. We estimated this regression over the 12 months after encouragement emails were sent out.

This approach is valid under standard instrumental variables (IV) assumptions: (i) relevance — encouragement must increase adoption, which we confirm with strong first-stage effects already shown in Table A2; (ii) independence — random assignment ensures encouragement is uncorrelated with unobserved determinants of outcomes, which holds given our assignment mechanism, as discussed in Section 2.5; (iii) exclusion — encouragement should affect electricity use only through tariff adoption, which we preserve by pooling IO Go and Octopus Go as the treatment; and (iv) monotonicity — no customers should be less likely to adopt when encouraged, which is plausible given the nature of the intervention. Under these conditions, the IV estimates identify the Local Average Treatment Effect (LATE): the causal effect of adoption for customers who take up a tariff if they receive the encouragement.[28]

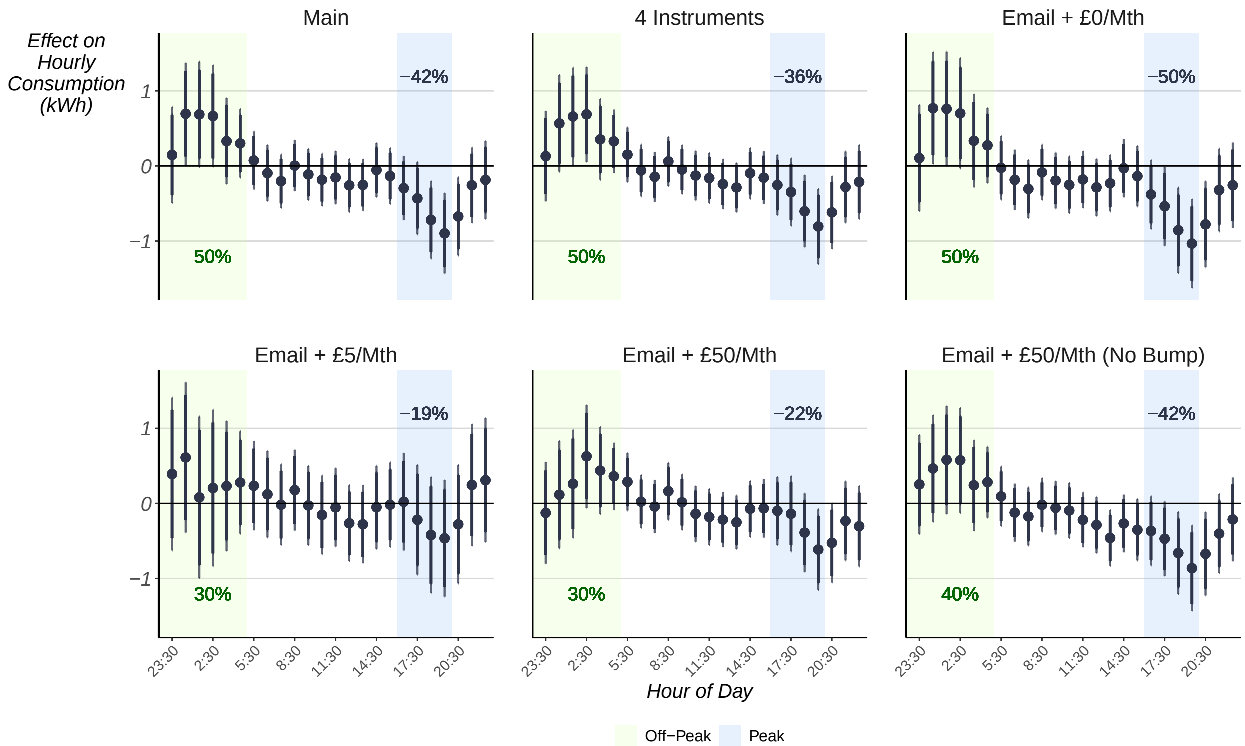
We found that adoption of EV tariff significantly shifted consumption from peak to off-peak hours. When we estimate Equation (8) over all hours within the specified peak and off-peak windows, our main specification shows that households' peak-period usage fell by 42% (0.581 kWh average hourly reduction), while their off-peak usage rose by 50% (0.481 kWh average hourly increase). This is a substantial reallocation of demand rather than an increase in overall consumption. Consistent with this interpretation, Table A9 reports no change in total electricity use. This load shifting pattern is further illustrated

[27] This deviates from our pre-analysis plan, reflecting our uncertainty at the outset of the trial about whether we would face this exclusion restriction violation. However, we find the coefficients are similar whether we include or not the Octopus Go customers into the D_{it} indicator.

[28] The LATE identifies the effect of IO Go for compliers, not the population average treatment effect. Our randomized encouragement generates exogenous variation in assignment, but treatment can only be estimated for those who comply with encouragement. Compliance itself involves multiple stages, such as opening and reading the email, and these decisions may correlate with unobserved characteristics.

in Figure 7, which shows consumption estimates, split by hour-of-day.

Figure 7: Impact of EV Tariff on Electricity Consumption

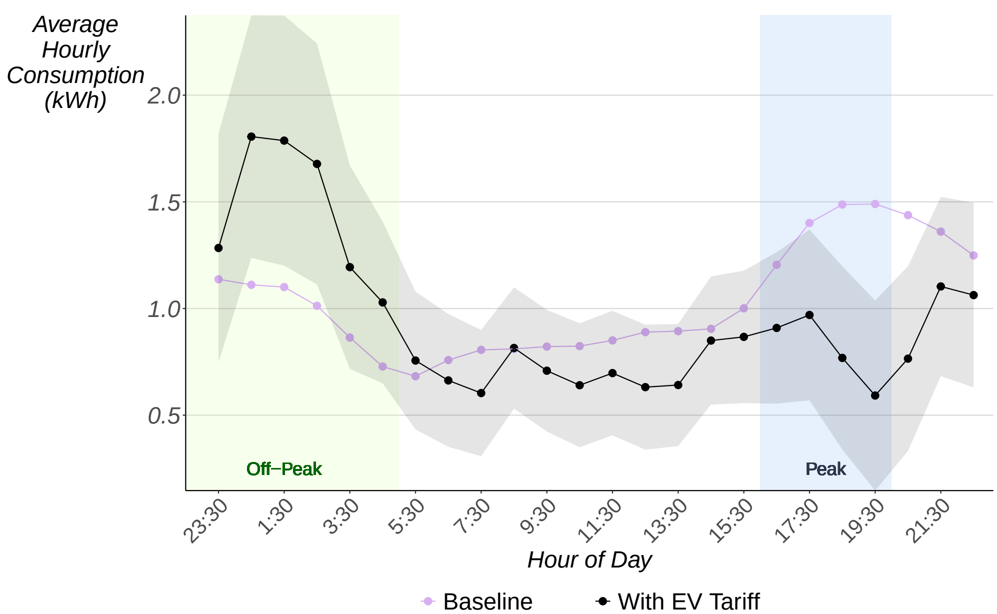


Notes: This figure reports IV estimates of the effect of adopting an EV tariff on electricity consumption (kWh), split by hour-of-day. The instrument is an indicator for assignment to any email-based encouragement. Each panel reports a different specification: (1) is our main specification (Equation (8)); (2) defines a separate instrument for each encouragement group; (3) restricts to just the £0/Month group and the control group; (4) restricts to just the £50/Month group and the control group; (5) just the £50/Month (No Bump) group and the control group. All specifications control for baseline consumption, and fixed effects for randomization block and week. Lines depict 90% (dark) and 95% (light) confidence intervals. Standard errors are clustered by participant and week. Green shaded box indicates IO Go off-peak hours (23:30–05:30), when electricity is charged at £0.07/kWh; all other hours are billed at the standard variable rate. Blue shaded box indicates typical system peak hours (16:30–20:30), which are highlighted to show times of heightened grid stress. Our outcome measure is hourly consumption, with hours defined as starting on the half-hour to align with IO Go's pricing structure. Percentages represent average treatment effects as a share of the control group trial participants who are not on an EV tariff, for off-peak (23:30–05:30, green) and peak (16:30–20:30, blue) periods. Estimates come from regressions pooling all hours in each window, as reported Tables A7 and A8.

Our preferred specification uses a single binary instrument for assignment to any of the four encouragements. Combining the four encouragement groups into a single binary instrument mitigates concerns raised in Mogstad et al. (2021), particularly the potential for negative weights being assigned to individual instruments when multiple instruments are used. To assess robustness, Figure 7 also shows (i) a specification that includes the four encouragement indicators as separate instruments and (ii) 4 additional specifications that use each encouragement indicator as a stand-alone instrument. The estimated effects are similar across specifications, indicating that our findings are not sensitive to the instrument definition.[29]

We show the consumption profile of the baseline group – control households who did not adopt an EV tariff (IO Go or its similar non-managed charging ToU version) – to contextualize these treatment effects. These customers' electricity consumption exhibits a pronounced peak beginning around 16:30, consistent with typical residential demand patterns and uncoordinated EV charging behavior (see Figure 8). Using the estimated treatment effects from the main specification in Figure 7, we overlay the causal impact of EV tariff adoption onto the baseline profile. This constructed profile illustrates how the EV tariff shifts electricity demand away from the evening peak and toward the designated overnight off-peak period. The resulting pattern flattens the peak-hour hump and concentrates usage during hours when electricity is less expensive and there is less grid stress, underscoring the potential of managed charging to reshape intraday load without increasing total consumption.

Figure 8: Electricity Consumption With and Without EV Tariff Adoption



Notes: This figure displays mean hourly electricity consumption (kWh) for trial participants in the control group who were not enrolled in the EV tariff. The purple line ("Baseline") plots their observed consumption. The black line ("With EV Tariff") adds the estimated hourly treatment effects of EV tariff adoption, recovered from the IV analysis displayed in the first panel of Figure 7. The shaded area denotes the 95% confidence interval for the treatment effect estimates.

[29] Note that this preferred specification deviates from our pre-analysis plan of using the encouragement arms as four separate instruments. However, results from the pre-specified analysis, which includes all four encouragement indicators in the regression, are presented in the second panel of Figure 7.

3.4 Heterogeneity

We explored heterogeneity in treatment effects across consumers by interacting the encouragement treatment with three key variables: (1) the Index of Multiple Deprivation (IMD), (2) property value of their home, and (3) baseline electricity consumption. In these specifications, we treat interaction terms (e.g., EV tariff $\times$ baseline covariate) as endogenous and instrument them with the corresponding interaction of the randomized encouragement and the covariate, consistent with standard IV practice.

The IMD is a composite measure that captures multiple dimensions of deprivation (e.g., crime, housing barriers, health) for small geographic areas, weighted to produce an overall deprivation score. For this analysis, we constructed IMD terciles by linking trial participants' meter-point-level postcodes to area-level IMD ranks.[30] Socioeconomic status shapes both the ability to adopt EVs and the potential financial gains from managed charging. Households facing financial hardship could in theory benefit most from cheaper charging, but structural barriers, such as lack of off-street parking or neighborhood safety concerns, may prevent uptake. Understanding these dynamics requires analyzing effects across the socioeconomic gradient.

We used property value as an additional indicator of socioeconomic status to complement our location-based Index of Multiple Deprivation (IMD) measure. Property value data were obtained from WhenFresh, a provider that aggregated daily listings from Zoopla, a major UK property portal. Each property was associated with a Unique Property Reference Number (UPRN), a persistent identifier that remained constant throughout the property's lifetime. Using customers' UPRNs, we matched each record to the most recent property value available in the WhenFresh dataset. This matching procedure yielded property value information for 98% of the sample. We used the full WhenFresh dataset (29,783,591 properties) to establish national cutoffs for property-value terciles. Thus, the heterogeneity bins in our analysis reflected thresholds derived from the national property distribution rather than from our sample alone.

For baseline electricity consumption, we computed the total kWh used per customer by aggregating all available half-hourly smart meter readings per day over the period from February 15, 2023, to August 31, 2023. Households with high baseline electricity consumption may have greater flexibility to shift charging, since larger batteries or multiple EVs provide more scope to delay without running short of range. At the same time,

[30] We discretize the IMD into terciles to avoid very small cell sizes in the most deprived group.

their greater and more time-sensitive energy needs can reduce flexibility, leaving less room to adjust without disrupting routines. This dimension highlights whether managed charging works chiefly for high-demand users or more broadly across households, shaping expectations about scalability and future generalizability.

We observed mixed evidence of heterogeneity in take-up. Adoption of IO Go was higher in IMD terciles 2 and 3 (Figure A8A), but we did not observe the same pattern in looking at heterogeneity by property values (Figure A9). Take-up also declined with baseline electricity use: households with higher pre-trial consumption were less likely to adopt, particularly under the £50/Month (No Bump) condition (Figure A10A). This pattern suggests that higher-consuming households may be more reluctant to accept restrictions on charging, perhaps due to greater perceived disruption to their routines.

Turning to impacts on electricity consumption, we found limited evidence of heterogeneous impacts. Across the three dimensions we considered, there was no evidence of differences in off-peak effects. For peak hours, the largest reductions occurred in the middle IMD tercile (Figure A8b) and the highest property value tercile (Figure A9b). However, these results should be interpreted with caution given the limited precision of subgroup estimates: only 9% of our sample resided in the most deprived tercile and 7% occupied properties in the lowest tercile of the national property-value distribution. This reflects the strong association between EV ownership and higher socioeconomic status (Figure A8c, Figure A9c). The small sample size in this group restricted our ability to detect precise effects among more deprived households. By contrast, baseline electricity use showed little role in shaping treatment impacts. Conditional on adoption, consumption effects are broadly similar across all terciles of pre-trial consumption (Figure A10b).[31]

3.5 Intelligent Octopus Go user behavior

For participants who adopted IO Go, we had telemetry data on their plug-in events and charging sessions, allowing us to examine their behavior more closely. On average, IO Go households consumed 1.02 kWh per hour (for their home overall, not just their EV), equivalent to 8,935 kWh per year. Of this total, 22.6% was attributable to EV charging, or around 2,020 kWh annually. If all charging occurred at home and the vehicle averaged about 3–4 miles per kWh, this implies 6,000–8,000 miles driven per year. These figures are consistent with the average annual car mileage in the UK (approximately 7,000 miles;

[31] We also looked at heterogeneity by weekday versus weekend, but found demand impacts to be similar regardless of the day of the week.

Department for Transport, 2024), suggesting that the mileage behavior in our sample is comparable to that of primary household vehicles in the broader UK population.

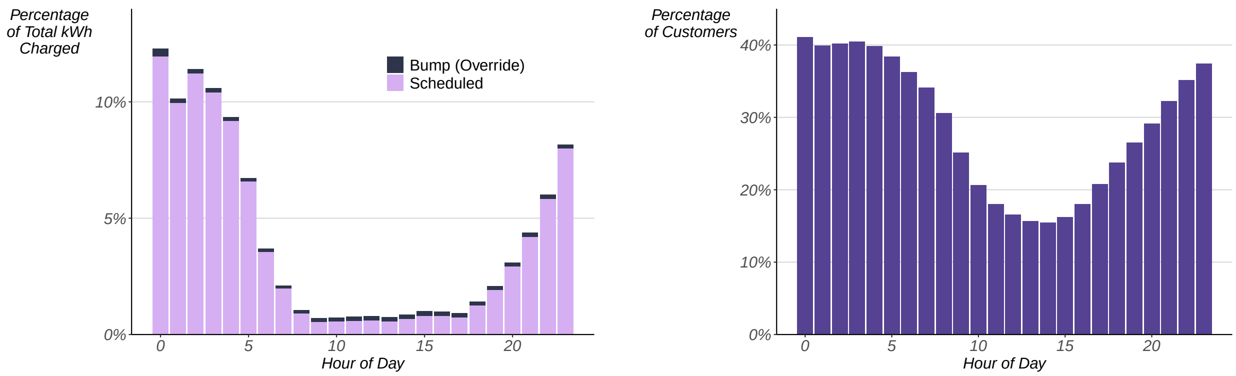
We further explored how users from our sample engaged with automated EV charging. Two key patterns emerged: strong adherence to the automation schedule and a high degree of consistency in treatment effects across users. Together, these findings suggested that managed charging was generally well integrated into users' routines with minimal disruption.

We begin with evidence of adherence to the automation schedule. By adherence, we mean their likelihood of overriding the AI charging algorithm. Adherence not only provides a useful proxy for satisfaction with the tariff. It also is the sufficient statistic, within the theoretical framework we develop in Appendix G, for understanding the efficiency of this tariff versus RTP. We analyzed 2,359 IO Go participants' use of the "bump" function, which allowed them to override the default schedule and initiate immediate charging. We found that 55% of users never used the bump feature at all (Figure A15), and bump events accounted for only 2.3% of total IO Go electricity consumption. These infrequent overrides, consistent across treatment groups, suggested that the AI managed charging schedule was generally accepted and rarely disrupted. This was also consistent with what we saw in Figure 5, where the sustained uptake of IO Go suggests that trial participants generally accepted the automated approach.

Overall, we found that the probability that a customer overrode on a given day was 1.9% – 0.3% probability of override in off-peak hours, 0.3% during peak times as we have defined them (16:30-20:30), and 1.3% during all other hours of the day (multiple overrides in a single day were extremely rare in our data). Given the simulations from the theoretical section, these numbers were well below the crossover override rate where RTP dominates an AI EV-managed tariff (i.e., 33%).

Although bump charging was infrequent, its implications for grid planning depend on whether charging is correlated across locations or time. To assess this, we used a fixed-effects regression to decompose the variance in bump charging across customer, temporal, spatial, and temperature components. As shown in Appendix E, most of the variation was driven by customer-specific differences, with little contribution from time, location, or temperature. This dominance of customer fixed effects indicated that bump charging was largely idiosyncratic rather than correlated, making it unlikely to generate coincident demand spikes that would strain the grid.

Figure 9: Hourly Patterns of EV Plug-In and Charging Behavior



(a) Consumption by Hour-of-Day

(b) Plugged In by Hour-of-day

Notes: Panel (a) shows, by hour of day, the percentage of electricity consumption that occurred during that hour. This consumption is further divided into charging triggered by bump (charging initiated by users overriding the schedule) versus charging scheduled by IO Go. Panel (b) plots, by hour of day, the percentage of active users who had their EVs plugged in. A user is considered active during a given hour if that hour falls between their first-ever and last-ever recorded plug-in event. The analysis is based on data from 2,359 IO Go customers.

Preferences for charging settings further reinforce this uniformity. 61% of trial participants preferred their vehicle to finish charging between 07:00 and 09:00, and 93% set their desired state-of-charge (SOC) at 80% or higher (Figure A12a). Plug-in patterns also followed a predictable rhythm: more than half of plug-in events occur within 24 hours of the previous one, typically following a post-work return home and preceding the next morning's commute (Figure A12b).

We also document that the IO Go mechanism operates effectively with sparse and infrequently updated private information. The median customer only changes their settings twice after the initial onboarding. The bulk of preference discovery occurs in the first week: one week after adopting IO Go, 54% of customers never updated their settings again (Figure A11). This stability underscores the mechanism's minimal information requirements. Once users report a small number of primitives, the algorithm can implement near-efficient charging schedules without ongoing preference elicitation or strategic interaction.

If customers were to begin charging as soon as they plugged in, typically between 5:00pm and 7:00pm, it would place significant strain on peak demand. Managed charging avoids this issue by decoupling plug-in time from charging time; customers still enjoy the convenience of plugging in when they arrive home, while AI managed scheduling shifts the actual charging to off-peak hours, as illustrated in Figure 9. Importantly, these behaviors are all consistent across our encouragement groups, suggesting that once trial participants opt into the tariff, they tend to use it in similar ways (see Figures A13, A14 and A16). This behavioral consistency is mirrored in the relatively homogeneous consumption impacts shown in Figure 7.

3.6 Heterogeneity in treatment effects based on telemetry data

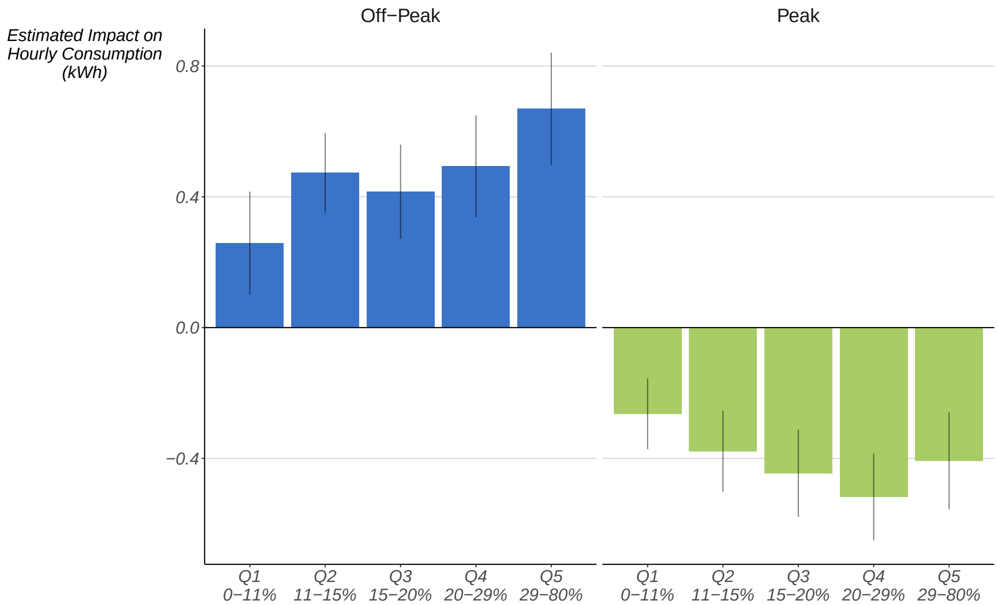
We next examined which dimensions of charging behavior drive heterogeneity in our estimated impacts. Specifically, we focused on behaviors that shape the potential for managed charging to shift load from peak to off-peak hours. Because IO Go specific telemetry is only available once customers have adopted the tariff, we cannot directly observe these behaviors for non-adopters. We therefore use a difference-in-differences design among participants in our field experiment who later adopted IO Go, following Callaway and Sant'Anna (2021). Approximately 55% of field experiment participants who adopted IO Go did so during the incentivized period, while the rest adopted in the nine months afterwards. This staggered adoption pattern produces sufficient variation to support a difference-in-differences style analysis.

We estimated heterogeneous treatment effects by plug-in rate — measured both as the proportion of all hours in which a customer's EV is plugged in and, separately, the proportion of peak hours with an active plug-in. We reason that these measures capture the degree of availability of the vehicle for managed charging. However, we interpret them as descriptive correlates rather than causal drivers of heterogeneity, since plug-in behavior is endogenous and also may be influenced by tariff adoption. We find substantial variation: households with higher plug-in rates experienced markedly larger shifts of electricity use from peak to off-peak periods (Figure 10). This pattern suggests that households providing the AI with more frequent charging opportunities enable greater flexibility in shifting demand away from system peaks.

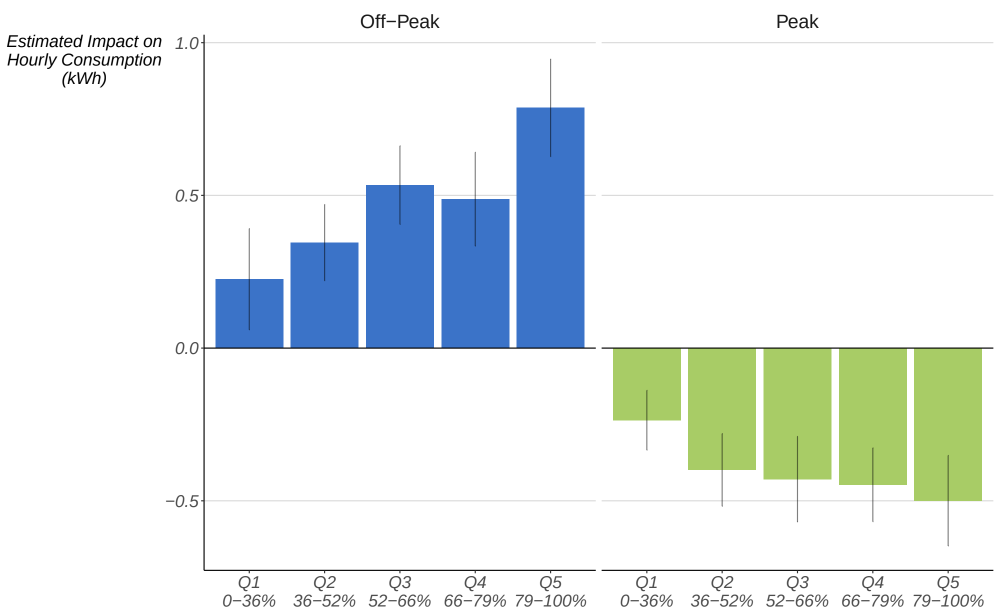
We also examined heterogeneity along other behavioral dimensions. First, we considered "bump" charging. Households that have ever used the bump function show a noticeably larger increase in off-peak consumption, but their reduction in peak-period usage is statistically indistinguishable from non-bump households (Figure A23). Second, we assess heterogeneity by users' preferred "ready-by" times and target charging levels, finding no systematic differences in either peak or off-peak effects (Figure A24). Taken together, these results suggest that, for reducing peak-hour consumption, plug-in availability is the dominant behavioral factor shaping treatment effects.

Figure 10: Heterogeneity in Impacts by Plug-in Behavior

(a) Impacts by Quintile of Plug-in Rate



(b) Impacts by Quintile of Plug-in Rate During Peak Hours



Notes: This figure reports staggered difference-in-differences estimates of the effect of adopting the IO Go tariff on hourly electricity consumption, for participants who were a part of our experiment. Panel (a) shows impacts by the quintile of the plug-in rate, and panel (b) shows impacts by the quintile of plug-in rate during peak hours (16:30-20:30). Estimates are also separately estimated by peak (16:30-20:30) and off-peak hours (23:30-5:30). Estimates are computed using the Callaway and Sant'Anna (2021) estimator. Error bars represent the 95% confidence interval.

3.7 Value of AI managed charging beyond time-of-use tariff structure

As we showed in Section 3.1, we encouraged customers onto a similar ToU tariff that was not AI managed because they have an incompatible charger or EV. The IO Go and Octopus Go tariffs both feature time-of-use pricing, but, importantly, IO Go includes AI managed charging, whereas Octopus Go relies on customers manually adjusting behavior in response to the tariff's day and off-peak rates or setting up their own automated charger schedules to align with the tariff (but note that we do not observe settings from these trial participants). To understand the added value of AI managed charging, we estimated how consumption under IO Go responded to the real-time system price, relative to consumption under Octopus Go.[32] While neither IO Go nor Octopus Go participants were directly exposed to system prices, the AI management of IO Go charging schedules responds to those prices on the retailer's behalf. Thus, the observed difference in elasticity reflects the supplier's algorithmic responsiveness, not household-level reactions to system costs.

Our analysis examined responsiveness to real-time system prices across three periods of the day: daytime (5:30–16:30), evening (16:30–23:30), and overnight off-peak (23:30–5:30). (Note that "system prices" in Great Britain refer to the real-time wholesale prices for each half-hour, rather than day-ahead prices.) We examined responsiveness using a Poisson regression, which accommodates zero consumption and yields coefficients that can be interpreted directly as elasticities of electricity consumption with respect to system prices. Table 1 reports results from the following specification:

$$log(\mathbb{E}[Y_{it}]) = \alpha + \beta_1 D_{it} + \beta_2 \log(P_t) + \beta_3 [D_{it} \times \log(P_t)] + \mu_d \tag{9}$$

where Y_{it} is electricity consumption for trial participant i at date-hour t, D_{it} is an indicator for IO Go participation, and P_t is the system price. The coefficient β_2 captures the price elasticity of demand among the baseline group (Octopus Go users), while β_3 captures the differential elasticity for IO Go users. We also included fixed effects for day d, so that coefficients are identified from within-day variation across participants. The sample in this analysis comprised the subset of our original trial sample (n = 13,233) who signed up for either IO Go or Octopus Go (n = 2,963).[33]

[32] In Great Britain, the system price is the market-wide wholesale price of electricity settled every half hour. It reflects the cost of balancing supply and demand on the grid and is published by the National Energy System Operator (NESO).

[33] This analysis was not pre-specified. We have included it as an exploratory analysis to help to isolate the automation-related mechanisms.

Given that tariff assignment was not random, this analysis should be interpreted as suggestive rather than causal. Many trial participants on Octopus Go were unable to enroll in IO Go due to compatibility constraints rather than personal preference, limiting selection into IO Go to some extent. However, even where this is the case, tariff choice remains mechanically correlated with vehicle and charger type; and, it is true that some Octopus Go customers may actively have chosen not to enroll in IO Go to retain autonomy over their charging schedule.

Overall, we found within-period differences in the price elasticity of demand from IO Go and Octopus Go customers – where, again, note that the price was the system price that consumers themselves never saw, but rather was a key input the AI used when scheduling charging. During the evening period that encompasses peak hours (16:30–23:30), trial participants on IO Go exhibited significantly greater price responsiveness than Go customers: a -0.044 additional price elasticity of demand. During the overnight off-peak window (23:30–05:30), the price elasticity of demand was again significantly greater for IO Go customers than Octopus Go customers (by -0.024). By contrast, during daytime hours (05:30–16:30), the interaction term is statistically indistinguishable from zero, suggesting no meaningful difference in price responsiveness between the two groups, possibly due to fewer vehicles being plugged in during these hours (and thus less available charge to shift).

When pooling hours across the full day, there is no evidence that IO Go participants systematically shifted more consumption from higher-priced periods toward lower-priced ones than Octopus Go customers (Column (4) of Table 1). If anything, Octopus Go customers appeared to consume relatively more when prices were low, a counterintuitive pattern. There are three potential reasons for this that we cannot disentangle. First, is selection: Octopus Go adopters may have been those who received the IO Go encouragement but opted into Octopus Go without any monetary incentive. These participants may have had more demand flexibility than the incentivized IO Go participants, which allowed them to be more able to respond to the ToU tariff structure.

Second, these estimates reflect total household consumption, not just EV charging. Thus, the apparent responsiveness of Octopus Go users could also reflect shifts in non-EV household load to overnight hours, rather than differences in automated charging. Unfortunately, we lack telemetry data for Octopus Go customers to directly test this mechanism. Third, differences in the composition of the two groups may also contribute to the pattern, with Octopus Go participants consuming more in low-price periods than IO Go customers, though the relative balance of baseline covariates between IO Go and Octopus Go users, shown in A13, helps alleviate concerns about compositional differences. Focusing on specific periods (Columns (1)–(3) of Table 1) reduced the influence of the issues noted above. In these periods, the AI was likely the main driver of when charging occurs, because households were not exposed to dynamic prices under their ToU structure and thus would have been unlikely to respond to half-hourly-varying market prices. This makes the within-period elasticities a clearer measure of the AI's responsiveness than the all-hours estimates.

Our synthesis of these results is that *within* defined periods, the AI managed scheduling feature in IO Go shifted consumption away from high-price hours more aggressively than trial participants exposed solely to Octopus Go's ToU pricing. This enhanced responsiveness likely reflects the supplier's ability to algorithmically optimize charging schedules when vehicles were most likely to be plugged in. By contrast, Octopus Go customers received a static overnight rate and had to determine their charging behavior manually, leading to flatter responsiveness within periods.

Table 1: Responsiveness to Systems Prices

period	5:30-16:30	16:30-23:30	23:30-5:30	All Hours
Model:	(1)	(2)	(3)	(4)
Variables				
IO Go	0.075**	0.077**	-0.119***	0.089***
	(0.030)	(0.031)	(0.036)	(0.025)
log(Price)	0.047***	-0.003	0.010	-0.125***
	(0.008)	(0.010)	(0.013)	(0.010)
IO Go × log(Price)	-0.004	-0.041***	-0.021*	0.017**
	(0.009)	(0.011)	(0.011)	(0.008)
Fixed-effects				
date	Yes	Yes	Yes	Yes
Fit statistics				
Octopus Go Mean	1.08	0.656	0.968	1.79
Observations	6,178,064	4,489,911	4,257,014	14,924,989

Clustered (User & date-hour) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Notes: This table reports estimates from Equation (9), examining how average hourly electricity consumption under IO Go and Octopus Go responds to system prices. Column (1) covers daytime hours (05:30–16:30), column (2) evening (16:30–23:30), column (3) overnight off-peak (23:30–05:30), and column (4) pools together hours across the whole day. All regressions included day fixed effects, so coefficients are identified from within-day variation across participants; standard errors were clustered by user and hour.

Taken together, these results indicate that managed charging under IO Go not only shifted demand away from the peak, but also redistributed it within periods in a way that was responsive to real-time system costs. These findings suggest that managed charging can outperform conventional time-of-use tariffs in aligning household electricity consumption with the dynamic needs of the grid, closer to real-time. We should also note that the AI algorithm was also responding to ancillary and other markets that we did not model here.

4 Alternative estimation strategy: difference-in-differences of a much larger and earlier sample

In addition to our experimental design, we leveraged a differences-in-differences (DiD) approach using observational data to estimate the impact of adopting IO Go on household electricity consumption. This strategy exploited the staggered, voluntary adoption of IO Go across Octopus Energy customers in 2023 (one year prior to our experiment). Everything that follows was pre-specified in our pre-analysis plan, except where deviations are explicitly noted.

Comparing estimates from the DiD to our field experiment allows us to contrast the behavior of voluntary early adopters of managed charging (captured by the DiD) with that of harder-to-recruit individuals who required external encouragement to adopt (captured by the field experiment). The field experiment sample may better reflect the future mainstream population, who are not proactively engaged, but require more pricing and marketing interventions to adopt managed charging. Some of the observed differences may also reflect other factors: changes to the Intelligent Octopus algorithm over time, evolving day-ahead price profiles (particularly as the volatility of the energy crisis subsided), or methodological differences, such as the potential for selection bias in the DiD estimates, as highlighted in critiques of observational methods (LaLonde, 1986; Imbens and Xu, 2024).

4.1 Empirical strategy

We began with a sample of 100,986 customers who adopted Intelligent Octopus Go (IO Go) at some point in 2023. To isolate the effect of IO Go from other contemporaneous changes, we further restricted the sample to customers who likely already owned an EV by December 2022, using the methodology outlined in Section 2.1. This ensured that observed changes in electricity demand patterns were due to changes in charging demand, rather than the initial uptake of EVs. These restrictions resulted in a sample of 20,249 customers.[34]

We implemented a standard event-study difference-in-differences estimator, allowing for staggered adoption and dynamic treatment effects, following Callaway and Sant'Anna (2021). We made a parallel trends assumption based on "not-yet-treated" units. For each group of units first treated in week g, and for each week $t \geq g$, we defined the group-time average treatment effect on the treated (ATT) as:

$$ATT(g,t) = \mathbb{E}[Y_t - Y_{g-1} \mid G = g] - \mathbb{E}[Y_t - Y_{g-5} \mid D_t = 0,\ G \neq g] \tag{10}$$

where $G = g$ denotes the cohort of units first treated at time g. We estimated both (1) aggregate group-time effects, and (2) a single post-treatment estimate, constructed as the weighted average of all group-time ATT estimates, with weights proportional to group size. To mitigate potential bias from anticipation effects, we excluded the four weeks prior to adoption from our estimation. Accordingly, our reference period, Y_{g-5}, was hourly consumption measured five weeks prior to the treatment (Roth, 2024).[35]

To enhance comparability, we limited control cohorts to those scheduled to adopt IO Go no later than twelve weeks after the treated group's anticipation period ends. This ensured treated units were compared only to future adopters with similar adoption timing. We chose a twelve-week window to balance comparability of treated and control groups against the length of the post-adoption estimation horizon. Comparability was assessed by examining pre-treatment trends, and we selected the longest horizon that yielded satisfactory pre-trend balance. Our treatment assessment therefore relied on the following parallel trends assumption: absent adoption, treated and not-yet-treated households would have experienced similar trends in electricity use.

In addition, to improve comparability with the field experiment estimates and probe underlying mechanisms, we estimated weighted versions of Equation 10, reweighting the

[34] This restriction was not specified in our pre-analysis plan. However, preliminary analysis revealed that failing to condition on EV ownership by December 2022 would conflate the effects of tariff adoption with those of initial EV uptake, thereby biasing our estimates of charging behavior.

[35] We assumed that once a customer first adopts IO Go, they remain "treated" in the sense that their experience with the tariff continues to shape their behavior, even if they subsequently switched to another Octopus tariff (Callaway and Sant'Anna, 2021). In practice, some customers did have more complex tariff histories. Under the irreversibility assumption, their electricity consumption patterns are considered to remain influenced by IO Go from the point of initial adoption. We view this as reasonable for two reasons: (1) 78% of customers who adopt IO Go subsequently remain on it for at least 12 months afterwards, and (2) it is plausible that IO Go induces some degree of habit formation, both in EV charging routines and in household electricity use more broadly.

DiD sample match the field experiment sample based on pre-treatment tariff type. The field experiment explicitly tried to exclude customers with any prior smart tariff usage.[36] Since smart tariffs incorporate time-of-use pricing structures, this exclusion disproportionately removed time-of-use tariff users from the field experiment sample. As a result, only 14% of the field experiment participants were on a time-of-use tariff at baseline, compared to 75% in the DiD sample.[37] To align the tariff composition across the two groups, we calculated the baseline shares of standard versus time-of-use tariff users in each sample. We then reweighted the DiD observations by the ratio of field experiment to DiD shares, ensuring that the reweighted DiD sample better reflected the field experiment's pre-treatment tariff distribution. [38]

Reweighting the DiD customers was important because their baseline consumption profiles differed substantially based on prior tariff. Figure Figure A17 shows that DiD customers on flat tariffs, DiD customers on ToU tariffs, and the field experimental control group had distinct consumption patterns before adopting IO Go, even though overall consumption was similar across the three groups. Customers who had been on ToU tariffs already displayed a consumption pattern closely aligned with IO Go's incentives: minimal afternoon peaking and high overnight usage, consistent with many having been on Octopus Go, which offered cheaper overnight rates from 00:30 to 05:30. In contrast, DiD customers who had been on flat tariffs still exhibited an afternoon peak, and while they also showed an overnight spike in consumption, its magnitude was only about one-third that of customers on ToU tariffs. Together, these patterns indicated that many DiD customers were already partially aligned with IO Go's incentivized charging profile prior to adoption.

[36] This exclusion was not perfect; a small number of customers who previously had smart tariffs *were* part of our trial sample. Smart tariffs are tariffs that require smart meters because their half-hourly unit rate changes.

[37] For historical reasons, there are a handful of time-of-use tariffs that are not "smart" tariffs; the most well-known of which is called "Economy 7", a tariff introduced in the 1970s to incentivize overnight electricity use, particularly for storage heaters by offering cheaper rates during a fixed seven-hour off-peak window. In recent years, some EV owners also adopted Economy 7 as a way to charge their vehicles at lower cost.

[38] We also implemented propensity-score reweighting. Specifically, we estimated the probability of being in the field experiment (vs. DiD) sample using a logit regression with the following covariates: (1) tariff type prior to treatment, (2) total electricity consumption in December before the study period (December 2022 for DiD, December 2023 for the field experiment), (3) the share of consumption occurring during peak hours, (4) Octopus tenure, (5) IMD rank, and (6) property value. The resulting propensity scores were then used as weights in the DiD regression. However, we found that only the pre-treatment time-of-use tariff indicator had a substantive effect on the results. Given this, we opted to show the results only of the simpler tariff-based reweighting approach described in the main text.

4.2 Difference-in-differences results

Figure 11 shows that the difference-in-differences estimates are notably smaller than those from the field experiment. In the unweighted specification (Column 1), IO Go adoption is associated with a 0.06 kWh (7%) average hourly reduction in peak-period consumption, and a 0.1352 kWh (8%) increase during off-peak hours. In contrast, the field experiment estimates imply much larger shifts: a 0.581 kWh (42%) decrease during peak periods and a 0.481 kWh (50%) increase off-peak.

This discrepancy appears to be largely explained by differences in baseline time-of-use tariff usage. Column 2 presents the results after reweighting the DiD sample to match the field experiment sample's pre-treatment tariff distribution. After reweighting, the estimated increase in off-peak consumption in the DiD analysis (0.451 kWh) matches the field experiment analysis (0.481 kWh). However, the reduction in peak consumption is still smaller in the DiD analysis: 0.24 kWh compared to 0.581 kWh in the field experiment. This weaker peak effect is compensated for by a decline in daytime, non-peak consumption in the DiD sample, as shown in Figure A18, resulting in no overall increase in consumption. We hypothesize that the observed differences in the timing of impacts throughout the day arise from compositional differences between IO GO participants and those in our field experiment sample, with DiD participants likely having greater daytime flexibility. We conclude from these results that the field experiment targeted a sample whose characteristics made their baseline consumption less aligned with IO Go's optimization – i.e., because their charging behavior was not previously responding to dynamic or off-peak pricing, leaving more scope for managed charging to change consumption in both peak and off-peak hours.

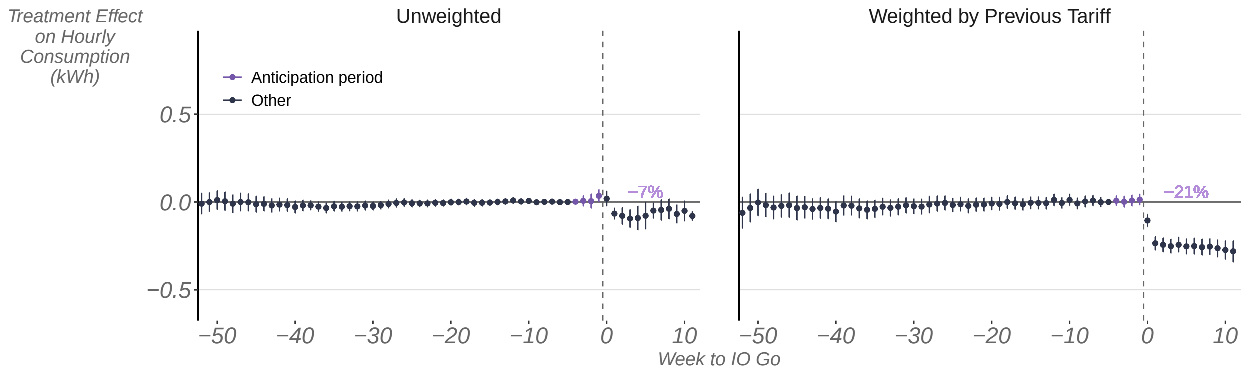
We also estimated cohort-specific treatment effects, defining cohorts by the week of adoption. We found that treatment effects were relatively homogeneous across cohorts, as shown in Figure A19.

Taken together, this homogeneity in treatment effects across cohorts, combined with the close alignment of field experiment and reweighted DiD estimates, suggests that the impact of IO Go is relatively stable across adopters. The primary source of variation appears to be baseline charging behavior, particularly whether customers were already on time-of-use tariffs prior to adoption, rather than any inherent heterogeneity in responsiveness to managed charging.

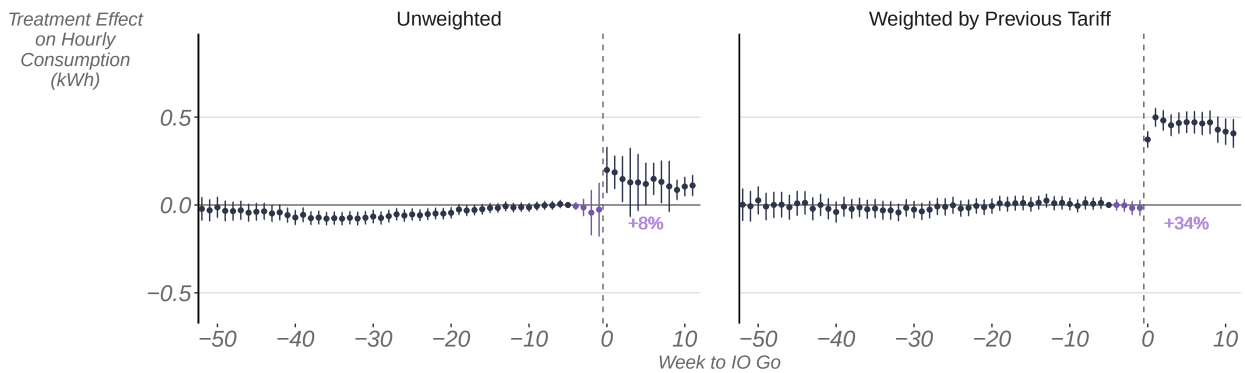
We also conducted a DiD analysis focusing on customers who adopted IO Go in 2024,

Figure 11: Difference-in-Differences Estimate of IO Go

(a) Peak



(b) Off-Peak



Notes: This figure reports staggered difference-in-differences estimates of the effect of adopting the IO Go tariff on hourly electricity consumption (in kWh) during (a) peak hours (16:30–20:30) and (b) off-peak hours (23:30–05:30), using a sample of 20,249 customers who first-ever enrolled in IO Go in 2023. Each panel plots treatment effects relative to the week before adoption. Estimates are reported under two specifications: (i) unweighted; (ii) and weighted by whether the trial participant was previously on a time-of-use tariff. Estimates are computed using the Callaway and Sant'Anna (2021) estimator. Percentages represent post-treatment effects as share of the pre-IO Go consumption levels. Post-treatment effects are estimated using average of all group-time average treatment effects, with weights proportional to the group size.

to make the comparison period more directly aligned with the field experiment timeframe. However, identifying which customers already owned an EV at the start of 2024 proved difficult. We believe this is likely due to increased uptake of low-carbon technologies (LCTs). Of particular note, there was a rise in heat pump installations at the end of 2023, driven by the UK Government's expansion of heat pump subsidies in October 2023.[39] This increase in LCT ownership likely produced additional consumption spikes resembling EV charging, but originating from other devices. Lacking direct observation

[39]Coverage on heat pump takeup can be found in the BBC and Guardian. Academic research on the electricity consumption impacts of heat pumps is documented in Bernard et al. (2024)

of LCT ownership, we cannot be sure whether changes in consumption among customers adopting IO Go in 2024 were driven by tariff adoption, adoption of EVs, or adoption of other LCT devices that encouraged more attention paid to which tariff they were on. [40]

We partially solve this problem by restricting the 2024 sample to customers who appeared to own an EV as of August 2023, which is (1) before the increase in the heat pump subsidy and (2) during summer months when heat pump use is minimal. This refinement improved the accuracy of EV identification but did not eliminate the possibility that some customers adopted heat pumps or other LCTs concurrently with switching to IO Go. The resulting 2024 DiD estimates closely matched the 2023 results for peak and off-peak periods but showed an 8 percent increase in total daily consumption, a pattern absent in 2023. We interpret this as reflecting contemporaneous adoption of other LCTs, such as heat pumps. Full results and further discussion are provided in Appendix F.

5 Welfare impacts

The large causal shifts in consumption from peak to off-peak hours from the AI managed EV tariff, as seen in Section 3.3, have four potential benefits: (1) benefits to consumers from lower electricity bills, (2) lower electricity procurement costs to the electricity supplier, (3) climate change mitigation benefits via CO_2e abatement,[41] and (4) reduced grid operation and stabilization costs. In this section, we first outline our methods for estimating these various benefits and costs of IO Go adoption. We then present the direct benefits accrued in 2024. Finally, we apply the Marginal Value of Public Funds (MVPF) framework developed by Hendren (2016) and Hendren and Sprung-Keyser (2020) to assess the welfare implications of subsidizing managed charging.

5.1 Estimation

To estimate the magnitude of these benefits, we combined administrative data from Octopus Energy with external inputs and applied outcome-specific methods. For con-

[40] This tariff-switching behavior is documented by Bernard et al. (2024), who examined households that received heat pump installations from Octopus Energy. They found that following the installation, two-thirds of these households adopted a smart tariff, with Intelligent Octopus being the most popular choice, possibly due to adoption of an EV at a similar time.

[41] The CO_2e abatement arises via two pathways: directly, through reduced consumption during high-emissions hours (a shift we identified using our randomized encouragement design); and indirectly, through potential substitution from internal combustion engine (ICE) vehicles to EVs caused by lower electricity bills (also identified through our randomized encouragement) that reduce the lifetime cost of EV ownership. We believe this latter behavioral response may be important, but we acknowledge that it is uncertain and speculative.

sumer benefits, we used administrative data on daily average electricity bills per kWh. We analyzed the impact on trial participants' bills using an instrumental variables specification analogous to Equation (8), but where Y_{iht} is the mean daily bill (in £) for user i on day t. Our outcome is the amount customers paid per kwh of electricity, and we weight the regression by the consumption each day.[42]

For supplier procurement costs, we applied an analogous IV analysis. Octopus Energy provided administrative data on the cost per kWh of electricity during each half-hour of the day. These costs include both the wholesale price of power and non-energy charges such as transmission and distribution fees, but exclude grid services, such as participation in ancillary markets. The wholesale price itself is a blend of hedged prices (days, weeks, and months ahead), day-ahead prices, intra-day prices, and the final system price for the half-hour. The exact weighting of these elements is somewhat subjective and may vary over time, but we believe this measure more closely approximates the supplier's actual procurement costs than using the system price alone. Here, the outcome is Octopus Energy's total daily procurement cost.

To assess direct emissions impacts from shifting consumption to lower-CO_2e-intensity hours, we multiplied trial participant electricity consumption in each half-hour interval by the corresponding average Marginal Operating Emissions Rate from WattTime, and aggregated these values to the daily level.[43] We analyzed the impact on CO_2e from electricity consumption using an instrumental variables specification analogous to Equation (8), but where Y_{iht} is CO_2e from electricity consumption (in grams of CO_2e) for trial participant i on day t.[44] To estimate the potential CO_2e abatement benefits from induced substitution of internal combustion engine (ICE) vehicles to EVs, we combined IO Go bill savings with existing estimates of EV price sensitivity.[45]

[42]Since IO Go customers typically charge only on a subset of days, a simple unweighted comparison across all days would dilute the treatment effect by including many days with no charging activity.

[43]As described in (as described in Section 2.6, the Marginal Operating Emissions Rate is the emissions associated with the marginal change in load on the grid (WattTime, 2022).

[44]We monetize this using the UK government's SCC, which is approximately £250 per tonne of CO_2e. This is calculated by estimating the marginal abatement cost (i.e., resource costs) per tonne of CO_2. The government estimates the amount of CO_2e that is needed to meet the UK's future CO_2e targets and walks up the marginal abatement cost curve until it hits that CO_2e target, which hits costs at £250 per tonne of CO_2e. This UK government approach is in contrast to how other countries, like the US, estimate the social cost of carbon. Those other countries use the marginal damage per tonne of CO_2e from integrated assessment models.

[45]We began by estimating the average lifetime electricity bill savings from IO Go over a ten-year vehicle lifespan, discounted to present value. We treated this as a reduction in the total cost of EV ownership. Applying a price elasticity of EV demand of -2.547 from Hahn et al. (2024), we inferred the corresponding proportional increase in EV adoption. To translate this into absolute uptake, we used estimates from Department for Transport (2024) that approximately 6.7% of UK households (1.95m of 28.8m households in the UK) purchase a new car each year. The resulting increase in EV uptake was multiplied by the difference in lifecycle CO_2e emissions between ICE vehicles (£8003.89) and EVs (£3259.66) (Hahn et al., 2024), yielding £4744.23 in CO_2e benefits per induced switch. We scaled these annual impacts using government projections of ICE vehicle sales, which decline over time (Department for Transport, 2023), and discounted future abatement to present value using a 3.5% rate recommended by HM Treasury (2020).

Finally, managed EV charging may contribute to avoiding broader electricity system costs, though the magnitude of these benefits is uncertain. Some top-down modeling studies suggest potentially large system-wide gains. We discuss the estimates from one such study in the contexts of our results in the next section.[46]

5.2 Benefits/savings from adoption of managed charging

AI managed charging presented large consumer bill benefits. Figure 12 presents the estimated benefits. We found that managed charging reduced consumer electricity bills by £0.04 per kWh, an approximately 18% decline relative to the control group. Applied to average consumption over the analysis period, this implies a total bill reduction of approximately £343 during our study.[47,48]

The electricity retailer saw similar savings in procurement costs (procurement costs include the wholesale price of power and non-energy charges such as transmission and distribution fees; 95% CI: £-261 to £766). Procurement savings should not be interpreted as additional to consumer bill savings; the similarity of their magnitudes suggests near 100% pass-through of savings to trial participants, at least based on the period of our analysis (2024).[49]

Managed charging also yielded positive environmental benefits. Direct CO_2e abatement from shifting electricity consumption to cleaner off-peak hours generated a decrease of 124kg CO_2e per trial participant in 2024, which translates to estimated benefits of £35 based on a carbon value of £287 per tonne of CO_2e emitted (Department for Energy Security and Net Zero, 2023), though with wide uncertainty (95% CI: £-253 to £287). Indirect CO_2e abatement, arising from substitution from ICE vehicles to EVs, due to lower operating costs of EVs,[50] generated a further decrease of 143kg CO_2e, or £40.9 (95% CI: £20.6

[46] In our pre-analysis plan, we pre-specified estimating CO_2e impacts using the ITT framework. We did not pre-specify the use of IV estimation for bills or CO_2e savings. We adopted the IV approach here because it more directly captures the causal effect of IO Go adoption — the quantity of substantive interest. While our pre-analysis plan focused on MVPF calculations rather than consumer bills, we now report bill savings as well, as they provide an important and policy-relevant measure of consumer benefits.

[47] Based on the average annual electricity consumption of control group trial participants not enrolled in an EV tariff (9,063 kWh).

[48] The counterfactual for our £343 annual savings estimate is the average electricity bill of the control group during the experimental period, which includes customers starting on and adopting a mix of tariffs. If instead we apply a counterfactual based on a standard flat tariff, the estimated annual savings increase to £650. This figure is calculated using our estimated consumption treatment effects from Figure 7, applied to the bill under the flat tariff. The £650 estimate represents a 34% bill reduction.

[49] The point estimate on procurement cost savings was smaller than the point estimate on bill savings, but the estimates' confidence intervals overlap each other's point estimates; also note that we have not included revenue from ancillary markets in the estimation of procurement cost savings.

[50] We are assuming that consumers value fuel efficiency in their decisions to buy a vehicle, which has some support

to £61.2), in estimated benefits. The confidence intervals for indirect CO_2e abatement given in Figure 12 reflect only the uncertainty from the regression of consumer bill savings on EV tariff adoption. This likely understates the true uncertainty, which is outside the scope of this research. Additional sources of uncertainty include the price elasticity of EV adoption, the future cost trajectory of EVs, and the relationship between EV adoption and net CO_2e damages.

We estimated a resource cost per tonne of approximately –£1,285, based on annual bill savings of £343 and total emissions reductions of 0.267 tonnes CO_2e per customer (this combines 0.124 tonnes CO_2e of direct and 0.143 tonnes CO_2e of indirect savings). Note that this calculation uses the annual savings as the cost, rather than taking the net present value of present and future savings. The £343 figure corresponds to average bill savings relative to the experimental control group; savings are £650 relative to the retailer's standard flat tariff. The latter implies a resource cost of –£2,434 per tonne (–650 ÷ 0.267). Using a more conservative measure based on supplier procurement savings (£237 per customer per year), the implied cost is –£888 per tonne (–237 ÷ 0.267). In summary, the tariff and technology combination was a very cost-effective way to reduce CO_2e emissions, much lower than the next best technology (Gosnell et al., 2020; Gillingham and Stock, 2018; Hahn et al., 2024).

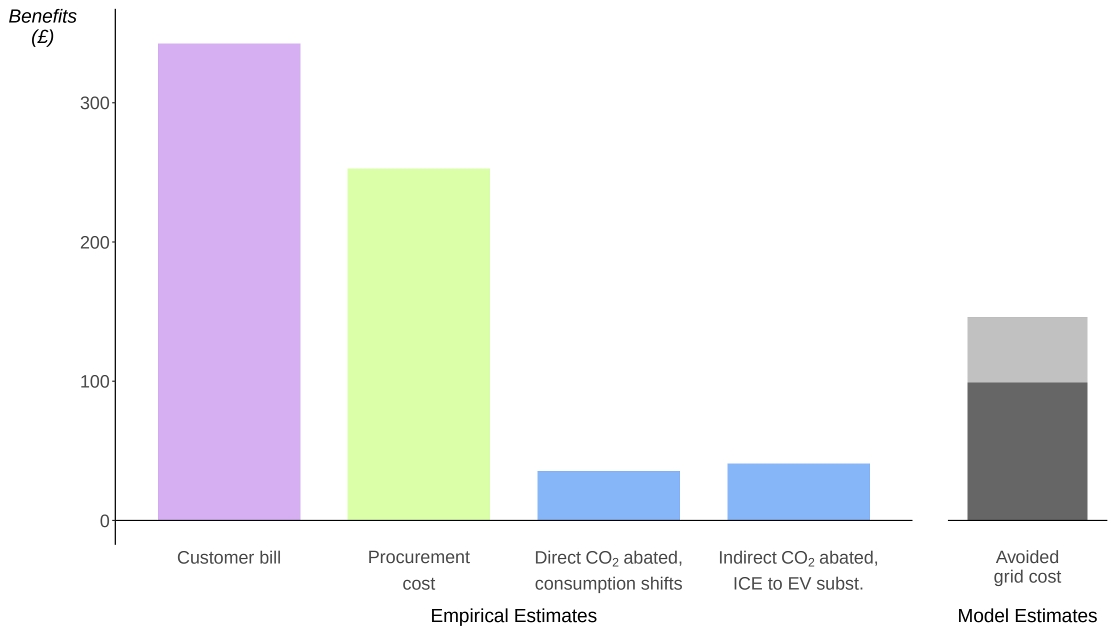
For savings from avoided system costs, we used the model in Franken et al. (2025), who estimated that fully flexible EV-related electricity demand, relative to a baseline with no flexible load, could reduce annual system costs in Great Britain by up to £0.25 billion in 2025 and as much as £4 billion by 2035.[51] These savings arose from both operational efficiencies, such as increased use of lower-marginal-cost renewable generation, and capital savings, including deferred investment in firm capacity and grid infrastructure. Notably, the study found diminishing marginal returns: the first 25% of EV users adopting managed charging account for 50% (£2 billion in 2035) of the projected savings.[52]

Grigolon et al. (2018); Forsythe et al. (2023).

[51] Their estimate was based on a whole-system linear cost optimization model, calibrated to assumptions used by the UK system operator. To the best of our knowledge, Franken et al. (2025) is the study most closely aligned with our context on two dimensions. First, it matches our outcome of interest: assessing the grid impacts of managed charging in monetary terms. By contrast, many other studies focus on EV deployment without more detailed modeling of managed charging (Heuberger et al., 2020), or in terms of electricity consumption, without translating to monetary benefits (Crozier et al., 2020). Second, it is aligned in the geographic focus on Great Britain. There are several relevant studies examining system benefits of managed charging in California (Li and Jenn, 2024) or the United States more broadly (Powell et al., 2022), but to the best of our knowledge, Franken et al. (2025) provides the most directly applicable evidence for our setting in Great Britain. For an overview of studies relevant to our setting, see Thornhill and Deasley (2018).

[52] As Franken et al. (2025) notes, "The first units of flexible EV charging tap into uncontested renewable generation, unlocking large benefits with a relatively modest flexibility rollout. However, beyond the 25% mark, excess renewable

Figure 12: Benefits of Adopting EV Tariff per Household in 2024



Notes: This figure presents estimated benefits in 2024 of adopting an EV tariff. Trial participant bill savings are derived from causal estimates using administrative data on daily average electricity bills per kWh; the estimated savings are £343 per vehicle (95% CI: £173 to £512). Procurement cost savings are estimated at £237 (95% CI: £-261 to £766). (These costs include both the wholesale price of power and non-energy charges such as transmission and distribution fees, but exclude revenue from grid services, such as participation in ancillary markets.) The similarity in the magnitudes of bill and procurement savings suggests that the supplier passed through nearly all cost reductions to customers over the 2024 analysis period. Direct CO_2e abatement reflects emissions reductions from shifting electricity consumption to off-peak periods; these are valued at £35 (95% CI: £-253 to £324), using half-hourly marginal emissions intensity data matched to observed load-shifting. Indirect CO_2e abatement from ICE to EV substitution is valued at £40.9 (95% CI: £20.6 to £61.2), estimated by combining observed IO Go bill savings with price elasticity-based projections of EV adoption and associated emissions reductions. Grid benefits are sourced from Franken et al. (2025), showing per-vehicle value under a scenario of 100% smart charging adoption. The shaded area illustrates the range of estimates across multiple modeled scenarios, where benefits range from £99 to £146 per year in 2035.

Our trial finds that nearly all peak EV load can be shifted through managed charging, and that this behavior can be sustained for over a year, consistent with the assumptions in Franken et al. (2025). This implies that system-level benefits in the range of £2-4 billion (from 2035 onward, in $£_{2024}$) could be feasible, conditional on widespread EV and tariff uptake, relative to a baseline with no flexible demand. When expressed on a per-vehicle basis, assuming 27 million EVs on the road in the UK by 2035 (National Energy System Operator, 2023), this equates to approximately £146 in system savings per vehicle in 2035 under universal managed charging.[53] Under "constrained" managed charging,[54] savings fall modestly to £120 per vehicle. When adding flexible heat demand to the

generation becomes more scarce slowing down further gains."

[53]In 2025, Franken et al. (2025) estimates the value is £94 per EV; the value is lower than in 2035 due to greater grid constraints in 2035 potentially solved by EV flexibility.

[54]In Franken et al. (2025), the constraints are: automation only between 12 am and 4 am and 12% of consumers opt out of flexible charging each day.

optimization, which cannibalizes some of the benefits from EV flexibility, the per-vehicle EV contribution falls to £99.[55]

To better understand how these reduced grid system costs might materialize, we estimated that if all EVs had used IO Go in 2024, 1.87 GW of load would have been shifted away from the highest demand hour in the country. Using demand data from the National Energy System Operator's Future Energy Scenarios, we added our estimated hourly consumption effects from Figure 7 to hourly load values, and constructed what the load duration curve would have looked like for 2024 with 100% adoption of managed charging (Panel A, Figure A6). Panel B zooms in on the top 100 consumption hours, illustrating how managed charging could significantly reduce extreme peaks.

Looking ahead to 2050, the precise hourly impacts of managed charging are uncertain, given expected changes in electricity pricing, renewable generation patterns, and climate conditions. To approximate potential effects, we modified the 2050 load duration curve by applying the estimated impacts from the four peak hours in Figure 7 to the top four hours of each day in the 2050 load profile, assuming that the displaced consumption would be evenly redistributed across the remaining hours. Under this scenario, the highest load hour in 2050 would experience approximately 5 GW less consumption (Panel C and D in Figure A6).[56]

5.3 Welfare impacts of subsidizing managed charging

We next evaluated the welfare implications of *subsidizing* managed charging, using the £50/month incentive (total £150 across three months) offered to trial participants as a proxy subsidy. We considered impacts over one, five, and 10 years (2024–2033), discounting future values at the 3.5% rate recommended by HM Treasury (2020). Ten years is a reasonable lower-bound estimate of vehicle lifetime (Bento et al., 2018; Held et al., 2021; Kolli, 2011). However, the duration of benefits attributable to the subsidy also depends on how long subsidized customers remain more likely than non-subsidized customers to adopt IO Go. This compliance advantage may decay more quickly than the

[55] These system benefits estimates are consistent with existing industry and research findings, and sit at the lower end of reported ranges. For instance, *The Utility Playbook: Turning EV Grid Risk into a $30 Billion Opportunity* (2025) projects system savings of between $145 and $575 per actively managed EV by 2035.

[56] The consumption profiles were obtained from the National Energy System Operator's Future Energy Scenarios, which provides baseline profiles for 2024 and projections for 2050 under the assumption of 0% unmanaged charging. To estimate the total load shift, our IV estimates were scaled by the total number of EVs on the road. Data on the number of battery electric vehicles (BEVs) and plug-in hybrid electric vehicles (PHEVs) in 2024 were obtained from the UK's Vehicle Licensing Statistics, indicating a total of 2,107,897 vehicles. Projections for 2050 were also taken from the Future Energy Scenarios, which forecast approximately 37.4 million EVs on UK roads by 2050.

vehicle lifetime, although in our data the encouragement effect remains relatively stable over the full year of observation.

The relevant welfare benefits in this case are (1) partial transfer of the subsidy to consumers and (2) CO_2e abatement. To estimate the share of the subsidy that constitutes a transfer, we inferred the proportion of marginal versus inframarginal adopters. Adoption was 6.98% in the email-only group and 9.31% in the £50/month group, implying that 25% of adopters were marginal.[57] We assumed inframarginal adopters (75%) valued the £150 incentive at its full face value (i.e., a 100% transfer of the total £150 to consumers). For marginal adopters (25%), we assumed an average valuation of 50% of the subsidy.[58] Combining these assumptions yields an average transfer of £0.875 per £1 of subsidy.

For welfare benefits from CO_2e abatement, we applied the estimated per-adoption benefit discussed above, but scaled it down to reflect that only 25% of adopters were induced by the subsidy.[59] We excluded reductions in trial participant bills from our welfare calculation, invoking the envelope theorem: these trial participants could have adopted IO Go without the subsidy but chose not to. We also assumed no producer surplus changes (e.g., procurement cost savings for Octopus Energy), under the assumption of a competitive retail electricity market. Finally, we excluded the indirect CO_2e benefits from increased EV adoption, since, again by the envelope theorem, a subsidy for managed charging should not affect EV uptake among trial participants who were already considering adoption.[60]

Changes in costs to the government include: (1) the subsidy itself, (2) lost VAT revenue from reduced electricity bills, (3) with greater uncertainty, a climate-related fiscal externality (increased tax revenue from higher economic growth due to the climate mitigation from CO_2e abatement), and, most speculatively, (4) avoided costs associated with electricity grid balancing. The VAT loss is calculated as 5% of the £343 annual reduction in electricity bills, totaling £12.42 less government revenue per customer. This translates to a fiscal cost of approximately £3.10 for each marginal IO Go adopter. The climate-related

[57] Calculated as $(9.31-6.98)/9.31 = 25\%$. We used this formulation of the marginals as opposed to using the elasticity approach as in Hahn et al. (2024) since we do not have a cost of adopting the tariff.

[58] For marginal adopters, we do not observe whether the first or last £1 of the subsidy induced adoption. If it were the first, the entire subsidy would be valued; if the last, the valuation would approach zero. Following the classic Harberger triangle approximation to deadweight loss (Harberger, 1964), and the approach in Hendren and Sprung-Keyser (2020) and Hahn et al. (2024), we assume a uniform distribution of latent subsidy valuations, consistent with a linear demand curve.

[59] We used the point estimate of the CO_2e impact, despite the statistical imprecision around that estimate.

[60] This welfare calculation deviates from our pre-analysis plan, but we believe these deviations more accurately reflect the full set of social returns that would accrue under real-world implementation. For more details on exact deviations, please see Appendix C.

fiscal externality is 1.07%[61] of the monetized value of CO_2e abatement attributable to IO Go adoption. This amounts to £3.22 in additional government revenue per marginal IO Go adopter.

In total, these costs and benefits imply an MVPF of 0.887 over 1 year, 0.933 over 5 years, and 0.982 over 10 years.[62] For the remainder of the discussion, we focus on the 10-year estimate, while estimates for 1 and 5 years are presented in A20.[63] This MVPF implies that for every £1 of fiscal cost to the government, the program generated £0.982 in societal benefits. The fiscal cost includes not only the direct subsidy but also the loss of VAT revenue due to lower energy bills, bringing the total cost per £1 of subsidy to £1.255. (i.e., $\frac{1.232}{1.255} = 0.982$). The ratio of these benefits to costs is then 0.982.[64]

The change in government costs from avoided grid balancing are uncertain. They may be zero. In a well-functioning electricity market, the benefits of shifting to lower-marginal-cost generation and deferring investment in generation, transmission, and distribution infrastructure should be internalized by market participants. Thus, while the overall system benefits of IO Go may be large, only a small share of those would accrue to the government. However, if there are market failures such that the value of EV flexibility is not fully internalized, there may be rationale for further policy intervention to incentivize managed charging, such as subsidies.[65] We find that if £70 (48%) of the £146 in per-customer benefits in 2035 identified by Franken et al. (2025) were borne by the government, the MVPF of the £150 subsidy would be infinite.[66]

---

[61] We assume that the UK accounts for 3.2% of global GDP (PwC, 2024), and that 33.5% of UK GDP accrues to the government as tax revenue (Office for Budget Responsibility, 2024). The PwC estimates are a weighted average of projections from national statistical authorities, EIKON from Refinitiv, IMF, Consensus Economics, the OECD, and Fitch Solutions. We therefore use the PwC estimate as it consolidates projections from these institutions into a single figure. The product of these shares yields 1.07%.

[62] As noted above, the duration of benefits depends not just on vehicle lifetime, but also on how long subsidized customers remain more likely than non-subsidized customers to adopt IO Go.

[63] We also present MVPFs using two alternative SCC values: (a) Bilal and Känzig (2024), which estimates an SCC of $1,367 for 2024, and (b) Interagency Working Group on Social Cost of Greenhouse Gases, United States Government, which estimates an SCC of $55.3 for 2024.

[64] We calculated this using:

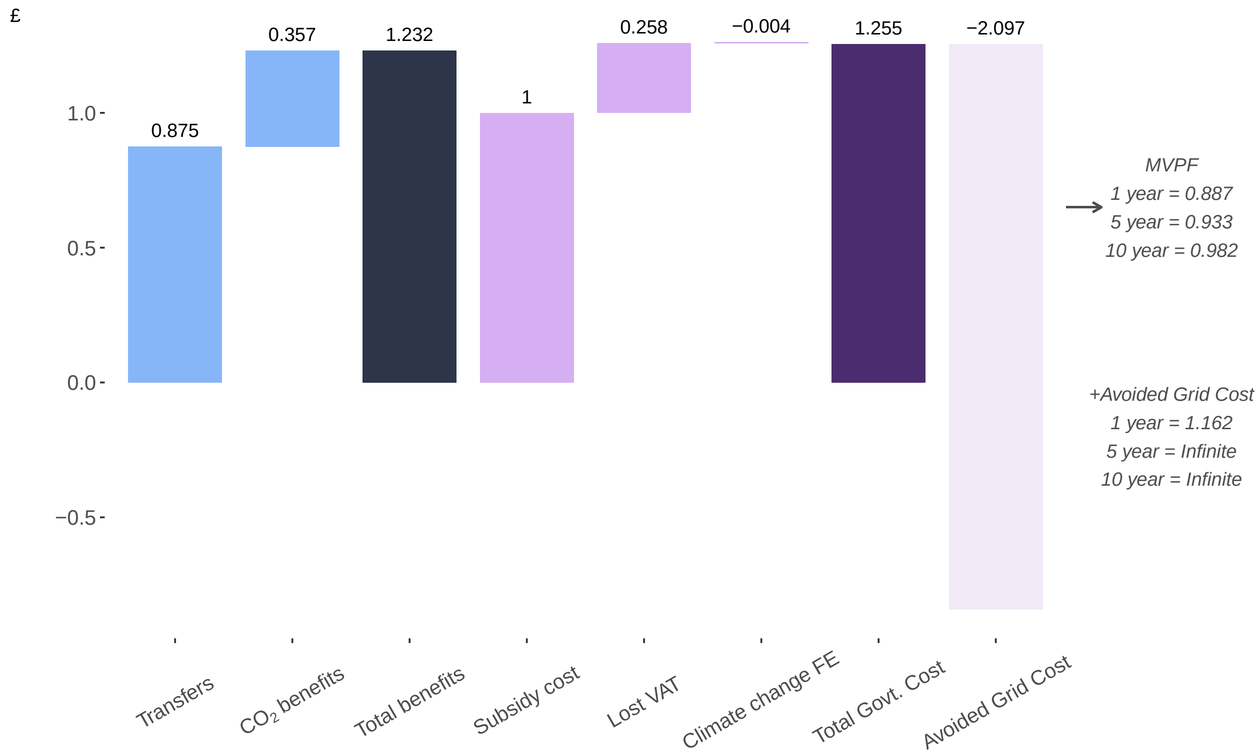
$$MVPF = \frac{xds + Edx}{xds + Vdx + Cdx + Gdx} \tag{11}$$

where x is quantity and s the subsidy. E represents CO_2 benefits to individuals; V, C and G represent VAT, climate change fiscal externalities, and avoided grid balancing costs respectively.

[65] This issue has some symmetry to the MVPF of healthcare subsidies, as the MVPF varies greatly depending on whether low-income individuals pay for their own healthcare, hospitals do, or the government does (Finkelstein et al., 2019).

[66] In this context, "infinite" does not mean literally unlimited welfare gains; rather, it is a formal term indicating that the government's net fiscal cost is negative, so, following Hendren and Sprung-Keyser (2020), the policy is treated as having an infinite MVPF (a Pareto improvement).

Figure 13: Marginal value of public funds of subsidizing managed charging over 10 years, 2024-2033



Notes: This figure presents the estimated costs and benefits of subsidizing adoption of the IO Go managed EV charging tariff over 10 years, from 2024-2033. Customer surplus is based on a decomposition of marginal and inframarginal adoption under the £50/month offer, following the approach of Hahn et al. (2024). Direct CO_2e benefits reflect emissions reductions from shifting electricity use to cleaner hours, scaled to marginal adopters. Indirect CO_2e benefits are excluded under the assumption that AI managed charging subsidies do not affect EV uptake among inframarginal adopters. Estimated costs to government include the subsidy, lost VAT revenue, and increased tax receipts from climate-related fiscal externalities. Grid balancing benefits are shown separately, based on Franken et al. (2025) estimates of per-vehicle system savings under three scenarios. Only a share of these may accrue to government.

6 Generalizability to other countries

A natural external validity question is whether the effects we estimated in the United Kingdom generalize to other markets. Octopus currently offers IO Go in seven countries, of which four — Germany, Spain, the United States (Texas), and the United Kingdom — have a substantial number of customers. Exact customer counts are commercially sensitive and cannot be disclosed, but all four markets have sufficient adoption to allow for meaningful comparisons of consumer behavior. Details on how IO Go is structured in each market are presented in Table A14.

To assess whether treatment effects are likely to hold beyond the UK setting, we begin by presenting descriptive statistics on customer behavior and charging demand patterns under IO Go in other countries. We then recalibrated our theoretical model to incorporate the behavioral patterns observed in the data, to assess the welfare margins of RTP versus IO Go.

6.1 Cross-country charging behavior

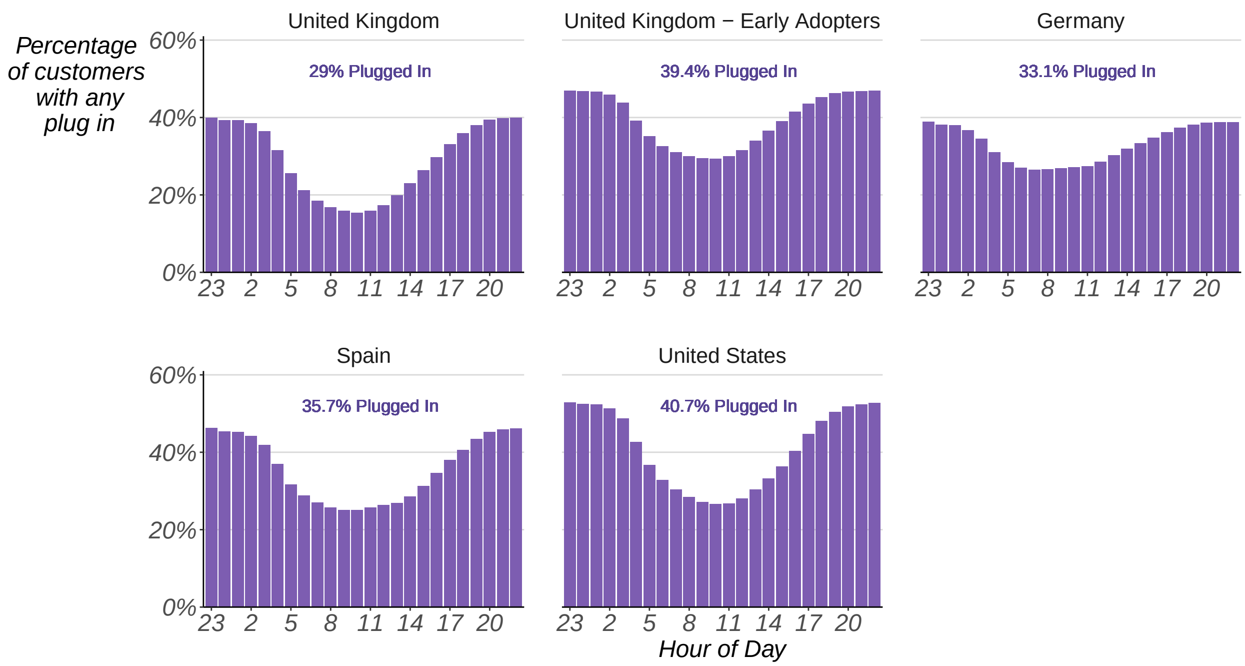
We examined three key EV charging dimensions in order to compare IO Go users across countries. First, plug-in behavior, referring to how often and when vehicles are connected to be ready to charge. Second, override ("bump") behavior, which measures how frequently users intervene in supplier managed schedules. Third, consumption profile throughout the day. Together, they show how the behavioral foundations of managed charging vary across markets.

We present these three dimensions for the four countries: UK, Germany, Spain, and United States. In addition, we also show the descriptive statistics for early adopters in the UK, as defined by those who took up IO Go in the first six months since it was rolled out. The UK has by far the most widespread usage of IO Go, with about two orders of magnitude more customers than the other countries we analyze; we therefore present statistics of early adopters, which could be more comparable to the customers from other countries.

Across countries, UK customers were plugged in the least overall, and exhibited the strongest variation in plug-in rates over the day (Figure 14). In contrast, customers in other countries kept their vehicles connected for a greater share of the day. This may be a pattern associated with earlier adopters rather than with the country itself; early adopters in the UK had higher plug-in rates than later UK adopters. The lower daytime plug-in rates among all but the earliest adopters in the UK suggest that customers there provided the AI with fewer opportunities to optimize charging, while more constant plug-in availability likely enabled greater algorithmic flexibility in Germany, Spain, and the United States.

Overrides ("bump charging") were more common outside the UK, although still infrequent overall. Many customers across all countries never used the bump function at all: 55% in the UK, 43% in Germany, 43% in Spain, and 56% in the United States (Figure 15). The highest share of bump-charged electricity occurred in Germany, where 3.3% of EV electricity consumption resulted from overrides (Figure A22), and there was an 10.7% probability of bumping for each charge-day (Table A15). The patterns suggest a behavioral substitution: customers who plug in less frequently also tend to intervene less often, whereas those who leave vehicles connected more consistently appear more likely to occasionally override the AI schedule, though note that we do not have causal evidence to support these hypotheses.

Figure 14: Plug-in Rate By Hour-of-Day

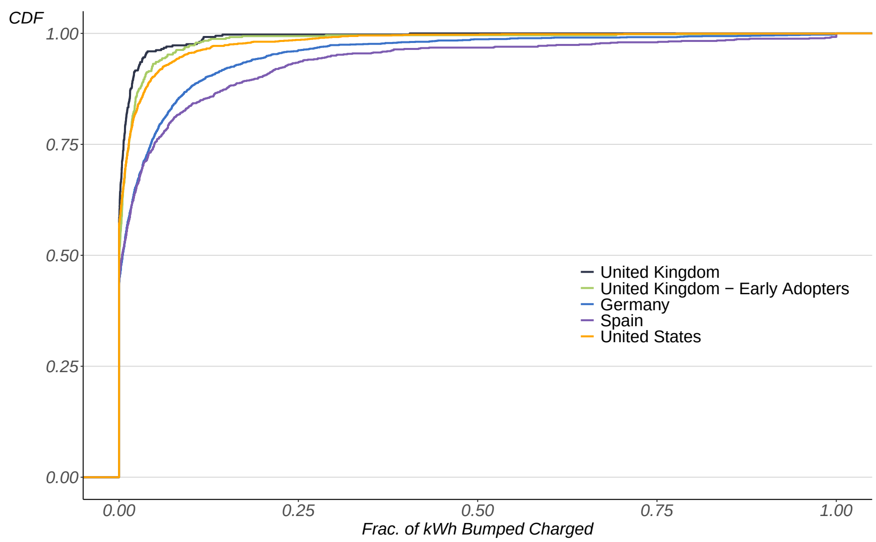


Notes: This figure plots, by hour of day, the percentage of active users who had their EVs plugged in. A user is considered active during a given hour if that hour falls between their first-ever and last-ever recorded plug-in event. We use a sample of 4,442 IO Go users across the four countries.

Despite higher rates of overriding outside the UK, the vast majority of charging still occurred during overnight periods, as shown in Figure 16. Using EV charging data, the figure plots average hourly EV consumption alongside average household electricity use for a random sample of non-EV households in each country. Across all settings, IO Go users displayed a pronounced overnight spike in charging — around 0.4 kWh per hour — indicating that most charging remains concentrated in off-peak night-time hours.

In the UK field experiment, adoption of IO Go shifted roughly 0.5 kWh per hour of consumption from peak to off-peak periods (i.e., our 42% reduction in peak consumption observed in Figure 7). The magnitude of the overnight 0.4 kWh spike observed across all four countries is broadly consistent with this shift, implying a similar volume of flexible load from EV charging. This consistency suggests that the consumption response

Figure 15: Proportion of Consumption Bump Charged Per Customer



Notes: This figure illustrates bump charging behavior, where trial participants manually overrode the AI managed charging schedule. The horizontal axis shows, for each customer, what share of their total consumption was through overriding the schedule. The vertical axis shows the cumulative proportion of customers. We use a sample of 4,442 IO Go users across the four countries.

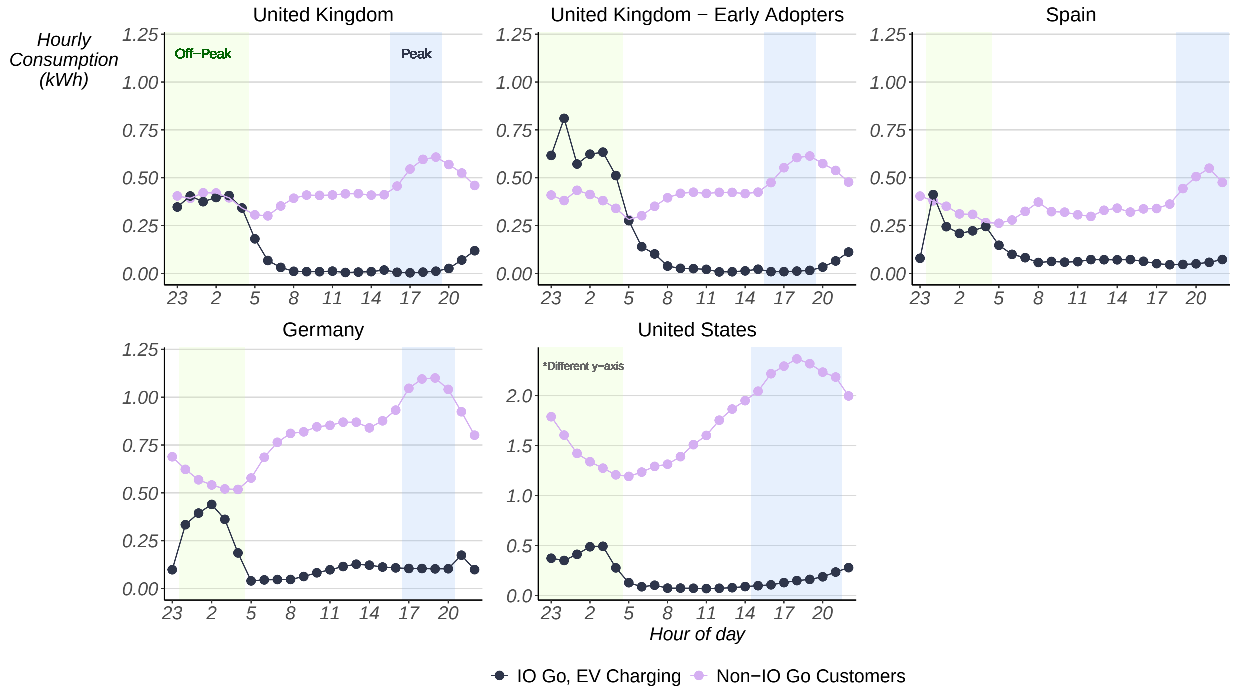
identified experimentally in the UK is likely to be similar across other markets.

In this section, we observed higher and more consistent plug-in rates in non-UK markets (Section 6.1). As we saw in Section 3.6, plug-in behavior as an important aspect of customer flexibility that influences when load shifting can occur. Thus, the observed plug-in patterns abroad align with the potential for similar flexibility conditions outside the UK. Taken together with the charging activity shown in Figure 16, these descriptive results indicate that flexibility gains associated with AI managed charging could plausibly extend across countries.

6.2 Cross-country robustness of the welfare gains of IO compared to RTP

We can now estimate welfare crossovers (for RTP vs. IO Go) across four major markets using observed override frequencies, tariff spreads, and empirical override magnitudes. Full calculations are in Appendix G.1. The expected welfare loss per override is $\hat{\lambda} = p_{\text{peak}}\, \Delta\tilde{c}_{\text{eff}}\, q^{O}$, combining each country's peak-landing probability, effective price differential, and typical override energy. With RTP elasticity $\varepsilon = -0.2$ and a baseline welfare

Figure 16: Cross-Country Consumption Profiles



Notes: The figure plots hourly electricity consumption for a random sample of customers across IO Go countries. Blue lines show EV charging consumption, based on telemetry data from IO Go users, while light purple lines show household consumption among non-EV households. All panels use 2024 data, except Panel 2, which presents 2023 consumption data for customers who adopted IO Go during its initial rollout period (September 2022 – March 2023).

gap $\Delta W_0 = 0.095$ per day,[67] we compute the crossover rate $\bar{\beta}^\star = \Delta W_0/\hat{\lambda}$ at which IO and RTP yield equal welfare. The results show that welfare losses per override vary widely – approximately £0.07 in the United Kingdom, £0.05 in Germany, £0.03 in the United States, and less than £0.01 in Spain – driven mainly by tariff spreads rather than user behavior.

Observed override rates (0.019-0.107 per day) are an order of magnitude below the corresponding partial-load crossovers. Consequently, IO remains strongly welfare-superior to RTP in all markets, even when only one-quarter of load is automated ($\alpha = 0.25$). Germany represents the tightest case, reflecting its combination of frequent overrides and moderate price differentials, but the observed behavior still lies well below the parity threshold. The United Kingdom and Spain exhibit especially large welfare margins, while the United States lies between the two extremes. Overall, these results indicate that in-

[67] The baseline welfare advantage of IO, ΔW_0, reflects the expected daily welfare gain from intelligent optimization relative to RTP in the absence of overrides, computed under an attention cost of £0.20 per session and average charging elasticity $\varepsilon = -0.2$. Where IO users allocate approximately 20% less flexible demand during peak hours compared with RTP users.

telligent charging is *behaviorally robust*: even with realistic user intervention rates, coordination gains from automated scheduling easily outweigh the welfare cost of overrides under current tariff and behavioral conditions.

7 Discussion

Our findings make three key contributions to the literature. First, we provided evidence on how real-time managed charging can reshape EV charging behavior at scale in the UK. Specifically, we found that managed charging led to a 42% reduction in household electricity demand during peak hours, with all of this demand shifted to low-cost, low-emission off-peak periods overnight. In contrast to prior observational studies or simulations, our randomized incentive design enabled clear causal inference about the elasticity of managed EV charging consumption to energy system price signals.

These consumption impacts occurred without requiring manual input or sustained behavior change from trial participants: over half of adopters never overrode the automated schedule, and override events comprised just 2.3% of electricity use. We showed that this override metric is the sufficient statistic to compare such AI traiff versus RTP, and thus highlights the potential of automation through AI to provide demand-side flexibility while respecting EV owners' preferences and constraints. The automation embedded in IO Go appeared to enhance responsiveness to grid signals, especially during the evening and overnight periods of the day.

Second, by comparing the outcomes of our randomized experiment with those from a standard difference-in-differences design, our results suggest that the impact of managed charging was relatively stable across different types of adopters. Impacts were homogeneous across cohorts of adopters in our difference-in-differences sample, and – after reweighting our difference-in-differences sample to match the field experiment sample on pre-adoption tariff, we found similar results between the two evaluations. Thus the primary source of variation in impact appears to be baseline charging behavior, particularly whether trial participants were already on smart or off-peak tariffs prior to adoption, rather than any inherent heterogeneity in responsiveness to managed charging.

Third, we quantified the welfare impacts across four dimensions – consumer costs, producer profits, environmental outcomes, and avoided costs associated with electricity grid balancing. We found that the managed charging reduced participants bills substantially. It also caused reductions in CO_2e emissions and retailer procurement costs. We generalize the findings to three other international markets with EV penetration.

References

Abel, G. A., Barclay, M. E. and Payne, R. A. (2016), 'Adjusted Indices of Multiple Deprivation to Enable Comparisons Within and Between Constituent Countries of the UK Including an Illustration Using Mortality Rates', *BMJ Open* **6**(11), e012750.

Agarwal, N., Moehring, A., Rajpurkar, P. and Salz, T. (2023), Combining human expertise with artificial intelligence: Experimental evidence from radiology, Technical report, National Bureau of Economic Research.

Agency, I. E. (2024), 'World Energy Outlook 2024'.

Aljbour, J., Wilson, T. and Patel, P. (2024), 'Powering intelligence: Analyzing artificial intelligence and data center energy consumption', *EPRI White Paper no. 3002028905* .

Allcott, H. (2011), 'Rethinking real-time electricity pricing', *Resource and Energy Economics* **33**(4), 820–842.

Antonopoulos, I., Robu, V., Couraud, B., Kirli, D., Norbu, S., Kiprakis, A., Flynn, D., Elizondo-Gonzalez, S. and Wattam, S. (2020), 'Artificial intelligence and machine learning approaches to energy demand-side response: A systematic review', *Renewable and Sustainable Energy Reviews* **130**, 109899.

Anwar, M. B., Muratori, M., Jadun, P., Hale, E., Bush, B., Denholm, P., Ma, O. and Podkaminer, K. (2022), 'Assessing the value of electric vehicle managed charging: a review of methodologies and results', *Energy and Environmental Science* **15**, 466–498.

Archsmith, J., Muehlegger, E. and Rapson, D. S. (2022), 'Future paths of electric vehicle adoption in the united states: Predictable determinants, obstacles, and opportunities', *Environmental and Energy Policy and the Economy* **3**, 71–110.

Asensio, O. I., Buckberg, E., Cole, C., Heeney, L., Knittel, C. R. and Stock, J. H. (2025), Charging uncertainty: Real-time charging data and electric vehicle adoption, Technical report, National Bureau of Economic Research.

Athey, S. and Segal, I. (2013), 'An efficient dynamic mechanism', *Econometrica* **81**(6), 2463–2485.

Bailey, J. and Axsen, J. (2015), 'Anticipating pev buyers' acceptance of utility controlled charging', *Transportation Research Part A: Policy and Practice* **82**, 29–46.

Bailey, M., Brown, D. P., Shaffer, B. and Wolak, F. A. (2024), 'Centralized vs decentralized demand response: Evidence from a field experiment'.

Bailey, M. R., Brown, D. P., Myers, E., Shaffer, B. C. and Wolak, F. A. (2025), 'Electric vehicles and the energy transition: Unintended consequences of time-of-use pricing'.

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.
URL: *https://pubs.aeaweb.org/doi/10.1257/pol.20230653*

Bento, A. M., Roth, K. and Zuo, Y. (2018), 'Vehicle lifetime and scrappage behavior: Evidence from the used car market', *The Energy Journal* **39**(1).

Bergemann, D. and Välimäki, J. (2019), 'Dynamic mechanism design: An introduction', *Journal of Economic Literature* **57**(2), 235–274.

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., Hackett, A., Metcalfe, R. and Schein, A. (2024), Decarbonizing Heat: The Impact of Heat Pumps and a Time-of-Use Heat Pump Tariff on Energy Demand, Technical report, National Bureau of Economic Research.

Bilal, A. and Känzig, D. R. (2024), 'The Macroeconomic Impact of Climate Change: Global vs. Local Temperature'.

Biswas, P., Rashid, A., Biswas, A., Nasim, M. A. A., Chakraborty, S., Gupta, K. D. and George, R. (2024), 'Ai-driven approaches for optimizing power consumption: a comprehensive survey', *Discover Artificial Intelligence* **4**(1), 116.

Björkegren, D., Choi, J. H., Budihal, D., Sobhani, D., Garrod, O. and Atherton, P. (2025), 'Could ai leapfrog the web? evidence from teachers in sierra leone', *arXiv preprint arXiv:2502.12397* .

Black, D., Panossian, N., Jingjing, L., Nordman, B., Farrell, J., Hodge, C., Meintz, A., Abdullah, M., Kisacikoglu, M., Bennett, J. et al. (2024), 'Survey and gap prioritization of us electric vehicle charge management deployments'.

Blonz, J., Palmer, K., Wichman, C. J. and Wietelman, D. C. (2025), 'Smart thermostats, automation, and time-varying prices', *American Economic Journal: Applied Economics* **17**(1), 90–125.

Bogmans, C., Gomez-Gonzalez, P., Melina, G., Miranda-Pinto, J., Pescatori, A. and Thube, S. (2025), Power hungry: How ai will drive energy demand, Technical report, International Monetary Fund.

Boopathy, P., Liyanage, M., Deepa, N., Velavali, M., Reddy, S., Maddikunta, P. K. R., Khare, N., Gadekallu, T. R., Hwang, W.-J. and Pham, Q.-V. (2024), 'Deep learning for intelligent demand response and smart grids: A comprehensive survey', *Computer science review* **51**, 100617.

Borenstein, S. (2005*a*), 'The long-run efficiency of real-time electricity pricing', *The Energy Journal* **26**(3), 93–116.

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

Borenstein, S. (2007), 'Customer risk from real-time retail electricity pricing: Bill volatility and hedgability', *The Energy Journal* **28**(2), 111–130.

Boyd, S. and Vandenberghe, L. (2004), *Convex Optimization*, Cambridge University Press.

Bradley, M. et al. (2018), 'Electric vehicle costbenefit analysis–plug-in electric vehicle cost-benefit analysis: Indiana'.

Brynjolfsson, E., Li, D. and Raymond, L. (2025), 'Generative ai at work', *The Quarterly Journal of Economics* p. qjae044.

Brynjolfsson, E., Rock, D. and Syverson, C. (2019), 'Artificial intelligence and the modern productivity paradox', *The economics of artificial intelligence: An agenda* **23**(2019), 23–57.

Budish, E. (2011), 'The combinatorial assignment problem: Approximate competitive equilibrium from equal incomes', *Journal of Political Economy* **119**(6), 1061–1103.

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.

Byrne, D. P., Martin, L. A. and Nah, J. S. (2022), 'Price discrimination by negotiation: A field experiment in retail electricity', *The Quarterly Journal of Economics* **137**(4), 2499–2537.

Callaway, B. and Sant'Anna, P. H. (2021), 'Difference-in-Differences with Multiple Time Periods', *Journal of Econometrics* **225**(2), 200–230.

Chen, S. (2025), 'How much energy will ai really consume? the good, the bad and the unknown', *Nature* **639**(8053), 22–24.

Colin Cameron, A. and Miller, D. L. (2015), 'A Practitioner's Guide to Cluster-Robust Inference', *Journal of Human Resources* **50**(2), 317–372.

Crozier, C., Morstyn, T. and McCulloch, M. (2020), 'The opportunity for smart charging to mitigate the impact of electric vehicles on transmission and distribution systems', *Applied Energy* **268**, 114973.

De Vries, A. (2023), 'The growing energy footprint of artificial intelligence', *Joule* **7**(10), 2191–2194.

Department for Energy Security and Net Zero (2023), 'Green Book Supplementary Guidance: Valuation of Energy Use and Greenhouse Gas Emissions for Appraisal'. Published by the UK Government.

Department for Transport (2023), Phasing out the sale of new petrol and diesel cars from 2030 and support for the zero-emission transition, Policy document, Department for Transport. Accessed 18 July 2025.

Department for Transport (2024), 'Vehicle licensing statistics, united kingdom: 2024', Official release via gov.uk. Accessed 18 July 2025.

Dorsey, J., Langer, A. and McRae, S. (2025), 'Fueling alternatives: Gas station choice and the implications for electric charging', *American Economic Journal: Economic Policy* **17**(1), 362–400.

Einav, L., Finkelstein, A., Ryan, S. P., Schrimpf, P. and Cullen, M. R. (2013), 'Selection on moral hazard in health insurance', *American Economic Review* **103**(1), 178–219.

Erdenesanaa, D. (2023), 'Ai could soon need as much electricity as an entire country', *The New York Times* .

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.

Few, J., Mckenna, E., Pullinger, M., Elam, S., Webborn, E. and Oreszczyn, T. (2022), Smart Energy Research Lab: Energy Use in GB Domestic Buildings 2021, Technical Report 1, University College London.

Finkelstein, A., Hendren, N. and Shepard, M. (2019), 'Subsidizing health insurance for low-income adults: Evidence from massachusetts', *American Economic Review* **109**(4), 1530–1567.

Forsythe, C. R., Gillingham, K. T., Michalek, J. J. and Whitefoot, K. S. (2023), 'Technology advancement is driving electric vehicle adoption', *Proceedings of the National Academy of Sciences* **120**(23), e2219396120.

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.

Frank, M. R., Autor, D., Bessen, J. E., Brynjolfsson, E., Cebrian, M., Deming, D. J., Feldman, M., Groh, M., Lobo, J., Moro, E. et al. (2019), 'Toward understanding the impact of artificial intelligence on labor', *Proceedings of the National Academy of Sciences* **116**(14), 6531–6539.

Franken, L., Hackett, A., Lizana, J., Riepin, I., Jenkinson, R., Lyden, A., Yu, L. and Friedrich, D. (2025), 'Power system benefits of simultaneous domestic transport and heating demand flexibility in great britain's energy transition', *Applied Energy* **377**, 124522.

Gabaix, X. (2019), Behavioral inattention, *in* 'Handbook of behavioral economics: Applications and foundations 1', Vol. 2, Elsevier, pp. 261–343.

Garcia-Osipenko, M., Kuminoff, N. V., Perry, S. and Vreugdenhil, N. (2025), Optimal second-best menu design: Evidence from residential electricity plans, Technical report, National Bureau of Economic Research.

Garg, T., Hanna, R., Myers, J., Tebbe, S. and Victor, D. G. (2024), Electric vehicle charging at the workplace: Experimental evidence on incentives and environmental nudges, Technical report, CESifo Working Paper.

Ghosh, A. and Aggarwal, V. (2017), 'Control of charging of electric vehicles through menu-based pricing', *IEEE Transactions on Smart Grid* **9**(6), 5918–5929.

Gillingham, K. and Stock, J. H. (2018), 'The cost of reducing greenhouse gas emissions', *Journal of Economic Perspectives* **32**(4), 53–72.

Gillingham, K. T., Ovaere, M. and Weber, S. M. (2025), 'Carbon policy and the emissions implications of electric vehicles', *Journal of the Association of Environmental and Resource Economists* **12**(2), 313–352.

Gosnell, G. K., List, J. A. and Metcalfe, R. D. (2020), 'The impact of management practices on employee productivity: A field experiment with airline captains', *Journal of Political Economy* **128**(4), 1195–1233.

Gravert, C. (2024), From intent to inertia: Experimental evidence from the retail electricity market, Technical report, CEBI Working Paper Series.

Green, M. (2025), 'Ai-powered vpp kraken reaches 2gw managed domestic assets', Current News. Online; accessed 17 July 2025.

Grigolon, L., Reynaert, M. and Verboven, F. (2018), 'Consumer valuation of fuel costs and tax policy: Evidence from the european car market', *American Economic Journal: Economic Policy* **10**(3), 193–225.

Hahn, R., Hendren, N., Metcalfe, R. and Sprung-Keyser, B. (2024), A Welfare Analysis of Policies Impacting Climate Change, Technical Report w32728, National Bureau of Economic Research, Cambridge, MA.

Harberger, A. C. (1964), 'The measurement of waste', *The American Economic Review* **54**(3), 58–76.

Harding, M. and Sexton, S. (2017), 'Household Response to Time-Varying Electricity Prices', *Annual Review of Resource Economics* **9**(1), 337–359.

Harrison, G. W. and List, J. A. (2004), 'Field experiments', *Journal of Economic literature* **42**(4), 1009–1055.

Heid, P., Remmy, K. and Reynaert, M. (2024), 'Equilibrium effects in complementary markets: Electric vehicle adoption and electricity pricing'.

Held, M., Ilgmann, C., Kuckshinrichs, W., Reuter, M. and Schneider, C. (2021), 'A new car fleet: From fossil-fuelled to fossil-free within 15 years?', *European Transport Research Review* **13**(4), 1–15.

Hendren, N. (2016), 'The policy elasticity', *Tax Policy and the Economy* **30**(1), 51–89.

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

Heuberger, C. F., Bains, P. K. and Mac Dowell, N. (2020), 'The EV-olution of the power system: A spatio-temporal optimisation model to investigate the impact of electric vehicle deployment', *Applied Energy* **257**, 113715.

Hildermeier, J., Burger, J., Jahn, A. and Rosenow, J. (2022), 'A review of tariffs and services for smart charging of electric vehicles in europe', *Energies* **16**(1), 88.

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.

HM Treasury (2020), 'The Green Book: Appraisal and Evaluation in Central Government'. Published by HM Treasury.

Holland, S. P. and Mansur, E. T. (2006), 'The short-run effects of time-varying prices in competitive electricity markets', *The Energy Journal* **27**(4), 127–156.

Hortaçsu, A., Madanizadeh, S. A. and Puller, S. L. (2017), 'Power to choose? an analysis of consumer inertia in the residential electricity market', *American Economic Journal: Economic Policy* **9**(4), 192–226.

Huang, Q., Jia, Q.-S., Wu, X. and Guan, X. (2023), 'Two-phase on-line joint scheduling for welfare maximization of charging station', *IEEE Transactions on Automation Science and Engineering* .

Imbens, G. and Xu, Y. (2024), 'LaLonde (1986) after Nearly Four Decades: Lessons Learned'. Issue: arXiv:2406.00827 _eprint: 2406.00827.
URL: *http://arxiv.org/abs/2406.00827*

Imelda, Fripp, M. and Roberts, M. J. (2024), 'Real-time pricing and the cost of clean power', *American Economic Journal: Economic Policy* **16**(4), 100–141.

International Energy Agency (2025*a*), 'Energy and AI'.

International Energy Agency (2025*b*), 'Global ev outlook 2025'.

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.

Jones, C. B., Vining, W., Lave, M., Haines, T., Neuman, C., Bennett, J. and Scoffield, D. R. (2022), 'Impact of electric vehicle customer response to time-of-use rates on distribution power grids', *Energy Reports* **8**, 8225–8235.

Joskow, P. L. and Tirole, J. (2006*a*), 'Merchant transmission investment', *The Journal of Industrial Economics* **54**(2), 233–264.

Joskow, P. and Tirole, J. (2006*b*), 'Retail electricity competition', *The RAND Journal of Economics* **37**(4), 799–815.

Kaack, L. H., Donti, P. L., Strubell, E., Kamiya, G., Creutzig, F. and Rolnick, D. (2022), 'Aligning artificial intelligence with climate change mitigation', *Nature Climate Change* **12**(6), 518–527.

Khanna, S., Martin, R. and Muûls, M. (2024), 'Shifting household power demand across time: incentives and automation', *Work. Pap., Imp. Coll. London* .

Knittel, C. R., Senga, J. R. L. and Wang, S. (2025), Flexible data centers and the grid: Lower costs, higher emissions?, Technical report, National Bureau of Economic Research.

Kolbert, E. (2024), 'The obscene energy demands of ai', *The New Yorker* **9**.

Kolli, K. (2011), 'Car longevity: a biometric approach', *Population* **66**(2), 307–323.

La Nauze, A., Friesen, L., Lim, K. L., Menezes, F., Page, L., Philip, T. and Whitehead, J. (2024), Can electric vehicles aid the renewable transition? evidence from a field experiment incentivising midday charging, Technical report, CESifo Working Paper.

LaLonde, R. J. (1986), 'Evaluating the Econometric Evaluations of Training Programs with Experimental Data', *American Economic Review* **76**(4), 604–620. JSTOR: 1806062.

Li, Y. and Jenn, A. (2024), 'Impact of electric vehicle charging demand on power distribution grid congestion', *Proceedings of the National Academy of Sciences* **121**(18), e2317599121.

Lowell, D., Jones, B. and Seamonds, D. (2017), 'Mjb&a analyzes state-wide costs and benefits of plug-in vehicles in five northeast and mid-atlantic states', *MJ Bradley & Associates* .

Mogstad, M., Torgovitsky, A. and Walters, C. R. (2021), 'The Causal Interpretation of Two-

Stage Least Squares with Multiple Instrumental Variables', *American Economic Review* **111**(11), 3663–3698.

Moore, R. T. and Schnakenberg, K. (2023), 'blockTools: Blocking, Assignment, and Diagnosing Interference in Randomized Experiments v.0.6.4'.

Moser, C. (2017), 'The role of perceived control over appliances in the acceptance of electricity load-shifting programmes', *Energy Efficiency* **10**(5), 1115–1127.

National Energy System Operator (2023), Future energy scenarios (2023), Technical report, National Energy System Operator.

Nicolson, M., Huebner, G. M., Shipworth, D. and Elam, S. (2017), 'Tailored Emails Prompt Electric Vehicle Owners To Engage With Tariff Switching Information', *Nature Energy* **2**(6), 17073.
URL: *https://www.nature.com/articles/nenergy201773*

Nicolson, M. L., Fell, M. J. and Huebner, G. M. (2018), 'Consumer demand for time of use electricity tariffs: A systematized review of the empirical evidence', *Renewable and Sustainable Energy Reviews* **97**, 276–289.

NYSERDA (2021), Benefit-cost analysis of electric vehicle deployment in new york state, Technical report, NYSERDA.

Office for Budget Responsibility (2024), 'The UK's Tax Burden in Historical and International Context'. Accessed: 25 July 2024.

Parkes, D. C. and Wellman, M. P. (2015), 'Economic reasoning and artificial intelligence', *Science* **349**(6245), 267–272.

Parsons, A. and mySociety (2021), 'Composite 2020 UK Index of Multiple Deprivation (England-Anchored) v.3.3.0'.

Pébereau, C. and Remmy, K. (2023), 'Barriers to real-time electricity pricing: Evidence from new zealand', *International Journal of Industrial Organization* **89**, 102979.

Poletti, S. and Wright, J. (2020), 'Real-time pricing and imperfect competition in electricity markets', *The Journal of Industrial Economics* **68**(1), 93–135.

Powell, S., Cezar, G. V., Min, L., Azevedo, I. M. and Rajagopal, R. (2022), 'Charging infrastructure access and operation to reduce the grid impacts of deep electric vehicle adoption', *Nature Energy* **7**(10), 932–945.

PwC (2024), 'Global Economy Watch: Projections'. Accessed: 25 July 2024.

Rapson, D. and Muehlegger, E. (2023*a*), 'Global transportation decarbonization', *Journal of Economic Perspectives* **37**(3), 163–188.

Rapson, D. S. and Muehlegger, E. (2023*b*), 'The economics of electric vehicles', *Review of Environmental Economics and Policy* **17**(2), 274–294.

Rigas, E. S., Ramchurn, S. D. and Bassiliades, N. (2014), 'Managing electric vehicles in the smart grid using artificial intelligence: A survey', *IEEE Transactions on Intelligent Transportation Systems* **16**(4), 1619–1635.

Roth, J. (2024), 'Interpreting Event-Studies from Recent Difference-in-Differences Methods'. Place: arXiv.

Sandalow, D. B., McCormick, C., Kucukelbir, A., Friedmann, J., Nachmany, M., Lee, H., Hill, A., Loehr, D., Wald, M., Halff, A. M. et al. (2024), 'Artificial intelligence for climate change mitigation roadmap'.

Schweppe, F. C., Tabors, R. D. and Kirtley, J. L. (1981), 'Homeostatic control: the utility-customer marketplace for electric power'.

Schweppe, F. C., Tabors, R. D., Kirtley, J. L., Outhred, H. R., Pickel, F. H. and Cox, A. J. (1980), 'Homeostatic utility control', *IEEE Transactions on Power Apparatus and Systems* (3), 1151–1163.

Seamonds, D. and Lowell, D. (2018), 'Electric vehicle cost-benefit analysis'.

SEPA (2024), The state of managed charging in 2024, Technical report, Smart Electic Power Alliance.

Si, C., Chau, K., Liu, W. and Hou, Y. (2025), 'Optimal scheduling for electric vehicle charging: A review of methods, technologies, and uncertainty management', *Journal of Energy Storage* **131**, 117500.

The Utility Playbook: Turning EV Grid Risk into a $30 Billion Opportunity (2025), Technical report, ev.energy.

Thornhill, C. and Deasley, S. (2018), 'My Electric Benefits Case: Peer Review by Frontier Economics'.

Turk, G., Schittekatte, T., Martínez, P. D., Joskow, P. L. and Schmalensee, R. (2024), *Designing Distribution Network Tariffs Under Increased Residential End-user Electrification: Can the US Learn Something from Europe?*, MIT Center for Energy and Environmental Policy Research.

Varian, H. R. (1995), 'Economic Mechanism Design for Computerized Agents'.

Wang, Z., Zheng, F. and Liu, M. (2025), 'Charging Scheduling of Electric Vehicles Considering Uncertain Arrival Times and Time-of-Use Price', *Sustainability* **17**(3), 1100.

WattTime (2022), Marginal Emissions Modeling: WattTime's Approach to Modeling and Validation, Technical report.

Yaghoubi, E., Yaghoubi, E., Khamees, A., Razmi, D. and Lu, T. (2024), 'A systematic review and meta-analysis of machine learning, deep learning, and ensemble learning approaches in predicting ev charging behavior', *Engineering Applications of Artificial Intelligence* **135**, 108789.

Yi, H., Lin, Q. and Chen, M. (2019), Balancing cost and dissatisfaction in online ev charging under real-time pricing, *in* 'IEEE INFOCOM 2019-IEEE Conference on Computer Communications', IEEE, pp. 1801–1809.

Zhou, Y. and Li, S. (2018), 'Technology adoption and critical mass: the case of the us electric vehicle market', *The Journal of Industrial Economics* **66**(2), 423–480.

Appendix

Table of Contents

A Tables

Table A1: Experimental Balance

Variable	Control Mean	Email + £0/Mth	Email + £5/Mth	Email + £50/Mth	Email + £50/Mth (No Bump)
Not Always Octopus Customer	0.26	0.00	0.00	0.00	0.00
		(0.01)	(0.02)	(0.02)	(0.02)
Octopus Tenure	2.72	-0.03	-0.01	-0.03	-0.01
	[1.63]	(0.04)	(0.06)	(0.06)	(0.06)
Proportion Total kWh Peak	0.23	0.00	0.00	0.00	0.00
	[0.08]	(0.00)	(0.00)	(0.00)	(0.00)
Smart Tariff Onboarding Processes	0.22	0.00	0.00	-0.01	0.00
	[0.47]	(0.01)	(0.02)	(0.02)	(0.02)
Structural Winnings (GBP/kWh)	752.64	-9.72	-17.76	-26.70	-18.12
	[684.94]	(16.50)	(25.08)	(23.84)	(25.02)
Total kWh	4,082.53	-42.42	-33.33	-65.43	-38.96
	[2,599.11]	(62.80)	(97.97)	(91.35)	(94.82)
Total kWh Stdev	0.77	-0.01	-0.01	0.00	0.00
	[0.25]	(0.01)	(0.01)	(0.01)	(0.01)
N	2,205	7,720	1,101	1,102	1,105

Note: The first column shows the mean and [standard deviation] for the control group. Each row and each subsequent encouragement column represents an individual regression of the row variable on an indicator for receiving the encouragement in the column. The encouragements appear to be balanced on baseline characteristics. The standard errors are in parentheses. Density plots of these covariates are shown in Figure A2. Detailed definitions of these covariates can be found in Appendix D.2.

Table A2: Impact of Encouragements on Take-up of IO Go

Dependent Variable:	IO GO
Model:	(1)
Variables	
Email + £0/Month	0.034***
	(0.004)
Email + £5/Month	0.033***
	(0.007)
Email + £50/Month	0.059***
	(0.008)
Email + £50/Month (No Bump)	0.057***
	(0.008)
Post-Incentive Period	0.004
	(0.004)
Email + £0/Month × Post-Incentive Period	-0.003
	(0.004)
Email + £5/Month × Post-Incentive Period	-0.006
	(0.007)
Email + £50/Month × Post-Incentive Period	-0.003
	(0.006)
Email + £50/Month (No Bump) × Post-Incentive Period	-0.016**
	(0.006)
Fixed-effects	
Block	Yes
Week of Year	Yes
Fit statistics	
Control Mean (Incentive Period)	0.0365
Test £0 = £50	0.002
Test £0 = £50 (No Bump)	0.003
Observations	661,891

Clustered (Participant & Week of Year) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Note: This table shows the effects of four email-based encouragements on adopting IO Go in the 12 months since emails were sent out. The outcome is a binary indicator for weekly use of the Octopus Go tariff. Encouragement indicators are interacted an indicator for whether the week is during the incentive period, three months after the start of the trial. The specification controls for fixed effects for randomization block and week-of-year. Standard errors, clustered by participant and week, are reported in the parentheses. Mean IO Go take-up rate during the incentive period is reported for the control group. “Test £0 = £50" is the p-value on the test of equality between the first and third coefficient; “Test £0 = £50 (No bump)" is the p-value on the test of equality between the first and fourth coefficient.

Table A3: Robustness, Impact of Encouragements on Take-up of IO Go, Daily Data

Dependent Variable:	IO GO
Model:	(1)
Variables	
Email + £0/Month	0.034***
	(0.004)
Email + £5/Month	0.034***
	(0.007)
Email + £50/Month	0.059***
	(0.008)
Email + £50/Month (No Bump)	0.057***
	(0.008)
Post-Incentive Period	0.004
	(0.005)
Email + £0/Month × Post-Incentive Period	-0.003
	(0.005)
Email + £5/Month × Post-Incentive Period	-0.006
	(0.007)
Email + £50/Month × Post-Incentive Period	-0.004
	(0.007)
Email + £50/Month (No Bump) × Post-Incentive Period	-0.017**
	(0.007)
Fixed-effects	
Block	Yes
Day	Yes
Fit statistics	
Control Mean (Incentive Period)	0.023
Test £0 = £50	0.002
Test £0 = £50 (No Bump)	0.003
Observations	4,580,025

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

Note: This table shows the effects of four email-based encouragements on adopting IO Go in the 12 months since emails were sent out. The outcome is a binary indicator for daily use of the Octopus Go tariff. Encouragement indicators are interacted an indicator for whether the date is during the incentive period, which is the three months after the start of the trial. The specification controls for fixed effects for randomization block and week-of-year. Standard errors, clustered by participant and day, are reported in the parentheses. Mean IO Go take-up rate during the incentive period is reported for the control group. "Test £0 = £50" is the p-value on the test of equality between the first and third coefficient; "Test £0 = £50 (No bump)" is the p-value on the test of equality between the first and fourth coefficient.

Table A4: Impact of Encouragements on Electricity Consumption

	Peak	Off-Peak	Overall	Peak	Off-Peak	Overall
Model:	(1)	(2)	(3)	(4)	(5)	(6)
Variables						
Email + £0/Month	-0.029***	0.018	-0.005			
	(0.011)	(0.012)	(0.007)			
Email + £5/Month	-0.011	0.018	0.004			
	(0.016)	(0.017)	(0.010)			
Email + £50/Month	-0.013	0.025	0.003			
	(0.016)	(0.017)	(0.010)			
Email + £50/Month (No Bump)	-0.033**	0.026	-0.006			
	(0.016)	(0.018)	(0.010)			
Any Encouragement				-0.026**	0.019*	-0.004
				(0.010)	(0.012)	(0.007)
Fixed-effects						
Week of Year	Yes	Yes	Yes	Yes	Yes	Yes
Block	Yes	Yes	Yes	Yes	Yes	Yes
Fit statistics						
Control Mean	1.4	1	1	1.4	1	1
Observations	2,646,712	4,631,639	15,879,842	2,646,712	4,631,639	15,879,842

Clustered (Participant & Week of Year) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Note: Each column reports coefficients from separate regressions of hourly electricity consumption (in kWh) on indicators for encouragement assignment. Columns (1), (2), and (3) estimate intention-to-treat effects for each arm; columns (4), (5), and (6) pool all treatment arms into a single binary indicator. The dependent variable is consumption during either the peak (16:30–20:30), off-peak (23:30–05:30), or overall hours. All regressions control for baseline consumption, and include fixed effects for week-of-year and randomization block. Standard errors, clustered by participant and week, are reported in parentheses. Means consumption for the control group are reported at the bottom of each panel.

Table A5: Impact of Encouragements on Octopus Go Take-up

Dependent Variable:	Octopus Go
Model:	(1)
Variables	
Email + £0/Month	0.0101***
	(0.0031)
Email + £5/Month	0.0099*
	(0.0052)
Email + £50/Month	0.0045
	(0.0049)
Email + £50/Month (No Bump)	0.0066
	(0.0051)
Fixed-effects	
Block	Yes
Week of Year	Yes
Fit statistics	
Control Mean	0.025
Observations	661,891

Clustered (Participant & Week of Year) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Note: This table shows the effects of four email-based encouragements on adopting Octopus Go in the 12 months since emails were sent out. The outcome is a binary indicator of for weekly use of the Octopus Go tariff. The specification controls for fixed effects for randomization block and week. Standard errors, clustered by participant and week, are reported in parentheses. Mean Octopus Go take-up rate is reported for the control group.

Table A6: Selection into IO Go and Baseline Characteristics

Variable	(1)	(2)	(3)	(4)
Email + £0/Month	0.044	0.043	0.044	0.041
	(0.005)	(0.005)	(0.005)	(0.005)
Email + £5/Month	0.042	0.042	0.042	0.039
	(0.008)	(0.008)	(0.008)	(0.008)
Email + £50/Month	0.060	0.059	0.059	0.055
	(0.009)	(0.009)	(0.009)	(0.009)
Email + £50/Month (No Bump)	0.062	0.061	0.062	0.057
	(0.009)	(0.009)	(0.009)	(0.009)
Structural Winnings (Z-Score)	-0.012	-0.005	-0.002	
	(0.002)	(0.004)	(0.005)	
Baseline TOU Tariff		0.023	0.025	0.018
		(0.007)	(0.009)	(0.008)
Baseline kWh		-0.008	-0.009	-0.007
		(0.004)	(0.005)	(0.004)
IMD Tercile 2		0.038	0.040	0.035
		(0.007)	(0.008)	(0.008)
IMD Tercile 3		0.034	0.039	0.034
		(0.007)	(0.008)	(0.007)
Octopue Tenure (Years)		-0.006	-0.006	-0.005
		(0.001)	(0.002)	(0.002)
Covariates			+ Covariates interacted	+ Nonparametric struc. winnings
Sq. Corr. Coef	0.0043	0.0072	0.0085	0.0093
p-score range	[0.035, 0.19]	[0.023, 0.22]	[0.0002, 0.32]	[0.006, 0.25]
p-score R2 with (1)		0.60	0.51	0.44

Note: This table shows the estimation results of a logit regression of takeup of IO Go on encouragement group and participants' baseline characteristics. We show the marginal effects at the means of the covariates.

Table A7: Robustness, Impact of EV Tariff on Peak Consumption (kWh)

	Main	4 Instruments	£0/Mth	£5/Mth	£50/Mth	£50/Mth (No Bump)	6 Months	No Baseline	No Week FE
Model:	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Variables									
EV Tariff	-0.581**	-0.498**	-0.697***	-0.269	-0.305	-0.587**	-0.473**	-0.554	-0.568**
	(0.224)	(0.209)	(0.247)	(0.338)	(0.230)	(0.236)	(0.214)	(0.336)	(0.224)
Fixed-effects									
Week of Year	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	
Block	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Fit statistics									
Control Mean	1.4	1.4	1.4	1.4	1.4	1.4	1.4	1.4	1.4
Observations	2,646,712	2,646,712	1,980,876	663,401	660,592	662,359	1,307,900	2,646,712	2,646,712
First Stage F-Stat	52.720	14.066	44.154	20.302	37.764	40.647	78.180	52.715	52.910

Clustered (Participant & Week of Year) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Note: This table reports IV estimates of the effect of adopting an EV tariff on hourly electricity consumption (kWh) during peak hours (16:30-20:30). The instrument is an indicator for assignment to any email-based encouragement. Columns (1) is our main specification, also reported in Figure 7; (2) defines a separate instrument for each encouragement group; (3) restricts to just the £0/Month group and the control group; (4) restricts to just the £5/Month group and the control group; (5) restricts to just the £50/Month group and the control group; (6) just the £50/Month (No Bump) group and the control group; (7) restricts to the 6 months after encouragement emails were sent out; (8) does not control for baseline consumption. All specifications control for baseline consumption (except column 6), and fixed effects for randomization block and week. Standard errors, clustered by participant and week, are reported in parentheses. The control mean is calculated over the same time periods for control group trial participants not enrolled in an EV tariff. The bottom row shows the first-stage Wald F-statistic.

Table A8: Robustness, Impact of EV Tariff on Off-Peak Consumption (kWh)

	Main	4 Instruments	£0/Mth	£5/Mth	£50/Mth	£50/Mth (No Bump)	6 Months	No Baseline	No Week FE
Model:	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Variables									
EV Tariff	0.481*	0.468*	0.501*	0.284	0.280	0.414	0.686***	0.118	0.488*
	(0.261)	(0.244)	(0.282)	(0.400)	(0.269)	(0.270)	(0.236)	(0.392)	(0.260)
Fixed-effects									
Week of Year	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	
Block	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Fit statistics									
Control Mean	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99
Observations	3,970,004	3,970,004	2,971,234	995,188	990,964	993,611	1,961,825	3,970,004	3,970,004
First Stage F-Stat	53.449	14.283	44.658	20.564	38.473	41.213	78.912	52.756	53.647

Clustered (Participant & Week of Year) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Note: This table reports IV estimates of the effect of adopting an EV tariff on hourly electricity consumption (kWh) during off-peak hours (23:30-05:30). The instrument is an indicator for assignment to any email-based encouragement. Columns (1) is our main specification, also reported in Figure 7; (2) defines a separate indicator for each encouragement group; (3) restricts to just the £0/Month group and the control group; (4) restricts to just the £50/Month group and the control group; (5) just the £50/Month (No Bump) group and the control group; (6) restricts to the 6 months after encouragement emails were sent out; (7) we do not control for baseline consumption. All specifications control for baseline consumption (except column 6), and fixed effects for randomization block and week. Standard errors, clustered by participant and week, are reported in parentheses. The control mean is calculated over the same time periods for control group trial participants not enrolled in an EV tariff. The bottom row shows the first-stage Wald *F*-statistic.

Table A9: Robustness, Impact of EV Tariff on Overall Consumption (kWh)

	Main	4 Instruments	£0/Mth	£5/Mth	£50/Mth	£50/Mth (No Bump)	6 Months	No Baseline	No Week FE
Model:	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Variables									
EV Tariff	-0.079	-0.060	-0.125	0.032	-0.038	-0.132	0.002	-0.203	-0.071
	(0.150)	(0.140)	(0.163)	(0.226)	(0.150)	(0.159)	(0.154)	(0.247)	(0.150)
Fixed-effects									
Week of Year	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	
Block	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Fit statistics									
Control Mean	1	1	1	1	1	1	1	1	1
Observations	15,879,842	15,879,842	11,884,810	3,980,493	3,963,619	3,974,217	7,847,164	15,879,842	15,879,842
First Stage F-Stat	52.809	14.090	44.257	20.465	37.912	40.763	78.281	52.747	53.001

Clustered (Participant & Week of Year) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Note: This table reports IV estimates of the effect of adopting an EV tariff on electricity consumption over the whole day. The instrument is an indicator for assignment to any email-based encouragement. Columns (1) is our main specification; (2) defines a separate indicator for each encouragement group; (3) restricts to just the £0/Month group and the control group; (4) restricts to just the £50/Month group and the control group; (5) just the £50/Month (No Bump) group and the control group; (6) restricts to the 6 months after encouragement emails were sent out; (7) we do not control for baseline consumption. All specifications control for baseline consumption (except column 6), and fixed effects for randomization block and week. Standard errors, clustered by participant and week, are reported in parentheses. The control mean is calculated over the same time periods for control group trial participants not enrolled in an EV tariff. The bottom row shows the first-stage Wald *F*-statistic.

Table A10: Heterogeneity of IO Go Take-up, by IMD and Baseline Electricity Consumption

Model:	x IMD (1)	x Baseline kWh (2)
Variables		
Email + £0/Mth × Tercile 1	0.021	0.029***
	(0.018)	(0.010)
Email + £0/Mth × Tercile 2	0.045***	0.050***
	(0.009)	(0.009)
Email + £0/Mth × Tercile 3	0.026***	0.016*
	(0.007)	(0.009)
Email + £5/Mth × Tercile 1	0.001	0.035**
	(0.026)	(0.016)
Email + £5/Mth × Tercile 2	0.052***	0.033**
	(0.017)	(0.015)
Email + £5/Mth × Tercile 3	0.022*	0.019
	(0.011)	(0.015)
Email + £50/Mth × Tercile 1	0.009	0.069***
	(0.031)	(0.018)
Email + £50/Mth × Tercile 2	0.062***	0.064***
	(0.018)	(0.017)
Email + £50/Mth × Tercile 3	0.059***	0.036**
	(0.013)	(0.016)
Email + £50/Mth (No Bump) × Tercile 1	0.050	0.080***
	(0.033)	(0.019)
Email + £50/Mth (No Bump) × Tercile 2	0.062***	0.045***
	(0.017)	(0.016)
Email + £50/Mth (No Bump) × Tercile 3	0.035***	0.009
	(0.012)	(0.015)
Tercile 2	-0.009	-0.013
	(0.017)	(0.012)
Tercile 3	0.010	-0.0005
	(0.017)	(0.013)
Fixed-effects		
Block	Yes	Yes
Week of Year	Yes	Yes
Fit statistics		
Observations	15,879,842	15,879,842

Clustered (Participant & Week of Year) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Note: This table presents heterogeneity in impact of encouragements on take-up of IO Go, interacting encouragement group with (1) Index of Multiple Deprivation (IMD) terciles and (2) baseline consumption terciles. The outcome is a binary indicator for weekly use of the Octopus Go tariff. The specification controls for baseline consumption, and fixed effects of randomization block and week of year. Standard errors, clustered by participant week, are reported in parentheses.

Table A11: Heterogeneity of Impact of EV Tariffs, by IMD and Baseline Electricity Consumption

	x IMD		x Baseline kWh	
	Off-Peak	Peak	Off-Peak	Peak
Model:	(1)	(2)	(3)	(4)
Variables				
(EV Tariff) x (Tercile 1)	0.924	1.16	0.295	-0.300
	(0.960)	(1.08)	(0.291)	(0.221)
(EV Tariff) x (Tercile 2)	0.604	-0.882**	0.504	-0.399
	(0.388)	(0.383)	(0.316)	(0.277)
(EV Tariff) x (Tercile 3)	0.508*	-0.207	0.998	-0.759
	(0.297)	(0.269)	(0.778)	(0.725)
(Other Tariff) x (Tercile 2)	0.078	0.327**	0.059	0.182***
	(0.127)	(0.140)	(0.072)	(0.060)
(Other Tariff) x (Tercile 3)	0.098	0.218	0.131	0.441***
	(0.127)	(0.138)	(0.134)	(0.125)
Fixed-effects				
Block	Yes	Yes	Yes	Yes
Week of Year	Yes	Yes	Yes	Yes
Fit statistics				
Observations	3,970,004	2,646,712	3,970,004	2,646,712

Clustered (Participant & Week of Year) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Note: This table presents heterogeneity in IV effects of EV tariff adoption on hourly electricity consumption during off-peak (23:30-5:30) and peak (16:30-20:30) periods, interacting EV tariff adoption with (1) Index of Multiple Deprivation (IMD) terciles and (2) baseline consumption terciles. All specifications control for baseline consumption, and fixed effects for randomization block and week. Standard errors, clustered by participant and week, are reported in parentheses.

Table A12: Bump Charging and Baseline Characteristics

Dependent Variable:	I(Ever Bumped)
Model:	(1)
Variables	
Constant	0.385***
	(0.050)
ZEmail+£0/Mth	0.004
	(0.032)
ZEmail+£5/Mth	-0.063
	(0.047)
ZEmail+£50/Mth	-0.002
	(0.045)
ZEmail+£50/Mth(NoBump)	-0.002
	(0.046)
Frac. of kWh During Peak (Z-Score)	0.039***
	(0.012)
Total Consumption (Z-Score)	-0.006
	(0.015)
Structural Winnings (Z-Score)	0.005
	(0.016)
IMD Tercile 2	0.029
	(0.043)
IMD Tercile 3	0.052
	(0.041)
Octopus Tenure (Years)	0.007
	(0.007)
I(Baseline TOU)	0.078**
	(0.033)
Fit statistics	
Observations	2,146
Dependent variable mean	0.45573

IID standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Note: This table reports regression results where the dependent variable is an indicator for whether a participant has ever engaged in bump charging. Explanatory variables include the encouragement group assignment and baseline characteristics. Detailed definitions of these baseline characteristics can be found in Appendix D.2.

Table A13: Baseline Differences Between IO Go and Octopus Go Trial Participants

Variable	IO Go Mean	Octopus Go Mean	IO Go - Octopus Go	p-value
Octopus Tenure	2.51	2.58	-0.06	0.34
	[1.65]	[1.64]		
Total kWh	3912.32	3753.20	159.12	0.076*
	[2343.05]	[2174.77]		
Total kWh Stdev	0.77	0.77	0.00	0.74
	[0.24]	[0.22]		
Not Always Octopus Customer	0.24	0.24	0.00	0.95
	[0.43]	[0.43]		
Smart Tariff Onboarding Processes	0.34	0.26	0.08	0.00025***
	[0.58]	[0.51]		
Structural Winnings (GBP/kWh)	687.70	684.59	3.11	0.9
	[630.75]	[592.95]		
Proportion Total kWh Peak	0.22	0.23	-0.01	0.057*
	[0.08]	[0.08]		

Note: Column (1) shows the mean and [standard deviation] for IO Go users; column (2) reports the same for Octopus Go participants; column (3) shows the difference in means between columns (1) and (2); column (4) reports the p-value from a two-sided t-test of equality of means. Detailed definitions of these covariates can be found in Appendix D.2.

Table A14: Overview of IO Go Tariff Features by Country

Country	**Cheap charging**	**Price during charging window**	**Additional Benefits**	**Launch Date**
UK	6 hour window: 23:30–05:30	7p/kWh	7p/kWh for the whole home between 23:30 and 05:30	May 2022
US	Dynamic based on forecasts	As per regular tariff price	Lower per kWh rate when connected to IO Go	December 2022
Germany (NB: changed as of June 2025)	5 hour window: 00:00–05:00	maximum of 0.20/kWh	Super-cheap electricity for the entire home between 00:00 and 05:00	August 2023
Spain	Dynamic based on forecasts	0.07/kWh	n/a	September 2024

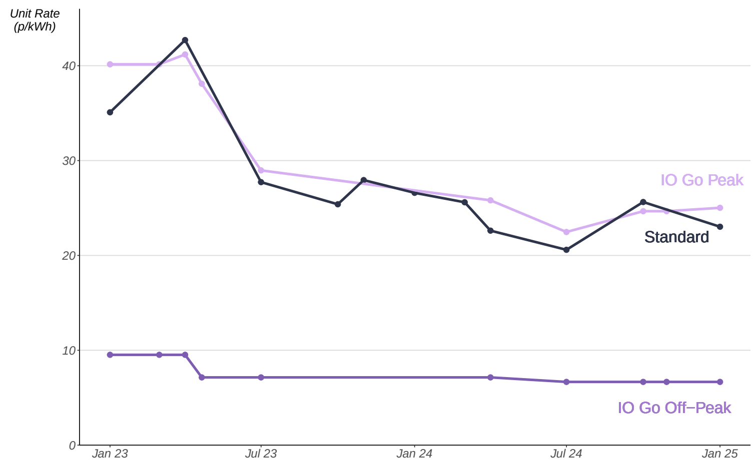
Table A15: Bump Charging per Charge Day

Country	Bumps/day	Peak bumps/day	Overnight bumps/day	Daytime bumps/day
United Kingdom	0.019	0.003	0.003	0.013
United Kingdom - Early Adopters	0.022	0.005	0.003	0.015
Germany	0.107	0.019	0.008	0.080
Spain	0.105	0.007	0.027	0.071
United States	0.035	0.011	0.004	0.020

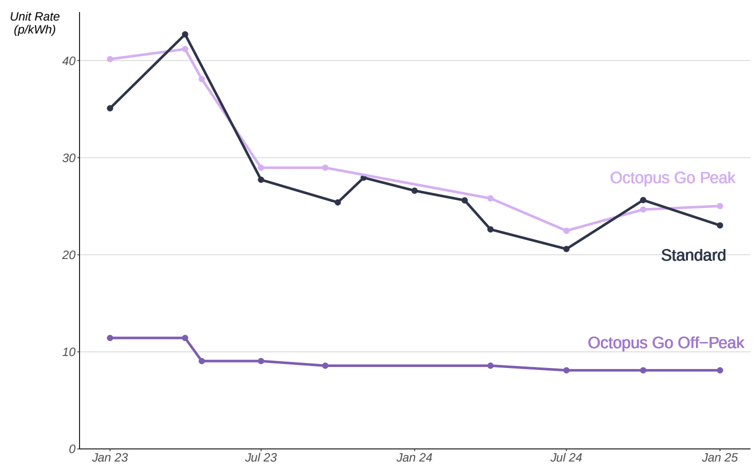
Note: This table presents the frequency of bumping each day there is a charge event, for each of the four major countries where IO Go is active. This uses data from a random sample of 4,442 users across the four countries. Peak times are defined as 16:30-20:30 for the UK, 17:00-21:00 for Germany, 20:00-00:00 for Spain, and 15:00-19:00 for the United States. Overnight times are defined as 23:00-5:00am, except in Germany, where IO Go there defines overnight off-peak as 00:00-05:00. Daytime are all other non-peak and non-evening hours.

B Figures

Figure A1: Tariff Rates in 2024



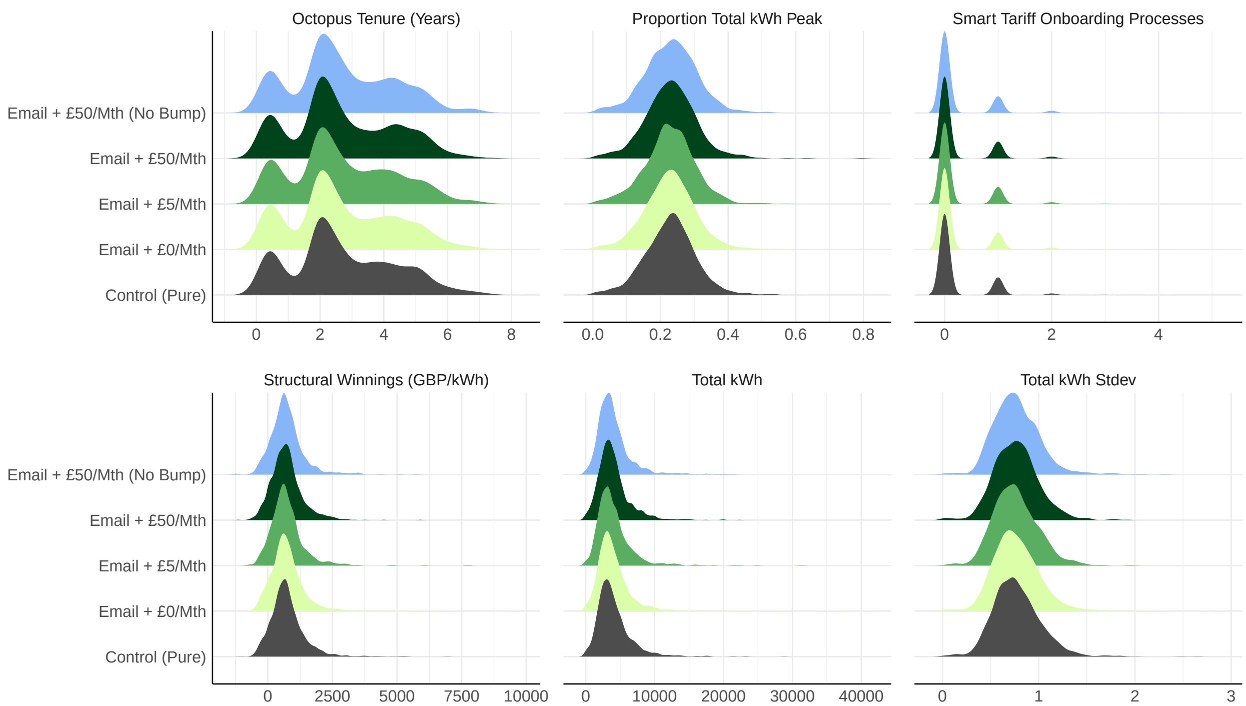
(a) Intelligent Octopus Go



(b) Octopus Go

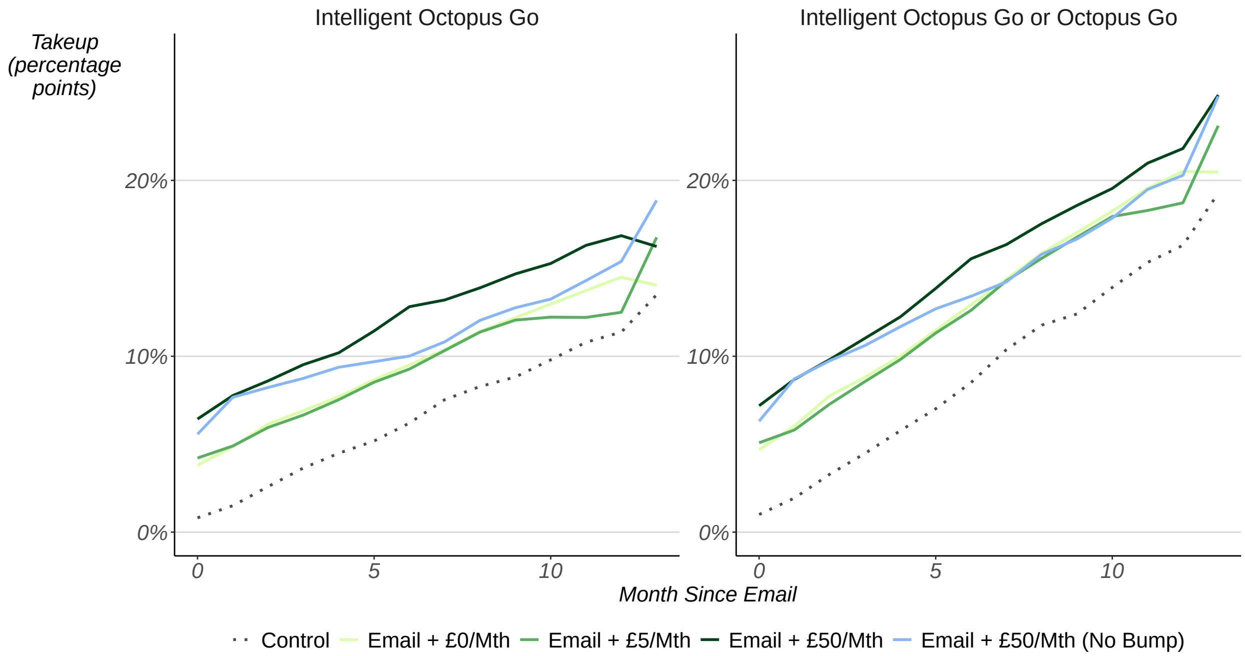
Notes: Panel (a) shows the tariff rates for Intelligent Octopus Go customers during the off-peak overnight period (23:30–05:30, dark purple) and the peak daytime period (05:30–23:30, light purple). For comparison, we also include the Flexible Octopus tariff from Octopus Energy, which maintains a flat rate throughout the day. Panel (b) shows analogous tariff rates for Octopus Go's off-peak overnight period (00:30-5:30) and peak daytime period (5:30-00:30).

Figure A2: Distribution of Baseline Covariates Across Encouragement Arms



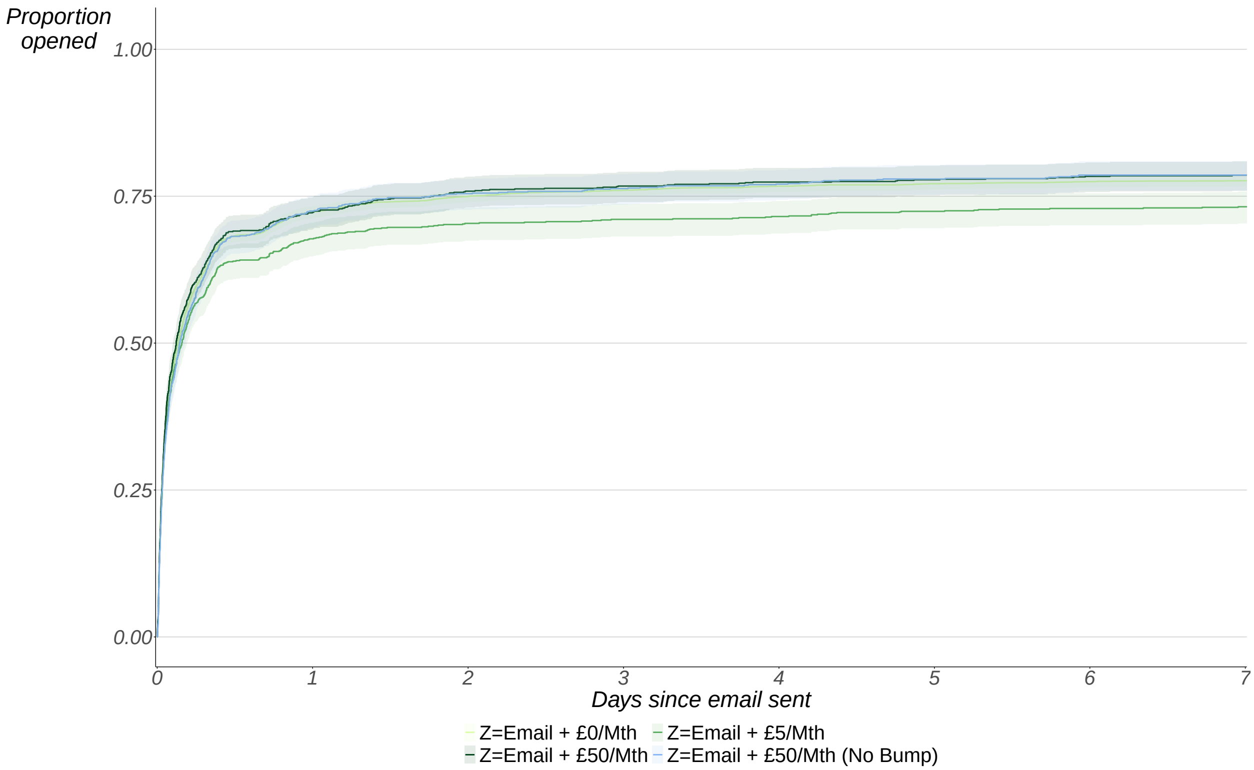
Notes: This figure plots kernel density estimates of baseline covariates across encouragement arms. The similarity of these distributions illustrates that block randomization achieved good covariate balance across arms, consistent with the regression-based balance tests reported in Table A1. Note that the variable 'Not Always Octopus Customer' has been excluded, as it is a binary variable that is unsuitable for a density plot. Detailed definitions of these covariates can be found in Appendix D.2.

Figure A3: Take-up of EV Tariffs Over Time by Trial Arms



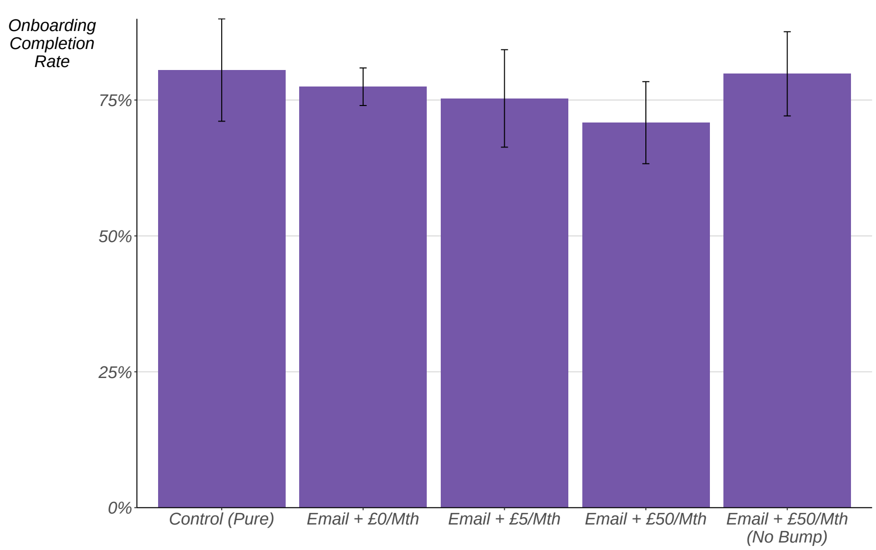
Notes: This figure plots the fraction of trial participants in each trial group that has taken up an EV tariff. The left panel shows take-up of the AI managed charging tariff, IO Go, which combines static time-of-use pricing with remote control of EV charging. The right panel shows take-up of either IO Go or Octopus Go, the latter being a static time-of-use tariff without supplier control.

Figure A4: Probability of First-Ever Opening Emailed Encouragement



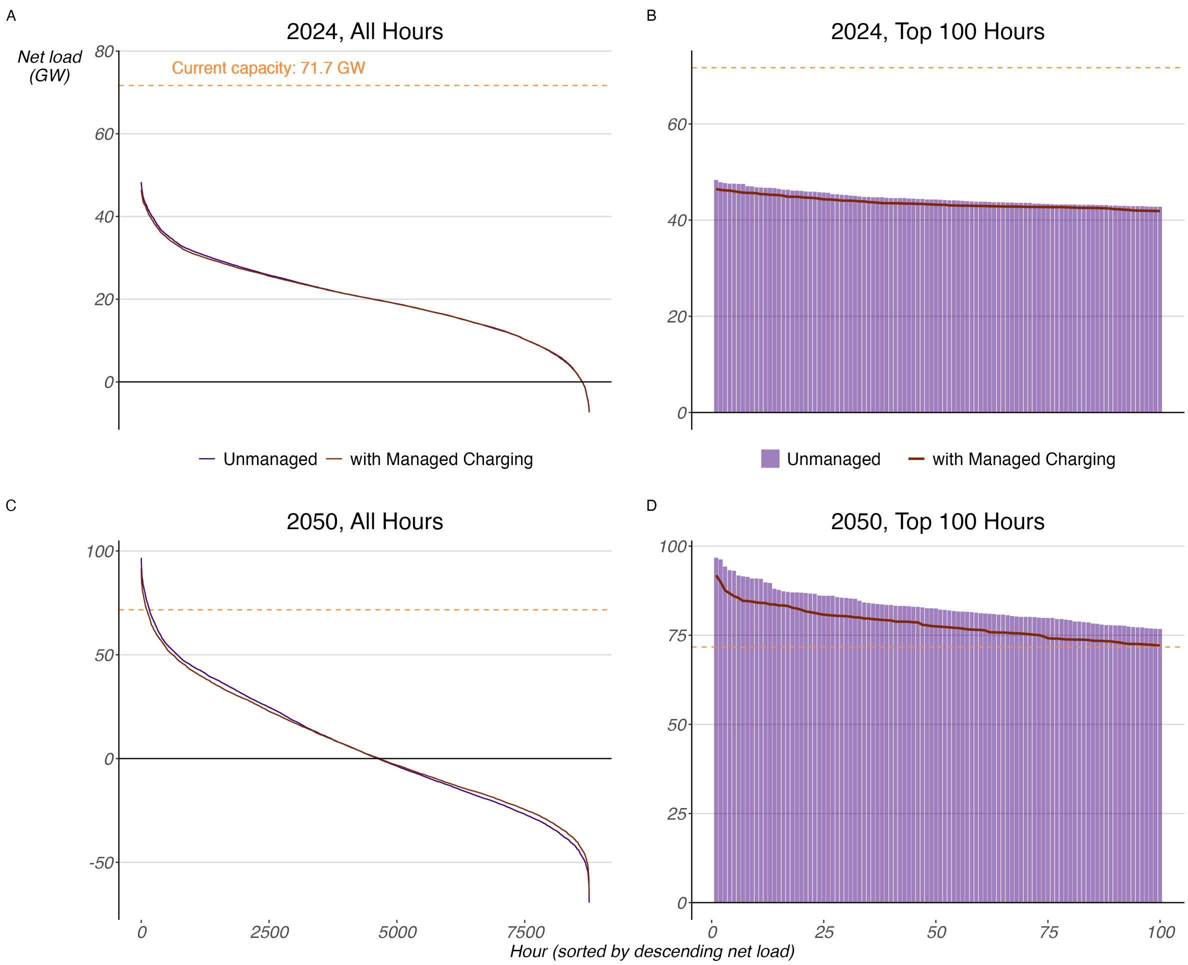
Notes: This figure plots the cumulative proportion of customers who first opened an encouragement email over time. Solid lines show Kaplan–Meier survival estimates of the probability of not yet opening, with shaded areas denoting 95% confidence intervals. Time is measured from the date of the encouragement email until the first observed open.

Figure A5: Completion Rate for Participants Signing Up for IO Go



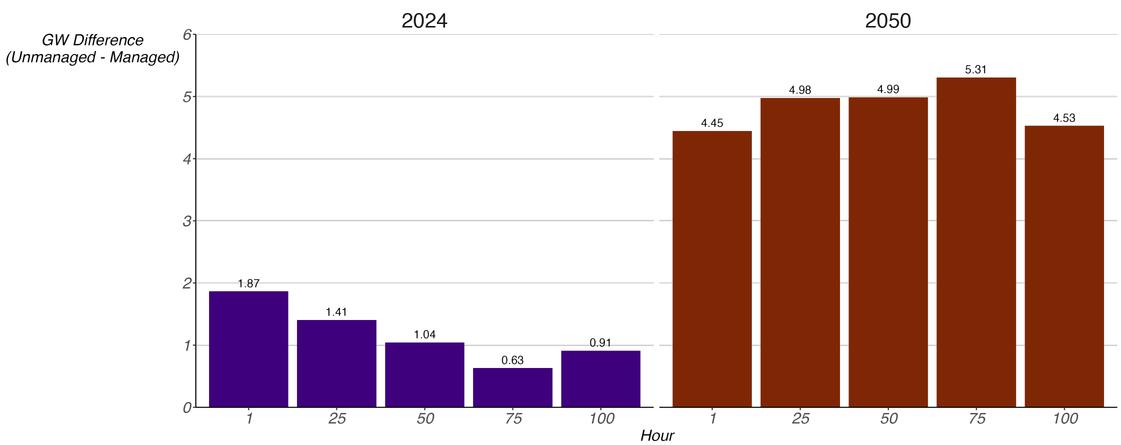
Notes: This figure shows the percentage of participants who, after attempting to sign up for IO Go during the incentive period, successfully completed the onboarding process. Completino of onboarding includes testing the compatibility of their vehicle or charger with the IO Go app.

Figure A6: Load Duration Curves



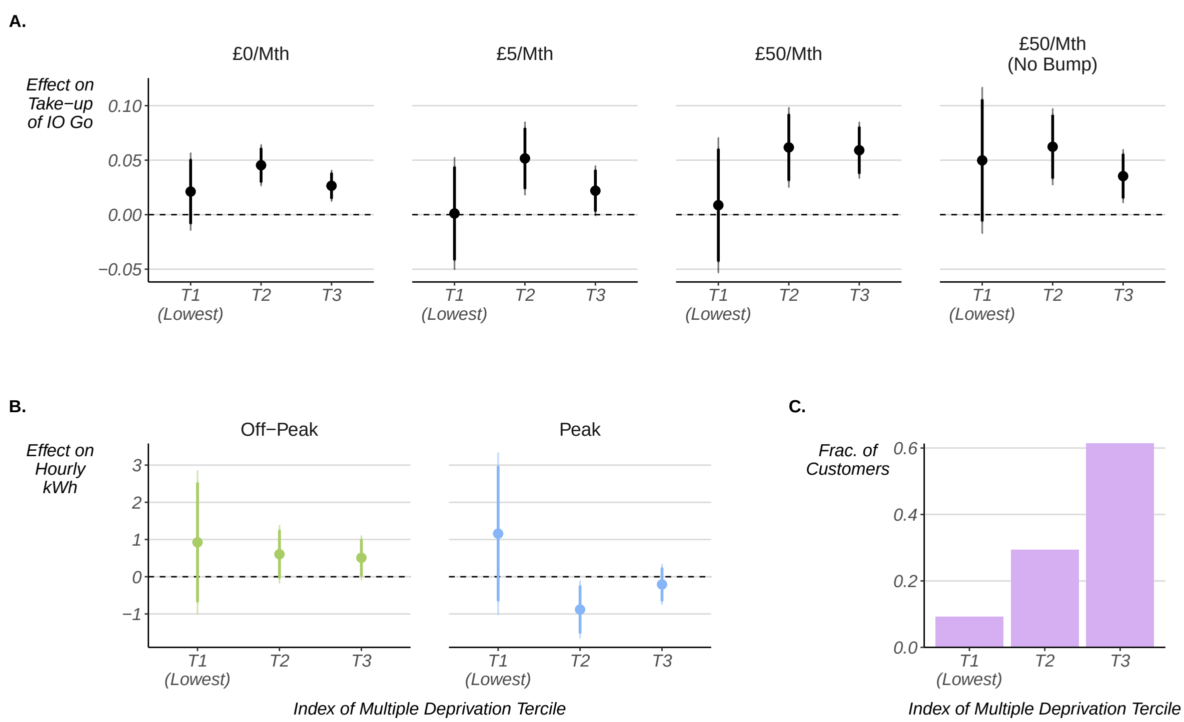
Notes: This figure shows load duration curves for 2024 and the forecasted load duration curve for 2050. The unmanaged charging consumption profile is based on demand modelling data from the UK National Energy System Operator's Future Energy Scenarios. To estimate the impact of managed charging in 2024, we applied the hourly coefficients from our IV estimates (Figure 7) to the corresponding hourly load values. For 2050, we assumed that managed charging would primarily affect the four daily peak hours. We therefore applied the estimated effects from the four peak hours in Figure 7 to the corresponding top four hours in each day of the 2050 load profile. The resulting reductions in peak consumption were assumed to be evenly redistributed across the remaining hours of the day.

Figure A7: Difference Between Unmanaged and Managed Load Duration Curves at Selected Hours



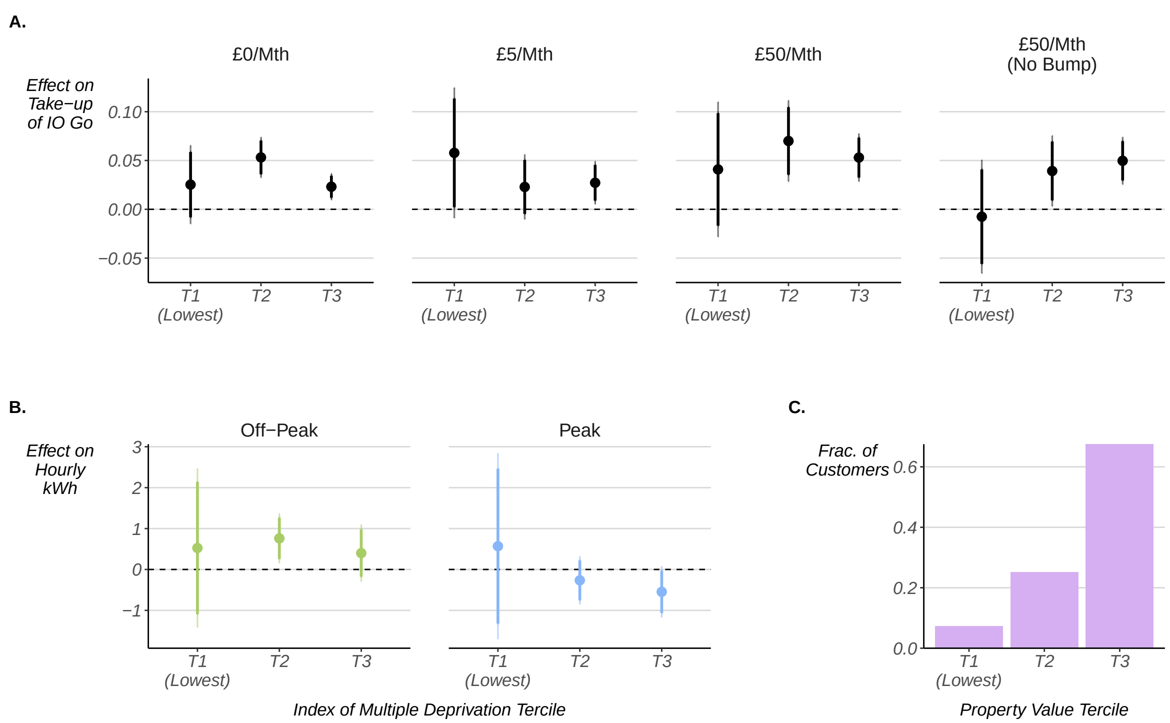
Notes: The bars show the difference in load (unmanaged minus managed) at several positions along the load-duration curve in Fig. A6. Values are taken at load hours 1, 25, 50, 75, and 100 for the years 2024 and 2050.

Figure A8: Heterogeneity by Index of Multiple Deprivation



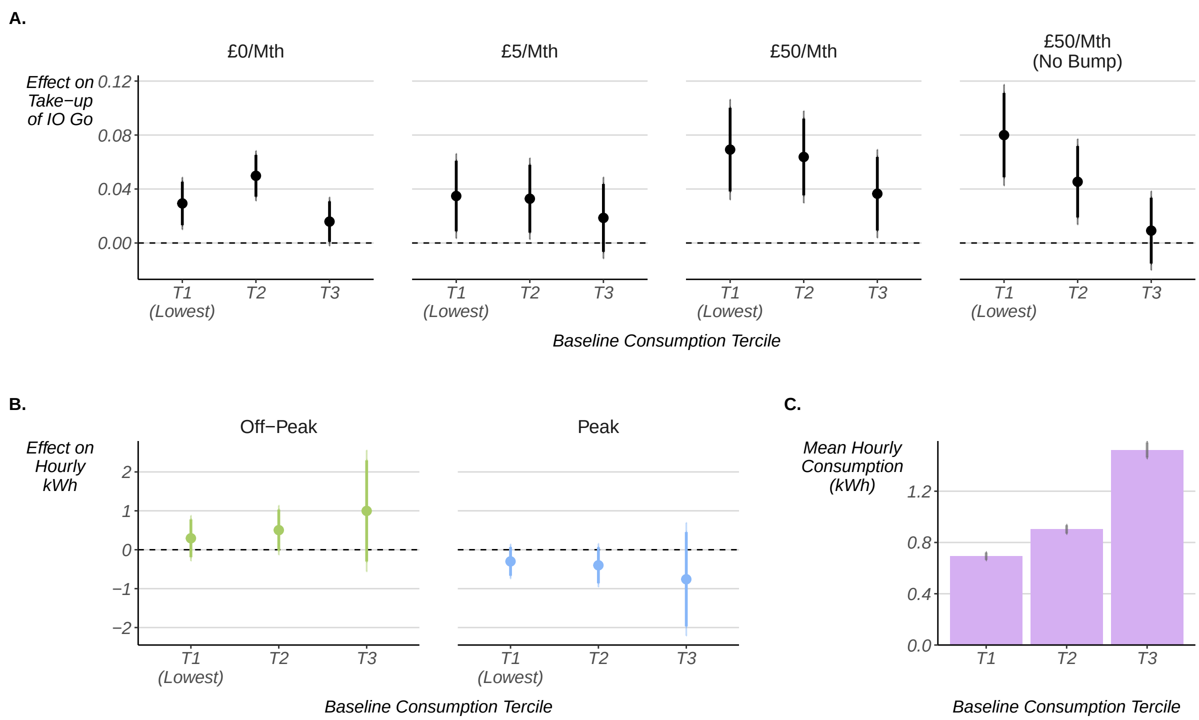
Notes: This figure presents heterogeneity in treatment effects by Index of Multiple Deprivation (IMD) terciles. Panel A shows the effect of each encouragement group on take-up of IO Go, interacting encouragement with deprivation tercile (T1: most deprived); regression results are also presented in Table A10. Panel B displays the estimated IV effects of EV tariff adoption on hourly electricity consumption during off-peak (23:30-5:30, green) and peak (16:30-20:30, blue) periods, again interacting EV tariff adoption with deprivation tercile; regression results are also presented in Table A11. Panel C shows the proportion of trial participants in the experimental sample in each tercile. Confidence intervals are shown at the 95% level. All specifications control for baseline consumption, and fixed effects for randomization block and week. Lines depict 90% (dark) and 95% (light) confidence intervals. Standard errors are clustered by participant and week.

Figure A9: Heterogeneity by Property Value



Notes: This figure presents heterogeneity in treatment effects by property value terciles. Panel A shows the effect of each encouragement group on take-up of IO Go, interacting encouragement with property value tercile (T1: most deprived); regression results are also presented in Table A10. Panel B displays the estimated IV effects of EV tariff adoption on hourly electricity consumption during off-peak (23:30-5:30, green) and peak (16:30-20:30, blue) periods, again interacting EV tariff adoption with property value tercile; regression results are also presented in Table A11. Panel C shows the proportion of trial participants in the experimental sample in each tercile. Confidence intervals are shown at the 95% level. All specifications control for baseline consumption, and fixed effects for randomization block and week. Lines depict 90% (dark) and 95% (light) confidence intervals. Standard errors are clustered by participant and week.

Figure A10: Heterogeneity by Baseline Electricity Consumption

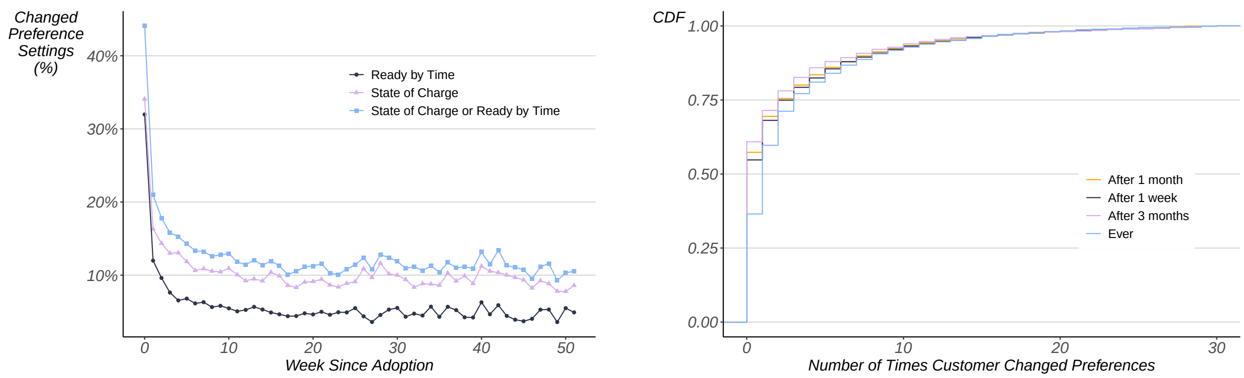


Notes: This figure presents heterogeneity in treatment effects by baseline total consumption (kWh). Panel A shows the effect of each encouragement group on take-up of IO Go, interacting encouragement with deprivation tercile (T1: lowest consumption); regression results are also presented in Table A10. Panel B displays the estimated IV effects of EV tariff adoption on hourly electricity consumption during off-peak (23:30-5:30, green) and peak (16:30–20:30, blue) periods, again interacting EV tariff adoption with baseline consumption tercile; regression results are also presented in Table A11. Panel C shows mean hourly consumption for each tercile, with 95% standard error bars. Confidence intervals are shown at the 95% level. All specifications control for baseline consumption, and fixed effects for randomization block and week. Lines depict 90% (dark) and 95% (light) confidence intervals. Standard errors are clustered by participant and week.

Figure A11: Customers' Preferences Over Time

(a) Changing Preferences Over Time

(b) Number of Times Customers Change Preferences

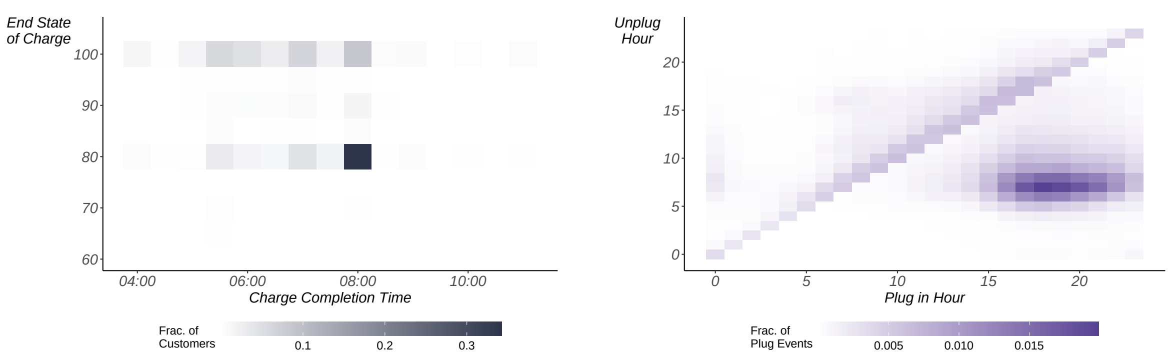


Notes: This figure summarizes how frequently IO Go customers change their preferences. Panel (a) shows what proportion of customers changed their preferences in the weeks after adopting IO Go; (b) displays the CDF of the number of times customers change their preferences, split by week since adoption of IO Go

Figure A12: Participant Preferences and Behaviors

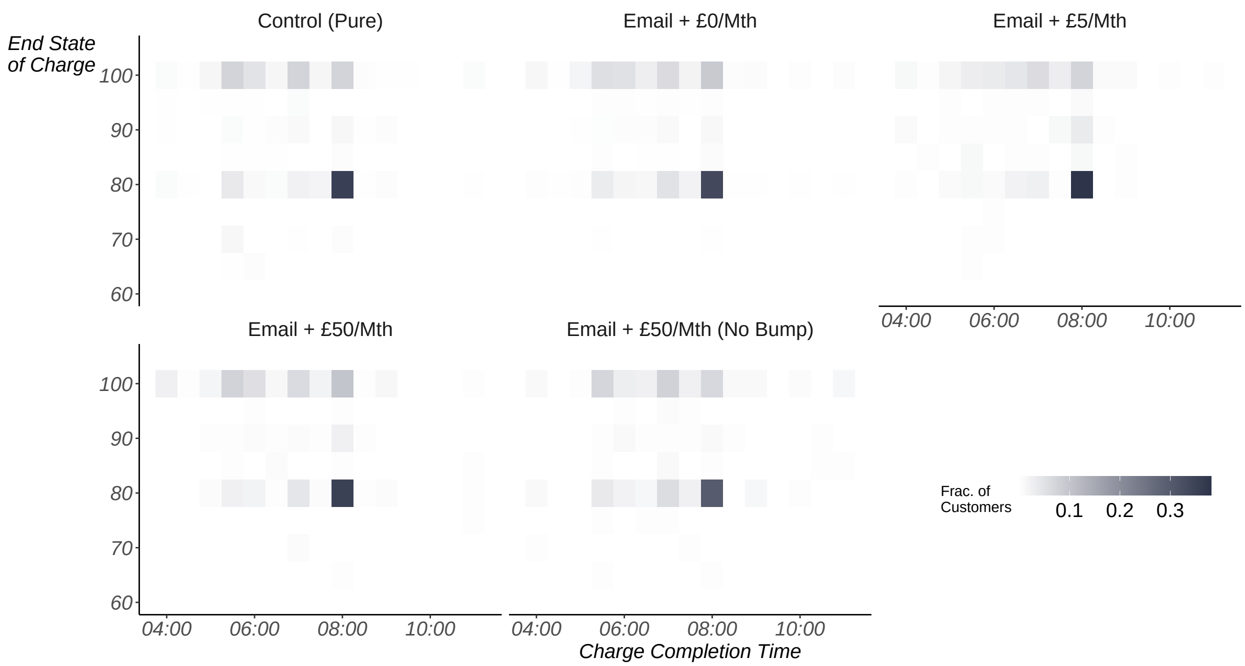
(a) Customer Settings

(b) Plug-in and Unplug Time



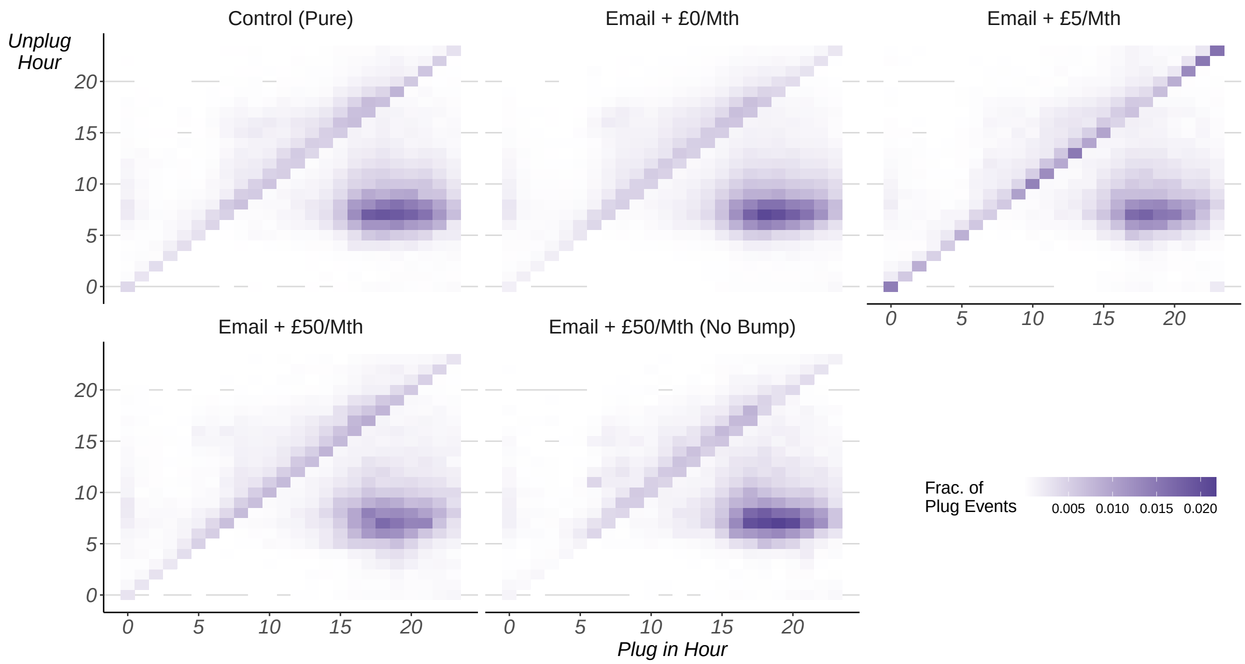
Notes: This figure summarizes IO Go customer preferences and behavioral patterns. Panel (a) shows the joint distribution of customer-selected end state of charge and charge completion time, as specified via the Octopus app. Panel (b) displays the frequency of plug-in and unplug times.

Figure A13: Preferences for End State of Charge and Completion Time - By Encouragement Group



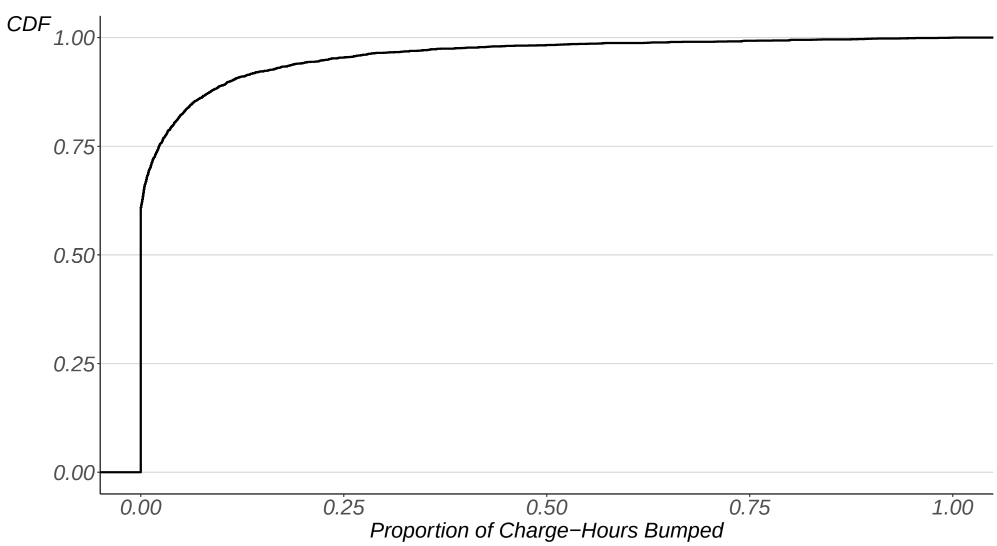
Notes: This figure shows the joint distribution of customer-selected end state of charge and charge completion time, as specified via the Octopus app. The trial participants are split by the encouragement group they were randomly assigned.

Figure A14: Preferences for End State of Charge and Completion Time - By Encouragement Group



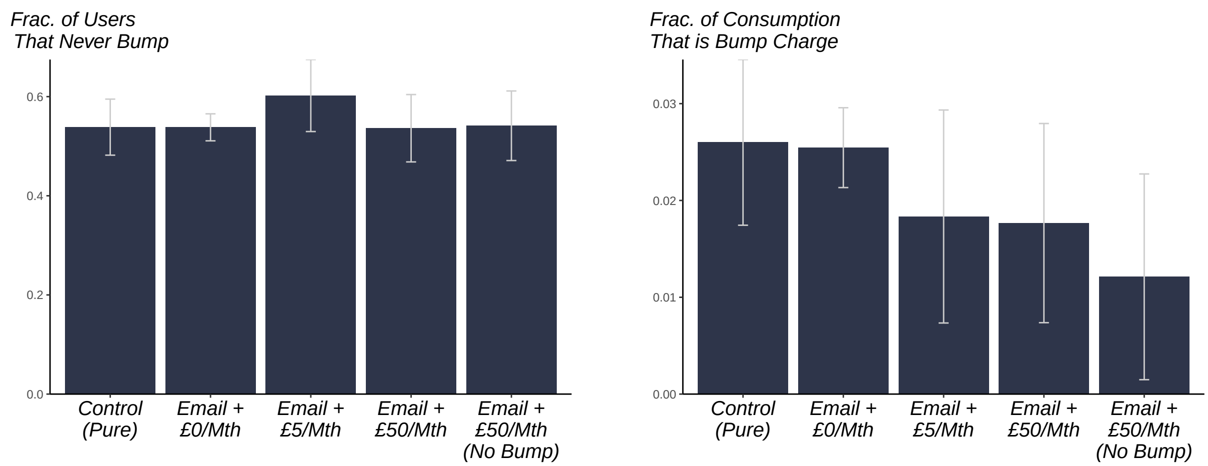
Notes: This figure shows the frequency of plug-in and unplug times. The results are shown separately by randomized encouragement group.

Figure A15: Distribution of Proportion of Charge-Hours Bumped Per Customer



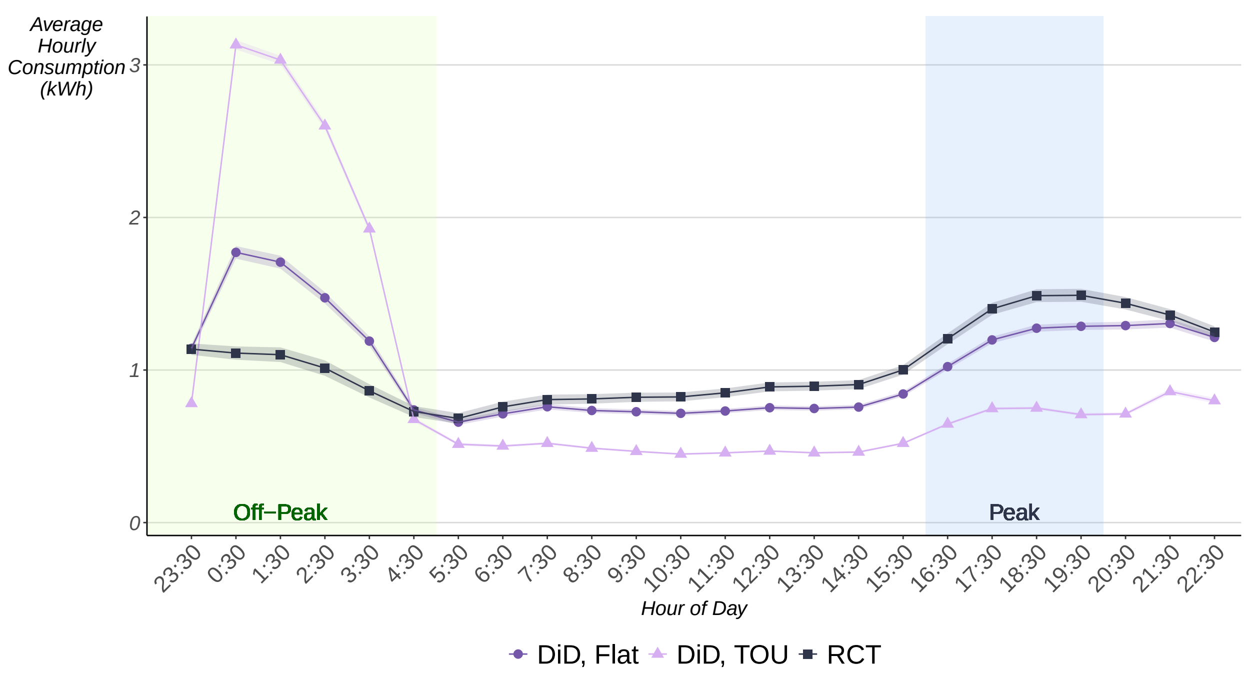
Notes: This figure illustrates bump charging behavior, where trial participants manually overrode the AI managed charging schedule. The horizontal axis shows, for each customer, what share of their total charge-hours was "bumped" (overridden). The vertical axis shows the cumulative proportion of customers. Charge-hours here are hours where any charging occurs.

Figure A16: Bump Charge Behaviors - By Encouragement Group



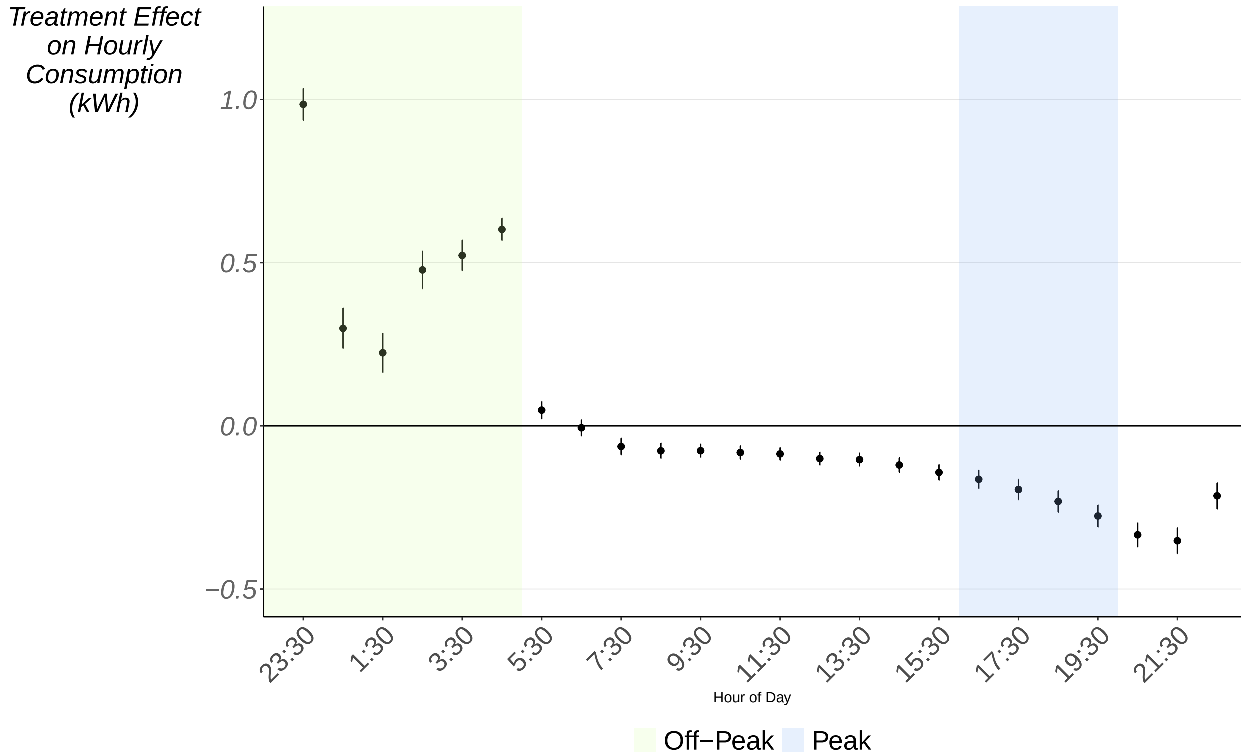
Notes: This figure illustrates bump charging behavior, where trial participants manually overrode the AI managed charging schedule. Results are shown separately by randomized encouragement group. The left panel shows the proportion of trial participants who never used bump charging. The right panel shows the share of total electricity consumption that came from bump charging.

Figure A17: Comparison of Baseline Consumption Profiles



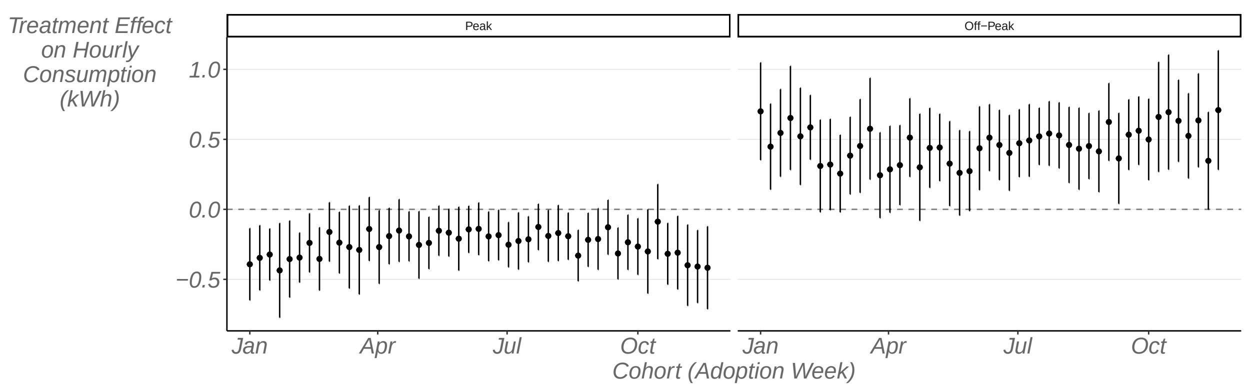
Notes: This figure shows baseline hourly electricity consumption patterns for customers included in the DiD and RCT analyses. For the DiD sample, baseline refers to consumption prior to adopting IO Go. For the RCT sample, baseline refers to control group participants who were not on an EV tariff. "Flat" indicates customers who were on a flat tariff before switching to IO Go, while "ToU" refers to those previously on a time-of-use tariff. The average daily consumption during the baseline period was 1.1 kWh for DiD–Flat customers, 1.02 kWh for DiD–ToU customers, and 1.0 kWh for the RCT control group. The shaded areas represent 95% confidence intervals, with errors clustered at the account user level.

Figure A18: Difference-in-Differences Estimate of IO Go, by Hour-of-day



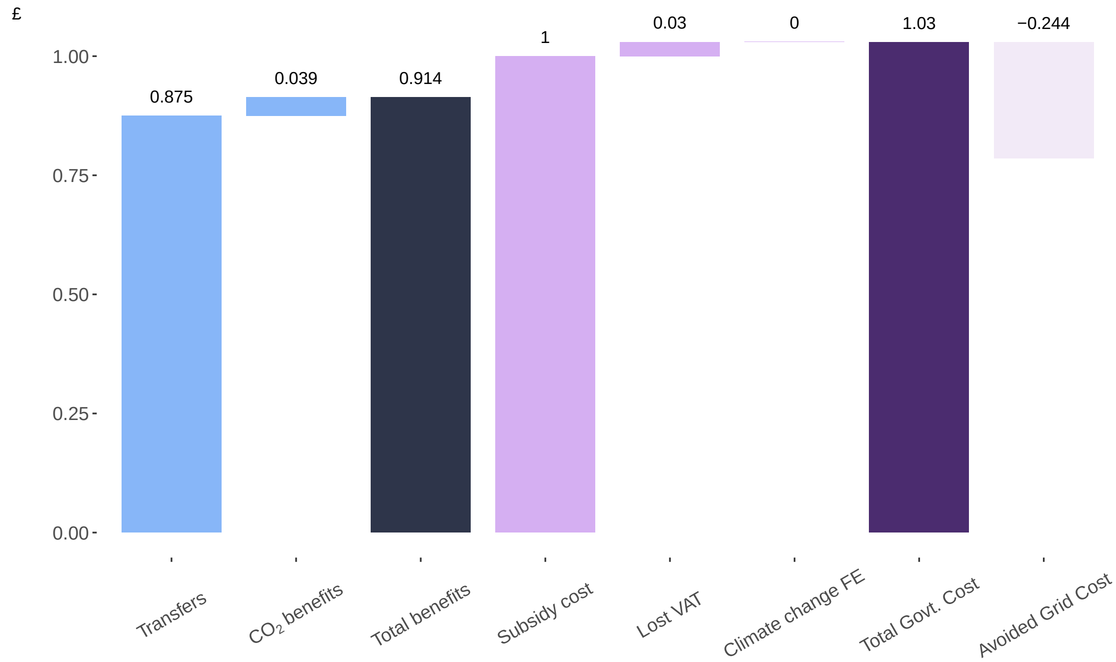
Notes: This figure reports effect of adopting the IO Go tariff on hourly electricity consumption (in kWh), split by hour of the day. These are estimated using a sample of 20,249 customers who first-ever enrolled in IO Go in 2023, weighted by whether the customer was previously on a time-of-use tariff. Estimates are computed using the Callaway and Sant'Anna (2021) estimator. Confidence intervals are shown at the 95% level.

Figure A19: Cohort Specific Difference-in-Differences Estimate of IO Go

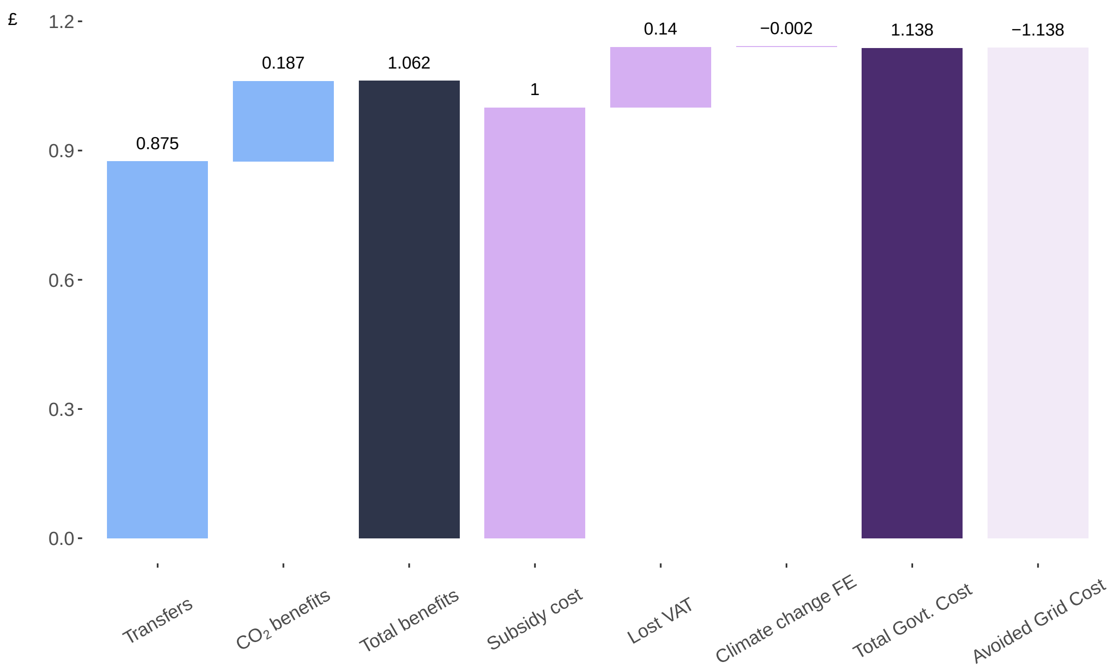


Notes: This figure reports cohort-specific estimates of the effect of adopting the IO Go tariff on hourly electricity consumption (in kWh) during (a) peak hours (16:30-20:30) and (b) off-peak hours (23:30-5:30). These are estimated using a sample of 9,317 customers who first-ever enrolled in IO Go in 2024, weighted by whether the customer was previously on a time-of-use tariff. Estimates are computed using the Callaway and Sant'Anna (2021) estimator. Confidence intervals are shown at the 95% level.

Figure A20: Marginal value of public funds of subsidizing AI managed charging



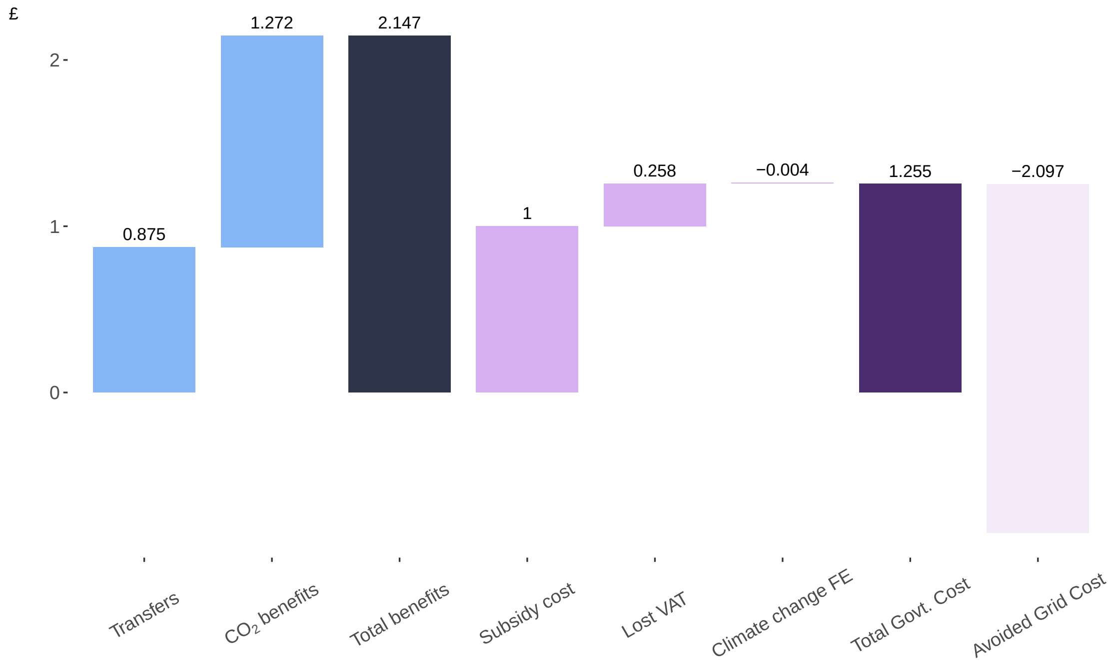
(a) 1 year, 2024



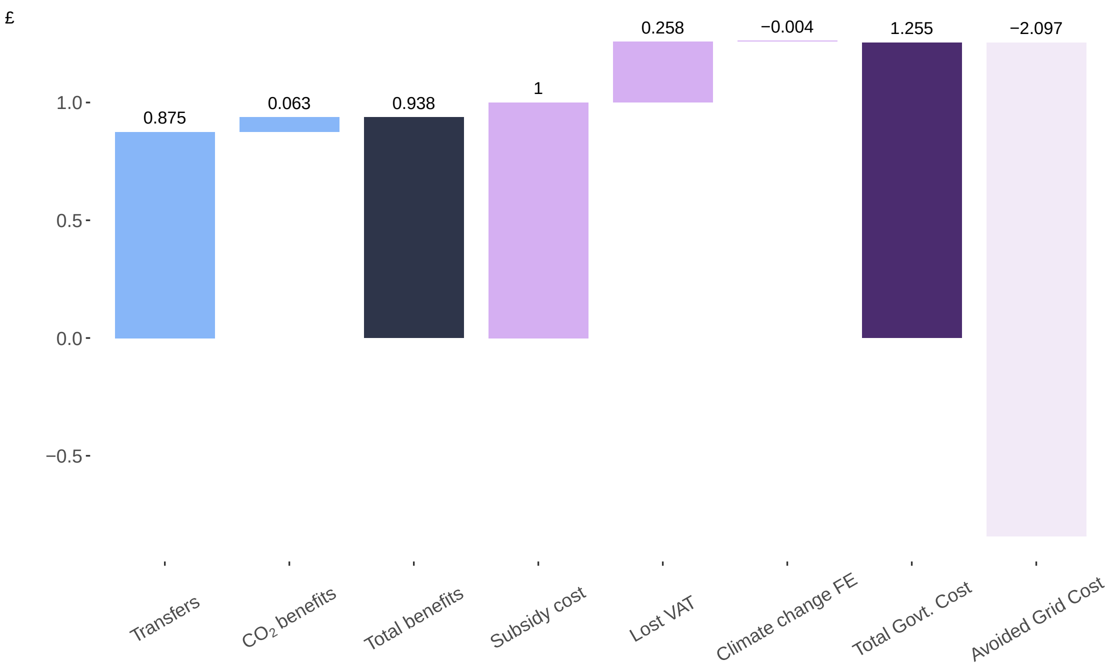
(b) 5 years, 2024-2028

Notes: This figure presents the estimated costs and benefits of subsidizing adoption of the IO Go managed EV charging tariff over the course of (a) 1 year, and (b) 5 years. Customer surplus is based on a decomposition of marginal and inframarginal adoption under the £50/month offer, following the approach of Hendren and Sprung-Keyser (2020). Direct CO_2e benefits reflect emissions reductions from shifting electricity use to cleaner hours, scaled to marginal adopters. Indirect CO_2e benefits are excluded under the assumption that managed charging subsidies do not affect EV uptake among inframarginal adopters. Estimated costs to government include the subsidy, lost VAT revenue, and increased tax receipts from climate-related fiscal externalities. Grid stabilization benefits are shown separately, based on Franken et al. (2025) estimates of per-vehicle system savings under three scenarios. Only a share of these may accrue to government.

Figure A21: Marginal value of public funds of subsidizing AI managed charging, Alternative SCC



(a) 10 years, SCC from Bilal and Känzig (2024)



(b) 10 years, SCC from Interagency Working Group

Notes: This figure presents the estimated costs and benefits of subsidizing adoption of the IO Go managed EV charging tariff over the course of 10 years. Panel (a) uses the SCC estimate from Bilal and Känzig (2024); panel (b) uses the SCC estimate from Interagency Working Group on Social Cost of Greenhouse Gases, United States Government. Customer surplus is based on a decomposition of marginal and inframarginal adoption under the £50/month offer, following the approach of Hendren and Sprung-Keyser (2020). Direct CO_2e benefits reflect emissions reductions from shifting electricity use to cleaner hours, scaled to marginal adopters. Indirect CO_2e benefits are excluded under the assumption that managed charging subsidies do not affect EV uptake among inframarginal adopters. Estimated costs to government include the subsidy, lost VAT revenue, and increased tax receipts from climate-related fiscal externalities. Grid stabilization benefits are shown separately, based on Franken et al. (2025) estimates of per-vehicle system savings under three scenarios. Only a share of these may accrue to government.

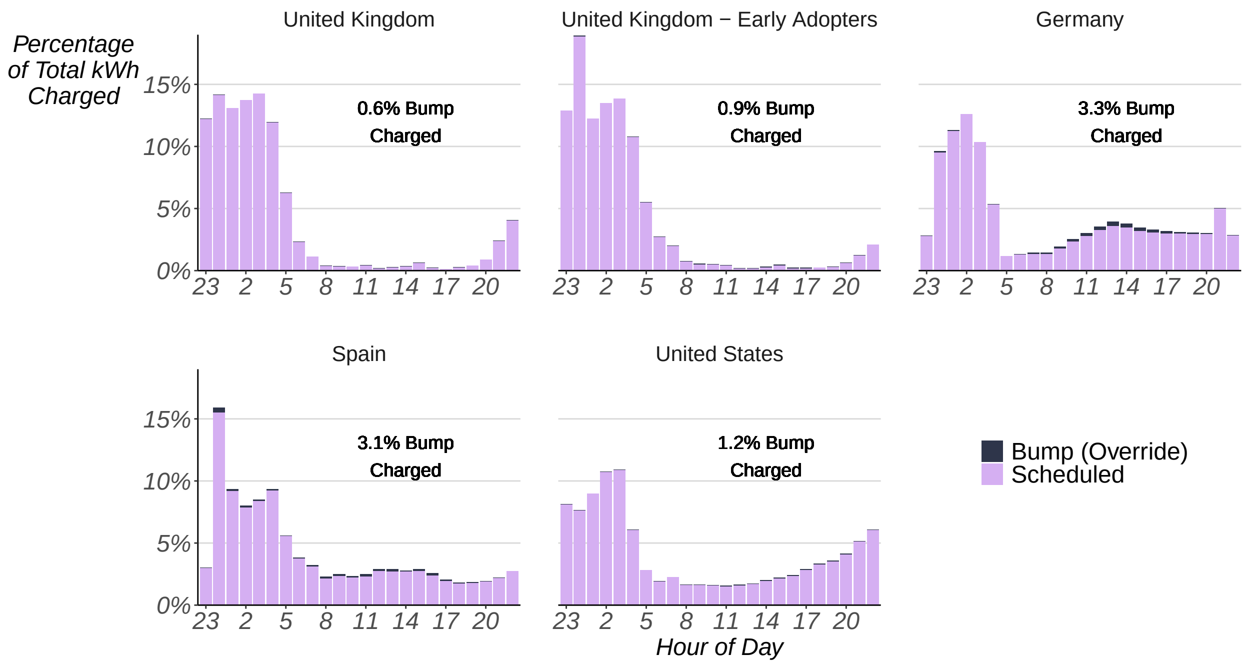


Figure A22: Scheduled and Bump Charging, By Hour-of-Day

Notes: This figure plots, by hour of day, the percentage of electricity consumption that occurred during that hour. This consumption is further divided into charging triggered by bump (charging initiated by users overriding the schedule) versus charging scheduled by IO Go. We use a sample of 4,442 IO Go users across the four countries.

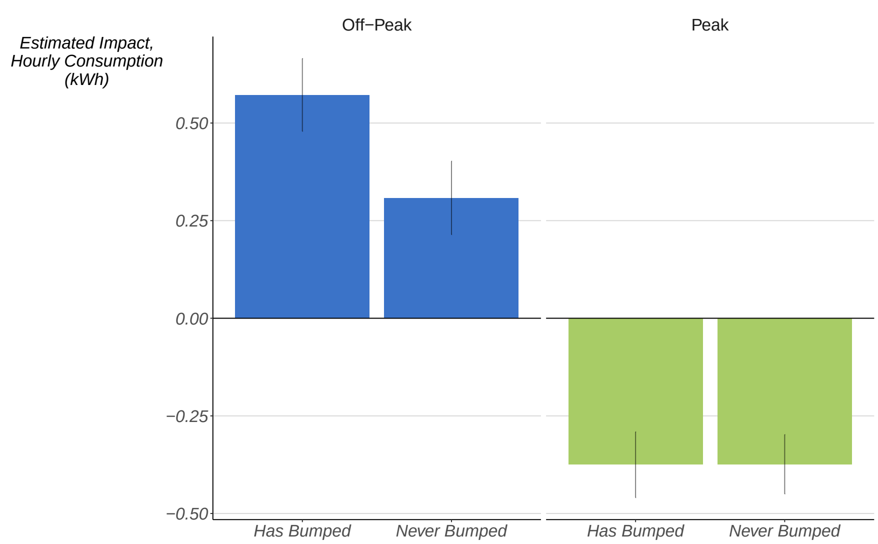


Figure A23: Estimated Impact of IO GO, by “Bump" (Override) Behavior

Notes: This figure reports staggered difference-in-differences estimates of the effect of adopting the IO Go tariff on hourly electricity consumption, for participants who were a part of our experiment. We report separate estimates for customers who (i) have “bumped", or overridden the supplier managed schedule, and (2) who have never bumped. Estimates are also separately estimated by peak (16:30-20:30) and off-peak hours (23:30-5:30). Estimates are computed using the Callaway and Sant’Anna (2021) estimator. Error bars represent the 95% confidence interval.

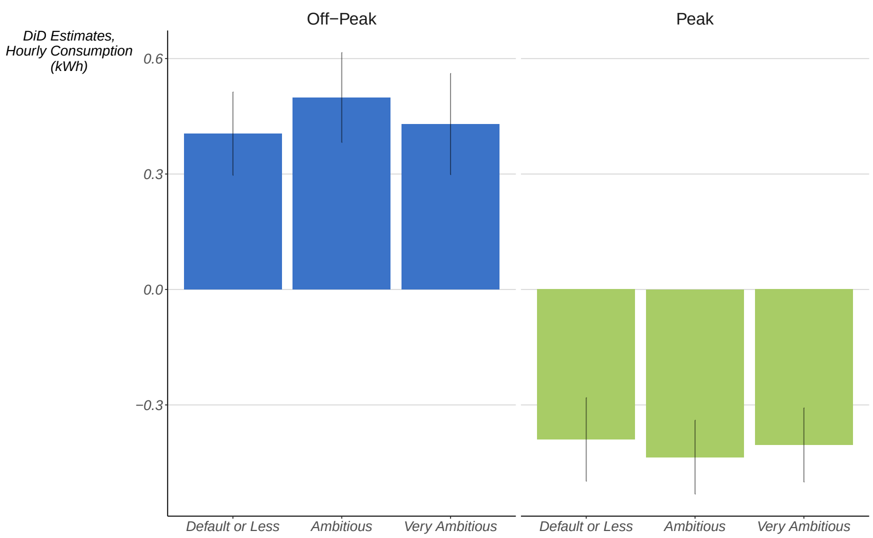


Figure A24: Estimated Impact of IO GO, by User Preferences

Notes: This figure reports staggered difference-in-differences estimates of the effect of adopting the IO Go tariff on hourly electricity consumption, for participants who were a part of our experiment. We report separate estimates for customers by their settings for when they need the car ready and how much charge is needed: (1) default or less ambitious - 8:00 a.m. ready-by time and 80% charge, or later/lower, (2) ambitious — either an earlier ready-by time or higher charge, and (3) most ambitious — both earlier and higher charge. Estimates are also separately estimated by peak (16:30-20:30) and off-peak hours (23:30-5:30). Estimates are computed using the Callaway and Sant'Anna (2021) estimator. Error bars represent the 95% confidence interval.

C Deviations from the Pre-Analysis Plan

This experiment was pre-registered with the American Economic Association (AEA) under Trial No. 0013037. While we aimed to follow the pre-analysis plan (PAP) as closely as possible, a number of deviations became necessary in the course of implementation and analysis. Below, we detail the key departures from the PAP. Where applicable, corresponding changes to regression specifications are noted via footnotes in the main text.

- **Data:** Originally, our consumption data was to be aggregated on the hour (e.g., 00:00, 01:00, etc.). However, IO Go tariff rates begin on the half-hour (e.g., 23:30 rather than 00:00), we adjusted our data aggregation accordingly. All half-hourly data was therefore aligned to begin on the half-hour mark.

- **Data:** We had pre-specified that electricity consumption data would be aggregated to the hour-day level. However, due to computational constraints in processing and analyzing high-frequency data, we instead aggregated consumption to the week × hour-of-day level.

- **Data:** We included 12 months of post-encouragement data, rather than the 6 months originally proposed, as we initiated our analysis later than anticipated and took advantage of the longer available data window. This change in data window does not substantively change our results, as shown in column (7) of Table A7, Table A8, and Table A9.

- **Analysis:** To improve precision and account for pre-existing consumption patterns, we included a control for baseline average hourly electricity consumption in our regression models. This adjustment was not pre-specified, but we show the version of the regression that does not include baseline consumption in column (8) of Table A7, Table A8, and Table A9. Coefficients for the peak consumption and overall consumption are similar, while the coefficient for the off-peak consumption changes from 0.515 (main specification) to 0.164 (no baseline consumption). We view this more as a loss of precision than of a substantive change in the underlying effect. As can be seen in Column (8), without baseline consumption as a control, the estimates become extremely noisy, and are statistically indistinguishable from the main specification.

- **Analysis:** Our pre-specified IV analysis planned to use the randomized encouragements as instruments for just IO Go adoption. However, given that the randomized encouragements increased enrollment in both Intelligent Octopus Go and Octopus Go, we defined treatment as adoption of either EV tariff, rather than IO Go alone. This adjustment was necessary to preserve the exclusion restriction in our instrumental variables analysis, a requirement we did not know would be necessary at the outset of the trial.

- **Analysis:** Following the concerns raised in Mogstad et al. (2021) regarding the interpretation of IV estimates with multiple instruments, we use a single binary instrument combining all encouragement groups. This approach helps avoid complications such as negative weights and improves interpretability under a common first-stage assumption. We report results from our pre-specified analysis using the four instruments separately in Figure 7, Table A7, Table A8, and Table A9. These include both the joint specification with all four instruments and separate regressions where each encouragement serves as an instrument individually.

- **Analysis:** Our pre-specified DiD sample included all customers who adopted IO Go in 2023. In the final analysis, our DiD analysis restricted the sample to customers who likely owned an EV by December 2022. This was to ensure that observed changes in consumption patterns are due to changes in charging behavior, rather than the initial uptake of EVs. We also re-weighted this restricted DiD sample by pre-adoption tariff to better understand the extent to which differences between RCT and DiD were due to pre-adoption tariff (this re-weighting was not pre-specified). We also added a separate DiD analysis of customers who adopted DiD in 2024, which was not pre-specified. However, this analysis is not a part of our main findings, as detailed in Appendix F.

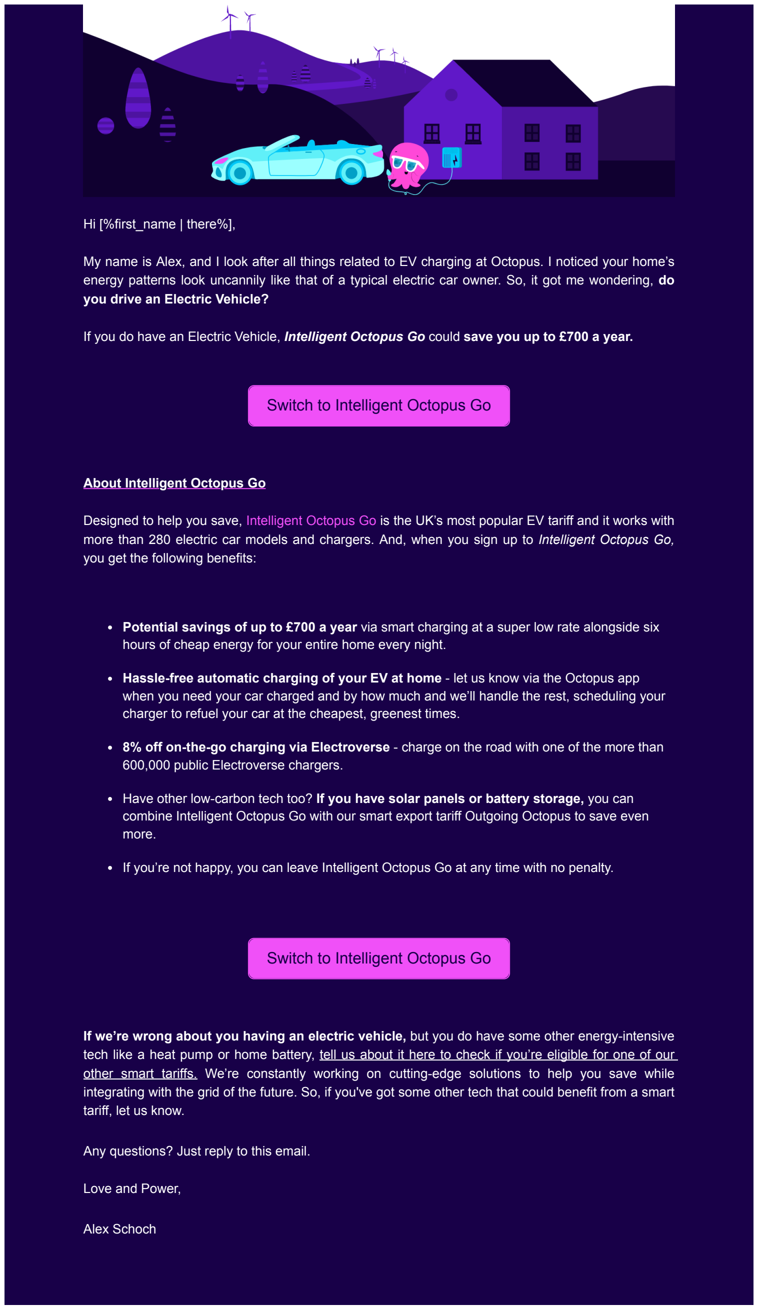
- **Analysis:** For our DiD analysis, we prespecified restricting control cohorts to customers who adopt IO GO within 30 days. However, our final analysis used a twelve-week window. This was to balance comparability of treated and control groups against the length of the post-adoption estimation horizon. Comparability was assessed by examining pre-treatment trends, and we selected the longest horizon that yielded satisfactory pre-trend balance. Additionally, we added an anticipation period, which was not part of the original specification. This decision followed a review of pre-trends, especially in the 2024 analysis, where we observed a rise in off-peak consumption in the four weeks before IO Go adoption (Figure A29b).

- **Welfare:** In keeping with the pre-analysis plan, we used the £150 incentive as a proxy subsidy, with J-PAL as proxy government. However, we refined our welfare analysis to better capture the long-term and system-wide implications of the intervention. To this end, we looked at CO_2e impacts over a 10-year time period; examined how the MVPF would change when we included avoided grid balancing costs; and included lost VAT as an extra cost to the government of the subsidy. While these additions increase the measured benefits, we believe they more accurately reflect the full set of social returns that would accrue under real-world implementation. In keeping with the pre-analysis plan, we assumed that trial participants who enrolled in response to the £0/month email reflected inframarginal participants, while the incremental take-up in the £50/month group represented marginal adopters with an average willingness to pay equal to 50% of the subsidy, consistent with standard MVPF assumptions.

- **Welfare**: We originally pre-specified estimating CO_2e impacts using the ITT framework. We did not pre-specify the use of IV estimation for bills or CO_2e savings. We adopted the IV approach here because it more directly captures the causal effect of IO Go adoption — the quantity of substantive interest. While our pre-analysis plan focused on MVPF calculations rather than consumer bills, we now report bill savings as well, as they provide an important and policy-relevant measure of consumer benefits.

D Additional details on design of field trial

D.1 Reproduction of email-based encouragements



Hi [%first_name | there%],

My name is Alex, and I look after all things related to EV charging at Octopus. I noticed your home's energy patterns look uncannily like that of a typical electric car owner. So, it got me wondering, **do you drive an Electric Vehicle?**

If you do have an Electric Vehicle, ***Intelligent Octopus Go*** could **save you up to £700 a year.**

Switch to Intelligent Octopus Go

About Intelligent Octopus Go

Designed to help you save, Intelligent Octopus Go is the UK's most popular EV tariff and it works with more than 280 electric car models and chargers. And, when you sign up to *Intelligent Octopus Go,* you get the following benefits:

- **Potential savings of up to £700 a year** via smart charging at a super low rate alongside six hours of cheap energy for your entire home every night.
- **Hassle-free automatic charging of your EV at home** - let us know via the Octopus app when you need your car charged and by how much and we'll handle the rest, scheduling your charger to refuel your car at the cheapest, greenest times.
- **8% off on-the-go charging via Electroverse** - charge on the road with one of the more than 600,000 public Electroverse chargers.
- Have other low-carbon tech too? **If you have solar panels or battery storage,** you can combine Intelligent Octopus Go with our smart export tariff Outgoing Octopus to save even more.
- If you're not happy, you can leave Intelligent Octopus Go at any time with no penalty.

Switch to Intelligent Octopus Go

If we're wrong about you having an electric vehicle, but you do have some other energy-intensive tech like a heat pump or home battery, tell us about it here to check if you're eligible for one of our other smart tariffs. We're constantly working on cutting-edge solutions to help you save while integrating with the grid of the future. So, if you've got some other tech that could benefit from a smart tariff, let us know.

Any questions? Just reply to this email.

Love and Power,

Alex Schoch

Figure A25: Randomized Encouragement Group 1 (Email + £0/Mth)

Note: The subject line read FYI: Do you drive an EV? You could save hundreds on your energy bills.

Hi [%first_name | there%],

My name is Alex, and I look after all things related to EV charging at Octopus. I noticed your home's energy patterns look uncannily like that of a typical electric car owner. So, it got me wondering, **do you drive an electric vehicle?**

If you do have an electric vehicle, ***Intelligent Octopus Go*** could **save you up to £700 a year.** Plus, if you **switch in March, we will pay you up to £15** over April, May and June for each day you stay on ***Intelligent Octopus Go.*** *This offer is valid for the next 11 days - through March 31.*

Switch to Intelligent Octopus Go

About Intelligent Octopus Go

Designed to help you save, Intelligent Octopus Go is the UK's most popular EV tariff and it works with more than 280 electric car models and chargers. And, when you sign up to *Intelligent Octopus Go,* you get the following benefits:

- **We will pay you up to £5 per month** during April, May and June for being on Intelligent Octopus Go.
- **Potential savings of up to £700 a year** via smart charging at a super low rate alongside six hours of cheap energy for your entire home every night.
- **Hassle-free automatic charging of your EV at home** - let us know via the Octopus app when you need your car charged and by how much and we'll handle the rest, scheduling your charger to refuel your car at the cheapest, greenest times.
- **8% off on-the-go charging via Electroverse** - charge on the road with one of the more than 600,000 public Electroverse chargers.
- Have other low-carbon tech too? **If you have solar panels or battery storage,** you can combine Intelligent Octopus Go with our smart export tariff Outgoing Octopus to save even more.
- If you're not happy, you can leave Intelligent Octopus Go at any time with no penalty.

Switch to Intelligent Octopus Go

If we're wrong about you having an electric vehicle, but you do have some other energy-intensive tech like a heat pump or home battery, tell us about it here to check if you're eligible for one of our other smart tariffs. We're constantly working on cutting-edge solutions to help you save while integrating with the grid of the future. So, if you've got some other tech that could benefit from a smart tariff, let us know.

Any questions? Just reply to this email.

Love and Power,

Alex Schoch

Figure A26: Randomized Encouragement Group 2 (Email + £5/Mth)

Note: The subject line read FYI: Do you drive an EV? You could save hundreds on your energy bills and get paid £15.

Hi [%first_name | there%],

My name is Alex, and I look after all things related to EV charging at Octopus. I noticed your home's energy patterns look uncannily like that of a typical electric car owner. So, it got me wondering, **do you drive an electric vehicle?**

If you do have an electric vehicle, ***Intelligent Octopus Go*** could **save you up to £700 a year.** Plus, if you **switch in March, we will pay you up to £150** over April, May and June for each day you stay on ***Intelligent Octopus Go.*** *This offer is valid for the next 11 days - through March 31.*

Switch to Intelligent Octopus Go

About Intelligent Octopus Go

Designed to help you save, Intelligent Octopus Go is the UK's most popular EV tariff and it works with more than 280 electric car models and chargers. And, when you sign up to *Intelligent Octopus Go,* you get the following benefits:

- **We will pay you up to £50 per month** during April, May and June for being on Intelligent Octopus Go.
- **Potential savings of up to £700 a year** via smart charging at a super low rate alongside six hours of cheap energy for your entire home every night.
- **Hassle-free automatic charging of your EV at home** - let us know via the Octopus app when you need your car charged and by how much and we'll handle the rest, scheduling your charger to refuel your car at the cheapest, greenest times.
- **8% off on-the-go charging via Electroverse** - charge on the road with one of the more than 600,000 public Electroverse chargers.
- Have other low-carbon tech too? **If you have solar panels or battery storage,** you can combine Intelligent Octopus Go with our smart export tariff Outgoing Octopus to save even more.
- If you're not happy, you can leave Intelligent Octopus Go at any time with no penalty.

Switch to Intelligent Octopus Go

If we're wrong about you having an electric vehicle, but you do have some other energy-intensive tech like a heat pump or home battery, tell us about it here to check if you're eligible for one of our other smart tariffs. We're constantly working on cutting-edge solutions to help you save while integrating with the grid of the future. So, if you've got some other tech that could benefit from a smart tariff, let us know.

Any questions? Just reply to this email.

Love and Power,

Alex Schoch

Figure A27: Randomized Encouragement Group 3 (Email + £50/Mth)

Note: The subject line read FYI: Do you drive an EV? You could save hundreds on your energy bills and get paid £150.

Hi [%first_name | there%],

My name is Alex, and I look after all things related to EV charging at Octopus. I noticed your home's energy patterns look uncannily like that of a typical electric car owner. So, it got me wondering, **do you drive an electric vehicle?**

If you do have an electric vehicle, ***Intelligent Octopus Go*** could **save you up to £700 a year.** Plus, if you **switch in February, we will pay you up to £150** over March, April, and May for each day you stay on ***Intelligent Octopus Go.*** *This offer is valid for the next 14 days – through Feb 29.*

Switch to Intelligent Octopus Go

About Intelligent Octopus Go

Designed to help you save, Intelligent Octopus Go is the UK's most popular EV tariff and it works with more than 280 electric car models and chargers. And, when you sign up to *Intelligent Octopus Go,* you get the following benefits:

- **We will pay you up to £50 per month** during March, April, and May for being on Intelligent Octopus Go **minus £2.00 for each day that you "bump charge"** by using the Octopus app to instantly charge your EV or otherwise suspend automated charging with Intelligent Octopus Go.
- **Potential savings of up to £700 a year** via smart charging at a super low rate alongside six hours of cheap energy for your entire home every night.
- **Hassle-free automatic charging of your EV at home** - let us know via the Octopus app when you need your car charged and by how much and we'll handle the rest, scheduling your charger to refuel your car at the cheapest, greenest times.
- **8% off on-the-go charging via Electroverse** - charge on the road with one of the more than 600,000 public Electroverse chargers.
- Have other low-carbon tech too? **If you have solar panels or battery storage,** you can combine Intelligent Octopus Go with our smart export tariff Outgoing Octopus to save even more.
- If you're not happy, you can leave Intelligent Octopus Go at any time with no penalty.

Switch to Intelligent Octopus Go

If we're wrong about you having an electric vehicle, but you do have some other energy-intensive tech like a heat pump or home battery, tell us about it here to check if you're eligible for one of our other smart tariffs. We're constantly working on cutting-edge solutions to help you save while integrating with the grid of the future. So, if you've got some other tech that could benefit from a smart tariff, let us know.

Any questions? Just reply to this email.

Love and Power,

Alex Schoch

Figure A28: Randomized Encouragement Group 4 (Email + £50/Mth, No Bump)

Note: The subject line read FYI: Do you drive an EV? You could save hundreds on your energy bills and get paid £150.

D.2 Block randomization implementation

We implemented block randomization on trial participant account identifiers using Mahalanobis distance calculated over the following pre-encouragement variables:

1. Tenure with Octopus: Years since a customer's earliest import tariff contract with Octopus Energy (as of August 31, 2023).
2. Total Consumption (kWh): Total electricity use (kWh) from February 15 to August 31, 2023, aggregated across all half-hourly smart meter readings.
3. Consumption Variability (kWh): Standard deviation of half-hourly consumption over the same period.
4. Not Always Octopus Energy Customer: Binary flag for whether a trial participant originally joined Octopus via acquisition (e.g., from Bulb, Co-op) or Supplier of Last Resort procedures.
5. Smart Tariff Onboarding Attempts: Count of historical attempts to enroll in Octopus smart tariffs (e.g., Intelligent Octopus Go) prior to February 15, 2024. Includes cases where customers initiated but did not complete the process.
6. DNO Region: Categorical variable for the customer's Distribution Network Operator region. For accounts with multiple active meter points, we used the region linked to the most recent tariff as of August 31, 2023.
7. Expected Structural Winnings: Estimated monetary difference between a customer's actual electricity cost (under observed tariff contracts) and a counterfactual IO Go contract from February 15 to August 31, 2023. We assumed an off-peak rate of £0.075/kWh and a peak rate of £0.30/kWh for IO Go, and ignored taxes, standing charges, and regional price variation.
8. Peak-Hour Consumption Share: Proportion of total electricity usage (Feb–Aug 2023) occurring during 16:00–20:00, a period of high grid constraint.

Full details of structural savings calculations and data preparation are available upon request.

E Variance decomposition of bump behavior

We conducted a variance decomposition to quantify the relative contributions of different factors to the overall variability in bump charging behavior. Using hourly charging data, we regressed an indicator for whether charging occurs via bump charging on a set of fixed effects and temperature controls. Specifically, we estimated the following specification:

$$I(\text{Bump}_{iht}) = \alpha_i + \mu_t + \gamma_h + \psi_r + \sum_k \beta_k Temp^k_{i,t+1} + \varepsilon_{iht}, \tag{12}$$

where $I(\text{Bump}_{iht})$ is an indicator equal to one if customer i engaged in bump charging during hour h on date t. The term α_i denotes customer fixed effects, μ_t date fixed effects, γ_h hour-of-day fixed effects, and ψ_r region fixed effects corresponding to the 14 UK Grid Supply Point (GSP) regions. The temperature controls, $Temp^k_{i,t+1}$, are indicators for ten bins of next-day temperature, included to capture anticipatory charging responses to expected weather conditions. For each customer, temperature was assigned using observations from the geographically closest UK Met Office weather station to the customer's postcode. The UK Met Office is the national meteorological service of the United Kingdom. We run this regression for our experimentla sample, as well as the random sample of UK IO Go users, described in Section 6.

The variance decomposition can be expressed as:

$$\text{Var}(Y_{it}) = \underbrace{\text{Var}(\alpha_i)}_{\text{Household FE}} + \underbrace{\text{Var}(\mu_t)}_{\text{Date FE}} + \underbrace{\text{Var}(\gamma_h)}_{\text{Hour FE}} + \underbrace{\text{Var}(\psi_r)}_{\text{Region FE}} \tag{13}$$

$$+ \underbrace{\text{Var}(\beta \cdot \text{Temp}_{i,t+1})}_{\text{Temperature effect}} + \underbrace{2 \cdot \text{Cov}(\cdot,\cdot)}_{\text{Covariance terms}} + \underbrace{\text{Var}(\epsilon_{it})}_{\text{Residual}} \tag{14}$$

As shown in Table A16, customer fixed effects accounted for the vast majority of the variation in bump charging behavior, explaining 0.274 of the total variance in the RCT sample. Hour-of-day and region fixed effects contributed an additional 0.088. This dominance of customer fixed effects indicates that bump charging is largely idiosyncratic rather than correlated, making it unlikely to generate coincident demand spikes that would strain the grid. The UK IO Go adopters sample show similar results, with customer fixed effects as the main source of variation.

Table A16: Decomposition of Variance in Bump Behavior

Component	Share (RCT Sample)	Share (UK Sample)
Customer FE	0.274	0.296
Date FE	0.003	0.001
Hour FE	0.060	0.008
Region (GSP)	0.031	0.001
Next day temperature	0.000	0.000

Note: This table reports the variance decomposition of bump charging behavior based on the regression specification in Equation (13). Each entry shows the share of total variance in the bump charging indicator explained by the corresponding component.

F 2024 difference-in-differences

As discussed in Section 4.1, we also conducted a DiD analysis focusing on customers who adopted IO Go in 2024. However, identifying which customers already owned an EV at the start of 2024 was difficult. This challenge likely reflected increased uptake of low-carbon technologies (LCTs). Of particular note, heat pump installations rose sharply at the end of 2023, following the UK Government's expansion of heat pump subsidies in October 2023.

To mitigate this issue, we restricted the sample to customers who appeared to own an EV as of August 2023. We chose August 2023 because it was both (1) before the increase in heat pump subsidies and (2) during summer months when heat pump use was minimal. Nonetheless, adoption of other LCTs may still have confounded our estimates by increasing customers' engagement with tariff selection. This confounder was documented by Bernard et al. (2024), who examined households that received heat pump installations from Octopus Energy. They found that following the installation, two-thirds of these households adopted a smart tariff, with Intelligent Octopus being the most popular choice, possibly due to adoption of an EV at a similar time.[68] We presented the results of the 2024 DiD analysis but advised caution in their interpretation.

We began with a sample of 146,143 customers who adopted IO Go at some point in 2024. To be included in the final sample, customers had to have been with Octopus Energy by August 2023, had smart-meter data available at that time, and likely owned an EV as of August 2023. We excluded 2,039 customers who were already part of our randomized controlled trial. After these restrictions, the final analysis sample consisted of 9,317 customers.

[68] Additionally, in an internal survey, 25% of customers with a heat pump reported being on IO Go; it is therefore plausible that heat pump adoption itself encouraged switching to IO Go.

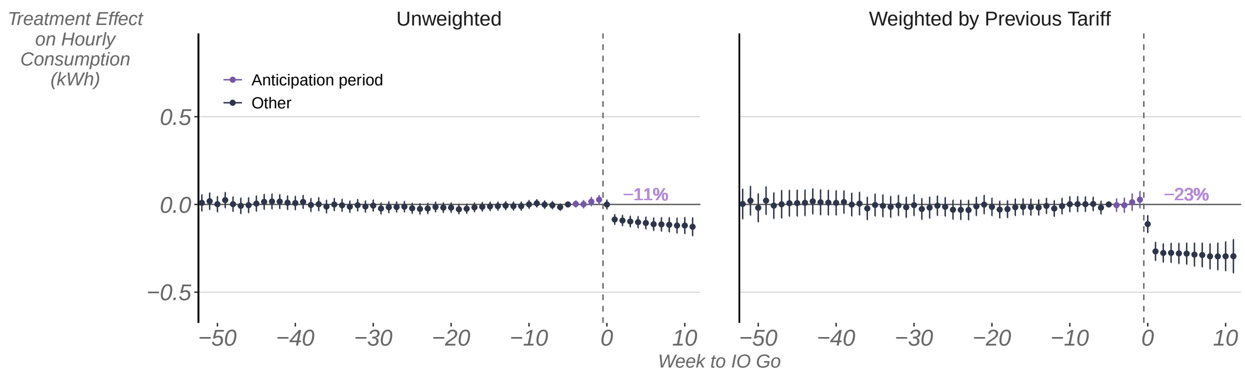
Our empirical strategy mirrored the 2023 DiD analysis. We used "not-yet-treated" customers as controls and defined an anticipation period of four weeks, using the five weeks prior to IO Go adoption as the reference period. We limited control cohorts to those scheduled to adopt IO Go no later than twelve weeks after the treated group's anticipation period ended. Finally, we estimated a weighted version of the model to match the 2024 DiD sample to the RCT sample based on pre-treatment tariff type.

Similar to the 2023 results, Figure A29 showed that the unweighted difference-in-differences estimates were smaller than those from the RCT. After reweighting the 2024 DiD sample to align with the RCT's pre-treatment tariff distribution, the estimated increase in off-peak consumption (0.548 kWh) closely matched the RCT estimate (0.481 kWh). The reduction in peak consumption was smaller in the DiD analysis: 0.269 kWh compared to 0.581 kWh in the RCT.

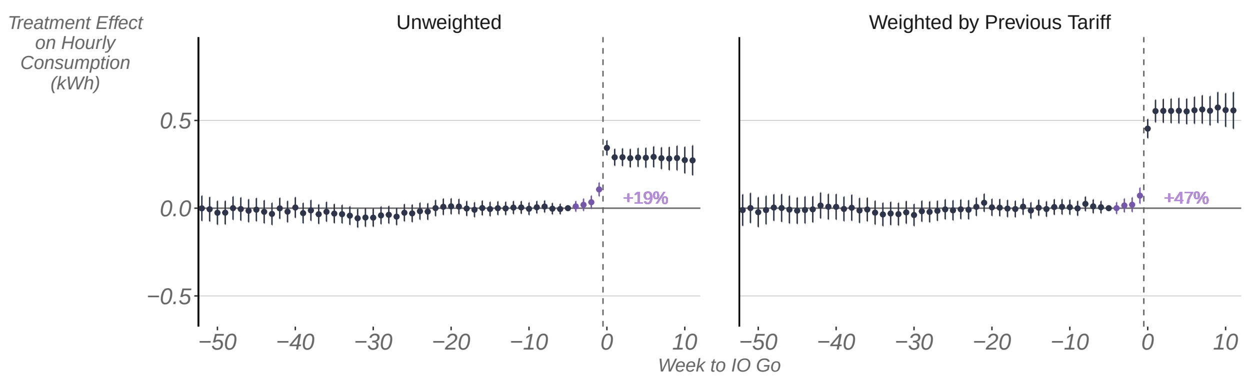
Diverging from the 2023 analysis, the decline in daytime, non-peak consumption did not fully offset the increase in off-peak consumption, resulting in an overall rise of 1.53 kWh per day, an 8% increase in total consumption. We believe this likely reflected concurrent adoption of other LCTs, such as heat pumps, during the same period. There was tentative evidence of this in the off-peak anticipation period in Figure A29b, although without direct data on LCT ownership, we could not confirm this hypothesis.

Figure A29: 2024 Difference-in-Differences Estimate of IO Go

(a) Peak

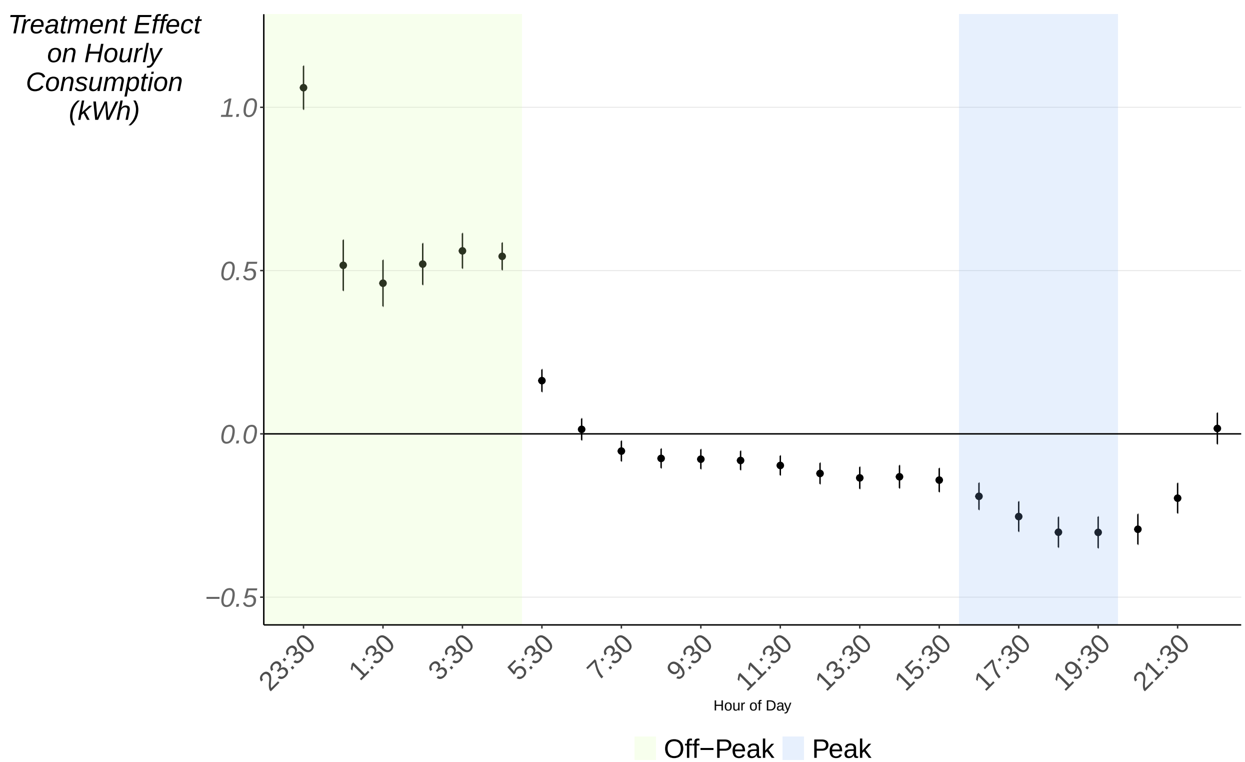


(b) Off-Peak



Notes: This figure reports staggered difference-in-differences estimates of the effect of adopting the IO Go tariff on hourly electricity consumption (in kWh) during (a) peak hours (16:30–20:30) and (b) off-peak hours (23:30–05:30), using a sample of 9,317 customers who first-ever enrolled in IO Go in 2024. Each panel plots treatment effects relative to the week before adoption. Estimates are reported under two specifications: (i) unweighted; (ii) and weighted by whether the trial participant was previously on a time-of-use tariff. Estimates are computed using the Callaway and Sant'Anna (2021) estimator. Percentages represent post-treatment effects as share of the pre-IO Go consumption levels. Post-treatment effects are estimated using average of all group-time average treatment effects, with weights proportional to the group size.

Figure A30: 2024 Difference-in-Differences Estimate of IO Go, by Hour-of-day



Notes: This figure reports effect of adopting the IO Go tariff on hourly electricity consumption (in kWh), split by hour of the day. These are estimated using a sample of 9,317 customers who first-ever enrolled in IO Go in 2024, weighted by whether the customer was previously on a time-of-use tariff. Estimates are computed using the Callaway and Sant'Anna (2021) estimator. Confidence intervals are shown at the 95% level.

G Theoretical Model

G.0.1 Environment and Notation

Time is divided into settlement periods $h = 1,\dots,H$. Each EV-owning household i requires E_i kWh by ready-by time T_i, has plug-in availability $\mathcal{A}_i \subseteq \{1,\dots,H\}$, maximum charge rate $\bar{x}_i$, and charging efficiency η_{ih}. Baseline (non-EV) load is b_{ih}, with aggregate $b_h \equiv \sum_i b_{ih}$. Total system load is

$$Q_h = b_h + \sum_i x_{ih}.$$

Retailer costs are $c_h(Q_h)$ (convex) and ancillary/avoided benefits $r_h(Q_h)$ (concave). Define the system shadow price

$$\tilde{c}_h = c_h'(Q_h) - r_h'(Q_h), \tag{15}$$

as in convex scheduling (Boyd and Vandenberghe, 2004; Joskow and Tirole, 2006*b*). Households derive mobility utility $U_i(E_i)$ (increasing, concave) and timing disutility $\psi_i(h) \ge 0$. Under RTP, they face attention/optimization cost $a_i \ge 0$ and risk penalty $\gamma_i \ge 0$ on bill variance (Borenstein, 2007). Under IO (AI scheduling), they can override at a hassle cost $\phi_i \ge 0$.

We compare four tariffs:[69]

1. **Flat**: price p^{flat}.
2. **ToU**: $p^{\text{peak}} > p^{\text{off}}$ across fixed windows.
3. **RTP**: hourly p_h^{RTP}; households self-schedule.
4. **IO (AI managed)**: centralized schedule $\{x_{ih}\}$ against $\{\tilde{c}_h\}$; users may override.

[69] Our framework builds on Joskow and Tirole (2006*b*), who study the welfare properties of RTP under convex scheduling. Their analysis contrasts RTP with the absence of RTP. Where we depart is in the introduction of (i) algorithmic intermediation (IO) as a distinct pricing/coordination regime, and (ii) frictions such as override behavior, execution costs (m_{ih}, ϕ_i), and aggregate override thresholds ($\bar{\beta}^\star$). These extensions allow us to bring the theoretical framework into closer alignment with experimental data, in particular by capturing behavioral deviations from the frictionless convex scheduling benchmark.

The feasible EV-charging schedules satisfy[70]

$$\sum_{h \le T_i, h \in \mathcal{A}_i} \eta_{ih} x_{ih} \ge E_i, \qquad 0 \le x_{ih} \le \bar{x}_i. \tag{16}$$

Elasticity by operative signal. Let π_h^k be the operative signal under regime $k \in \{\mathrm{IO}, \mathrm{RTP}, \mathrm{ToU}\}$ with $\pi_h^{\mathrm{IO}} = \tilde{c}_h$, $\pi_h^{\mathrm{RTP}} = p_h^{\mathrm{RTP}}$, $\pi_h^{\mathrm{ToU}} = p_h^{\mathrm{ToU}}$. Define

$$\epsilon_h^k = \frac{\partial \ln Q_h^k}{\partial \ln \pi_h^k}, \tag{17}$$

evaluated in high plug-in hours ($Q_h^k > 0$). Let e_h denote marginal CO_2 emissions (kg/kWh).

G.0.2 Household and Aggregator Problems

RTP household problem. A household chooses $\{x_{ih}\}$ to minimize expected cost plus frictions:[71]

$$\begin{aligned} \min_{\{x_{ih}\}} \quad & \mathbb{E}\Big[\sum_h p_h^{\mathrm{RTP}} x_{ih}\Big] + \gamma_i \operatorname{Var}\Big(\sum_h p_h^{\mathrm{RTP}} x_{ih}\Big) + \sum_h \psi_i(h) x_{ih} + a_i \cdot \mathbf{1}\{\text{actively scheduling}\} \\ \text{s.t.} \quad & \sum_{h \le T_i, h \in \mathcal{A}_i} \eta_{ih} x_{ih} \ge E_i, \quad 0 \le x_{ih} \le \bar{x}_i. \end{aligned} \tag{18}$$

At interior hours, first-order conditions (FOC) take the (mean–variance) form

$$\mathbb{E}[p_h^{\mathrm{RTP}}] + 2\gamma_i \, Cov\Big(p_h^{\mathrm{RTP}}, \sum_\ell p_\ell^{\mathrm{RTP}} x_{i\ell}\Big) + \psi_i(h) = \mu_i \eta_{ih}, \tag{19}$$

with $\mu_i \ge 0$ the KKT multiplier on the energy requirement. The attention cost $a_i > 0$ drives corner solutions by discouraging small reallocations.[72]

[70] We model EV-only RTP for comparability to IO (which optimizes EV load). If non-EV end uses are price-responsive under whole-household RTP, the welfare gap ΔW_0 generalizes accordingly; our empirical mapping focuses on the EV subproblem.

[71] This functional form is inspired by Borenstein (2007); Gabaix (2019) who modeled variations of an expected expenditure minimization augmented with cognitive and risk terms.

[72] Derivation of (19): Rewrite the energy requirement as a KKT-friendly inequality $g_i(x) \equiv E_i - \sum_{h \le T_i, h \in \mathcal{A}_i} \eta_{ih} x_{ih} \le 0$ with multiplier $\mu_i \ge 0$. Let $S_i(x,p) \equiv \sum_h p_h^{\mathrm{RTP}} x_{ih}$ be the (random) daily bill. Ignoring the fixed attention cost a_i (which

IO aggregator problem. The IO scheduler minimizes social cost plus timing disutility:

$$\begin{aligned} \min_{\{x_{ih}\}} \quad & \sum_h \Big[c_h(Q_h) - r_h(Q_h)\Big] + \sum_{i,h} \psi_i(h)\, x_{ih} \\ \text{s.t.} \quad & \sum_{h \le T_i,\, h \in \mathcal{A}_i} \eta_{ih} x_{ih} \ge E_i, \quad 0 \le x_{ih} \le \bar{x}_i. \end{aligned} \tag{20}$$

The objective is convex and the constraints are affine; under Slater's condition (capacity slack), KKT are necessary and sufficient (Boyd and Vandenberghe, 2004).

G.0.3 Core Lemmas and Propositions

Standing assumptions for this subsection. (i) $c_h(\cdot)$ convex, $r_h(\cdot)$ concave, so $\tilde{c}_h$ is nondecreasing in Q_h; (ii) price-taking (a single household does not affect $\tilde{c}_h$); (iii) Slater's condition holds for Eq. (20)–Eq. (16) (there exists feasible slack).

Lemma G.0.1 (Peak shaving under IO (merit-order via KKT)). *Under assumptions (i)–(iii), define the adjusted hourly cost*

$$\kappa_{ih} \equiv \frac{\tilde{c}_h + \psi_i(h)}{\eta_{ih}}.$$

(a) For each i, an IO optimum allocates charging to hours with the lowest available κ_{ih} subject to Eq. (16). (b) If there exist feasible peak and off-peak hours p, o with $\kappa_{ip} > \kappa_{io}$, then $x_{ip} > 0$

---

does not depend on $\{x_{ih}\}$ and thus only affects the extensive/corner decision), the Lagrangian is

$$\mathcal{L} = \mathbb{E}[S_i] + \gamma_i \operatorname{Var}(S_i) + \sum_h \psi_i(h) x_{ih} + \mu_i \left(E_i - \sum_{\substack{h \le T_i \\ h \in \mathcal{A}_i}} \eta_{ih} x_{ih} \right) + \sum_h \alpha_{ih}(x_{ih} - \bar{x}_i) - \sum_h \beta_{ih} x_{ih},$$

with box-constraint multipliers $\alpha_{ih}, \beta_{ih} \ge 0$. Since prices are exogenous to the household, $\frac{\partial}{\partial x_{ih}} \mathbb{E}[S_i] = \mathbb{E}[p_h^{\mathrm{RTP}}]$ and $\frac{\partial}{\partial x_{ih}} \operatorname{Var}(S_i) = 2\big(p_h^{\mathrm{RTP}}, S_i\big)$ because $\operatorname{Var}(S_i) = \mathbb{E}\big[(S_i - \mathbb{E}[S_i])^2\big] \Rightarrow \partial \operatorname{Var}(S_i)/\partial x_{ih} = 2\,\mathbb{E}\big[(S_i - \mathbb{E}[S_i])(p_h^{\mathrm{RTP}} - \mathbb{E}[p_h^{\mathrm{RTP}}])\big]$. Stationarity w.r.t. x_{ih} gives

$$\mathbb{E}[p_h^{\mathrm{RTP}}] + 2\gamma_i \Big(p_h^{\mathrm{RTP}}, \sum_\ell p_\ell^{\mathrm{RTP}} x_{i\ell}\Big) + \psi_i(h) - \mu_i \eta_{ih} + \alpha_{ih} - \beta_{ih} = 0.$$

At interior hours ($0 < x_{ih} < \bar{x}_i$), $\alpha_{ih} = \beta_{ih} = 0$, yielding

$$\mathbb{E}[p_h^{\mathrm{RTP}}] + 2\gamma_i \Big(p_h^{\mathrm{RTP}}, \sum_\ell p_\ell^{\mathrm{RTP}} x_{i\ell}\Big) + \psi_i(h) = \mu_i \eta_{ih},$$

which is (19). Complementary slackness for α_{ih}, β_{ih} covers corner hours.

implies all lower-cost hours $\{o : \kappa_{io} < \kappa_{ip}\}$ are saturated or infeasible. Aggregating over i, IO (weakly) lowers peak load and (weakly) raises off-peak load while daily kWh per EV is unchanged.

Proof. Let $Q_h = b_h + \sum_i x_{ih}$. The Lagrangian is

$$\mathcal{L}(\{x_{ih}\},\{\mu_i\},\{\mu_{ih}^{\pm}\}) = \sum_h \Big[c_h(Q_h) - r_h(Q_h)\Big] + \sum_{i,h} \psi_i(h) x_{ih} \\ + \sum_i \mu_i \left(E_i - \sum_{\substack{h \le T_i \\ h \in \mathcal{A}_i}} \eta_{ih} x_{ih} \right) + \sum_{i,h} \mu_{ih}^{+}(x_{ih} - \bar{x}_i) - \sum_{i,h} \mu_{ih}^{-} x_{ih}. \quad (21)$$

Stationarity w.r.t. x_{ih} gives

$$\frac{\partial \mathcal{L}}{\partial x_{ih}} = \underbrace{c_h'(Q_h) - r_h'(Q_h)}_{\tilde{c}_h} + \psi_i(h) - \mu_i \eta_{ih} + \mu_{ih}^{+} - \mu_{ih}^{-} = 0.$$

If $0 < x_{ih} < \bar{x}_i$ then $\mu_{ih}^{\pm} = 0$ and $\mu_i = \frac{\tilde{c}_h + \psi_i(h)}{\eta_{ih}} \equiv \kappa_{ih}$. If $x_{ih} = 0$ then $\mu_{ih}^{-} \ge 0$ and $\kappa_{ih} \ge \mu_i$; if $x_{ih} = \bar{x}_i$ then $\mu_{ih}^{+} \ge 0$ and $\kappa_{ih} \le \mu_i$. Hence x_{ih} is nonincreasing in κ_{ih}: the IO scheduler fills the lowest adjusted-cost hours first, up to feasibility. If $x_{ip} > 0$ while some feasible o has $\kappa_{io} < \kappa_{ip}$ and $x_{io} < \bar{x}_i$, then $\kappa_{ip} \le \mu_i \le \kappa_{io}$, a contradiction. Summing over i yields load shifted from high-$\tilde{c}_h$ to low-$\tilde{c}_h$ hours, conserving daily energy by Eq. (16).

Lemma G.0.2 (Elasticity ordering)**.** *In high plug-in hours,*

$$|\epsilon_h^{\mathrm{IO}}| \;\ge\; |\epsilon_h^{\mathrm{RTP}}| \;\ge\; |\epsilon_h^{\mathrm{ToU}}|.$$

Proof. Step 1 (IO upper-bound response). Consider a small perturbation $d\tilde{c}$ concentrated in hour h. Differentiating the KKT system in Lemma G.0.1 yields a linear system in $\{dx_{ih}, d\mu_i, d\mu_{ih}^{\pm}\}$ whose solution preserves the merit order: $dx_{ih} \le 0$ at the perturbed hour and $dx_{i\ell} \ge 0$ at some lower-$\kappa_{i\ell}$ hours, with $\sum_{\ell \le T_i, \ell \in \mathcal{A}_i} \eta_{i\ell} dx_{i\ell} = 0$ (intra-day reallocation). Aggregating over i,

$$\frac{\partial Q_h^{\mathrm{IO}}}{\partial \tilde{c}_h} \;\le\; 0, \qquad \left| \frac{\partial Q_h^{\mathrm{IO}}}{\partial \tilde{c}_h} \right| \text{ is maximal subject to Eq. (16).}$$

Thus $|\epsilon_h^{\mathrm{IO}}|$ is an upper bound among feasible reallocations.

Step 2 (RTP attenuation by risk/attention). From the household FOC Eq. (19), totally differentiate across hours. Stacking in vector form,

$$H_i\,dx_i \;=\; -\Big(I + 2\gamma_i\,\Sigma_p\,S_i\Big)\,dp,$$

where H_i is the Hicksian substitution matrix (negative semidefinite), Σ_p is the covariance matrix of prices, and S_i maps p to $\sum_\ell p_\ell x_{i\ell}$. The matrix $(I + 2\gamma_i \Sigma_p S_i)$ is positive semidefinite for $\gamma_i \geq 0$; multiplying a negative semidefinite H_i by such a factor weakly *shrinks* the response (Loewner order). Furthermore, $a_i > 0$ creates inactive hours (corners), reducing the response support. Aggregating over i preserves attenuation:

$$\left\|\frac{\partial Q^{\mathrm{RTP}}}{\partial p}\right\| \;\leq\; \left\|\frac{\partial Q^{\mathrm{IO}}}{\partial \tilde{c}}\right\|.$$

Step 3 (ToU as projection). Let P be the block-averaging operator mapping hourly prices to ToU blocks ($P^2 = P$, $\|P\|_2 \leq 1$). For small perturbations, $dQ^k \approx H^k\,d\pi^k$ with H^k negative semidefinite. Under ToU, $d\pi^{\mathrm{ToU}} = P\,d\pi^{\mathrm{RTP}}$, hence

$$\|dQ^{\mathrm{ToU}}\|_2 \;=\; \|H^{\mathrm{ToU}}\,P\,d\pi^{\mathrm{RTP}}\|_2 \;\leq\; \|H^{\mathrm{RTP}}\,d\pi^{\mathrm{RTP}}\|_2,$$

so $|\partial Q_h^{\mathrm{ToU}}/\partial p_h| \leq |\partial Q_h^{\mathrm{RTP}}/\partial p_h|$. Combining Steps 1–3 yields the stated ordering.[73]

Lemma G.0.3 (Welfare ranking with frictions)**.** *With $a_i, \gamma_i > 0$,*

$$W^{\mathrm{IO}} \;\geq\; W^{\mathrm{RTP}} \;\geq\; W^{\mathrm{ToU}} \;\geq\; W^{\mathrm{Flat}}.$$

Proof. Let W^k be per-EV social welfare (net of pure transfers). IO maximizes W for the

[73]We do not analyze a bill-variance ordering in our empirical set-up, but with $\mathrm{Bill}^k = \sum_h p_h^k X_h^k$, the law of total variance gives

$$\mathrm{Var}(\mathrm{Bill}^k) = \mathbb{E}\left[\mathrm{Var}\left(\sum_h p_h^k X_h^k \,\middle|\, \mathbf{p}^k\right)\right] + \mathrm{Var}\left(\mathbb{E}\left[\sum_h p_h^k X_h^k \,\middle|\, \mathbf{p}^k\right]\right).$$

Moving from RTP to ToU replaces $\mathbf{p}$ by $P\mathbf{p}$, where P is a contraction ($\|P\|_2 \leq 1$) and a Blackwell coarsening. Both the within-state term and the between-state term weakly fall. Under IO (EV sub-load), consumers face a flat off-peak retail rate; wholesale volatility is internalized by the aggregator, yielding

$$\mathrm{Var}(\mathrm{Bill}^{\mathrm{RTP}}) \;\geq\; \mathrm{Var}(\mathrm{Bill}^{\mathrm{ToU}}) \;\geq\; \mathrm{Var}(\mathrm{Bill}^{\mathrm{IO}}).$$

EV sub-load given $\{\tilde{c}_h\}$ (Lemma G.0.1 and KKT optimality). Relative to IO,

$$W^{\mathrm{RTP}} = W^{\mathrm{IO}} - \underbrace{\Delta W_{\mathrm{coord}}}_{\geq 0} - \underbrace{\mathbb{E}[a_i]}_{>0} - \underbrace{\mathbb{E}[\gamma_i \, \mathrm{Var}(\mathrm{Bill}^{\mathrm{RTP}})]}_{>0}.$$

ToU coarsens the signal (projection loss $\Delta W_{\mathrm{gran}} \geq 0$) and Flat removes intertemporal incentives entirely. Hence the stated ordering.

Proposition G.0.1 (Emissions ordering)**.** *If e_h is (weakly) lower off-peak, then relative to baseline: IO achieves the largest emissions reduction, followed by RTP, then ToU, then Flat.*

Proof. By Lemma G.0.1, IO shifts the most load into low-$\tilde{c}_h$ hours, which coincide with low-e_h by assumption; RTP and ToU shift less; Flat does not reallocate. Summing $e_h x_{ih}$ across hours yields the ordering.

G.0.4 Overrides ("Bump Charging"): Behavior and Welfare

Override decision and probability. A household overrides in hour h iff

$$v_{ih} > m_{ih} + \phi_i, \tag{22}$$

where v_{ih} is the immediate utility from charging now and m_{ih} the marginal benefit from deferring to the IO-planned hour. If F_{ih} is the CDF of v_{ih}, the override probability is

$$\beta_{ih} \; = \; \Pr\Big[v_{ih} > m_{ih} + \phi_i\Big] \; = \; 1 - F_{ih}(m_{ih} + \phi_i). \tag{23}$$

Lemma G.0.4 (Override monotonicity)**.** *If F_{ih} is nondecreasing, then β_{ih} is weakly decreasing in ϕ_i and in m_{ih}.*

Proof. Differentiate Eq. (23) (where densities exist): $\partial\beta_{ih}/\partial\phi_i = -f_{ih}(m_{ih} + \phi_i) \leq 0$ and similarly for m_{ih}. Without densities, monotonicity follows from the CDF order.

Lemma G.0.5 (Welfare effect of an override)**.** *Let h' be the IO-planned hour and h the override hour for EV energy $q_{ih} > 0$. Private net benefit is $v_{ih} - m_{ih} - \phi_i > 0$ when overriding. Social cost changes by $(\tilde{c}_h - \tilde{c}_{h'})\,q_{ih}$; if $\tilde{c}_h > \tilde{c}_{h'}$ the override raises system cost.*

Proof. Private part is by Eq. (22). Social part: the IO plan equalizes marginal costs across used hours; deviating to higher-$\tilde{c}_h$ increases procurement net of r_h by $(\tilde{c}_h - \tilde{c}_{h'})\,q_{ih}$.

Proposition G.0.2 (Aggregate override rate). *If v_{ih} are i.i.d. with CDF F_v and $m_{ih} \in \{m^{\mathrm{chg}}, m^{\mathrm{def}}\}$ depending on whether IO planned charging in h, and if ρ is the share of hours planned to charge, then*

$$\bar{\beta} \;=\; (1-\rho)\Big[1 - F_v(m^{\mathrm{def}} + \bar{\phi})\Big] \;+\; \rho\Big[1 - F_v(m^{\mathrm{chg}} + \bar{\phi})\Big],$$

where $\bar{\phi}$ is the (mean) hassle cost.

Proof. Law of total probability conditioning on planned status; apply Eq. (23) in each state and average.

Welfare with overrides and crossover. Let W^{IO} and W^{RTP} be per-EV welfare absent overrides and define the baseline gap

$$\Delta W_0 \;=\; W^{\mathrm{IO}} - W^{\mathrm{RTP}}. \tag{24}$$

If overrides occur at rate $\bar{\beta}$ with per-override loss λ, IO welfare becomes

$$W^{\mathrm{IO+O}} \;=\; W^{\mathrm{IO}} - \bar{\beta}\,\lambda. \tag{25}$$

We calibrate λ by

$$\hat{\lambda} \;\approx\; p_{\mathrm{peak}} \cdot \Delta\tilde{c}_{\mathrm{eff}} \cdot q^{O}, \qquad \lambda^{\mathrm{UB}} \;=\; (\max_h \tilde{c}_h - \min_h \tilde{c}_h) \cdot q^{O}_{\max}, \tag{26}$$

where p_{peak} is the probability an override lands in peak, $\Delta\tilde{c}_{\mathrm{eff}}$ the peak–off-peak spread (£/kWh), and q^O kWh shifted per override.

Lemma G.0.6 (IO–RTP crossover threshold). *If $\Delta W_0 > 0$ and $\lambda > 0$, the override rate at which IO with overrides equals RTP is*

$$\bar{\beta}^{\star} \;=\; \frac{\Delta W_0}{\lambda}. \tag{27}$$

Proof. Equate Eq. (25) to $W^{\mathrm{RTP}} = W^{\mathrm{IO}} - \Delta W_0$ and solve for $\bar{\beta}$.

Proposition G.0.3 (Welfare ordering with overrides). *If $\bar{\beta} < \bar{\beta}^{\star}$, then $W^{\mathrm{IO+O}} > W^{\mathrm{RTP}}$; if $\bar{\beta} > \bar{\beta}^{\star}$, RTP can dominate.*

Proof. From Eq. (25) and Lemma G.0.6, the sign of $W^{\mathrm{IO+O}} - W^{\mathrm{RTP}} = \Delta W_0 - \bar{\beta}\lambda$ is determined by $\bar{\beta}$ relative to $\bar{\beta}^{\star}$.

G.1 Cross-country calibration and crossover analysis

The experiment (IO vs. alternatives) identifies (i) peak-to-off-peak reallocation magnitudes (Lemma G.0.1); (ii) override frequencies $\bar{\beta}$ and their timing relative to peak/off-peak (for $\hat{\lambda}$ via Eq. (26)); and (iii) the baseline welfare gap ΔW_0 via cost/benefit accounting under no overrides. These map directly into the decision rule induced by Lemma G.0.6.

G.1.1 Objective and overview

This section develops a cross-country calibration of the *override–welfare* relationship in IO systems. The goal is to identify, for each major IO market, the *crossover override rate* – that is, the frequency of user overrides that would eliminate IO's welfare advantage over real-time pricing (RTP).

The analysis combines three empirical components: (1) the observed probability of an override per charge day, from a random sample of users across the United Kingdom, Germany, Spain, and the United States; (2) the conditional peak-landing probability $p_{\text{peak}} = \Pr(\text{peak bump/day})/\Pr(\text{bump/day})$; and (3) country-level parameters describing tariff spreads and average energy per override q^O (kWh per bump event), drawn from IO Go pricing data.

G.1.2 Estimating ΔW_0 for the United Kingdom

This appendix documents how we estimate the baseline welfare gap $\Delta W_0^{(\text{UK})}$ – the per-EV daily welfare advantage of IO relative to RTP in the absence of overrides. The welfare gap is defined as the expected daily difference in total surplus between IO and RTP regimes:

$$\Delta W_0 = \left[U^{\text{IO}} - C^{\text{IO}}\right] - \left[U^{\text{RTP}} - C^{\text{RTP}}\right],$$

where U^k is the household mobility utility net of timing disutility and C^k the expected system procurement cost. Under RTP, households face optimization and attention frictions that reduce responsiveness; under IO, scheduling is automated, internalizing system prices without these frictions.

For the UK, we set the representative elasticity of RTP demand at $\varepsilon = -0.2$ and assume a daily attention cost of $a = 0.20$ per day. Using observed load and tariff data (off-peak

£0.07/kWh, peak £0.27/kWh; spread $\Delta\tilde{c}_{\text{eff}} = 0.20$/kWh), we simulate hourly charging schedules under both IO and RTP. Automated IO coordination reallocates approximately 0.375 kWh of energy per day from peak to off-peak hours, yielding procurement-cost savings of about $0.20/\text{kWh} \times 0.375 = 0.075$ per day. In addition, IO reduces timing disutility and bill-variance penalties by a further £0.020, as households experience fewer inconvenient charging hours and lower price volatility exposure. Together these components yield: $\Delta W_0^{(\text{UK})} = 0.075 + 0.020 = 0.095$ per day.[74]

G.1.3 Empirical calibration of override behavior

Table A15 reports observed override behavior across countries. The United Kingdom exhibits a low overall bump frequency (1.7% per charge day) and correspondingly low peak-landing probability (0.3%), yielding $p_{\text{peak}} = 0.176$. Early adopters show slightly higher values ($p_{\text{peak}} = 0.23$). In Germany and Spain, users override much more frequently (10% per day), while in the United States roughly one-third of overrides fall in peak windows. Average energy per override q^O ranges from 1.25 kWh in Spain to around 2.4 kWh in the UK and US (see Table 17 below). Tariff spreads $\Delta\tilde{c}_{\text{eff}}$ vary widely, from £0.17/kWh in the UK to only £0.037/kWh in the US. These differences strongly affect the expected welfare loss per event.

Table 17: Tariff and override characteristics by country

Country	Off-peak rate	Standard rate	q^O (kWh)	$\Delta\tilde{c}_{\text{eff}}$ (£/kWh)	% bumped in peak
UK	£0.24	£0.07	2.41	0.17	0.3
Germany	€0.27	€0.39	2.06	0.12	1.9
Spain	€0.07	€0.128	1.25	0.058	0.7
US	\$0.11	\$0.147	2.45	0.037	1.1

G.1.4 Per-override welfare loss

Using the expression $\hat{\lambda} = p_{\text{peak}} \Delta\tilde{c}_{\text{eff}} q^O$, we obtain the average loss per override (in £). Table 18 presents results under $\varepsilon = -0.2$ and $\Delta W_0 = 0.095$. Price spreads are expressed in pounds for comparability.

[74] For the remaining countries, we scale ΔW_0 proportionally to the observed effective price spread $\Delta\tilde{c}_{\text{eff}}^{(c)}$, following $\Delta W_0^{(c)} = \Delta W_0^{(\text{UK})} \times \left(\Delta\tilde{c}_{\text{eff}}^{(c)} / \Delta\tilde{c}_{\text{eff}}^{(\text{UK})}\right)$. This assumes similar elasticities and flexible load shares across markets, attributing cross-country differences in welfare to the strength of the intertemporal tariff gradient. Empirically, this yields

Table 18: Cross-country crossover analysis (RTP $\varepsilon = -0.2$)

Country	p_{peak}	$\Delta\tilde{c}_{\text{eff}}$ (£/kWh)	q^O (kWh)	$\hat{\lambda}$ (£/override)	Full-load $\bar{\beta}^\star$	Partial ($\alpha = 0.25$)	Observed rate
UK	0.1	0.17	2.41	0.072	1.32	0.33	0.0186
Germany	0.205	0.12	2.06	0.051	1.87	0.47	0.107
Spain	0.083	0.058	1.25	0.006	15.83	3.96	0.105
US	0.367	0.037	2.45	0.033	2.88	0.72	0.035

Notes: The expected welfare loss per override is $\hat{\lambda} = p_{\text{peak}}\Delta\tilde{c}_{\text{eff}}q^O$, where p_{peak} is the probability an override lands in peak, $\Delta\tilde{c}_{eff}$ the peak-off-peak spread (£/kWh), and q^O kWh shifted per override. Crossover rates $\bar{\beta}^\star$ (events/EV–day) are computed as $\Delta W_0/\hat{\lambda}$, with $\Delta W_0 = 0.095$ and $\alpha = 0.25$ for partial-load automation. Observed rates reflect empirical daily override frequencies. The US is from Texas.

We can see that for the four countries, the observed override rate is well below the crossover override rate. In the UK, there is a low override frequency (1.9% per day) and a moderate price spread yield $\hat{\lambda} \approx 0.072$. IO remains welfare-superior up to $\bar{\beta}^\star(0.25) = 0.33$ per day, an order of magnitude above observed behavior. For Germany, more frequent overrides and a nontrivial tariff gradient narrow IO's margin, but the observed rate (0.107) remains below $\bar{\beta}^\star(0.25) = 0.47$. Germany is the tightest case, where welfare parity could emerge if overrides doubled. For Spain, flat tariffs imply tiny welfare penalties ($\hat{\lambda} \approx 0.006$). Even frequent overrides barely affect welfare: the crossover is 15.8/day (full-load) and 3.96/day (α=0.25), far above observed 0.105. In the US (Texas), high p_{peak} but very small $\Delta\tilde{c}_{\text{eff}}$ yield $\hat{\lambda} = 0.033$ and a partial crossover of 0.72/day. Observed 0.035/day lies safely below this threshold, though steeper future tariffs could tighten the margin. To note, the $\bar{\beta}^\star$ are expressed as override events per EV–day rather than probabilities. A value above 1 therefore indicates the rate at which Intelligent Octopus would lose its welfare advantage if users could, on average, override more than once per day. In practice, each vehicle can generate at most one override per plug-in episode, so observed rates (0.02–0.11 per day) are far below this threshold. Thus, $\bar{\beta}^\star > 1$ should be interpreted as meaning that the tariff remains welfare-superior even if every vehicle overrode on every charging day.

Across all countries, observed rates (0.019–0.11 per day) are an order of magnitude below the corresponding α=0.25 crossovers, confirming that IO remains welfare-superior even with only 25% of load automated.[75]

$\Delta W_0^{(\text{DE})} \approx 0.067$, $\Delta W_0^{(\text{ES})} \approx 0.032$, and $\Delta W_0^{(\text{US})} \approx 0.021$.

[75] Higher elasticity (e.g. $\varepsilon = -0.5$) reduces ΔW_0 and proportionally lowers $\bar{\beta}^\star$, narrowing IO's advantage by roughly 20%. Yet even under $\varepsilon = -0.5$, IO dominates in all countries except possibly Germany.

Limitations. Our theoretical framework abstracts from several important considerations. First, the welfare thresholds depend on calibrated parameters for attention costs (a_i), hassle costs (ϕ_i), and risk-aversion penalties (γ_i) that are not directly observed in our experiment; while we draw on the literature and perform sensitivity analysis, these inputs remain assumption-driven.

Second, we model overrides as independent events with constant per-override welfare loss λ, whereas in reality override behavior and costs may vary systematically across hours, days, and households.

Third, we treat plug-in behavior as exogenous to the tariff, assuming it remains fixed across pricing regimes. In practice, customers may adapt when and how they charge in response to tariff signals. For instance, RTP could encourage shifting plug-in times toward periods of lower prices and across days, whereas IO Go's ToU pricing is fixed, and cannot flexibly incentivize plug-in. We lack data on plug-in behavior for customers who are not on IO Go, but future work could incorporate these behavioral responses to better capture the interaction between tariff design and charging behavior.

Finally, our analysis is partial-equilibrium: prices $\{\tilde{c}_h\}$ are taken as given and households are price-takers, omitting possible feedback effects if IO Go or RTP adoption is widespread. These simplifications are deliberate, allowing a tractable link between model predictions and experimental data, but they should be borne in mind when interpreting the welfare rankings.