Centralized vs Decentralized Demand Response: Evidence from a Field Experiment*

Megan Bailey† David P. Brown‡ Blake Shaffer§
Frank A. Wolak¶

January 2023

Preliminary and Incomplete.
Please do not cite or circulate.

Abstract

We report on a large-scale experiment to evaluate the effectiveness of centrally-managed (utility-controlled) versus customer-controlled demand response to dynamic prices by residential customers. Both groups of households have wireless communication devices installed in their dwelling to control as many as three types of appliances using a mobile phone app: (1) baseboard thermostats, (2) hot water heaters, and (3) electric vehicle chargers. Centrally-managed households can only use the mobile app to op-out of a utility-controlled demand response event, whereas customer-controlled households must use the mobile app to reduce the electricity use of any of these appliances. We find that the centrally-controlled households reduce demand by 27% on average during 3-hour demand response events and customer-controlled households reduce their demand on average by 6% during 3-hour demand response events, even though both sets of households receive the same financial compensation for a same percentage demand reduction. The percent demand reduction customer controlled households is not statistically different from that of third group of experimental households that have no ability to remotely control their appliances and only receive the demand response event signal on the mobile phone app and receive the same compensation for their actions during demand response events as the other two groups.

*We are grateful to our partner electric utility company for sponsoring and managing the field experiment on which this study is based. We are grateful to attendees at the 2022 Berkeley Energy Camp, 2022 Electricity Camp in the Rockies, and 2022 Canadian Resource and Environmental Economics annual conference for useful feedback. This work is supported by excellent research assistance from Mallika Sharma. We acknowledge funding from our utility partner, the University of Calgary, and the Canada First Research Excellence Fund as part of the University of Alberta's Future Energy Systems research initiative and the University of Calgary's Global Research Initiative. This experiment is registered in the AEA RCT Registry under ID AEARCTR-0008680. All errors are are own.

†Haskayne School of Business and the School of Public Policy, University of Calgary megan.bailey@ucalgary.ca

‡Department of Economics, University of Alberta dpbrown@ualberta.ca

§Department of Economics and the School of Public Policy, University of Calgary blake.shaffer@ucalgary.ca

¶Department of Economics, Stanford University and NBER wolak@stanford.edu

1 Introduction

A fundamental challenge of electricity markets is the need to balance supply and demand at every instant. This challenge is exacerbated by limited storability of electricity and inelastic demand. Historically, this balance has been met by forecasting demand and adjusting (dispatching) supply. Going forward, however, as policymakers work to curb greenhouse gas emissions from electricity and other sectors, we will see an increase in intermittent renewable generation in the electricity supply as well as a new suite of flexible sources of demand (such as electric vehicles). This means that electricity market operators may benefit from flipping the prior adage upside down, forecasting supply and dispatching demand.

Economists have long recognized the benefits of price-responsive demand to improve the efficiency of electricity markets (Boiteux, 1951; Kahn, 1970), often advocating for the first-best policy prescription of dynamic pricing (Joskow and Wolfram, 2012). However, given the presence of information acquisition costs and adjustment costs—or more broadly transaction costs—as well as behavioral biases, such dynamic pricing may actually be sub-optimal (Fabra et al., 2021; Sallee, 2014; Schneider and Sunstein, 2017). Residential electricity demand offers a setting where the conditions are ripe for such issues to be material. Electricity consumers are generally inattentive to both their electricity price and usage (perhaps rationally so), and even when made aware, the task of acting on this information can be onerous.

In this paper, we focus on the role of transaction costs in inhibiting flexible demand response from residential electricity customers. Working with a large electric utility, we recruited over 1,000 households to participate in a field experiment and randomized participants into treatment and control groups. Participants in our "Central" treatment group have the electricity consumption of some of their home devices (thermostats, water heaters, and/or electric vehicle [EV] chargers) centrally managed by the utility during "peak events" for compensation. Households in this group must opt-out of such management, so their default (or passive) action is to have their electricity consumption of these devices reduced on their behalf. Our Central group therefore experiences elimination of in-the-moment transaction costs for responding to electricity prices. In contrast, our Tech group received the same remote control capability for these devices but must actively reduce electricity consumption during events. We compare our Central and Tech groups to understand the extent to which active dynamic effort prevents consumption adjustments to prices. Our Manual treatment group receives compensation for reducing electricity consumption during peak events but received no utility management or remote control capability of devices. A comparison of our Tech and Manual groups yields a measure of how much transaction costs are resolved with enabling technology in this setting.[1]

We estimate the causal effect of receiving peak event incentives for each group by randomizing events at the household-day-time of day-incentive level *within* treatment groups. This ensures that the events are not correlated with other drivers of household electricity consumption. In total, we analyze approximately 10 million hourly household-level consumption observations with over 15,000 household-event days. These data are supplemented by including data on approximately 500 devices with installed load controllers (that allow remote control of devices) in the Central and Tech groups, yielding over 8 million hourly observations.

The main findings are stark. We find the Central group reduces demand by an average of 27% during events, as compared to only 4% for the Tech group. Perhaps more surprisingly, we see no meaningful difference between the Tech group's performance and the Manual group's, which has a 6% reduction on average during events. These main results speak to the importance of minimizing transaction costs in settings where both inattention is rife and the incentives per event are relatively small.

One possible trade-off for the superior performance of the Central group is that households will be less inclined to accept this managed treatment. Despite our expectations, we find relatively limited differences in the take-up rates of the Central and Tech group programs (which were 42% and 48%, respectively). This surprising result indicates that consumers were not deterred by the idea of the electricity utility managing the electricity consumption of their devices.

The difference in responsiveness to incentives between the Central and Tech groups remains in our intention-to-treat (ITT) estimates (15% vs. 3%), which encompass both take-up rates and per-event electricity consumption responses by pro-

[1] Households in all of our treatment groups as well as an additional "Info" group receive real-time household-level electricity consumption data via a mobile phone app. Households in the Central and Tech groups observe device-specific consumption for devices with remote control capability. Therefore, we resolve part of the electricity usage consumption information barriers households face in responding to prices; consumers in our experiment remain unaware of other appliance- or device-specific consumption unless they seek that information. Households in the Central, Tech, and Manual treatment groups receive randomized event notifications via their mobile phone app whereby monetary rewards could be earned for reducing electricity demand during peak events. The rewards per event range from $1 to $6 Canadian dollars, designed to mimic observed electricity prices during electricity market "peak" conditions when matching supply and demand has been challenging. This corresponds to approximately $1.18 to $2.35 per kWh as compensation for consumption reduction for the average household in our sample.

gram/group. These estimates indicate that centralized utility management of household device electricity use during peak events is the clear program winner (of those considered here) for utility companies considering demand-side measures.

A novel feature of our experiment is the inclusion of EV chargers and electric hot water heaters.[2] These two devices in the home generate large shares of household electricity demand. We see stark differences between the event treatment effect estimates by homes with various devices across the Central and Tech groups. For example, we find that households in the Central group with hot water heaters reduce their demand by 24.5% during events, as compared to 0.3% for similar households in the Tech group.[3] This stands in contrast to households with EV chargers, where reductions during events are similar at 15% for the Central and 11% for Tech group. These differences across the groups for different appliances speak to the greater understanding and/or salience of EV charging as compared to other devices, such as hot water heaters, in consumers' willingness to take action. It provides an important caveat to extrapolating our findings to other demand response programs and time-varying price incentives to motivate EV charge responsiveness. At the same time, the difference in estimates across groups for hot water heaters highlights potential undiscovered sources of electricity demand flexibility.

Demand response, or more generally flexible electricity demand, is of heightened contemporary importance for several reasons. First, the growth of variable renewable energy has increased the value of flexible demand as a resource to help maintain grid reliability. Second, electrification adds to the need for flexibility to avoid costly infrastructure upgrades from large demand changes, such as from EV chargers. At present, the paradigm of increasing supply capacity to meet inflexible demand involves a social cost that compensates for demand-side market failures, and this stands to become a worsening problem due to large demand changes on the horizon. These changes also, however, mean the magnitude of potential savings to be had from flexible demand increases for consumers, making it more likely to overcome the fixed hurdle costs preventing action. Third, Internet-of-Things (IoT) technology improvements have lowered the cost and effort for consumers to respond to price signals and other events.

[2]Much of the literature on automated electricity demand response has largely focused on smart thermostats. See, for example, Blonz et al., 2021, Brandon et al., 2022, Bollinger and Hartman, 2020.

[3]We do not obtain these estimates for the Manual group because these households do not have load controllers installed.

The importance of flexible electricity demand was on display in California in Summer of 2022, where statewide appeals for conservation narrowly avoided rolling blackouts (Balaraman, 2022). This event, as well as the February 2021 blackout in Texas during a cold snap (King et al., 2021), illustrates the urgency of the problem. A critical policy question is whether greater consumer demand flexibility could be cheaper to facilitate than costly system upgrades to accommodate changing patterns in electricity demand and supply. The behavioural question addressed in this paper is how best to overcome impediments to otherwise beneficial demand response.

Our paper builds on several strands of the literature. First, we add to the rich set of empirical research estimating household responsiveness to time-varying pricing in electricity.[4] Our experimental design is most similar to the critical peak pricing (CPP) strand of this literature. The results from our Manual group with no technology nor automation is most inline with the settings observed in prior CPP studies. Our main estimate of a 6% reduction in the Manual group falls within the range of past findings.

Second, our paper contributes to a growing strand of literature that explores automation options for consumers to overcome transaction costs and other barriers to demand response. There is evidence that automation of smart thermostats can assist in facilitating short-run demand responsiveness when combined with pricing (Harding and Lamarche, 2016; Bollinger and Hartmann, 2020; Blonz et al., 2021). However, some papers suggest that consumers may override important settings with such technology (Brandon et al., 2022). Somewhat consistent with these papers, we find that our Tech group performs no better than the Manual group in altering electricity consumption in response to events. That is, given the ability to remotely control large appliances as well as automate some aspects of their electricity usage (such as thermostat settings and turning back on EV chargers and water heaters after events), they fare no better than consumers who require a more manual action to change electricity consumption. In our setting, assistive technology is not resolving demand-side failures in price coordination.

Our key contribution beyond the existing literature is the finding of significantly greater responsiveness when demand response is made the default, or passive, action. Requiring customers to actively opt-in—even with the provision of control technology that makes the transaction cost as minimal as pushing a single button on an app—is no match for the power of switching the default action to one that opts the customer in to demand response events.

[4] See Faruqui and Sergici (2010) and Yan et al. (2018) for surveys of this literature.

The policy analog of our Central group is a version of demand response whereby the utility centrally initiates infrequent demand reductions, with the user needing to actively opt-out of an event to not participate. This is sometimes called active load management or dynamic load management, with most adjustments going unnoticed by the household. The benefits of this approach are likely to increase as localized demand response coordination will be required to reduce costs on the distribution system, not just the broader energy system, as households electrify larger loads, such as electric vehicles. A managed system can resolve local coordination challenges in ways broader price signals may not. Our research finds that, at least at the intensive margin of demand response, such an approach is likely to elicit the largest and most consistent response.

Our analysis proceeds as follows. Section 2 describes our experimental design. In Sections 3 and 4, we describe the data we obtain through the experiment and provide descriptive statistics of our sample. We then lay out our empirical framework, including our estimation strategies and robustness checks, in Section 5. In Section 6, we present our results, and in Section 7, we conclude.

2 Experimental Design

In this section, we provide an overview of our experimental design. We then summarize our various experimental groups and event categories. We describe our sample selection criteria and recruitment strategy, and we detail how we randomized households into the different experimental groups.

2.1 Overview

We explore how centralized decision-making can overcome barriers in customers expressing demand preferences in electricity markets via a large-scale field experiment. We partnered with a large regulated Canadian electric utility (hereafter referred to as the "Utility") to randomly offer customers one of several programs. One of these programs involves the Utility remotely controlling customers' select devices during "peak events", times in which customers are offered compensation for reducing electricity demand. Households that were offered and accepted each program (described below) subsequently comprise each of our treatment groups, and our sample consists of all treatment and control groups.[5]

Notably, we recover causal treatment effects, the average effect of an event household consumption by treatment group, by randomizing events at the household-day-time of day-event "type" level. That is, average treatment effects are recovered by randomizing a treatment (events) *within* treatments groups. Households within each treatment group that receive events (described below) receive a random, independent schedule of events. In our main specification, identification of our parameters of interest that give us the average household electricity consumption response to events, by group, relies on comparing event-time consumption from households that randomly receive events to non-event consumption from all households that do not receive events at a particular time, conditional on controls. With this design, we do not need to be concerned about selection affecting our identification.[6]

2.2 Treatment Groups

In this section, we provide an overview of our experimental groups as they relate to our research questions and empirical design. We provide more information about the customer pool from which these groups were recruited, the recruitment process, technology installation and features, and electricity consumption reduction offers below.

Table 1 summarizes our experimental groups. The *Central* group is our centralized decision-making group. This group received load controller equipment that is controlled by the Utility to turn off the select devices during "peak events". This equipment was available for several major electricity-consuming devices: electric baseboard heaters, electric hot water tanks, and/or level 2 electric vehicle chargers. This equipment allows for enhanced control of the devices via the Utility's App, allowing households to turn on and off the devices remotely. Peak events (hereafter "events") are times when the Utility offers consumers compensation for reducing electricity consumption, which are framed as "rewards" or "earnings" in our experiment.

Our *Tech* group also received load controller equipment. However, households in this group differ from the Central group in that the Utility does not control the load

[5] We consider participants in each group as similar to those that would accept many utilities' offers for "demand-side management" programs. Indeed, they give us more conservative estimates of consumer responsiveness to these types of programs, as utilities often select and recruit customers for these programs that they anticipate will be the most responsive.

[6] Selection into our treatment groups is only an issue for any external validity concerns. Conditional on customers enrolling in one of the offered programs or being included in the control group, the treatment events are randomized. Therefore, we do not need to recover a LATE estimate or use other selection-correction techniques to obtain the effect of events on household consumption.

controller devices during events. Both the Central and Tech groups can remotely control the devices on which they have load controllers. However, while the Utility controls the Central groups' devices on their behalf, the Tech group must take action to turn off or adjust their devices.[7]

The *Manual* group earns financial rewards for demand reductions during events, but does not have load controller equipment. If there is an asymmetric response to events between the Manual and Tech groups, this captures the impact of the load controller technology in facilitating demand reductions–potentially resolving various informational limitations and transaction costs that consumers have in responding to changes in electricity prices.

All participants in the Central, Tech and Manual groups receive offers for financial compensation for reducing electricity consumption during peak events. The events are randomized at the day-household-event time-event type level.[8] All participants in these groups plus a fourth group, our *Info* group, receives real-time electricity usage information via the App.[9] The Info group does not receive offers to receive compensation for reducing electricity consumption during events or load controller equipment and allows us to observe and estimate the impact of these offers independently from the impact of providing consumers with real-time price information.

Finally, we have a *Control* group, a group of Utility customers whose electricity consumption we just observe. This group has never been contacted, and they can only view their electricity usage at a one-day lag through the same App.

In addition to the five groups above, we assign a subset of households to a *Choice* treatment group. Households that were allocated to this group had the opportunity to chose among the Central, Tech, or Manual groups. Households were provided information on the potential financial rewards, equipment, and experiences of each group and self-selected which group they'd like to join. (Appendix B.3 describes the recruitment materials sent to this group.) We asked these households to rank their preferences across these three offers. After we observed their preference ranking, we

[7] In general, the Tech group must respond to event offers in-the-moment by adjusting their devices; however, they do have the capability to pre-set their thermostats and make settings that ensure any device turns back on after events. As described below, all experimental groups receive the first notifications regarding events 21 hours in advance.

[8] As described below in Section 2.3, event times were constrained to two periods in the day in which electricity demand is typically high. Event types reflect two magnitudes of compensation households could receive for demand reduction, during evening events only.

[9] Households that received load-controllers also receive real-time information about electricity consumption from these specific devices on the App.

Table 1. Summary of Treatment Groups

Groups	DR Control	Control Tech	Price Incentive	Usage Info
Central	Utility	✓	✓	✓
Tech	Household	✓	✓	✓
Manual	Household		✓	✓
Info	Household			✓
Control	Household			

Notes. DR Control represents whether demand response to events is controlled entirely by the household (decentralized) or by the Utility for the load-controlled devices (centralized). Control Tech denotes whether the household has load controller equipment installed. Price Incentive reflects if households receive peak events and rewards for reduced demand during events. Usage Info denotes whether households receive real-time household-level consumption information. ✓indicates categories that are applicable to each group.

allocated these households to their most preferred group.[10] These households provide valuable insights into individuals' willingness to self-select into different treatment groups.

Throughout the discussion below, we will reference that we have six different groups. However, it is important to note that the Choice group is allocated to one of the three price incentive groups summarized in Table 1. The Central, Tech, and Manual groups have, respectively, 19, 13, and 6 participants each–10.1%, 7.1%, and 2.5% of their total household counts–that were given a choice between programs from being assigned to the Choice group. Section 6.4 provides a robustness check where we remove the Choice group participants to address concerns that these households and their responses to events are different from the randomly assigned households.

2.3 Events

We use peak events to evaluate households' responses to financial incentives, and determine if this response differs across the treatment groups. There are two event times and two event types. Events occur either in the morning (7 AM - 10 AM) or evening (5 PM - 8 PM). We consider two event types that have different reward levels: a "normal" and a "high" peak event. The high peak event receives elevated compensation for larger demand reductions. We designed high peak events for several

[10]We did not tell the households that they would be allocated to their most preferred group with certainty upfront. Rather, we asked for their preference ranking and indicated that we would take their preference ranking into account when we allocate them to the different potential groups and stated that the groups were subject to enrollment limits.

reasons: (1) Incentives for consumers to reduce electricity consumption, even during peak times, are relatively low on the margin. The peak and high peak event incentive structures capture the range of peak time pricing that consumers would face if they had real-time pricing. The peak price schedule allows us to observe behavior under common peak pricing, and the high peak schedule gives us a view of consumer behavior under high peak pricing incentives that may occur when electricity demand approaches maximum supply available. (2) We hypothesized that while the Central group would not need relatively large per-event incentives to make consumption changes, participants in other groups might. (3) Infrequent, large rewards may keep consumers engaged in a peak pricing rewards program. All morning events are normal peak events. Evening events are permitted to be either normal or high peak events.

Households receive financial rewards for demand reductions relative to their household-specific baseline consumption. Baselines reflect the average consumption during the relevant event time window of the last five weekdays prior to the event. We designed this rolling baseline to evade customer efforts to "game" their baseline by under-consuming during time when peak events occur. Customers do not know how the baseline is calculated.[11] We consider three reward levels that depend on the reduction in consumption relative to the household's baseline. During normal peak events, households receive \$1 for a 10% reduction, \$2 for a 30% reduction, or \$3 for a 50% reduction. High peak evening events elevate the incentives to \$1 for 10%, \$3 for 30%, or \$6 for 50% reductions. These incentives translate to incentive payments ranging from approximately \$1.18 to \$2.35 per KWh for the average household.[12] These incentives are in the range of wholesale price caps that are used to limit scarcity pricing in a number of jurisdictions in North America.[13]

For households that receive events, we randomly allocate events to households across all weekdays, excluding holidays. Consequently, households are unable to predict when they will be exposed to a peak event throughout the month. The two

[11]The baseline calculation was adjusted on March 3, 2022 (after 7 treatment days) from the average of the highest 3 of the past 5 weekdays to the average of the last 5 weekdays. This change was made to mitigate the impact of outliers on the baseline calculation.

[12]The average household consumes 1.7 KWhs in each hour in our sample. A 10%, 30%, and 50% reduction translates to a 0.51, 1.53, and 2.55 KWh reduction over the three-hour event, respectively. Consequently, for a 1 KWh reduction in consumption, we are compensating households on the lower bound $\frac{3}{2.55} = 1.18$ per KWh (i.e., \$3 for a 2.55 KWh reduction over three hours) for a 50% reduction during a normal peak event to a higher bound of $\frac{6}{2.55} = 2.35$ for a 50% reduction during a high peak event. The other percent reductions lie between these two cases.

[13]Examples include the wholesale price cap of \$1.00/KWh in Alberta (Brown and Olmstead, 2017), \$3.50/KWh in the Mid Continent Independent System Operator that operates in the Midwest United States (IRC, 2007), and \$5/KWh in Texas (Smith, 2022).

event times and types are randomly assigned to these event slots throughout the month. In each month, households experience one to five events, with an average of three events. There is an average of one normal evening event, one high peak evening event, and one morning event per month. This schedule was altered in the summer months of July and August, when the likelihood of peak events is low in Canada. During these months, households experience no "high peak" events. Events started on February 22, 2022 and continue for 18 months.

This randomization approach allows us to estimate treatment effects by household, using households' own consumption during non-event times as a control. In contrast to other common experimental approaches in the literature estimating consumer demand responses to electricity prices, we have unique identifying variation: Households do not all receive the events at the same time.

Households that are exposed to events receive event notifications 21 and 2 hours prior to the event. These event notifications provide information on the time of the event and the available financial incentives for the different demand reduction levels. When consumers receive the 21-hr notifications, they are also able to see event details in the App itself. See Appendix C.1 for examples of the notification and in-App event messages.

Households' rewards for consumption reduction during events are displayed in the App at a two-to-three day lag. The App also provides households with a summary of their total rewards to date. See Appendix C.2 for details.

2.4 Sample Selection, Recruitment, and Assignment

2.4.1 Phase I Recruitment: Onto the App

The study sample was drawn from the population of residential customers in the Utility's service territory in and near a large metropolitan city. We employed a two-step recruitment strategy. First, starting in August 2021, the Utility invited households to join a home electricity consumption management App operated by a third-party company, in partnership the Utility. The App provides households with household-level hourly consumption posted at 11 AM the following morning. The App can be coupled with other devices to provide more detailed information on real-time usage and device control. Households were recruited to the App using several marketing strategies, including advertisements on the Utility's website, social media posts, the Utility's newsletter, website notifications when users logged into their

Utility accounts, and emails sent to approximately 306,000 residential households.[14]

The recruitment onto the Utility's App provided us with a pool of 9,020 households to draw from. When households signed up to join the App, they were required to answer a six question survey. The survey asked households about their motivation for joining the App and whether the household rents or owns their home. It asked about devices eligible for load control in our experiment, including whether the household has an electric hot water tank, an electric vehicle (EV), air conditioning, and electric baseboard heaters as the primary heat source. Households with EVs were asked what type of charger (level 1 or 2) they use. [15]

We applied several selection criteria to this pool of households. Customers had to be in and near a large metropolitan city in the province for which it was feasible for Utility-partnered electricians to install load control equipment, as needed. Only homeowners were permitted to participate. Condos and apartments were removed, leaving primarily single-family homes, duplexes, and row homes as eligible. Households must have at least one month of historical consumption data as of September 2021, and the customers must have at least one controllable electric device. The set of controllable electric devices include a level 2 electric vehicle charger, electric baseboard heaters used as the primary heat source, and an electric hot water heater tank. This left us with a sample of 1,763 potential households that we used for our randomized assignment to experimental groups.

2.4.2 Phase II Recruitment: Group Offers

We describe the procedure we used to assign households to groups below. After the households were assigned to one of the experimental groups detailed in Section 2.2, we entered the second phase of recruitment. The Control group that receives no

[14]Emails were sent only to customers who had indicated prior to the recruitment that they could be contacted via email. One portion of the App Recruitment campaign targeted a set of customers in an energy-savings program consisting of regular communication from the Utility regarding energy-saving tips and incentives. We consider these customers sophisticated users. Approximately 64% of the households in our final sample (all groups combined) consist of these sophisticated users (639 out of 1005). The share of these users across groups is fairly even: 70% of each of the the Central and Tech groups, 63% of the Manual group, and 57% of the Info and Control groups (combined). We conduct a robustness check when estimating Equation 1 that interacts our coefficients of interest with sophisticated user status. We find no evidence that participants sophisticated users in the Central or Tech groups have a different response to events than others. We find limited evidence (statistically significant at the 10% level) that the Manual group participants who are sophisticated users reduce electricity demand more during events than non-sophisticated users. We are not concerned that this set of customers is driving our main treatment effects.

[15]Appendix B provides the text used in the initial survey.

equipment, price incentives, or real-time usage information (recall Table 1) received no further communication beyond joining the App in the first phase of recruitment. Starting in October 2021, we sent group-specific emails to households inviting them to join a new "Trial" program. These emails provided details about the specific experiences households would face in the group to which they were being invited, including a summary of the expected rewards they could earn over the course of the Trial, equipment they would receive and its estimated value, and future peak events. Households were also randomly offered a small sign-on incentive of the amounts $10 or $20, or no incentive. See Appendix B for a copy of these emails. Recall that all households faced a yes/no decision regarding whether they would accept our group-specific offer, except for the relatively small Choice group. The Choice group received an email asking households to rank their preferences between the Central, Tech, and Manual group offers. Households were also sent reminder emails regarding their offer, as per the schedule documented in Appendix B.2.

Each group recruitment email to the Central, Tech and Manual groups contained a link to a group-specific, informational website about the "Trial" to which customers were being invited. The websites provided as place for households to find complete information about the Trial and technology involved and an additional place to sign-up. This website was custom-designed by the research team in partnership with the Utility. Details about the website can be found in Appendix B.4. Additionally, invited household were given a Utility contact email for asking about the group offers. Emails responses were managed by responsive internal Utility staff managing the Trial.

Households had to actively accept the invitation to join the relevant experimental group. After selecting to join, households were either mailed a device called the "Hub" that facilitates monitoring real-time energy usage via the App or, for households in the Central and Tech groups, were contacted by installers to install the load controller equipment.[16]

This two-phase recruitment process occurred over the months August 2021 - February 2022. The second phase of recruitment occurred in fives waves starting in October 2021. As additional households joined the App, we collected the survey responses, identified eligible households, randomized households into groups, and sent the second-phase recruitment emails. This process was used to facilitate the time required to schedule and install the load controllers, as well as to achieve the targeted

[16]Households that were in the Manual and Info groups were mailed the Hub and did a self-installation, while households that received the Central and Tech group offers had to schedule installations of the equipment with professional electricians in coordination with the Utility.

sample size. [17]

2.4.3 Phase II Recruitment Randomization Procedure

We now summarize our randomization procedure. Our objective is to randomize households to the different treatment groups ensuring the characteristics of households are balanced across groups. To help ensure this is achieved, we first clustered households into groups using k-means clustering. This aims to create clusters of households with similar characteristics. We then randomized households within each cluster to one of the six groups outlined in Section 2.2. This process helps ensure that we do not end up with an unbalanced mix of households in any given group that have certain characteristics.

We use several variables to cluster including the cumulative consumption (in KWhs) and load factor by season (Fall, Spring, Winter, and Summer), variables that indicate if a household has an electric vehicle, electric baseboard heating, or air conditioning, and census data on median household income.[18] In each wave, we form 8 clusters providing a balance for the similarity of the households within each cluster while ensuring the clusters are not too small. For a subset of the eligible households (379, 21%), we have less than a full years worth of historical electricity consumption.[19] For this subset, we use the cumulative consumption and load factor for the month of September 2021 in the clustering procedure; this represents the last month of data prior to the second phase of our recruitment.

We randomized households into each group, placing a different weight on the number of households allocated to the groups based on our *ex-ante* expectations of the likelihood of acceptance of the treatment. A higher proportion of households were randomized into the Central, Tech, and Manual groups because we anticipated a relatively lower likelihood of acceptance given the higher level of commitment imposed on the households from these treatments. This was followed by the Info group which only receives an offer to accept the Hub device and then the Control which is automat-

[17] Note that participants in the Choice group were only invited during the first wave of invites. As households in each wave differed along observables, Choice group invitees were different than the entire sample pool. Note, however, that our final sample of households had balanced observable characteristics, as discussed in Section 4.1 and shown in Appendix Table A1. Also, we conduct a robustneess check in which we drop Choice group participants from each of their ultimate groups, as discussed in Section 5.4.

[18] The load factor represents the average electricity consumption divided by the maximum demand over a specific time period. This measure aims to capture the relative utilization rate of consumption at the household-level.

[19] These households represent recent movers who opened a new account with the Utility.

ically accepted because they receive no additional treatment in the second phase of recruitment (i.e., we just observe their consumption). Finally, we randomized a small proportion of our sample to the Choice group to provide us with an understanding of households' preferences over several key treatment categories.[20]

After randomized group assignment, we randomized upfront incentives that household would receive if they signed up to receive treatment. We randomly allocated households $0, $10, or $20 with equal probability. This is used to assess if the likelihood of treatment acceptance varies with the upfront incentive payment.

3 Data and Validation of Randomization

3.1 Data Description

Our analysis uses hourly household-level consumption (in KWhs) for all households in our experiment. Specifically, we have hourly electricity consumption data starting on October 1, 2020 and continuing through to June 30, 2023 representing the end of our experiment. We complement this household-level data with device-level consumption data on devices that receive load controllers during our trial. These data will allow us to identify changes in consumption behaviour during events and evaluate if there are specific changes to the use of particular devices.

We have information on a number of household characteristics that were provided through survey responses as a necessary condition to enter the first phase of our recruitment process. In addition, the Utility provided supplementary household information including the type of household (e.g., single-family/duplex, row home) and whether the household is enrolled in any other Utility programs. We are also provided time-stamped information on household interactions with the Utility's App. We can use this information to differentiate active and inactive households.

We complement the detailed household-level data with information from the 2016 Canadian Census (Statistics Canada, 2021). We are provided a household's Census Dissemination Area (CDA) identifier; the CDA is the most granular geographical unit for which all census information is provided publicly. We collected hourly weather information to control for environmental factors that impact electricity consumption, including temperature and humidity at three stations that are geographically rep-

[20]More formally, we randomized 22% to 24% of our sample to the Central, Tech, and Manual groups, 15% to the Info group, 11% to the Control group, and 6% to the Choice group.

resentative of the households located in our study.[21] These data were accessed at Environment and Climate Change Canada.

3.2 Validation of Randomization

We evaluate if there are differences in pre-treatment data across our various groups to assess the quality of our randomization. Table 2 provides summary statistics by group for a number of variables, including those used in the clustering procedure during randomization (recall Section 2.4). The sample presented in this Table represents all 1,763 households invited to participate in the experiment. For all variables, we report the results of a one-way ANOVA test to evaluate if there are statistical differences in means across the groups.[22] The only variable that shows statistically significant differences is cumulative consumption in the winter, which is marginally statistically significant at the 10% level.[23]

Table 2 demonstrates that the majority of households in our sample have electric hot water heating and use baseboard heating as the primary heat source, while electric vehicles are less common representing approximately 30% of households. The majority of households are single-family homes or duplexes, with the remainder being primarily row homes. The households consume considerably more electricity during the winter, with the lowest consumption arising during the summer. This demonstrates the potential for larger opportunities for load shifting during these months.

4 Descriptive Results

In this section, we present information on program adoption rates by group. We provide initial descriptive evidence that households reduce their electricity consumption during peak events and demonstrate that this response differs across our treatment groups. We also use detailed device-level data to illustrate how households adjust their load controlled devices during peak events. These descriptive results will be supplemented by a formal econometric model below.

[21] We match the households in our sample with their closest weather station.

[22] The seasonal cumulative consumption and load factor data only contain households with a full years worth of historical consumption. We computed analogous statistics focusing only on data from September 2021, the month in which all households have complete pre-treatment consumption data. We find no evidence of statistically significant differences in means across the groups using this data.

[23] The statistical significance in this variable is driven by the Choice group which has lower cumulative consumption on average than the remaining groups. We will demonstrate that our results are robust to the exclusion of these households below.

Table 2. Comparison of Means by Group

	Central	Tech	Manual	Choice	Info	Control	ANOVA
Cumul. KWhs							
Winter	5,279	5,268	5,442	4,588	4,859	5,265	2.04*
	(2,694)	(3,032)	(3,076)	(2,474)	(2,748)	(2,950)	
Spring	3,760	3,773	3,818	3,282	3,503	3,712	1.67
	(1,924)	(2,112)	(1,911)	(1,739)	(2,116)	(1,974)	
Summer	2,845	2,836	2,708	2,591	2,614	2,729	0.79
	(1,742)	(1,872)	(1,539)	(1,784)	(1,861)	(1,710)	
Fall	3,633	3,670	3,700	3,276	3,458	3,623	1.18
	(1,663)	(1,945)	(1,974)	(1,586)	(1,796)	(1,860)	
Load Factor							
Winter	24.66	24.98	25.41	24.02	24.73	24.67	0.54
	(8.20)	(8.15)	(8.80)	(8.51)	(8.29)	(8.63)	
Spring	19.52	20.12	20.01	18.01	19.28	19.91	1.66
	(7.25)	(6.97)	(6.70)	(6.77)	(7.73)	(7.41)	
Summer	16.82	16.55	16.73	16.19	16.12	16.32	0.34
	(7.89)	(6.30)	(5.93)	(8.59)	(8.11)	(8.29)	
Fall	18.56	18.90	19.34	17.82	18.42	19.06	1.33
	(5.89)	(6.23)	(6.00)	(5.77)	(6.48)	(6.50)	
Electric Vehicle	0.27	0.27	0.27	0.35	0.33	0.27	1.27
	(0.44)	(0.45)	(0.45)	(0.48)	(0.47)	(0.45)	
Baseboard Heating	0.61	0.64	0.61	0.67	0.63	0.63	0.30
	(0.49)	(0.48)	(0.49)	(0.47)	(0.48)	(0.48)	
Air Conditioning	0.52	0.51	0.50	0.61	0.51	0.54	0.81
	(0.50)	(0.50)	(0.50)	(0.49)	(0.50)	(0.50)	
Electric Hot Water	0.70	0.66	0.70	0.62	0.66	0.72	1.28
	(0.46)	(0.47)	(0.46)	(0.49)	(0.47)	(0.45)	
House/Duplex	0.77	0.76	0.81	0.74	0.78	0.84	1.60
	(0.42)	(0.43)	(0.39)	(0.44)	(0.42)	(0.37)	
Median Income	86,376	88,291	85,931	85,954	87,470	85,948	0.73
	(19,503)	(22,227)	(19,255)	(21,004)	(21,574)	(21,541)	
Households	423	382	409	102	259	188	

Notes. This table compares pre-treatment average values across the six different groups. Parentheses contain the standard deviations. Cumul. KWhs and Load Factor represents the cumulative household-level consumption and load factor by season. Electric Vehicle, Baseboard Heating, Air Conditioning, and Electric Hot Water are indicator variables denoting the presence of each device. House/Duplex is a indicator variable that equals one if the home type is a single-family home or duplex and zero otherwise. ANOVA reports the F-statistic from one-way ANOVA tests for differences in means across groups. Statistical Significance $^{*}p < 0.10$, $^{**}p < 0.05$, and $^{***}p < 0.01$.

4.1 Program Adoption and Comparability Between Groups

Table 3 summarizes household acceptance rates and details the attrition we faced between the initial and final acceptance by group. We experienced attrition between households initially accepting treatment and the final acceptance where households had the equipment installed and received events. This attrition arose for a number of reasons including households never responded to subsequent inquiries to receive and install equipment, equipment installers arrived but households were not compliant with local electrical codes, or there were meter connection issues between the equipment and the household meter.

Table 3. Program Acceptance and Withdraws by Group

	Central	Tech	Manual	Choice	Info	Control
Invited	423	382	409	102	259	188
Accepted - Initial	245	261	273	58	198	188
	(58%)	(68%)	(67%)	(57%)	(77%)	(100%)
Accepted - Final	177	184	242	38	177	188
	(42%)	(48%)	(59%)	(37%)	(68%)	(100%)
Accepted - Withdraw	68	77	31	20	21	0
	[28%]	[30%]	[11%]	[34%]	[11%]	[0%]

Notes. Invited reflects the households invited to participate in the experiment. Accepted - Initial are households that accepted our invitation during the second phase of recruitment. Accepted - Final are households that made it through the onboarding and equipment installation process. Accepted - Withdraw is the difference between Accepted - Initial and Accepted - Final. Numbers in parentheses reflects percentages relative to the number of households Invited. Numbers in brackets reflect percentages of Accepted - Withdrawn relative to the number of households that Accepted - Initial.

Table 3 demonstrates that we have high rates of initial acceptance, all exceeding 50%. The Control group receives no intervention, so their acceptance rate is 100%. The Info group had the next highest rate of initial acceptance. Interestingly, the Manual group, which only differs from the Info group in that these households receive events and can earn rewards, had a statistically significantly lower acceptance rate than the Info group. This suggests that households had an aversion to being subject to events despite the fact that households were only able to earn money as a result.

The Tech group had approximately the same initial acceptance rate as the Manual group. The Tech group receives the same events as the Manual group, but also receives load controller technology. Alternatively, the Central group had a statistically significantly lower acceptance rate than the Tech and Manual groups. The Central group only differs from the Tech group in that the load controllers are controlled by the Utility during events. This suggests that households are less likely to accept treatment as a result of having the Utility control the household's load controlled devices. Finally, the Choice group has approximately the same rate of initial acceptance as the Central group. This is despite the fact that households in this group can freely chose among the Central, Tech, or Manual groups. It is possible that the relatively lower rate of acceptance in this group is driven by the perceived complexity associated with making this decision.

Table 3 reports the sizes of final samples that receive the treatments outlined in Section 2.2. The difference between the initial and final acceptance reflects any withdraws. There were two distinct levels of household withdraws. The Manual and Info groups both had 11% of households withdraw. This largely reflects households that either did not follow-up with the Utility to receive the Hub device or households that had connection issues between the Hub and their meter. The Central, Tech, and Choice groups had a higher rate of withdraws. This is consistent with the fact these households had to both successfully receive and connect the Hub and have load controllers installed. (As noted below, the Choice group systematically chose the Central and Tech groups, which required installation of load controllers. They were therefore exposed to the same technical challenges as these groups.)

We take the similarity among final acceptance rates between the Central and Tech groups as the first set of evidence that we can confidentially compare our estimated treatment effects between these groups. While the Manual group has a lower attrition rate and therefore higher final acceptance rate, we are unconcerned that the Manual group systematically differs from the other two groups, based on a comparison of observables across groups in our final groups.

Table A1 in the Appendix compares average pre-treatment household characteristics by group for households that were in the final treatment groups. Consistent with Table 2, we observe no statistically significant differences in these characteristics across groups. The only exception is in the proportion of households that live in single-family homes/duplexes. There is a larger proportion of households in this building type in the Manual group than other groups, in particular. We observed differential attrition across groups between households' initial acceptance of group offers and final participation in each group program that was driven by technical challenges of participants connecting the smart Hubs, which transmit real-time meter data to participants' App, to their meters. Participants in row homes, the other house type, often could not connect Hubs to meters at a higher rate than participants in detached homes or duplexes. The reason for this is nearness of one's unit to the meter area in a row home setup, which is likely random. Overall, these results suggest that the balance on observables that arose as a result of the initial randomization remains in the final treated sample.

Before proceeding to the estimated treatment effects, two additional results warrant emphasis. Recall, the households that were invited to join the Choice group were given the opportunity to state their preference rankings across the Central, Tech, or Manual groups. Specifically, households were asked to rank their first, second, and third choices. In the second and third entries, we allowed households to select "none of the above". The first choices were systematically to join either the Central or Tech groups at 43% and 45%, respectively. This was followed by 12% for the Manual group.[24] This suggests that these households have a strong preference to receive the load controller equipment, and a sizable proportion preferred to have their consumption automatically adjusted by the Utility. However, there was a subset of households (26%) that appear to have a preference to avoid being part of the Central group; these households chose the Tech and/or Manual group followed by none of the above.

Finally, during the invitations to join each group, we randomized upfront incentives. While we find a higher rate of initial acceptance with higher upfront incentive payments, the differences are small and not statistically significantly different.[25]

4.2 Descriptive Consumption Patterns

Figures 1 – 3 provide average hourly household-level consumption for the Central, Tech, and Manual groups for March 2020 during non-event and event days. The shaded regions reflect the relevant event hours. These figures are illustrative of the broader patterns we observe throughout our sample.

Figure 1 demonstrates that the Central group has a large reduction in average consumption during events regardless of the event type. After each event, we observe a large spike in consumption. This "snap-back" is consistent with the devices turning on immediately after the event (e.g., to reheat the water tank and/or home).[26]

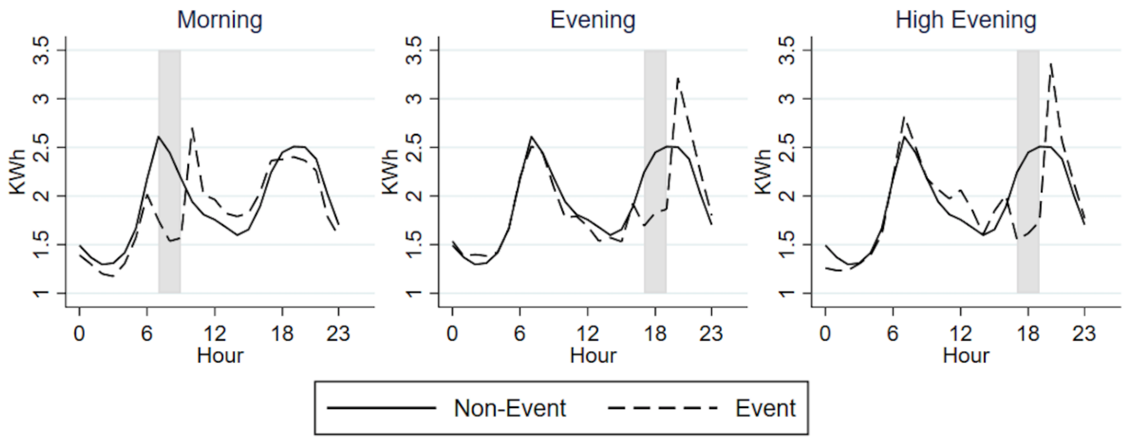
Figures 2 and 3 show that the Tech and Manual groups exhibit substantially

[24] See Table A2 in the Appendix for a detailed summary of the preference rankings.

[25] Households that received a $0, $10, and $20 upfront incentive accepted the initial invitation with a 63%, 67%, and 68% probability, respectively.

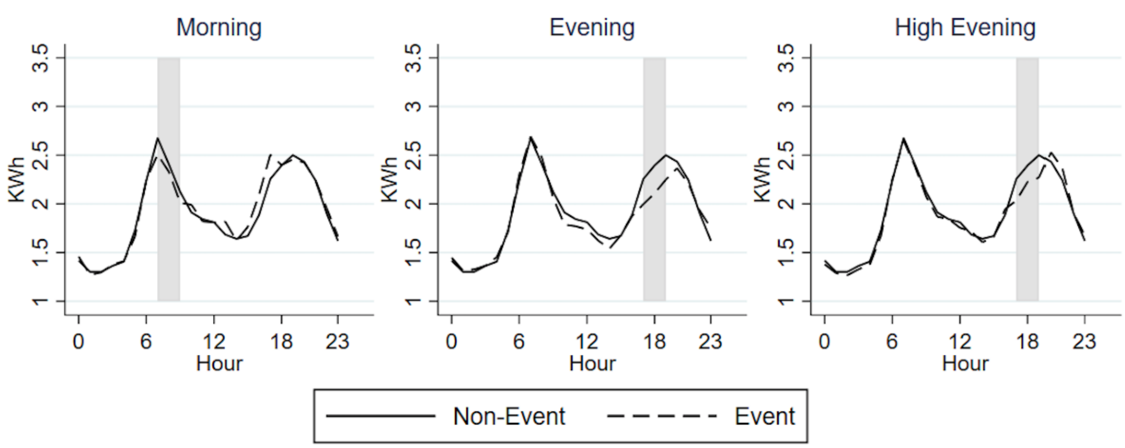
[26] The observed snap-back could be mitigated by the Utility by staggering the beginning and/or end of the load controlled event.

Figure 1. Average Household Consumption - Central (March 2022)



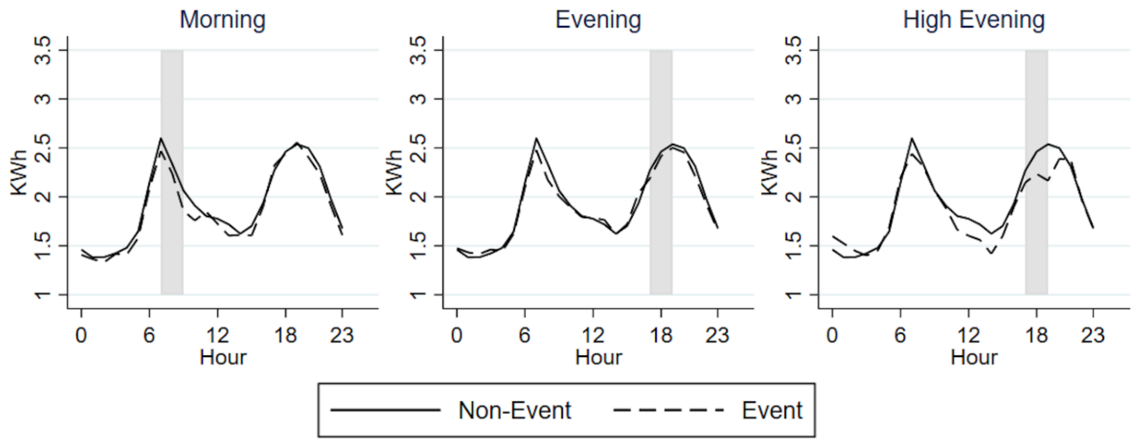
smaller responses to events. Despite the fact that the Tech group has access to the same equipment as the Central households, its average consumption changes by a considerably smaller amount during events. The Manual group shows a relatively limited change in its average consumption outside of the high evening events for this month. Neither group experience the same snap-back effect immediately following the events. Taken together, these results suggest that the Central group has a considerably larger response to each event type.

Figure 2. Average Household Consumption - Tech (March 2022)



We use the device-level data for devices that have load controllers to begin to evaluate how and if households in the Central and Tech groups use the devices during events. Figures 4 and 5 present the average hourly consumption of hot water heaters

Figure 3. Average Household Consumption - Manual (March 2022)



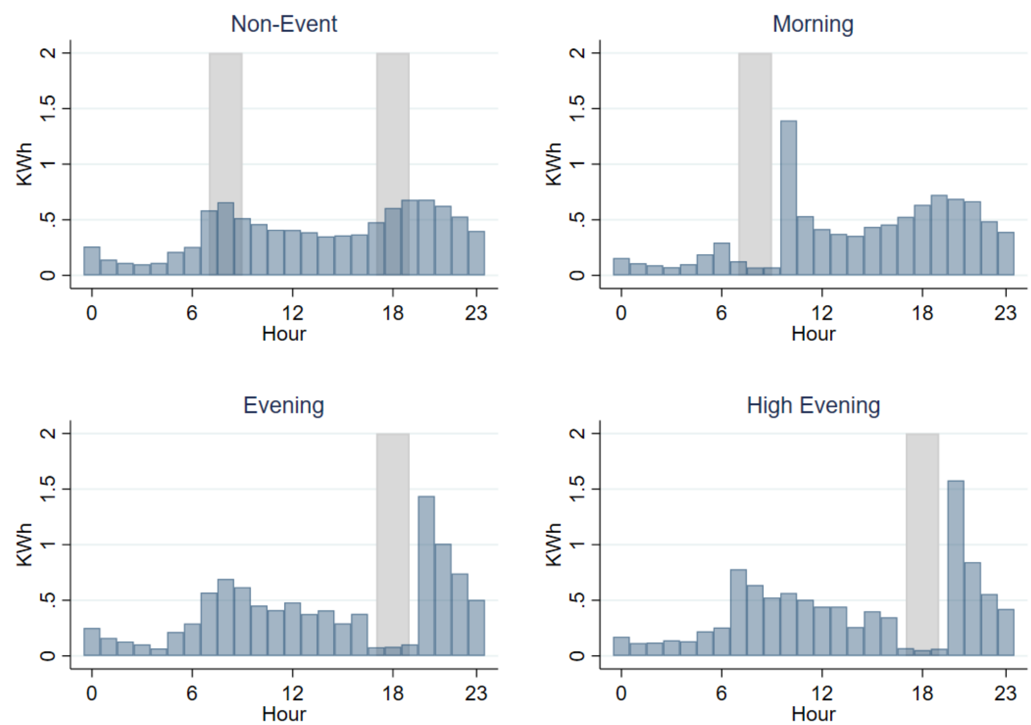
on non-event days and on the three different event day types for the Central and Tech groups in March 2022. We observe a large reduction in average hot water use during the event windows for the Central group. Consistent with the household-level results above, there is a sizable snap-back after the event. Alternatively, Figure 5 provides limited evidence that households in the Tech group are using the hot water heater load controllers during events. It is important to note that for the Tech group, households can adjust their hot water usage during an event by pushing a button in the Utility's App.

Figures A1 – A4 in the Appendix provide analogous figures for the load controllers installed on electric baseboard heaters and level 2 EV chargers.[27] We continue to observe a distinct reduction in consumption during events for the Central group for both the baseboard heater(s) and level 2 EV chargers.[28] In particular, average consumption from level 2 EV chargers decrease to essentially zero during events. Alternatively, for households in the Tech group, we observe relatively limited differences in consumption patterns on baseboard heaters during events. In contrast, there is evidence that these households are adjusting their level 2 EV consumption during events. This provides suggestive evidence that these households may be using the EV load controllers.

[27]For the level 2 EV chargers, we present the average hourly EV consumption (i.e., charging) across all months in 2020. We take this approach because of the relatively small number of households that have level 2 EV chargers in our sample. Focusing on one month can lead to higher variability in charging patterns because there is (typically) only one day of each event type for each household.

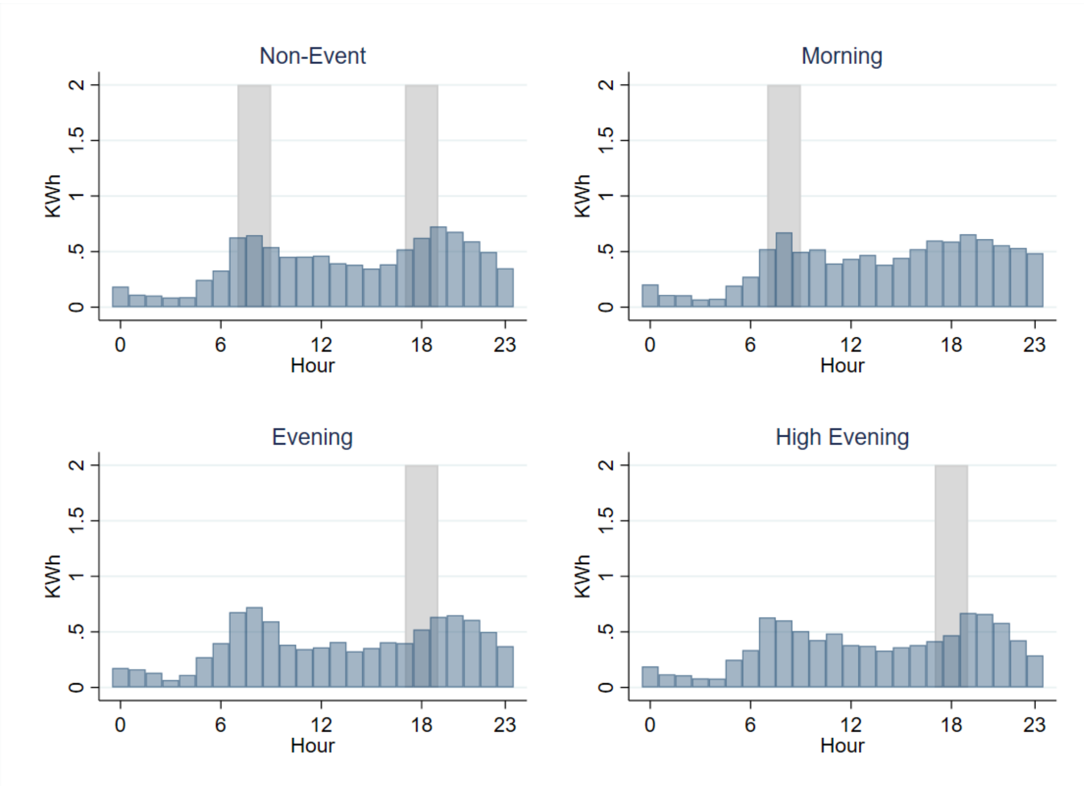
[28]For the baseboard heaters, the consumption does not decrease to zero during the events. This likely reflects the fact that the thermostats do not completely turn off these devices during events, but the temperature set points are reduced.

Figure 4. Average Hot Water Consumption - Central (March 2022)



These descriptive results provide initial evidence that the Central group experiences large reductions in consumption during events. This contrasts with households in the Tech and Manual groups which have considerably smaller adjustments in hourly consumption patterns. The device-level data provide suggestive evidence that the Tech group does not use the full capability of the load controllers to achieve larger demand reductions during events, with exception to the use of the level 2 EV load controllers. We undertake a formal empirical analysis to quantify these effects and control for potentially confounding factors.

Figure 5. Average Hot Water Consumption - Tech (March 2022)



5 Empirical Framework

5.1 Treatment effects by group

We begin by estimating the following specification to identify the average treatment effect of events on electricity consumption, by group, for the population of customers that participate in our experiment.

We model the effect of peak events on a household i's consumption c_{it} (in log KWhs) in hour t using the following model:

$$ln(c_{it}) \;=\; \alpha_i + \beta_t + \gamma\, D_i\, E_{it} + X_t\, \delta + \varepsilon_{it} \qquad (1)$$

where D_i is a vector of treatment group dummies that each equal one if household i is in the Central, Tech, or Manual groups and zero otherwise. E_{it} is the household-specific event indicator that equals one if the household is (randomly) assigned an event in hour t. We use the log of household electricity consumption on the left hand side to normalize the right-skewed variable. [29] We include α_i, household fixed effects, which control for time-invariant household characteristics. We also include β_t, a vector of time fixed effects that includes hour, day-of-week, and year-month, which control for time-varying factors that impact consumption. In our main specification, we include household consumption data on all groups (experimental groups and control groups).

Consumer responsiveness to events (especially with thermostat settings) may vary in local weather conditions. To capture this variation, we include X_t, a vector of hourly weather controls that include the relative humidity and cooling degrees and heating degrees above and below 65° F (18.33° C). Since consumer responsiveness may vary in weather conditions in a nonlinear way, we include a flexible functional form with a polynomial up to the third degree for each weather-related covariate. ε_{it} is the error term. We cluster standard errors at the household-level.

Our parameters of interest are γ, which measure the change in household-level electricity consumption during peak events for each of the Central, Tech, and Manual groups. Identification of our parameters of interest relies on comparing event-time consumption from households that randomly receive events to non-event consumption from all households (treatment and control households) that do not receive events at a particular time, conditional on average consumption in each hour of day, average

[29] Our results are robust to functional form; we observe qualitatively similar results with a linear-linear specification...

household-level consumption over all hours, and other controls. The identifying assumption behind Equation (1) recovering a causal effect of events on consumption is that events are not correlated with other drivers of household electricity consumption. This is met via our randomization of events.[30] As discussed in Section 2.1, our randomization of events within treatment groups allows us to causally identify treatment effects without employing common selection-correction techniques, such as estimation of local average treatment effects (LATE) parameters.

We report the average marginal effect of an event on households' electricity consumption by group, which is a group-specific function of $\widehat{\gamma}$, $f(\widehat{\gamma})$. Because of our log-linear specification, $f(\widehat{\gamma})$ is a semi-elasticity. We transform this function to report the percentage change in hourly consumption during an event via $100 \times (exp(f(\widehat{\gamma})) - 1)$.

We include data from before the events started as well as after (from September 1, 2021 through October 31, 2022), which gives more precision to our estimates. (Recall events started in February, 2022.)

5.2 Treatment effects by group: Heterogeneous treatment effects by event type and installed devices

We consider additional specifications to test for the presence of heterogeneous treatment effects across multiple dimensions. First, we consider a specification that interacts our treatment group indicator D_i by a vector of Event Type$_{it}$ indicator variables for each of the morning, evening, and high evening event types. This allows us to evaluate if there are asymmetric responses to events both by the treatment group and event-type. In particular, we anticipate a large response during the high evening events, which offer consumers greater incentives for demand reductions.

[30] One may be concerned, for the sake of external validity, that non-event consumption among treatment group households is not a good control for event-time behavior, if participation in a program with events leads people to alter non-event time behavior. To some extent, non-event consumption among treated households changing after people enter such a program is expected and part of the overall program effect we want to capture; in particular, people shifting consumption out of event windows may lead to shifting consumption to other times. Additionally, we imagine that our program, including the frequency of events, mimics the experience households may have in coming years due to the variety of factors impacting electricity markets. Event-level estimates are net of consumption shifting to other hours, which is ultimately of interest to a utility concerned about event-time consumption only. Additionally, we crafted the baseline calculations so that customers could not "game" their baseline to earn rewards, which may result in over-estimates of treatment effects. Ultimately, however, we run a version of Equation (1) without the Info and Control groups and obtain similar results. This suggests that non-event consumption of the treatment groups during the months after events started is not appreciably different than that of households who do not receive events at all.

Second, we focus on the Central and Tech groups to investigate how the responses to events vary by the installed load controller devices. For these groups, we run separate regressions that interact the event indicator E_{it} with a vector of indicator variables that are device-specific and equal one if the household has a load controller installed on a given device and zero otherwise.[31]

5.3 Intention-to-treat

We estimate a regression specification that allows us to estimate an effect similar to an intention-to-treat (ITT) estimate. A classic ITT estimate is a treatment effect that includes all households (or other units) that were invited to participate in an experiment, yielding an "overall program" effect estimate (Athey and Imbens, 2017). In our setting, this is of particular interest, as we anticipate a low take-up rate of the Central program offer but, conditional on take-up, large household-level reductions in electricity demand during events. In contrast, we anticipate a much larger take-up of the Tech program offer with smaller household-level reductions in electricity consumption during events. The net result of these extensive- and intensive-margin effects determines the ultimate event-level demand flexibility from the pool of customers that were offered each program. This is of particular interest to electric utility companies and balancing authorities which are weighing options for demand flexibility in the changing electricity landscape with system upgrade costs. From a general economic point of view, the ITT difference between these two groups in our study highlights the net effect of customers' appetite for an alternative market arrangement (quantity control for compensation) with the change in their demand that results from the barriers to the price coordination in the market being lifted.

Unlike many experimental settings, our treatment (events) are randomized within experimental groups and are not assigned globally by group. Therefore, unlike standard ITT specifications, we cannot have a binary treatment indicator that turns on for all households assigned to a specific experimental group, regardless of whether

[31] Note that these indicators do not precisely capture consumption reductions due to load controllers on devices; they capture household consumption reductions from households during events with particular installed devices. However, we observe a low rate of opt-outs among the Central group households: Households opt-out from about 1.8% of device-event observations. Opt-outs, when done, are for thermostats about 88% of the time and for hot water heaters about 9% of the time. Additionally, households in our sample can have a combination of one or more of the eligible load controlled devices. We run additional specifications where we interact the event indicator with each possible load controller device combinations. This allows us to evaluate if there are potential "synergies" associated with having different device combinations. For additional details, see Appendix D.1.

they accept that program. Because we have periodic events that are randomly assigned to only those households that chose to participate in our trial, we create an analogous environment in our setting for all households invited to each group. To do this, we assign households that were randomized to receive the Central, Tech, and Manual group offers but did not accept the offer a distribution of randomized events that is the same as the households that participated. This creates a new variable $\widehat{E}_{it}$ that includes the observed events for households in our experiment and a synthetic distribution of events for households that were invited to the Central, Tech, or Manual groups but did not accept our offer. For Choice group participants who did not respond to our invitation and did not indicate their preference for one of the Central, Tech, and Manual group offers, we randomly assign them one of these groups based on the preference frequencies for these groups (as a household's first choice) from the Choice group participants that did respond. If a Choice group participant expressed an offer preference but later dropped out of our experimental pool, they are assigned their top preference.[32]

We specify the following estimating equation:

$$ln(c_{it}) \; = \; \alpha_i + \beta_t + \omega \, I_i \, \widehat{E}_{it} + X_t \, \delta + \varepsilon_{it} \tag{2}$$

where I_i represents a vector of indicator variables that equal one if household i was invited to participate in the Central, Tech, or Manual groups and zero otherwise. $\widehat{E}_{it}$ is the household-specific event indicator that equals one if household i experiences an event in hour t (or is assigned a synthetic event). This regression is estimated on the full sample of household hourly consumption c_{it}, including the households that did not participate in our experiment. Similar to our main specification in (1), we include household fixed effects, a vector of time fixed effects, and X_t hourly weather controls. We cluster standard errors at the household-level.

5.4 Robustness checks

We run a number of robustness checks to evaluate the sensitivity of our results.

[32] Specifically, we do not have preferences from 60 out of 102 people invited to the Choice group. These invited households were assigned to the Central, Tech, and Manual groups with probabilities of 43%, 45%, and 12%, respectively, for the purpose of assigning them to treatment groups for the ITT calculations.

5.4.1 Removing pre-event data

We estimate Equation 1 on post-event data only to alleviate any concerns that changes in pre- and post-event behavior not absorbed by our year-month fixed effects could impact our estimates. We estimate versions with group-by-event interactions as well as group-by-event type interactions.

5.4.2 Removing Choice group participants

Households in the Choice group were asked to rank their preferences between the Central, Tech, and Manual group offers and were ultimately given their top choice preference. One may be concerned that participants in each of the Central, Tech, or Manual groups that come from the Choice group may be deferentially selected than the participants in those groups that opt-ed in from a yes/no participation offer. (Note that the Central, Tech, and Manual groups have, respectively, 19, 13, and 6 participants each–10.1%, 7.1%, and 2.5% of their total household counts–that were given a choice between programs.) Additionally, one could imagine an electric utility company either giving households a yes/no offer for a program akin to our Central, Tech, or Manual groups. Or, it could offer customers a choice. In an applied setting, it would be rare to do both, and utilities would like to know the flexibility of electricity demand from customers that take-up offers from each paradigm.

To address these concerns, we estimate the parameters of Equation 1 while excluding participants from the Choice group. We estimate versions with group-by-event interactions as well as group-by-event type interactions.

5.5 Information for Tech group

One may be concerned that the Tech group may not sufficiently capture the ability of assistive technology to overcome barriers to consumers' ability to respond to price, if consumers are uninformed about how to use it or face other challenges with its use that other technology may overcome. In particular, we were concerned that consumers' lack of familiarity with hot water heaters as an electricity-consumption shifting device, a device that one could turn off (and turn off temporarily without cold showers), would prevent them from utilizing their hot water heater load controllers in response to events. To address these concerns, we sent out a set of emails (detailed in Appendix C.4) to all experimental groups in fall 2022 tailored to provide the Tech group with (1) information about the electricity-consumption reduction potential of using their specific load controllers, (2) a statement to alleviate fears about turning off hot water heaters, and (3) step-by-step guides that provided the group with information about how to automate the turn-on of their devices after events. We provided symmetrical emails to the other experimental groups that included (1). We based the facts in (1) on the average per-event reductions in consumption estimated for households in the Central group that had each respective load controller (heterogeneous treatment effects by installed device, as described above), using data from the start of events (February 2022) through August 2022. This gave all groups a sense of the potential of consumption reductions possible through altering operation of their hot water heater, EV charger, or thermostat.

6 Empirical Results

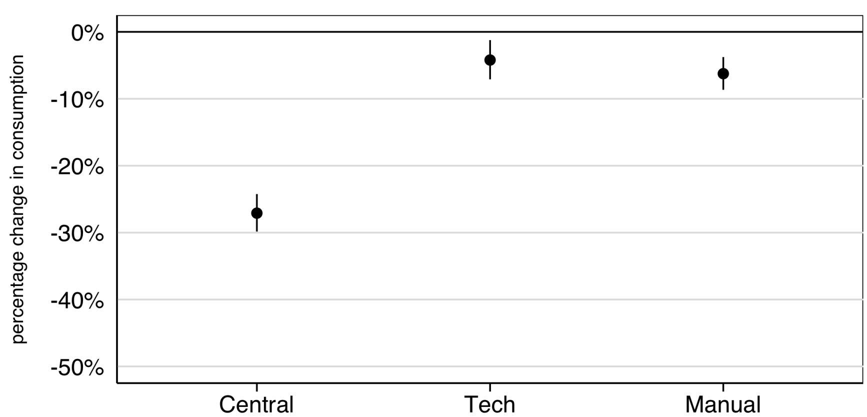
In this section, we present the results of our econometric model. We begin by presenting the average treatment effect for the population of households that participate in our experiment. We then describe heterogeneous treatment effects by event type and installed household devices as well as our ITT results. Finally, we present the results of a series of robustness checks.

6.1 Treatment effects by group

Figure 6 provides the estimated average response to events by group as a percentage change in household-level consumption. (See Appendix Table A3 for more detail.)

We observe an average 27% reduction in consumption during events for the Central group. In contrast, the Tech and Manual group have an approximate 4% and 6% reduction in demand, respectively. All of these effects are statistically significantly different from zero. Despite the fact that the Tech group has the same equipment as the Central group, it has a statistically significantly lower response to events. In particular, the average response for the Tech and Manual group are not statistically significantly different. These results are consistent with the descriptive data analysis presented in Section 4.2 that suggest that households in the Tech group were not broadly using the load controller equipment to the same extent as the Central group.

Figure 6. Average Treatment Effect of Participants by Group



Notes: The reported results are group-specific marginal effects calculated from estimating $\widehat{\gamma}$ in (1). We adjust marginal effects $f(\widehat{\gamma})$ to be a percentage change in consumption using the transformation $100 \times (exp(f(\widehat{\gamma})) - 1)$. Vertical lines indicate 95% confidence intervals. Standard errors are clustered at the household level.

6.2 Treatment effects by group: Heterogeneous treatment effects by event type and installed devices

Figure 7 presents the estimated response to events allowing for differential responses by event type and group. (See Appendix Table A4 for more detail.) For the Central group, we see a large demand reduction for all event types, with an approximate 26% reduction during both morning and evening events and a 31% average reduction during high evening events. The estimated high evening event effect is statistically significantly different from normal evening events. The Central group may be responding more to the larger offer for compensation.[33]

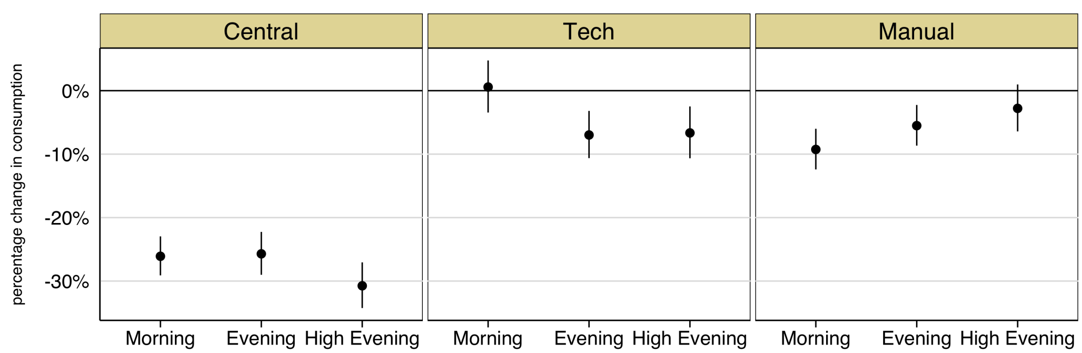
The Tech and Manual groups exhibit different patterns of event-time consumption changes across event type, compared to the Central group. The Tech group has a response to morning events that is not statistically significantly different from zero. This differs (statistically significantly) from the Manual group, which has an average estimated reduction of 9% during the morning events. This is a counter-intuitive result, as the Tech group has all the same information, incentives, and abilities as the Manual group in making electricity consumption reductions during events, with the added ability to remotely control thermostats, EV chargers, and hot water heaters on which they have load controllers installed. One potential explanation could lie in the Tech group being distracted by the tech: If participants have a kind of myopia to utilize their load-controlled devices only to reduce consumption but are unwilling to do so at certain times such as during the morning, then one might see results such as these. Indeed, the descriptive plots of the Tech group's utilization of load controllers in

[33]We observe the Central group opting out of events at a rate of 1.7% during morning events and at a rate of 2.4% for high evening and evening events. Therefore, differential opt-out rates by event type is not the likely explanation for the patterns seen here. We do observe Central group participants making greater reductions in electricity usage during events when they interact with the App then when they do not. (Precisely, when we estimate a version of Equation 1 with group-event-app usage indicator interactions. We then calculate marginal effects analogous to those in Figure 6, with different effects by group, depending on whether participants engaged with the App. We observe the Central group reducing electricity consumption by 23.1% and 31.7% on average, per event, when they do not and do engage with the App, respectively. If App engagement is tied to high peak events as well as electricity consumption beyond controlled devices, the Central group could be taking extra steps to reduce consumption when presented with greater incentives. However, we observe high rates of connection errors that lead to failed device control during events: rates of 25.1%, 23.2%, and 20.0% among morning, evening, and high evening events, respectively. A simple logistic regression reveals that the probability of these errors occurring is lower during high evening than other events. We do not know at this time whether this pattern of errors or the Central group's behavior with non-controlled devices during events is leading to greater reductions in electricity consumption during high events compared to regular evening events. We are currently discussing resolution of these errors with the Utility and App partner.

Figure 5, Figure A2, and Figure A4 suggest that the Tech group is making adjustments of EV chargers during evening and high evening events. Adjustments of other devices are small. Overall, the Tech and Manual group effects are not significantly different from each other, when compared within event type. This suggests that the Tech group's tech is not facilitating a response to the price incentive.

During the evening and high evening events, the Tech group reduces its demand by about 7%, while the Manual group has about 5.5% and 3% estimated reductions during these event types, respectively. These effects are statistically different from zero, except for the Manual group's behavior during high evening events. The Tech and Manual group do not reduce consumption more during high evening than regular evening events, which suggests the increased financial incentives for reduced electricity consumption during these times is not enough for participants in these groups to make greater reductions in usage.

Figure 7. Average Treatment Effect of Participants by Group and Event Type



Notes: The reported results are group- and event type-specific marginal effects calculated from estimating $\widehat{\gamma}$ in (1), adjusted to allow for event-type interactions with the group indicator variables D_i. We adjust marginal effects $f(\widehat{\gamma})$ to be a percentage change in consumption using the transformation $100 \times (exp(f(\widehat{\gamma})) - 1)$. Vertical lines indicate 95% confidence intervals. Standard errors are clustered at the household level.

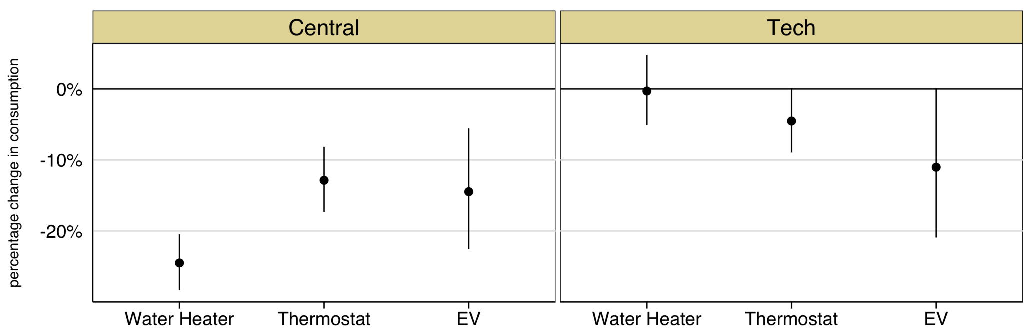
Figure 8 presents the results of the device-specific treatment effects for the Central and Tech groups.[34] We observe no responses to events from households in the Tech group that have load controllers on their hot water heater. These results indicate that these households are not using the load controller technology on this device, during events. In contrast, we observe statistically significant evidence (at the 10% level) of a response by Tech group households with thermostats or electric vehicle

[34] Appendix D.1 presents supplemental results for the device-specific analysis.

load controllers, with 4.5% and 11% reductions in demand during events from these devices. This could reflect the fact that electricity consumption for thermostats and EV charging and their impact on a household's bill is, in comparison to that for electric water heaters, relatively familiar to consumers. Additionally, consumers may have unresolved fears related to changing hot water heater settings.

Figure 8 shows that the households in the Central group respond to events with statistically significant reductions in consumption across all device types. The largest effect arises from households with load-controlled hot water heaters, resulting in an estimated reduction in consumption of 24.5%. Households with thermostats and EVs have reductions equal to about 13% and 14.5%, respectively. The results in Figure 8 are consistent with the descriptive results presented in Section 4.2.

Figure 8. Average Treatment Effect of Participants by Group and Technology

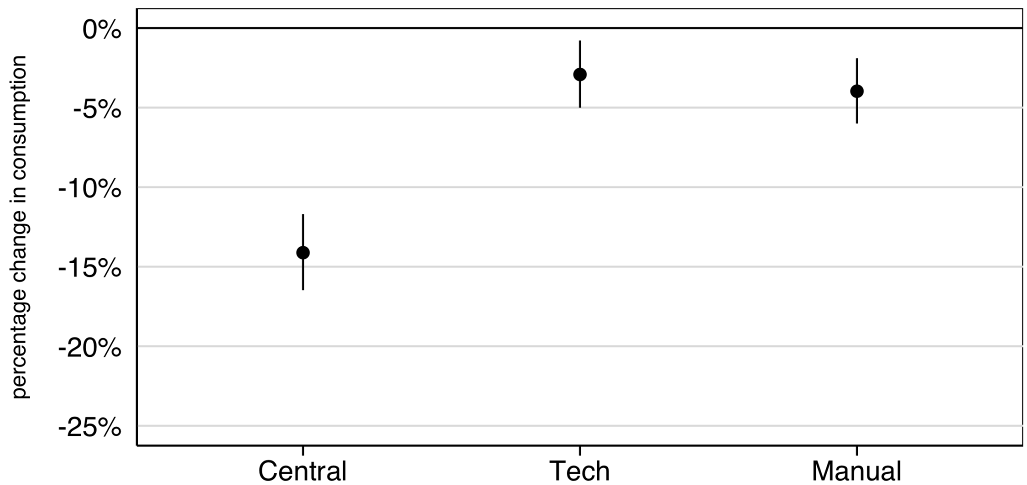


Notes. The reported results are group- and household-type specific marginal effects, obtained from estimating specification (1) run separately for the Central and Tech groups and interacting an indicator for whether a household has each device type with the event indicator E_{it}. We adjust marginal effects $f(\widehat{\gamma})$ to be a percentage change in consumption using the transformation $100 \times (exp(f(\widehat{\gamma})) - 1)$. Vertical lines indicate 95% confidence intervals. Standard errors are clustered at the household level.

6.3 Intention-to-treat

Overall, we find statistically significant event-time reductions in electricity consumption among households in the Central, Tech, and Manual groups, when including households who declined the offer to participate. Figure 10 shows group-specific effects: On average, households in the Central invitee group have reduced consumption by about 14% during events. (See Tables A5 and A6 for details.) The Tech and Manual invitee groups have reduced consumption during events by about 3% and 4%, respectively. The Central group response is significantly different from the responses of the Tech and Manual groups, as in the effects estimated from participants. The Tech and Manual group responses do not significantly differ from each other.

Figure 9. ITT Results by Group



Notes. The reported results reflect group-specific marginal effects computed from estimates of $\widehat{\omega}$ from specification (2). We adjust marginal effects $f(\widehat{\omega})$ to be a percentage change in consumption using the transformation $100 \times (exp(f(\widehat{\omega})) - 1)$. Vertical lines indicate 95% confidence intervals. Standard errors are clustered at the household level.

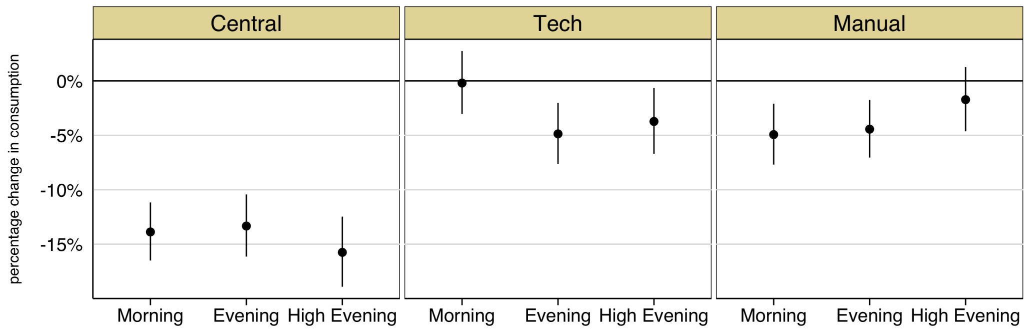
Figure 10 shows ITT results by group and event type. As with the treatment effect estimates from participants, effects are statistically significant, except for the Tech group during morning events, and the Manual group during high evening events.

6.4 Robustness Checks

6.4.1 Removing pre-event data

We explore whether results are sensitive to inclusion of pre-event data by estimating Equation 1 on post-event data only. We estimate versions with group-by-event interactions as well as group-by-event type interactions. Appendix Table A7 presents the results. Comparing it to Table A3, we see that the estimated average marginal effects of being in an event on households' consumption by group reveal a similar story: the Central group makes a large reduction in consumption, and the other groups lag behind in responding to the incentives. The estimated marginal effect on the Tech group households' change in consumption is statistically indistinguishable from zero, however, in these results. Table A8, which is comparable to Table A4, reveals that

Figure 10. ITT Results by Group and Event Type



Notes. The reported results reflect group- and event-type specific marginal effects computed from estimates of $\widehat{\omega}$ from a version of specification (2) that included event-type interactions with the event indicator $\widehat{E}_{it}$. We adjust marginal effects $f(\widehat{\omega})$ to be a percentage change in consumption using the transformation $100 \times (exp(f(\widehat{\omega})) - 1)$. Vertical lines indicate 95% confidence intervals. Standard errors are clustered at the household level.

this could be due to the distribution in behavior of the Tech group across events types. Reassuringly, the Tech group's responses to events by event type is broadly similar when we include and do not include pre-event data in our regressions.

6.4.2 Removing Choice group participants

To address concerns regarding differential selection of participants in each of the Central, Tech, and Manual groups via a yes/no offer vs. via a selection between them (i.e, being a part of the small Choice group), we estimate the parameters of Equation 1 while excluding participants from the Choice group. We estimate versions with group-by-event interactions as well as group-by-event type interactions.

Tables A9 and A10 show the results. These are comparable to Tables A3 and A4, respectively. Results are remarkably similar, indicating that the behavior of participants in each group that arrived there via a choice between programs are not driving event treatment effects.

7 Conclusion

We explore the potential for centralized decision-making to resolve consumers' ability to respond to incentives in complex market environment. Retail electricity markets are one such context, and we conduct a novel, large-scale field experiment in partnership with an electricity Utility company. We find several important results.

First, though we did not expect consumers to be as comfortable with Utility-centralized decision-making about electricity consumption of devices, we find no difference in the take-up rate of such a program with a comparable program that offers consumers the same consumption reduction incentive structure; technology; self-directed, remote control of device consumption; and device-specific electricity consumption information (programs that resulted in our Central and Tech groups, respectively). This surprising result indicates that even consumers may recognize the potential that centralized decision-making offers in reducing in transaction costs, cognitive load, and other barriers for consumers as they make market decisions and act on those decisions.

Second, we find that the Central group in our experiment reduces electricity consumption by a large 27% (on average) when incentivized to do so with "peak events" from the Utility. In contrast, the Tech group only makes, on average, about a 4% reduction in consumption during events. We find that the Tech group responds to events on par with our Manual group, which receives the same incentives as the Central and Tech groups but does not have the ability to remotely control device electricity consumption or view device-specific consumption data. This striking result indicates that the tech adopted by the Tech provides little, if any, resolution of consumers' difficulty of making decisions about and taking actions toward responding to incentives to reduce electricity consumption.

The difference in responsiveness to incentives between the Central and Tech groups remains in our intention-to-treat (ITT) estimates (15% vs. 3%), indicating that centralized Utility management of household device electricity use during peak events is the clear program winner for utility companies looking to address potential peak market conditions with demand-side measures. It demonstrates that, when considering program take-up rates and per-event electricity consumption responses by program/group, central decision-making in the winner.

We also uncover an important device-specific result: Hot water heaters are very flexible in terms of their electricity consumption in time, and we observe the Central group households with hot water heater load controllers making large (24.5%) per-event reductions in electricity usage during events, on average. (The Tech group households with hot water heater load controllers, in contrast, do not reduce consumption during events.) We suspect there is "untapped potential" in consumer electricity consumption flexibility that is embedded in similar devices, those that are flexible but about which there is little consumer familiarity of electricity usage flexibility.

Centralized decision-making may overcome a variety of consumer barriers in shifting usage of these devices, including informational barriers and fears over adjusting their usage.

We interpret the large difference in the response to incentives by the Central and Tech group as indicating that in-the-moment actions needed to respond to the incentives involve large transaction costs for consumers, and centralized decision-making can resolve these transaction costs. In particular, consumers in these groups face a difference in the default electricity consumption of the devices they have with load-controllers: The Central group, by default, has the Utility managing these devices' consumption during events (thought they can opt-out), and the Tech group, by default, has their normal usage of these devices that they must take action to alter during events. We suspect that changing default states of product allocations in markets that require consumers to make complex and/or quick decisions and/or overcome large transaction costs may be key to overcoming these barriers, and centralized decision-making is one option to this end.

References

Athey, Susan and Guido W Imbens, "The econometrics of randomized experiments," in "Handbook of economic field experiments," Vol. 1, Elsevier, 2017, pp. 73–140.

Blonz, Joshua, Karen Palmer, Casey J Wichman, Derek C Wietelman et al., "Smart thermostats, automation, and time-varying prices," 2021.

Bollinger, Bryan K and Wesley R Hartmann, "Information vs. Automation and implications for dynamic pricing," *Management Science*, 2020, *66* (1), 290–314.

Brandon, Alec, Christopher M Clapp, John A List, Robert D Metcalfe, and Michael Price, "The Human Perils of Scaling Smart Technologies: Evidence from Field Experiments," Technical Report, National Bureau of Economic Research 2022.

Fabra, Natalia, David Rapson, Mar Reguant, and Jingyuan Wang, "Estimating the elasticity to real-time pricing: evidence from the Spanish electricity market," in "AEA Papers and Proceedings," Vol. 111 2021, pp. 425–29.

Harding, Matthew and Carlos Lamarche, "Empowering consumers through data and smart technology: Experimental evidence on the consequences of time-of-use electricity pricing policies," *Journal of Policy Analysis and Management*, 2016, *35* (4), 906–931.

Joskow, Paul L and Catherine D Wolfram, "Dynamic pricing of electricity," *American Economic Review*, 2012, *102* (3), 381–85.

Sallee, James M, "Rational inattention and energy efficiency," *The Journal of Law and Economics*, 2014, *57* (3), 781–820.

Schneider, Ian and Cass R Sunstein, "Behavioral considerations for effective time-varying electricity prices," *Behavioural Public Policy*, 2017, *1* (2), 219–251.

Statistics Canada, "Census of Population. Data Products, 2016 Census," 2021. https://www12.statcan.gc.ca/census-recensement/2016/dp-pd/index-eng.cfm, Accessed: 2021-07-15.

A Supplementary Tables and Figures

Table A1 shows observables across groups for the final set of households in each group. Recall that our groups are comprised of households that both accepted our offer and then did not leave the experiment due to technical reasons, as described in Section 4.1.

Importantly for our ability to compare treatment effects between groups, we observe no statistically significant difference between means of observable characteristics across groups. The one exception to this is...

Table A1. Comparison of Means by Group - Final Accepted Households

	Central	Tech	Manual	Choice	Info	Control	ANOVA
Cumul. KWhs							
Winter	5,507	5,302	5,422	4,771	5,037	5,265	0.69
	(2,706)	(2,737)	(3,240)	(2,475)	(2,768)	(2,950)	
Spring	3,900	3,739	3,797	3,318	3,642	3,712	0.65
	(1,934)	(1,791)	(1,939)	(1,575)	(2,159)	(1,974)	
Summer	2,851	2,672	2,766	2,470	2,702	2,729	0.36
	(1,869)	(1,759)	(1,547)	(1,589)	(1,849)	(1,710)	
Fall	3,754	3,550	3,677	3,291	3,547	3,623	0.52
	(1,733)	(1,659)	(1,992)	(1,497)	(1,788)	(1,860)	
Load Factor							
Winter	24.62	25.56	24.93	24.07	24.93	24.67	0.29
	(8.68)	(8.21)	(9.04)	(8.24)	(7.76)	(8.63)	
Spring	19.33	20.48	19.80	17.59	19.59	19.91	1.19
	(7.43)	(6.30)	(6.45)	(5.59)	(7.05)	(7.41)	
Summer	16.33	16.87	16.95	14.93	16.61	16.32	0.57
	(8.54)	(6.02)	(5.95)	(8.53)	(7.80)	(8.29)	
Fall	18.17	18.97	19.11	17.31	18.53	19.06	0.92
	6.27	5.78	6.29	5.03	6.11	6.50	
Electric Vehicle	0.25	0.21	0.30	0.32	0.34	0.27	1.81
	(0.43)	(0.41)	(0.46)	(0.47)	(0.47)	(0.45)	
Baseboard Heating	0.68	0.70	0.60	0.71	0.59	0.63	1.63
	(0.47)	(0.46)	(0.49)	(0.46)	(0.49)	(0.48)	
Air Conditioning	0.46	0.46	0.51	0.58	0.51	0.54	0.99
	(0.50)	(0.50)	(0.50)	(0.50)	(0.50)	(0.50)	
Electric Hot Water	0.75	0.74	0.68	0.66	0.65	0.72	1.41
	(0.43)	(0.44)	(0.47)	(0.48)	(0.48)	(0.45)	
House/Duplex	0.82	0.77	0.89	0.74	0.84	0.84	2.75**
	(0.39)	(0.42)	(0.32)	(0.45)	(0.37)	(0.37)	
Median Income	84,978	88,274	86,718	83,535	89,504	85,948	1.35
	(19,647)	(20,432)	(19,494)	(19,717)	(21,079)	(21,541)	
Households	177	184	242	38	177	188	

Notes. This table compares pre-treatment average values across the six different groups for households that were in our final treatment groups. Parentheses contain the standard deviations. Cumul. KWhs and Load Factor represents the cumulative household-level consumption and load factor by season. Electric Vehicle, Baseboard Heating, Air Conditioning, and Electric Hot Water are indicator variables denoting the presence of each device. House/Duplex is a indicator variable if the home type is a single-family home or duplex. ANOVA reports the F-statistic from one-way ANOVA tests for differences in means across groups. Statistical Significance $^{*}\ p < 0.10$, $^{**}\ p < 0.05$, and $^{***}\ p < 0.01$.

Table A2. Choice Group Preference Ranking

	Central	Tech	Manual	None
First Choice	25	26	7	0
	(43%)	(45%)	(12%)	(0%)
Second Choice	10	29	14	5
	(17%)	(50%)	(24%)	(9%)
Third Choice	8	1	18	31
	(14%)	(2%)	(31%)	(53%)

Notes. 58 households assigned to the Choice group accepted the initial invitation and provided their preference ranking. The numbers represent the count of Choice households with a specific preference ranking. Percentages are presented in the parentheses.

Figure A1. Average Baseboard Heater Consumption - Central (March 2022)

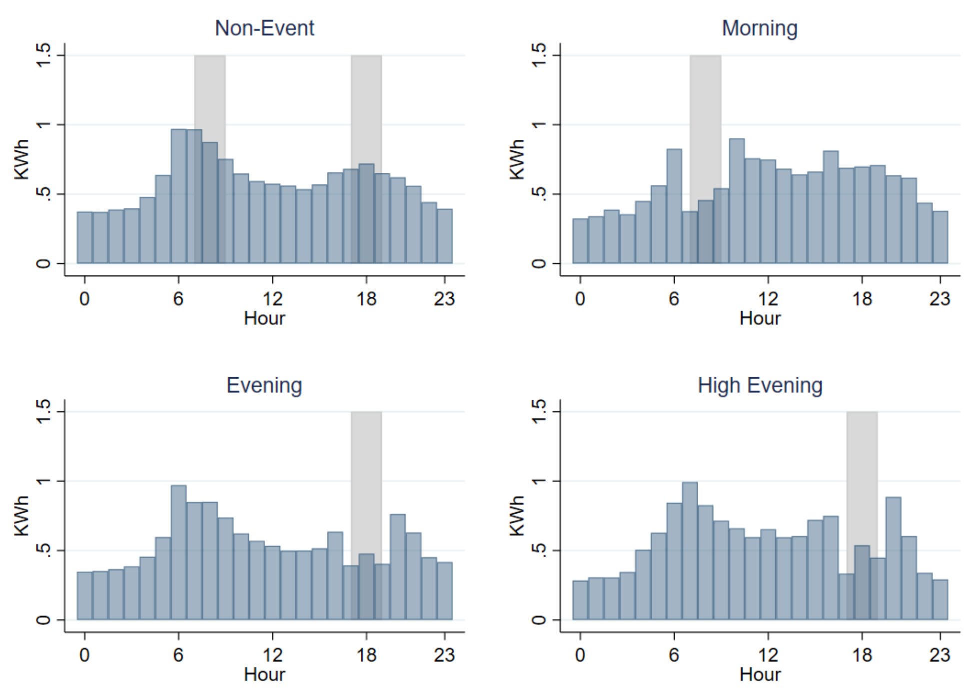


Figure A2. Average Baseboard Heater Consumption - Tech (March 2022)

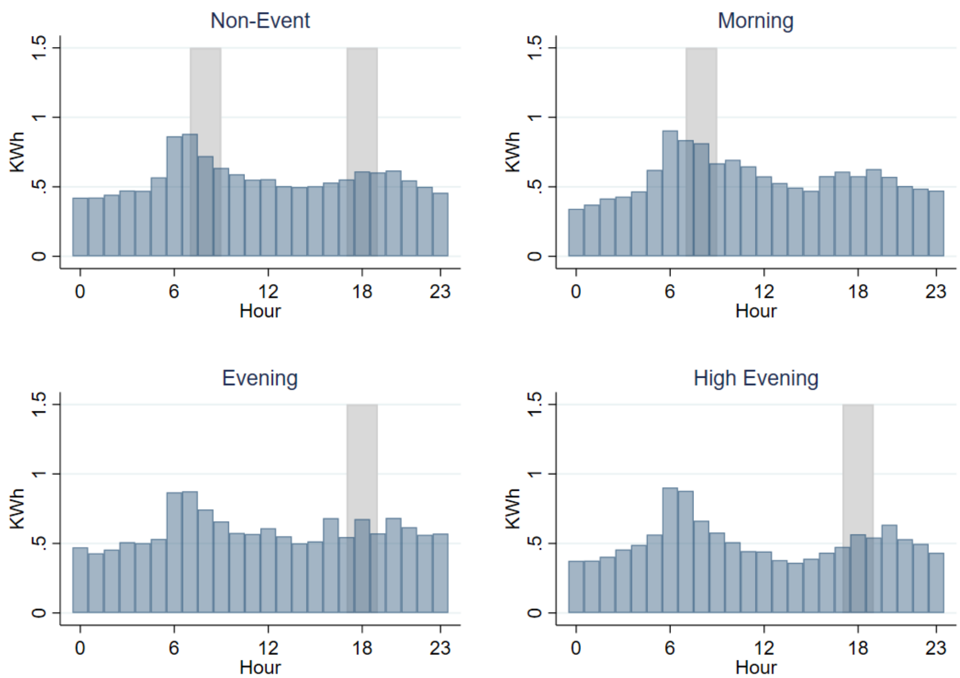


Figure A3. Average Level 2 EV Consumption - Central (Year 2022)

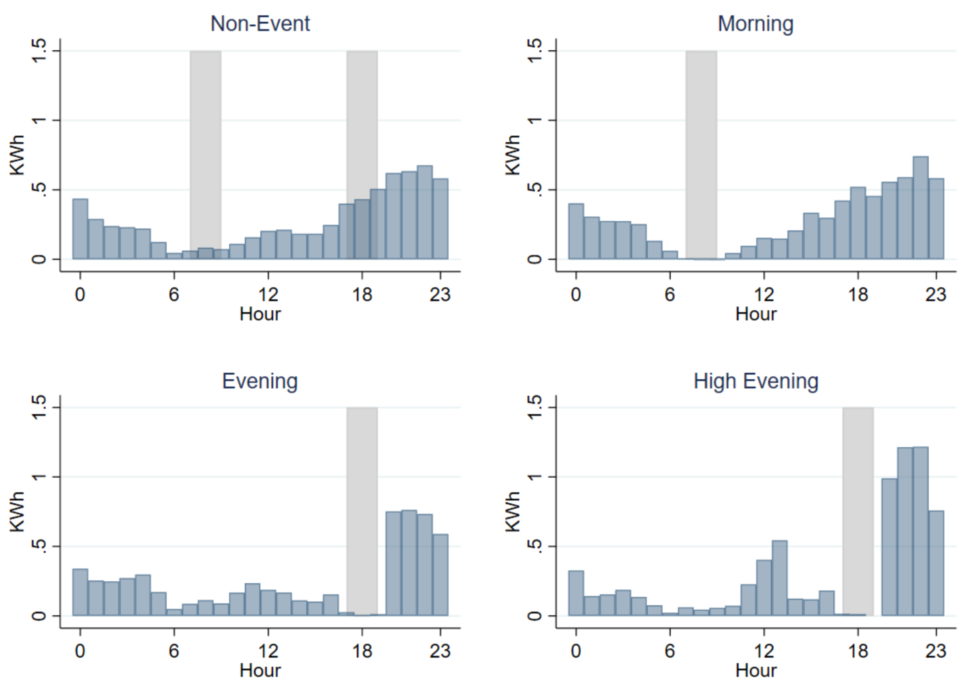


Figure A4. Average Level 2 EV Consumption - Tech (Year 2022)

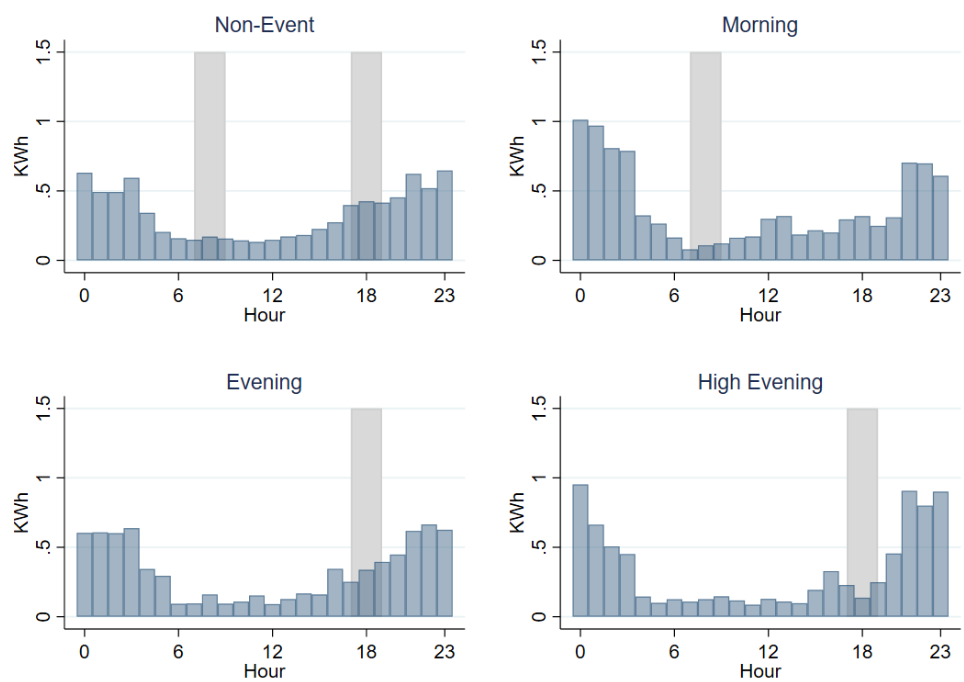


Table A3. Regression Results: Event - Group Interactions

	Marginal Effect	Standard Error	P-Value	Marginal Effect (%)
Central	-0.3161	0.0196	0.0000	-27.1044
Tech	-0.0430	0.0156	0.0059	-4.2117
Manual	-0.0645	0.0133	0.0000	-6.2458
Adj. R^2	0.4625			
N	10,057,077			

Notes. The reported results are group-specific marginal effects calculated from estimating $\widehat{\gamma}$ in (1), with their associated standard errors and p-values to indicate statistical significance. Standard errors are clustered at the household level. We adjust marginal effects $f(\widehat{\gamma})$ to be a percentage change in consumption using the transformation $100 \times (exp(f(\widehat{\gamma})) - 1)$. Adj. R^2 reflects the adjusted R-squared value and N denotes the number of observations.

Table A4. Regression Results: Event Type - Group Interactions

		Marginal Effect	Standard Error	P-Value	Marginal Effect (%)
Central	Morning	-0.3022	0.0212	0.0000	-26.0836
	Evening	-0.2973	0.0234	0.0000	-25.7161
	High Evening	-0.3678	0.0264	0.0000	-30.7778
Tech	Morning	0.0057	0.0208	0.7849	0.5702
	Evening	-0.0724	0.0204	0.0004	-6.9840
	High Evening	-0.0689	0.0223	0.0021	-6.6566
Manual	Morning	-0.0972	0.0180	0.0000	-9.2663
	Evening	-0.0567	0.0173	0.0011	-5.5110
	High Evening	-0.0283	0.0194	0.1438	-2.7922
Adj. R^2	0.4625				
N	10,057,077				

Notes. The reported results are group- and event type-specific marginal effects calculated from estimating $\widehat{\gamma}$ in (1), adjusted to allow for event-type interactions with the group indicator variables D_i. The associated standard errors and p-values are reported to indicate statistical significance. Standard errors are clustered at the household level. We adjust marginal effects $f(\widehat{\gamma})$ to be a percentage change in consumption using the transformation $100 \times (exp(f(\widehat{\gamma})) - 1)$. Adj. R^2 reflects the adjusted R-squared value and N denotes the number of observations.

Table A5. Regression Results: Event - Group Interactions - Intention-to-Treat

	Marginal Effect	Standard Error	P-Value	Marginal Effect (%)
Central	-0.1522	0.0142	0.0000	-14.1183
Tech	-0.0296	0.0111	0.0076	-2.9214
Manual	-0.0405	0.0109	0.0002	-3.9682
Adj. R^2	0.4701			
N	15,483,368			

Notes. The reported results are group-specific marginal effects computed from estimates of $\widehat{\omega}$ from specification (2), as well as their associated standard errors and p-values to indicate statistical significance. Standard errors are clustered at the household level. We adjust marginal effects $f(\widehat{\omega})$ to be a percentage change in consumption using the transformation $100 \times (exp(f(\widehat{\omega})) - 1)$. Adj. R^2 reflects the adjusted R-squared value and N denotes the number of observations.

Table A6. Regression Results: Event Type - Group Interactions - Intention-to-Treat

		Marginal Effect	Standard Error	P-Value	Marginal Effect (%)
	Morning	-0.1494	0.0158	0.0000	-13.8816
Central	Evening	-0.1431	0.0168	0.0000	-13.3369
	High Evening	-0.1714	0.0195	0.0000	-15.7484
	Morning	-0.0020	0.0148	0.8926	-0.1994
Tech	Evening	-0.0499	0.0150	0.0009	-4.8697
	High Evening	-0.0380	0.0160	0.0179	-3.7320
	Morning	-0.0506	0.0150	0.0008	-4.9344
Manual	Evening	-0.0454	0.0141	0.0013	-4.4393
	High Evening	-0.0174	0.0153	0.2567	-1.7232
Adj. R^2	0.4701				
N	15,483,368				

Notes. The reported results reflect group- and event-type specific marginal effects computed from estimates of $\widehat{\omega}$ from a version of specification (2) that included event-type interactions with the event indicator $\widehat{E}_{it}$. The associated standard errors and p-values are reported to indicate statistical significance. Standard errors are clustered at the household level. We adjust marginal effects $f(\widehat{\omega})$ to be a percentage change in consumption using the transformation $100 \times (exp(f(\widehat{\omega})) - 1)$. Adj. R^2 reflects the adjusted R-squared value and N denotes the number of observations.

Table A7. Regression Results: Event - Group Interactions Without Pre-Treatment Data

	Marginal Effect	Standard Error	P-Value	Marginal Effect (%)
Central	-0.2973	0.0186	0.0000	-25.7200
Tech	-0.0243	0.0153	0.1127	-2.4040
Manual	-0.0441	0.0127	0.0005	-4.3127
Adj. R^2	0.4353			
N	5,881,668			

Notes. Regressions only include data starting on February 22, 2022, the first peak events in our experiment. The reported results are group-specific marginal effects calculated from estimating $\widehat{\gamma}$ in (1) and their associated standard errors and p-values to indicate statistical significance. Standard errors are clustered at the household level. We adjust marginal effects $f(\widehat{\gamma})$ to be a percentage change in consumption using the transformation $100 \times (exp(f(\widehat{\gamma})) - 1)$. Adj. R^2 reflects the adjusted R-squared value and N denotes the number of observations.

Table A8. Regression Results: Event Type - Group Interactions Without Pre-Treatment Data

		Marginal Effect	Standard Error	P-Value	Marginal Effect (%)
Central	Morning	-0.2890	0.0207	0.0000	-25.0993
	Evening	-0.2764	0.0221	0.0000	-24.1488
	High Evening	-0.3440	0.0258	0.0000	-29.1057
Tech	Morning	0.0196	0.0210	0.3505	1.9837
	Evening	-0.0526	0.0198	0.0080	-5.1275
	High Evening	-0.0447	0.0221	0.0437	-4.3692
Manual	Morning	-0.0829	0.0176	0.0000	-7.9562
	Evening	-0.0342	0.0163	0.0361	-3.3574
	High Evening	-0.0022	0.0197	0.9105	-0.2216
Adj. R^2	0.4354				
N	5,881,668				

Notes. Regressions only include data starting on February 22, 2022, the first peak events in our experiment. The reported results are group- and event type-specific marginal effects calculated from estimating $\widehat{\gamma}$ in (1), adjusted to allow for event-type interactions with the group indicator variables D_i. The associated standard errors and p-values are reported to indicate statistical significance. Standard errors are clustered at the household level. We adjust marginal effects $f(\widehat{\gamma})$ to be a percentage change in consumption using the transformation $100 \times (exp(f(\widehat{\gamma})) - 1)$. Adj. R^2 reflects the adjusted R-squared value and N denotes the number of observations.

Table A9. Regression Results: Event - Group Interactions Without Choice Group

	Marginal Effect	Standard Error	P-Value	Marginal Effect (%)
Central	-0.3180	0.0206	0.0000	-27.2401
Tech	-0.0461	0.0164	0.0050	-4.5052
Manual	-0.0633	0.0135	0.0000	-6.1363
Adj. R^2	0.4639			
N	9,677,926			

Notes. These regression results do not include the Choice group households in their ultimate experimental groups. The reported results are group-specific marginal effects calculated from estimating $\widehat{\gamma}$ in (1) and their associated standard errors and p-values to indicate statistical significance. Standard errors are clustered at the household level. We adjust marginal effects $f(\widehat{\gamma})$ to be a percentage change in consumption using the transformation $100 \times (exp(f(\widehat{\gamma})) - 1)$. Adj. R^2 reflects the adjusted R-squared value and N denotes the number of observations.

Table A10. Regression Results: Event Type - Group Interactions Without Choice Group

		Marginal Effect	Standard Error	P-Value	Marginal Effect (%)
Central	Morning	-0.3093	0.0225	0.0000	-26.6061
	Evening	-0.3009	0.0243	0.0000	-25.9847
	High Evening	-0.3590	0.0278	0.0000	-30.1652
Tech	Morning	0.0043	0.0217	0.8415	0.4347
	Evening	-0.0770	0.0211	0.0003	-7.4112
	High Evening	-0.0720	0.0235	0.0023	-6.9462
Manual	Morning	-0.0980	0.0183	0.0000	-9.3369
	Evening	-0.0537	0.0175	0.0022	-5.2320
	High Evening	-0.0270	0.0197	0.1707	-2.6656
Adj. R^2	0.4639				
N	9,677,926				

Notes. These regression results do not include the Choice group households in their ultimate experimental groups. The reported results are group- and event type-specific marginal effects calculated from estimating $\widehat{\gamma}$ in (1), adjusted to allow for event-type interactions with the group indicator variables D_i. The associated standard errors and p-values are reported to indicate statistical significance. Standard errors are clustered at the household level. We adjust marginal effects $f(\widehat{\gamma})$ to be a percentage change in consumption using the transformation $100 \times (exp(f(\widehat{\gamma})) - 1)$. Adj. R^2 reflects the adjusted R-squared value and N denotes the number of observations.

B Recruitment and Assignment

B.1 In-app survey

In the first phase of recruitment, when households accepted the invitation to install the app, they were required to answer the following six-question survey.

1. Out of these options, which factor most influenced you to use/test the app?
    - Convenience
    - Better understand my energy use
    - Save money
    - Help the environment
2. Do you have an electric vehicle?
    - No
    - Yes, I charge with a Level 2 charger at home
    - Yes, I charge with a Level 1 charger at home
3. Do you have an electric hot water tank (not gas)?
    - Yes
    - No
4. Do you primarily heat your home with electric baseboard heaters? If so, how many do you have?
    - Yes, I have 1 - 3 baseboards
    - Yes, I have 4 or more baseboards
    - No, I do not have electric baseboard heating
5. Do you have air conditioning?
    - Yes, window or room units
    - Yes, central air conditioning
    - No, I do not have air conditioning
6. Do you rent or own your home?
    - Rent
    - Own

B.2 Recruitment into treatment groups

As discussed in Section 2.4, customers using the App and eligible to participate in our experiment were randomized into treatment groups. Households assigned to each group were sent the email below corresponding to their treatment assignment.

Email invites went out in five waves in fall 2021, as eligible participants became available on the App. Though the pool of customers in each wave differed along observable characteristics, our randomization and ultimate household acceptance resulted in no concerning differences between ultimate treatment groups along observables. (See Table A1.) Households were sent one email reminders regarding their invites, always on the date of the next wave of invites.

Table B2 shows the dates invites and reminders were sent to households invited to participate in each treatment group:

Table B1. Email Group Invitations and Reminders

Date	**Communication**
October 13, 2021	Wave I invite email
October 28, 2021	Wave II invite email + Wave I non-responder reminder
November 15, 2021	Wave III invite email + Wave II non-responder reminder
November 22, 2021	Wave IV invite email + Wave III non-responder reminder
January 18, 2022	Wave V invite email + Wave IV non-responder reminder

Emails sent to households are included on subsequent pages. Red text varied by customer name or randomized up-front offer ($10 or $20) given to customers. If customers received no up-front offer, offer text was absent.

*****Central group Offer Email #1**

Email Subject: Free smart home devices, rewards for reducing your electricity use

Mobile preview text: Join the Peak Rewards trial today

Dear name,

As a [redacted] app user, you're invited to a join a new Peak Rewards trial. By participating, you'll receive free smart home devices and have the opportunity to earn rewards for reducing your electricity use to help manage load on our system.

What's in it for me?

- **Rewards on your bill:** A typical household can expect to earn $100 over the course of the trial—the more you reduce during peak times, the more you can earn. As an additional reward for joining, you'll receive $20 when you sign up.
- **Free [redacted] Hub** (valued at $50): Get live electricity usage feedback through the [redacted] app.
- **Free smart home devices** (valued at $120 each)**:** Control the electricity use of up to five electric baseboard heaters, your electric hot water tank, and/or your electric vehicle charger through the [redacted] app.
- **Professional installation**: One of our contractors will install your smart home devices at no cost to you.

How do I earn rewards?

1. Before each peak time "event", you'll be notified in the [redacted] app about the event day and time, and the financial rewards you could earn for reducing your usage.
2. Take advantage of the convenient option of allowing us to remotely manage your electricity use during the event using the installed smart home devices. You can opt-out of events at any time.

Learn more about Peak Rewards. (Note: You'll be taken to a trial webpage external to [redacted], managed by trial personnel in partnership with [redacted].)

Apply now

*****Tech group Offer Email #1**

Email Subject: Free smart home devices, rewards for reducing your electricity use

Mobile preview text: Join the Peak Rewards trial today

Dear name,

As a [redacted] app user, you're invited to a join a new Peak Rewards trial. By participating, you'll receive free smart home devices and have the opportunity to earn rewards for reducing your electricity use to help manage load on our system.

What's in it for me?

- **Rewards on your bill:** A typical household can expect to earn $100 over the course of the trial—the more you reduce during peak times, the more you can earn. As an additional reward for joining, you'll receive $20 when you sign up.
- **Free** [redacted] **Hub** (valued at $50): Get live electricity usage feedback through the [redacted] app.
- **Free smart home devices** (valued at $120 each)**:** Control the electricity use of up to five electric baseboard heaters, your electric hot water tank, and/or your electric vehicle charger through the [redacted] app.
- **Professional installation**: One of our contractors will install your smart home devices at no cost to you.

How do I earn rewards?

1. Before each peak time "event", you'll be notified in the [redacted] app about the event day and time, and the financial rewards you could earn for reducing your usage.
2. During the event, use the [redacted] app to remotely adjust your appliances with the installed smart home devices to reduce your electricity usage.

Learn more about Peak Rewards. (Note: You'll be taken to a trial webpage external to [redacted], managed by trial personnel in partnership with [redacted].)

Apply now

***Manual group Offer Email #1**

Email Subject: Free [redacted] Hub, rewards for reducing your electricity use

Mobile preview text: Join the Peak Rewards trial today

Dear name,

As a [redacted] app user, you're invited to a join a new Peak Rewards trial. By participating, you'll receive a free [redacted] Hub and have the opportunity to earn rewards for reducing your electricity use to help manage load on our system.

What's in it for me?

- **Rewards on your bill:** A typical household can expect to earn $100 over the course of the trial—the more you reduce during peak times, the more you can earn. As an additional reward for joining, you'll receive $20 when you sign up.
- **Free [redacted] Hub** (valued at $50): Get live electricity usage feedback through the [redacted] app.

How do I earn rewards?

1. Before each peak time "event", you'll be notified in the [redacted] app about the event day and time, and the financial rewards you could earn for reducing your usage.
2. During the event, use the live electricity usage feedback in the [redacted] app to help you reduce your electricity usage.

Learn more about Peak Rewards. (Note: You'll be taken to a trial webpage external to [redacted], managed by trial personnel in partnership with [redacted].)

Apply now

*****Info group Offer Email #1**

Email Subject: Free [redacted] Hub for live electricity usage feedback

Mobile preview text: Sign up today

Dear name,

As a [redacted] app user, you're invited to receive a **free** [redacted] **Hub** (valued at $50). The Hub gives you live electricity usage feedback through the [redacted] app to help you manage your electricity usage.

As an additional reward for adopting the Hub, you'll receive $20 when you sign up.

Learn more about the [redacted] Hub. (Note: You'll be taken to a webpage external to [redacted] managed by personnel in partnership with [redacted].)

Apply now

B.3 Choice group preference elicitation

A random subset of eligible participants on the App were presented with the choice to join the Central group, Tech group, or Manual group using the following email and online questionnaire as part of Wave I only.

*****Choice group Offer Email #1**

Email Subject: Free smart home devices, rewards for reducing your electricity use

Mobile preview text: Join the Peak Rewards trial today

Dear name,

As a [redacted] app user, you're invited to a join a new Peak Rewards trial. By participating, you'll receive free smart home devices and have the opportunity to earn rewards for reducing your electricity use to help manage load on our system.

What's in it for me?

- **Rewards on your bill:** A typical household can expect to earn $100 over the course of the trial—the more you reduce during peak times, the more you can earn. As an additional reward for joining, you'll receive $20 when you sign up.
- **Free [redacted] Hub, smart home devices, and professional installation:** Depending on the option below you choose (and availability):
  - Option 1: Receive a free [redacted] Hub (valued at $50) to get live electricity usage feedback through the [redacted] app.
  - Option 2: Receive a free [redacted] Hub and free smart home devices (valued at $120) installed for you to control your electricity use through the [redacted] app.
  - Option 3: Receive a free [redacted] Hub and smart home devices installed for you, plus take advantage of the convenient option of allowing us to remotely manage your electricity use during the event using the installed smart home devices. You can opt-out of events at any time.

How do I earn rewards?

1. Before each peak time "event", you'll be notified in the [redacted] app about the event day and time, and the financial rewards you could earn for reducing your usage.
2. During the event, use the live electricity usage feedback in the [redacted] app to help you reduce your electricity usage by adjusting your appliances manually or by using [redacted] app, or by allowing us to adjust on your behalf (depending on the option selected).

Learn more about Peak Rewards. (Note: You'll be taken to a trial webpage external to [redacted], managed by trial personnel in partnership with [redacted].)

Apply now

Peak Rewards Trial Registration – "Choice" group

Thank you for your interest in our Peak Rewards offer, as part of the [redacted] trial. Please complete the form below to register your interest. There are a limited number of spaces for this offer; we will contact you if we are able to confirm your participation.

Once confirmed, we will arrange for the delivery of your equipment, and we will be in touch to explain what happens next.

The peak time events are expected to start in December; however, we will contact you shortly to get you set up so you are ready to go.

Date October 1, 2021 – March 31, 2023

[redacted] is collecting your information for the purposes of administering and evaluating the [redacted] trial ("the Trial") in furtherance of its mandates under the [redacted], the Clean Energy Act, and the [redacted] Electric Tariff regulated by the [redacted] Utilities Commission under the Utilities Commission Act. [redacted] may also collect personal information in the course of administering the Trial.

If you have any questions about how [redacted] collects, uses, or discloses any personal information with regards to this Trial or if you are having technical problems with this site or the registration process, please contact [redacted]

Contact Information

* First Name	
* Last Name	
* Postal Code	
* Email	

Choose your Peak Rewards Trial Preference

Let us know your preferences for the Peak Rewards options below.

[redacted] may remotely adjust your installed smart devices, only during peak events; you can control your devices at all times through the app.

Only you can control your installed smart devices, either manually or through the app.

There are no installed smart devices, you adjust your usage manually.

NONE

Confirm Ranking	I am happy with my ranking of my Peak Rewards options.
Confirm participation	I agree to participate in this Peak Rewards offer.

Registration Successful

Thanks for registering your interest. Subject to availability, you may receive a [redacted] Smart Hub in the mail in next few weeks. We will be in touch about any additional installation as required.

B.4 Treatment group websites

We crafted custom websites to facilitate recruitment into treatment groups and to provide households with information throughout the experiment.

During recruitment, recruitment emails linked to these websites. They give an overview of the "Trial" into which the Utility has invited households to participate, provide the sign-up link, highlight the financial rewards participants could earn by participating and reducing electricity consumption, explain the details of the technology involved, and give information relevant to specific groups, such as the fact that the Central group can opt-out of Utility control of their devices. Text and images on these websites is identical except for information that is group-specific.

We used very unique URLs for each group website to reduce the probably of participants invited to one group seeing an alternative group site. We additionally ensured that likely internet searches would not lead invited households to find these websites outside of the provided links. We did not indicate that this Trial would be used for academic research; however, as per the Utility's request, we included a statement at the bottom of each page saying, "This website is developed and managed by Peak Rewards Trial personnel at the University of Calgary and the University of Alberta in partnership with" the Utility.

Throughout the experiment, participating households see a "Learn More" link in the App's description of upcoming events. (See Appendix C.1.) This link takes participants to an FAQs page on their group-specific site that details events, incentives, electricity consumption reduction options, and group-specific information related to events. Manuals on how participants can use the technology related to their group are available on this page.

Throughout the experiment, PDF documents sent to participants (such as updated technology guides) were hosted on these sites.

B.5 Pre-launch communication

Before events began in February 2022, the following series of emails were sent to customers in the final Central, Tech, and Manual groups, according the the schedule below. The purpose of these emails was to remind households of upcoming potential events and to make sure they were ready with the App. In particular, given the amount of time needed to recruit households and install their equipment (if in the Central or Tech groups) meant that households could have been waiting for the Trial to start for a number of months, depending on their recruitment wave (see Appendix B.2).

Table B2. Pre-launch Email Schedule

Date	**Communication**
February 15, 2022	Pre-launch email, 1 week before
February 18, 2022	Pre-launch email, 4 days before
February 22, 2022	Launch day, events start

[redacted]

From: [redacted]
Sent: 2022, February 15 8:31 PM
To: [redacted] <[redacted]>
Subject: [redacted] Peak Rewards starts next week -get ready to earn rewards

[redacted]
Your smart home application

[redacted] Peak Rewards

Peak Rewards starts February

Thank you for participating in the [redacted] Peak Rewards Trial. Peak events will start shortly.

Here's what to expect.

- You will be notified the day before a Peak Rewards event through the [redacted] app. Please ensure your app settings are set to allow notifications.
- Your electricity usage dial should turn blue and your usage during an event will be displayed in a different colour.
- The events vary in severity and provide different financial rewards.
- To earn financial rewards, reduce your household electricity consumption during the event period.

The more you reduce during these peak events, the more you can earn.

To assist you in reducing electricity usage, [redacted] will remotely adjust your installed smart home devices during peak events.

You have full control over your devices – before, during, and after the event. Make the remote adjustments yourself on the "My Devices" page in the app, or by adjusting your smart devices directly.

For more information on what to expect during an event please click here.

If you have any questions please contact [redacted]

Regards,

The [redacted] Team

You've received this email because you're a participant in the Peak Rewards trial through the [redacted] app.

If you would like to unsubscribe please email- [redacted]

collect call or return email to notify us of this error. Thank you for your cooperation.

[redacted]

From: [redacted]
Sent: [redacted]
To: [redacted]
Subject: [redacted]

From: [redacted]
Sent: 2022, February 15 8:30 PM
To: [redacted] <[redacted]>
Subject: [redacted] Peak Rewards starts next week -get ready to earn rewards

[redacted]
Your smart home application

[redacted] Peak Rewards

Peak Rewards starts February

Thank you for participating in the [redacted] Peak Rewards Trial. Peak events will start shortly.

Here's what to expect.

- You will be notified the day before a Peak Rewards event through the [redacted] app. Please ensure your app settings are set to allow notifications.
- Your electricity usage dial should turn blue and your usage during an event will be displayed in a different colour.
- The events vary in severity and provide different financial rewards.
- To earn financial rewards, reduce your household electricity consumption during the event period.

The more you reduce during these peak events, the more you can earn.

To reduce electricity usage, you can remotely adjust your installed smart home devices during peak events on the "My Devices" page in the app, or by adjusting your smart devices directly.

For more information on what to expect during an event please click here.

If you have any questions please contact [redacted]

Regards,

The [redacted] Team

You've received this email because you're a participant in the Peak Rewards trial through the [redacted] app.

If you would like to unsubscribe please email- [redacted]

[redacted]

From: [redacted]
Sent:
To:
Subject:

From: [redacted]
Sent: 2022, February 15 8:30 PM
To: [redacted] <[redacted]>
Subject: [redacted] Peak Rewards starts next week -get ready to earn rewards

[redacted]
Your smart home application

[redacted] Peak Rewards

Peak Rewards starts February

1

Thank you for participating in the [redacted] Peak Rewards Trial. Peak events will start shortly.

Here's what to expect.

- You will be notified the day before a Peak Rewards event through the [redacted] app. Please ensure your app settings are set to allow notifications.
- Your electricity usage dial should turn blue and your usage during an event will be displayed in a different colour.
- The events vary in severity and provide different financial rewards.
- To earn financial rewards, reduce your household electricity consumption during the event period.

The more you reduce during these peak events, the more you can earn.

For more information on what to expect during an event please click here.

If you have any questions please contact [redacted]

Regards,

The [redacted] Team

[redacted]

From: [redacted]
Sent: [redacted]
To: [redacted]
Subject: [redacted]

From: [redacted]
Sent: 2022, February 18 7:14 PM
To: [redacted] <[redacted]>
Subject: [redacted] Peak Rewards starts next week -get ready to earn rewards.

[redacted]
Your smart home application

[redacted] Peak Rewards

Peak Rewards starts February

Let the rewards begin.

Peak Rewards events will start to appear in your [redacted] app as early as next week.

To ensure you don't miss out on the opportunity to earn rewards, please:

- Log out of your app and log back in to refresh your data.

- Ensure your in- app notifications are turned on
  - In the app main menu at the bottom right of your screen, go to "Settings", and enable the three "Push Notifications".

The more you reduce during peak events, the more you can earn!

For more information on what to expect during an event please click here.

If you have any questions please contact [redacted]

Regards,

The [redacted] Team

C Treatment details

C.1 Group-specific event notifications

Each treatment group experienced event notifications tailored to their treatment. We describe these below. Each group received a notification 21 hours before an event as well as two hours before the event. As a function of the App company's load control system, the Central group was sent an additional, generic notification at two hours.

The "Short notification" is the notification all participants would see on their mobile device, according to their device and in-app notification settings. (Note that the Utility encouraged participants to make sure that their notifications were enabled so they would see these event notifications, right before events started in February 2022 and as well as in August 2022. See Appendices B.5 and C.3.) If participants touched and pressed the notification, they were shown the "Long notification" with event incentive details.

[Short event notifications will be included in subsequent version of draft]

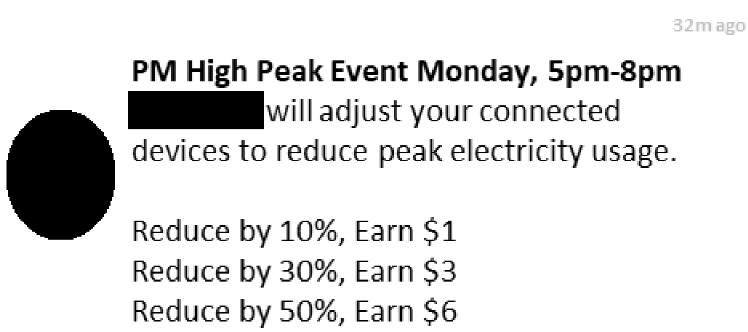


Figure C1. Long notification for Central group

[Extra notification to Central group from load control system will be included in subsequent version of draft]

Note that all group participants in the three groups below are able to locate event details in the "Advisor" tab of the App, a centralized location for information from the App company, once they receive an event notification. The "Learn More" button

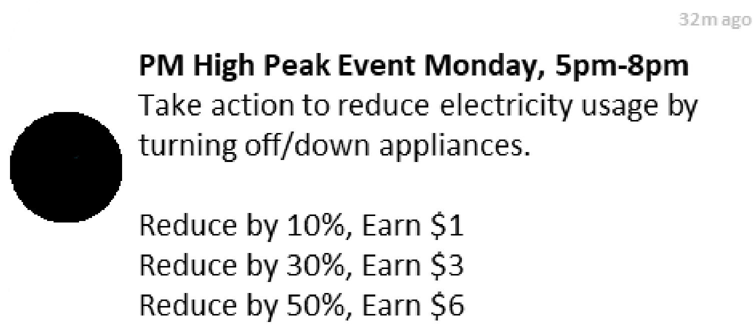


Figure C2. Long notification for Tech group

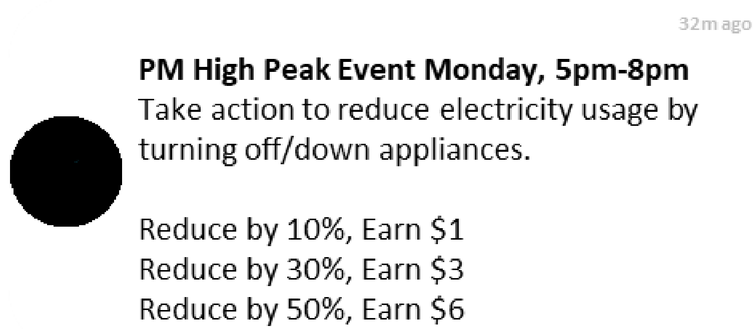


Figure C3. Long notification for Manual group at the bottom right of this information card takes participants to the "FAQs" section of the group-specific experiment website.

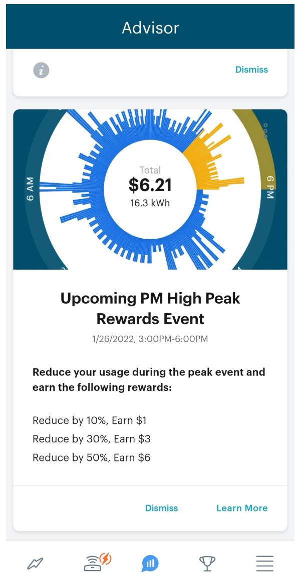


Figure C4. Event info in App

C.2 Treatment group-specific app functionality

Each group in our experiment had an App experience and functionality that differed according to their group assignment. We detail that here and walk through how participants in each group could have responded to peak events, given the options in the App.

C.2.1 Central Group

The Central group participants receive 21-hr and 2-hr notifications regarding upcoming events, as described in Appendix C.1. These notifications allow them to see the timing of the event and the magnitude of rewards for electricity consumption reductions. They also remind participants that their devices with load controllers would be altered by the Utility to reduce consumption, unless they opted-out of the event.

There are several ways that Central group participants can opt-out of events. Before an event starts, they can push an "Opt-out" button in the "My Devices" tab of the App. (This tab is a central App location that allows App users to remotely control devices that have load controllers and see the individual electricity consumption of those devices.) This button removes the participant from the event globally by removing all of their load-controlled devices from the event.

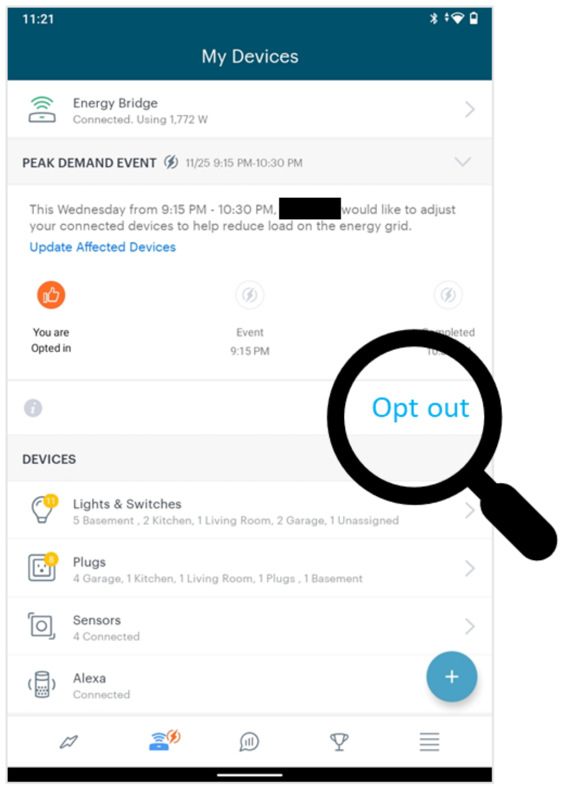


Figure C5. Central group Opt-out functionality

If they do not opt-out in this way, they see a series of screens in the "My Devices" tab. These indicate the progression of the event to the participant and signal when their devices' electricity consumption is being controlled by the utility. The first screen has an orange icon above the text "You are Opted in", as shown below in Figure C6. This occurs before an event starts. When the event is underway and participants' devices are being controlled, they see the icon above "Event" turn orange. The icon above "Completed" turns orange after an event is completed.

During an event, participants can cancel Utility device control in an device-specific way. For EV chargers and hot water heaters, they need to physically turn off load controller at the device itself. For thermostats, participants can opt-out of load control by adjusting them physically or remotely through the App, during an event.

Note that the Central group has remote and manual control of all devices with load controllers, just like the Tech group. Central group households can also change anything else in the house to alter their electricity consumption during events.

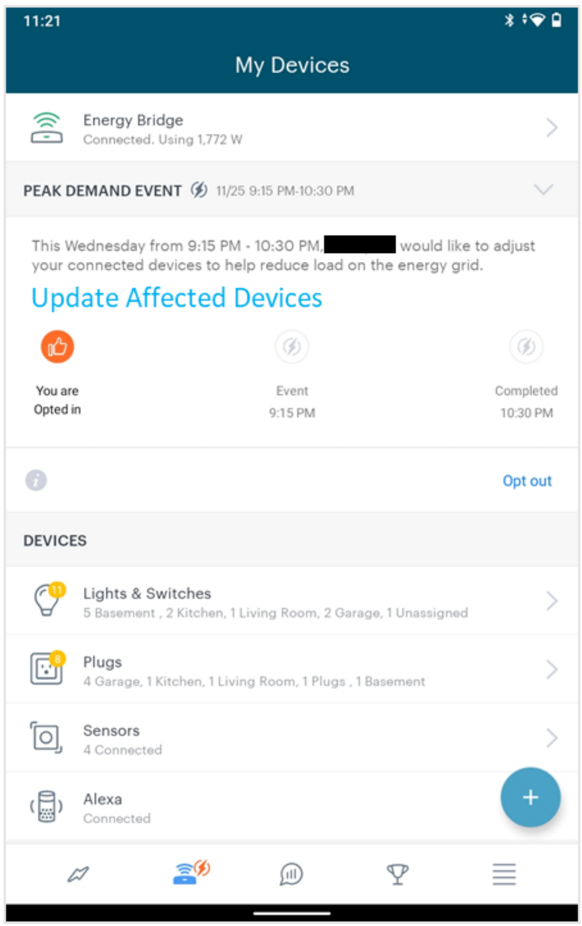


Figure C6. Central group event experience

C.2.2 Tech Group

The Tech group participants receive 21-hr and 2-hr notifications regarding upcoming events, as described in Appendix C.1. These notifications allow them to see the timing of the event and the magnitude of rewards for electricity consumption reductions. They also remind participants that they need to "take action" to make consumption changes to receive the rewards offered.

The Tech group can remotely control any device that has an installed load controller through the App. Tech group participants can "opt-in" these devices, in ways that differ by device. For EV chargers and hot water heaters, they can turn them off via two clicks from the My Devices section of the App. (See Figure C7 below for the instructions sent to participants and available on the trial website that explain these actions.) Tech group participants cannot make a schedule to turn off these devices before events start and must turn them off before or during events to reduce consumption this way. (They must also remember to turn them on unless they set up a turn-on schedule, as described in Appendix C.4.)

For thermostats, the Tech group can set up schedule for their thermostat set-point before events, using the App. They can also adjust their thermostats remotely during events with the App.

C.2.3 Manual Group

The Manual group participants receive 21-hr and 2-hr notifications regarding upcoming events, as described in Appendix C.1. These notifications allow them to see the timing of the event and the magnitude of rewards for electricity consumption reductions. They also remind participants that they need to "take action" to make consumption changes to receive the rewards offered.

Manual group participants do not load controllers given to them as part of this experiment or Utility control of any devices. They therefore only observe these notifications as well their aggregate, real-time household consumption through the App. If Manual group participants install their own smart home devices, they may be able to link them to the smart electricity consumption technology ecosystem used in this experiment. If so, they may have the capabilities of the Tech group to observe the real-time consumption of those devices/devices individually and adjust them remotely through the App. (So far we only see two households in the Manual Group that have done this, with smart thermostats.)

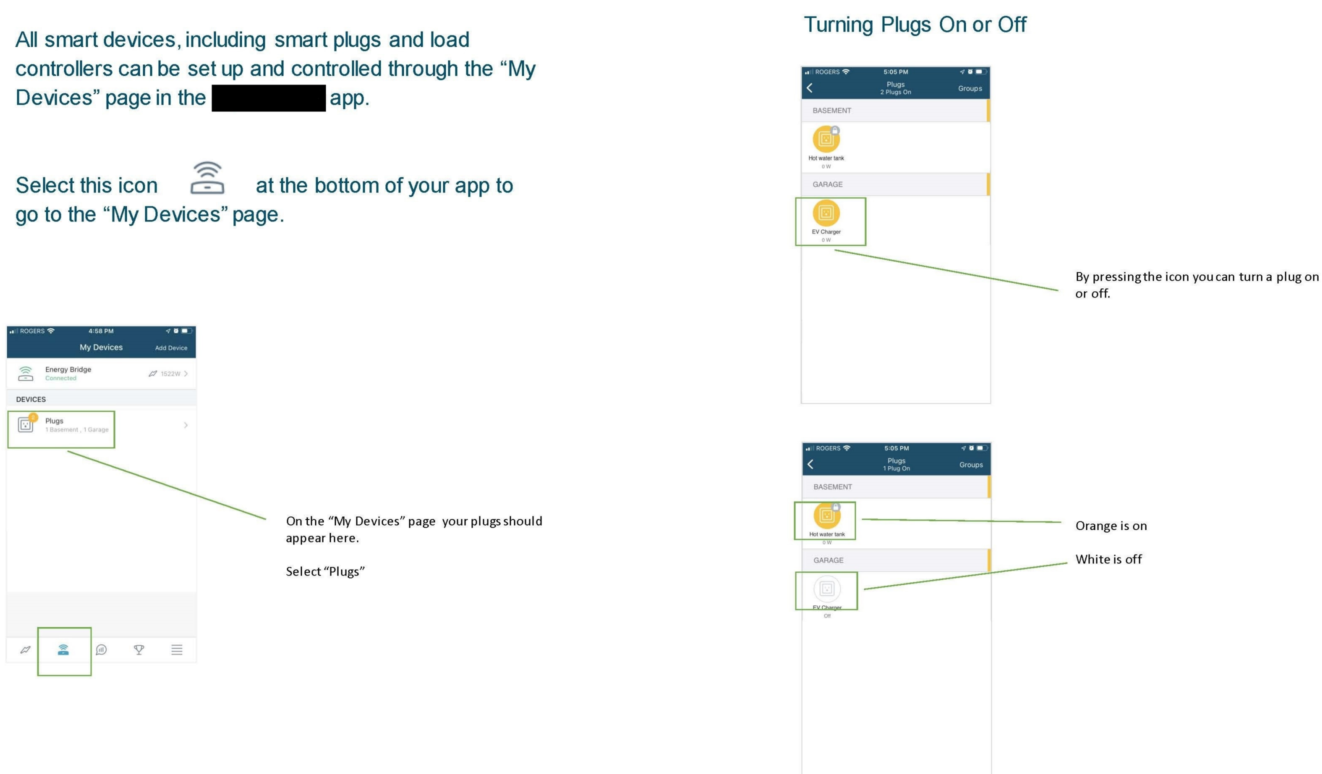


Figure C7. Controller guide for Tech group

C.2.4 Central, Tech, and Manual groups

After each event, all three of the Central, Tech, and Manual groups receive a result on their performance, as depicted below. This appears in the "Advisor" tab of the App, a central location for information from the App company. This result card reminds participants of the event type (reward magnitudes being "high" or not) and the day and time of the event. It shows the incremental reward the participant earned from the event as well as their cumulative rewards throughout the entire experiment, including the reward from the prior event. The text below the reward for the last event is variable and depends on whether a participant met one of the reward tiers. The rewards screen with one of these text options is shown below in Figure C8.

From this rewards screen, participants can select "Event History" and see their recent history of event rewards, as shown in Figure C9.

C.3 Reminder communication fall 2022

On August 25, 2022, all participants in the Central, Tech, and Manual groups were sent a reminder about the program in which they are enrolled. This reminder was sent at request of the Utility, which knew of another attractive incentive program that participants could learn about in subsequent months and wanted to proactively avoid participant attrition from programs in our experiment. We view this reminder as a "salience bump" for participants. The email text is provided below.

[redacted]

From: [redacted]
Sent:
To:
Subject:

From: [redacted] <[redacted]>
Sent: 2022, August 25 8:16 AM
To: [redacted] <[redacted]>
Subject: Peak Rewards with [redacted]

[redacted]
Your smart home application

[redacted] Peak Rewards

Peak Rewards Trial Update

Hi everyone.

The Peak Rewards trial with [redacted] has been running for 6 months now, we really appreciate you choosing to participate in this trial and your efforts to minimize your usage during the peak events.

As we move into fall and beyond, the opportunity to earn more rewards will grow. A typical household can earn **over $100** during the trial.

Keep up the good work- make sure you use the [redacted] app to keep alerted of the upcoming events.

The more you reduce during peak events, the more you can earn!

If you have any questions please contact [redacted]

Regards,

The [redacted] Team

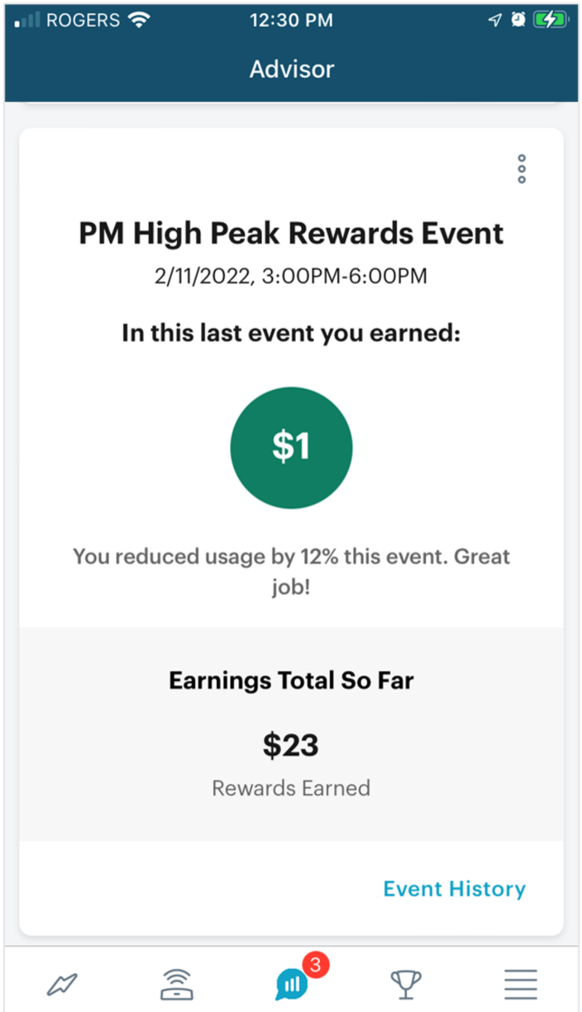


Figure C8. Rewards screen

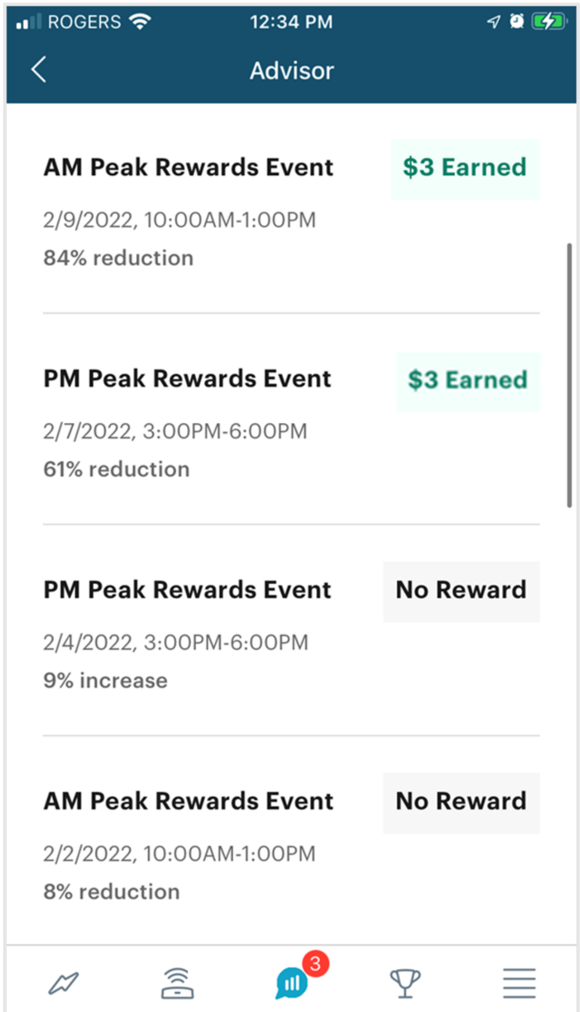


Figure C9. Event history

C.4 Robustness communication: Information for Tech group

As described in Section **??**, we sent a series of emails to participants in the Central, Tech, and Manual groups in fall 2022 to ensure that the Tech group is adequately informed about how to use its load controller technology. We also endeavored to alleviate any Tech group fears regarding moving the electricity usage of hot water heaters to after events. To both of these ends, the Utility constructed device-specific guides that offer participants instructions on how to automate devices to turn back on after any events. (Participants still need to actively adjust thermostats before/during events or hot water heaters and EV chargers during events to reduce their electricity consumption.) This "failsafe" setting may assist the Tech Group participants in adjusting their hot water heaters, in particular.

At the advice of the Utility, we first sent a generic reminder email to participants in the these groups on October 17, 2022. We then sent the main informational email on November 3, 2022. All emails are provided in subsequent pages.

[redacted]

From: [redacted] <[redacted]>
Sent: October 17, 2022 10:42 AM
To: [redacted]
Subject: Get ready to Earn Peak Rewards this Winter.

[△EXTERNAL]

[redacted]
Your smart home application

[redacted] Peak Rewards

Peak Rewards Notifications

Hi everyone.

Be on the lookout for notifications on your [redacted] app that indicate the rewards you can earn by reducing electricity usage during specific hours.

For a reminder of Peak Rewards Trial details and how you can reduce electricity usage, click here.

The more you reduce during peak events, the more you can earn!

If you have any questions please contact [redacted]

Regards,

The [redacted] Team

You've received this email because you're actively enrolled in the in the Peak Rewards trial through the [redacted] app.

If you would like to unsubscribe please email- [redacted]

[redacted]

From: [redacted]
Sent: October 17, 2022 10:43 AM
To: [redacted]
Subject: Get ready to Earn Peak Rewards this Winter..

[△EXTERNAL]

[redacted]
Your smart home application

[redacted] Peak Rewards

Peak Rewards Notifications

Hi everyone.

Be on the lookout for notifications on your [redacted] app that indicate the rewards you can earn by reducing electricity usage during specific hours.

For a reminder of Peak Rewards Trial details and how you can reduce electricity usage, click here.

The more you reduce during peak events, the more you can earn!

If you have any questions please contact [redacted]

Regards,

The [redacted] Team

You've received this email because you're actively enrolled in the in the Peak Rewards trial through the [redacted] app.

If you would like to unsubscribe please email- [redacted]

[redacted]

From: [redacted]
Sent: November 3, 2022 1:19 PM
To: [redacted]
Subject: Ready to earn more Peak Rewards this winter?

[△EXTERNAL]

[redacted]

[redacted] Peak Rewards

Peak Rewards Earnings

Hi everyone,

Winter is coming—get ready to maximize your earnings!

Did you know:

- Households that shifted **water heater usage till after peak evening events** saw a **26% drop in electricity usage**, during events.
- What's that number for **EV charging**? A **15% drop.**

- How about for **thermostats**? A **15% drop**.

For a refresher on how to keep your earnings automated, click here.

The more you reduce during peak events, the more you can earn!

If you have any questions please contact [redacted]

Regards,

The [redacted] Team

You've received this email because you're actively enrolled in the in the Peak Rewards trial through the [redacted] app.

If you would like to unsubscribe please email- [redacted]

[redacted]

From: [redacted]
Sent: November 3, 2022 1:18 PM
To: [redacted]
Subject: Ready to earn more Peak Rewards this winter?

[△EXTERNAL]

[redacted]
Your smart home application

[redacted] Peak Rewards

Peak Rewards Earnings

Hi everyone,

Winter is coming—get ready to maximize your earnings!

Did you know:

- Households that shifted **water heater usage till after peak evening events** saw a **26% drop in electricity usage**, during events.
    - Hot water heaters can retain heat for hours after shut-off.

- To learn how to *automate water heater turn-on* after events and for other tips, see this guide.

- What's that number for **EV charging**? A **15% drop.**
- How about for **thermostats**? A **15% drop**.

Our **new** EV and thermostat guides also have great tips on how to automate your earnings.

The more you reduce during peak events, the more you can earn!

If you have any questions please contact [redacted]

Regards,

The [redacted] Team

[redacted]

From: [redacted]
Sent: November 3, 2022 1:16 PM
To: [redacted]
Subject: Ready to earn more Peak Rewards this winter?

[△EXTERNAL]

[redacted]
Your smart home application

[redacted] Peak Rewards

Peak Rewards Earnings

Hi everyone,

Winter is coming—get ready to maximize your earnings!

There are many ways to generate rewards during peak events, including through some “big user” appliances.

Did you know:

- Households that shifted **water heater usage till after peak evening events** saw a **26% drop in electricity usage**, during events.
- What's that number for **EV charging**? A **15% drop.**
- How about for **thermostats**? A **15% drop**.

For a refresher on how to maximize earnings, click here.

The more you reduce during peak events, the more you can earn!

If you have any questions please contact [redacted]

Regards,

The [redacted] Team

D Supplementary Results

D.1 Device-Specific Regressions

This section provides a detailed analysis of the device-specific regression results. Figure 8 presents the regression results for the specification that includes device-type by event interactions. Households in our sample can have multiple load controller devices. Consequently, this specification implicitly assumes that there are no "synergies" associated with having multiple devices. For example, this assumes that a household with an EV and hot water heater load controller does not have an elevated response to events above the summation of the response of an EV only and hot water heater only household. We can empirically test this assumption.

Table D1 reports the results of a regression specification that interacts the event indicator with a vector of indicator variables for each possible load controller device combination. Each regression is run separately for the Central and Tech groups and compares each treated group to the two control groups (i.e., the Info and Control).

For households with only one load controlled device in the Central group, we observe larger effects for households with hot water heater load controllers compared to thermostats and EVs. Households with more devices have larger estimated treatment effects. We perform t-tests to evaluate if there are synergies associated with having multiple load controlled devices and find no statistically significant evidence in the Central group.

For the Tech group, we find households with EV only load controllers have the largest response to events, while thermostat and hot water heater only households have weaker responses to events. Unlike the Central group, there is no consistent relationship between additional devices and demand reductions. There is a small sample of households (n=2) with thermostats and EVs that appear to have a considerably elevated response to events. However, it is difficult to draw broad conclusions from these results given the small sample size. We perform t-tests to evaluate if households with multiple installed load control devices respond more to events than the summation of the individual devices and find no statistically significant evidence of such synergies at the 5% level.

Table D1. Treatment Effects of Events - Device-Specific Combinations

Device Combination	Central	Tech
Thermostat Only	-0.1856***	-0.0592**
	(0.0360)	(0.0299)
	[-16.9413]	[-5.7443]
EV Only	-0.1327*	-0.1478**
	(0.0801)	(0.0664)
	[-12.4306]	[-13.7383]
Water Only	-0.3422***	-0.0654*
	(0.0380)	(0.0353)
	[-28.9812]	[-6.3302]
Thermostat & EV	-0.5789***	-0.6200***
	(0.1250)	(0.2172)
	[-43.9460]	[-46.2061]
Thermostat & Water	-0.4583***	-0.0462
	(0.0379)	(0.0282)
	[-36.7622]	[-4.5106]
Water & EV	-0.4719***	0.1219
	(0.1457)	(0.0929)
	[-37.6202]	[12.9683]
Thermostat & EV & Water	-0.4597***	-0.1715***
	(0.1069)	(0.0089)
	[-36.8528]	[-15.7563]

Notes. The results represent the coefficients for the interaction of the event indicator and each device combination indicator variable. The values in the parenthesis represent the standard errors. The bracket represents the percentage change in consumption using the transformation $100 \times (exp(\widehat{\gamma}) - 1)$.