How Do Consumers Respond to Price Complexity? Experimental Evidence from the Power Sector

Grant D. Jacobsen and James I. Stewart*

November 2021

Abstract

Spurred in part by growing production from renewable sources and adoption of electric vehicles, time-variant pricing programs for electricity are increasingly being used to influence the shape of residential demand. The most common time-variant prices are time-of-use (TOU) prices, which vary by hour of day, and event-based prices, which take effect during idiosyncratic "critical" events. We present evidence on the effects of TOU prices and event-based prices when implemented in isolation versus simultaneously. The key finding is that time-variant prices reduce demand during critical events by 19% when event-based pricing is implemented in isolation, but only 5% when TOU and event-based prices are implemented together, despite both price schemes creating similar financial incentives. The results suggest that price complexity may dull consumer responsiveness to price signals.

JEL Codes: Q21, Q54, D12
Keywords: price complexity, time-variant pricing, critical peak rebates, time-of-use pricing, electricity, dynamic prices

*Jacobsen is at the University of Oregon and Stewart is at the Cadmus Group. Jacobsen gratefully acknowledges funding received from the *Alfred P. Sloan Foundation*. We are thankful for comments received from Ed Rubin, Eric Zou, Grant McDermott, Glen Waddell, Kenneth Gillingham, and participants at the Workshop on Environmental Economics and Data Science and the Advances with Field Experiments Conference.

1 Introduction

Electric utilities are increasingly using time-variant pricing programs, in which households pay different prices depending on when electricity is consumed, to shape residential demand. Shifting demand is an important element in the transitional path toward further integrating low-carbon renewable resources into the electric system because it can help align consumer demand with the supply of intermittent renewable sources of power. For example, time-variant pricing programs can be used to shift consumption toward periods of elevated solar and wind generation and away from periods of elevated coal and natural gas-based generation.[1] While time-variant pricing has historically been used sparingly, the installation of automatic metering infrastructure ("smart meters"), the increased prevalence of intermittent renewables, and increased adoption of electric vehicles have sparked broad interest in recent years. Many utilities across the country are considering or actively implementing broad deployment of time-variant pricing programs. The number of consumers on time-variant pricing globally is expected to rise to 75 million by 2025, in part because jurisdictions, such as California and New York, are requiring utilities to shift toward time-variant rates (Feldman, 2018).[2]

Time-variant prices can be implemented in a variety of ways. Real-time pricing, in which prices continuously reflect the costs of generation, would best align prices with the costs of generation, thereby creating large efficiency gains (Borenstein, 2005). However, due partly to fears about price volatility, utilities and consumers have been reluctant to embrace such programs. In contrast, time-of-use (TOU) pricing, in which consumers pay different prices for electricity depending on the time of day when electricity is used regardless of system conditions, has received relatively greater acceptance. TOU programs typically set a different per kWh price for electricity depending on whether the electricity is consumed during "peak" (e.g., 3:00 p.m.- 8:00 p.m.) or "off-peak" periods. A shortcoming of TOU pricing is that it is not responsive to idiosyncratic events that affect electricity supply or demand, such as un-

[1] See Holland and Mansur (2008) for a discussion of how the environmental effects of time-variant pricing depend on regional generation patterns.

[2] Matisoff et al. (2020) discuss the appeal of time-vary electricity prices, as well as potential barriers that could slow its widespread adoption.

usual weather. Event-based pricing schemes, including critical peak prices (CPP) or critical peak rebates (CPR),[3] are designed to address such events. Event-based programs provide a large per-kWh incentive for customers to conserve electricity during "critical" events, which utilities have traditionally called when demand is predicted to be unusually high. With critical peak prices, consumers are charged a higher price for electricity consumed during critical events. With critical peak rebates, consumers receive a rebate for each kWh they conserve during critical events relative to their reference usage.[4]

TOU pricing and event-based pricing have been studied in isolation,[5] but little research has examined how these two types of pricing programs interact when implemented simultaneously. This is an important shortcoming because there are potential benefits to implementing the programs simultaneously. TOU prices are oriented toward creating consistent shifts in daily demand patterns, whereas event-based programs are designed to shift demand in response to idiosyncratic events. Both types of demand changes are helpful for decreasing peak loads and reducing reliance on generators at the end of the dispatch curve, which tend to be inefficient and powered by fossil fuels, and are potentially important tools as the supply of intermittent renewable generation expands. Recognizing the possible benefits of using TOU and event-based programs in tandem, utilities have begun implementing time-variant pricing programs that include both elements (e.g., Consumers Energy's "Peak Power Savers" program in Michigan).

While there is theoretical appeal to using TOU and event-based pricing in combination, it is unclear how well the two pricing schemes will work in combination in practice, in part because consumers have been shown to respond in unpredictable ways in the face of multiple financial incentives. For example, Chetty et al. (2009) show that consumers do not fully account for sales taxes when purchasing goods for which the full price is both the posted price and the sales tax. Similarly, Hossain and Morgan (2006) present evidence that consumers are less sensitive to shipping costs than to posted prices in eBay auctions. As described in DellaVigna (2009), this behavior is consistent with consumers having limited

[3]Critical peak rebates are sometimes referred to as "peak time rebates."

[4]Reference levels that are used for calculating savings and rebates represent an estimate of the customer's consumption that would have occurred if the event had not been called and are typically based on weather and customer-specific historical usage patterns.

[5]We discuss the literature on time-variant pricing in more detail in the next section.

attention and being forced to develop simplifying heuristics for decision-making.

Predicting consumer decision-making related to financial incentives in the power sector has proven particularly difficult. For example, while only indirectly related to time-variant prices, there is evidence that consumers are weakly motivated by financial incentives that operate through investments in energy efficiency, potentially because these investments require consumers to undertake relatively complex decision-making that involves consideration of many factors (e.g., upfront costs, future prices, projected usage, and non-monetary factors, such as the inconvenience of a new installation).[6] The layering of TOU and event-based pricing programs may be a case in the power sector where consumer decision-making is especially challenging because there are multiple price signals and the signals are intermittently in effect across time.[7]

This paper evaluates how consumers respond to time-variant pricing, focusing especially on the effectiveness of layering time-of-use pricing with event-based pricing. The specific types of time-variant pricing that are evaluated include TOU pricing (in isolation), critical peak rebates (in isolation), and TOU and critical peak rebates offered simultaneously ("hybrid pricing"). The analysis is based on data from a field experiment run by a vertically-integrated electric utility in the western U.S.[8] The treatments were initiated in the summer of 2016 and included about 3,500 households.[9] The key finding is that, during summer critical events, the use of rebates in isolation is highly effective and reduces consumption by 19 percent. In contrast, hybrid pricing schemes that create a nearly identical incentive to conserve electricity during events are much less effective, only reducing consumption by about

[6]Fowlie et al. (2015) present evidence that the uptake of free energy efficiency programs is extremely low even when the programs have sizable monetary benefits. Jacobsen (2015) presents evidence that the willingness of consumers to purchase high efficiency products does not change in the face of elevated energy prices. Allcott and Greenstone (2012) provide an overview of the "energy efficiency gap," which refers to low consumer take-up of seemingly high-return investments in energy efficiency.

[7]While not the focus of our paper, layering price incentives and non-pecuniary incentives may also create unanticipated behavioral responses. Ferraro and Price (2019) present some evidence along these lines, showing that less price-sensitive consumers are more responsive to norm-based messaging in the context of residential water demand.

[8]We coordinated with the utility on the design of the experiment and have received permission from the utility to use the data for an academic article.

[9]The main analysis focuses predominantly on summer-time effects, but winter-time effects are also discussed in the Supplementary Material (SM). Winter-time effects are less pronounced and more sensitive to modeling assumptions.

5 percent.[10,11]

How should our findings—especially the reduced effectiveness of hybrid pricing—be interpreted? We posit that the key results are driven by the complexity of hybrid pricing and the response of consumers to this complexity.[12] Hybrid pricing can be considered the most complex form of time-variant pricing we evaluate based on either of the two primary adjective definitions for "complex," which are "involving a lot of different but related parts" and "difficult to understand or find an answer to because of having many different parts" (Cambridge Dictionary, 2020). In the context of these definitions, hybrid pricing involves the most "parts" relative to TOU-only or event-only versions of time-variant pricing because it includes two time-variant pricing components, as opposed to one. Relatedly, as we discuss and show graphically in Section 3, hybrid pricing creates the most changes in the marginal price of electricity across hours of the day. Households on the hybrid price plans we evaluate experienced four to six changes in the marginal price of electricity on event days; whereas households on the CPR-only price plans experienced only two changes in the marginal price of electricity on event days. With respect to being "difficult to understand," we present survey evidence in Section 5 that customers on hybrid pricing had weaker comprehension of how time-variant prices operate than customers on either type of stand-alone pricing.

If customers indeed found hybrid pricing to be more complex, then our findings, while initially surprising, are consistent with predictions from the literature on behavioral economics. Specifically, this literature has found that, when forced to make even mildly complex decisions, consumers sometime choose to simply maintain the status quo (Kahneman et

[10] We also find that, outside of critical events, none of the pricing interventions have a statistically significant effect on consumption. In the next section, we discuss the literature on TOU pricing and why TOU pricing may not have triggered a change in usage in our empirical setting. In short, the literature suggests that household electricity demand is very inelastic to TOU pricing when the ratio of peak to off-peak price ratios is low and enabling technology (e.g., an in-home display or a programmable communicating thermostat) is not installed.

[11] We examine several different critical rebate levels—both as stand-alone and hybrid offerings—and there is evidence supporting the reduced relative effectiveness of hybrid pricing during events in all cases.

[12] An alternative explanation for our results, rather than complexity, is that the different effects are driven by selection in the type of customer that opted into each experimental arm (i.e. each type of time-variant pricing we investigate) before the recruit-and-deny randomization procedure was administered. We present evidence that this is unlikely to drive our results in Section 6. The key reason is that we document much stronger effects from rebate-only pricing relative to hybrid pricing even across experimental arms that do not have significant differences in pre-experiment energy consumption or other observable covariates, such as age, income, household size, enrollment in auto-pay billing, etc., suggesting a limited role for a selection-based mechanism.

al., 1991; Samuelson and Zeckhauser, 1988; Clerides and Courty, 2017; Earl, 1990). In our context, the status quo would manifest as reduced effectiveness at changing consumption patterns. Our findings complement existing work from laboratory experiments that shows that complex prices can confuse consumers (Kalayci and Potters, 2011), that complex marketplaces narrow the focus of consumers (Huck et al., 2016), that complex prices negatively impact consumer decision making (Huck and Wallace, 2015); and that consumers dislike complexity in choice settings with uncertainty (Sonsino et al., 2002).

While the findings are not a general endorsement of any specific type of time-variant pricing and caution should be taken when applying the results to other settings, they support the notion that price interventions in the electricity sector should value simplicity and that the optimal pricing structure may be different than what might be prescribed absent behavioral considerations.[13] This conclusion is in line with existing qualitative work that indicates that simplicity is an important attribute in designing interventions in the energy sector (Hobman et al., 2016). It is also related to findings from other settings that show that simplicity can enhance the effectiveness of policy interventions for obtaining desirable outcomes, such as increasing the uptake of tax benefits (Bhargava and Manoli, 2015), improving behavior related to public health (Matjasko, 2016), increasing application rates for college financial aid (Bettinger et al., 2012; Dynarski and Scott-Clayton, 2006), and improving understanding of financial products (Brown et al., 2020). An implication of the findings is that policymakers and utility managers in the energy sector should weigh the benefits of adding price complexity—which can enable financial incentives to better align with the social costs of generation—against the costs, which may include dampened consumer responsiveness. More broadly, there is increasing interest among policymakers in employing or mandating new types of prices in various sectors (e.g., congestion pricing during periods of elevated traffic, dynamic curb pricing to address the increasing use of ride-hailing services, changes in the prices observed and paid by consumers in the health sector, initiatives to reduce the "hidden" fees in cable and wireless bills). Consideration of price complexity may be important as policymakers and other actors consider alternative types of pricing in

[13] See Schneider and Sunstein (2017) for a discussion of the role of behavioral considerations in the design of time-varying electricity prices.

a wide array of new settings.

2 Findings in the Context of Existing Research on Time-Variant Pricing

A large body of research now exists on the effects of time-variant pricing.[14] Much of the empirical evidence is comprised of pilot studies funded through Department of Energy grants or initiated internally by utilities as part of resource plans submitted to utility commissions or other regulatory bodies. Most of these pilots have not been peer-reviewed, but they still represent a potentially valid source of information. Faruqui et al. (2017) provide a meta-analysis of these pilots, focusing on 63 studies that took place between 1997 and 2017. They find that time-variant pricing changes consumption and that the effect is larger when the peak-to-off-peak price ratio is larger and if enabling technology, such as a programmable communicating thermostat or an in-home display, is present. For example, absent enabling technology, a consumer facing TOU pricing with a 2:1 peak-to-off-peak price ratio is expected to decrease peak demand by 5% relative to peak demand for consumers on standard pricing, whereas a consumer facing a 4:1 peak-to-off-peak price ratio is expected to reduce peak consumption by 10%.

A variety of time-variant pricing experiments have also been studied in the academic literature. For example, in a relatively early study, Allcott (2011) evaluates the response of Chicago households to real-time pricing and estimates a demand elasticity of -.06 to -.08. Subsequent studies have frequently focused on identifying how the response to time-variant prices differs depending on whether certain complementary types of technology are present in the home. Jessoe and Rapson (2014) evaluate a CPP program based on an experiment in Connecticut. They find that households without in-home displays (IHD) communicating real-time feedback on consumption only reduce their consumption by 0-7% (depending on how long in advance households were notified of the critical event). For households with IHDs, effects are larger at 8-22%. Bollinger and Hartman (2019) and Harding and

[14] There is also a substantial literature that examines the price elasticity of demand for non-time-variant rate schedules. For example, Deryugina et al. (2019) present evidence that consumers on standard electricity rates are mostly inelastic to price changes, with estimated elasticities falling at -0.27. Auffhammer and Rubin (2018) estimate a similar elasticity for residential natural gas (-0.23 to -0.17).

Lamarche (2016) evaluate the effect of TOU and variable peak pricing (VPP) based on an experiment implemented with a utility in the southern United States.[15] Both price interventions are effective at changing consumption. Households are most responsive to short-term price fluctuations induced through VPP if they are equipped with programmable communicating thermostats that can automatically turn-off air conditioning systems in response to price changes. Gillan (2018) examines incentives similar to critical peak rebates and finds that the incentives are effective at reducing consumption, especially if programmable communicating thermostats are present in the household. Most recently, Burkhardt et al. (2019) present evidence from Texas that critical peak pricing reduces electricity consumption by 14% and that three quarters of the response can be attributed to reduced air-conditioning.

Fowlie et al. (2021) provide an additional study related to time-variant pricing. They examine the effects of time-of-use pricing and critical peak pricing when each type of program is offered in isolation. The study focuses specifically on how the effects of the pricing programs vary depending on whether the programs are structured as "opt-in" programs, in which participants must voluntarily enroll, or "opt-out" programs, in which participants are automatically enrolled. They find that, while all interventions are effective, opt-out programs create the largest aggregate demand reductions because they have higher levels of participation.

How do our estimates fit into the existing literature? Our average treatment effect for critical peak rebates (19%) is in the expected range based on critical peak rebate interventions, which have fallen in the 0-40% range (see Figure 11 of Faruqui et al., 2017). Our TOU effects, which are insignificant, are a bit more surprising, at least upon initial inspection, considering that most studies have found a significant effect. However, the TOU prices in our intervention used a relatively low peak-to-off-peak ratio of about 2:1. For interventions of this size, the expected reduction in peak consumption is only about 5% (Faruqui et al., 2017), which is inside the confidence interval for most of the models that we estimate. Several other pilots using low peak-to-off-peak ratios have also found insignificant effects.[16] Additionally, the TOU pricing intervention we evaluate may have been muted or

[15]Variable peak pricing is similar to critical peak pricing, with the exception that the price can vary across critical events.

[16]See Faruqui (2018), which reports a more detailed version of Figures 1 and 2 from Faruqui et al. (2017).

ineffective in our setting because it was not paired with an enabling technology. A final possible explanation for the absence of TOU effects is that the experiment was implemented in a geographic region with a relatively low peak summer demand due to a mild summer climate and low air-conditioning demand, which mitigates that extent to which households can achieve savings by adjusting their thermostats.

The key contribution the present study makes relative to existing work is that we focus on how layering price incentives alters their effectiveness. In this regard, our study evaluates how adding complexity influences the effectiveness of time-variant pricing, where complexity is defined by the number of components in a price plan, the number of changes in the marginal price of electricity embedded in a price plan, or how difficult a price plan is to understand.

3 The Experiment

The experiment was implemented with the utility's residential customers using a recruit-and-deny randomization procedure.[17] The sampling frame was limited to households on the utility's standard residential rate who had resided at the same address for the previous twelve months, who were not participating in other demand response programs, who had an interval consumption meter, and who had a valid email address on record. Among this group, a random set of slightly over 100,000 households were included in the pilot. All of these households were randomly pre-assigned to one of nine time-variant pricing programs. Additionally, all households were randomly pre-selected as either "treatment" households, who would be enrolled upon opting into the program, or "control" households, who would not be enrolled even if they opted into the program. After assigning households to an experimental group and designating them as either treatment or control households, the utility then encouraged the households to select into their pre-assigned time-variant pricing plan through emails, post cards, and business letters. Households who opted into a time-variant pricing program were either placed into the program or denied enrollment according to

[17] Recruit-and-deny experimental studies are sometimes termed lottery or oversubscription methods. They are used commonly in the academic literature (e.g., Fowlie et al., 2021). See Gandhi et al. (2016) for a discussion of the application of recruit-and-deny field experiments in the energy setting.

their randomized pre-selection.[18] Control households remained on standard block-rate pricing schedule, which had a marginal price of about 11 cents/kWh regardless of the time of consumption.[19]

There were nine different pricing treatments included in the experiment: three CPR-only, three TOU-only, and three TOU + CPR hybrids. The three CPR-only treatments differed in the rebate amount. The rebates under each plan were as follows: CPR-1: 80 cents/kWh; CPR-2: 155 cents/kWh; and CPR-3: 225 cents/kWh. Households on CPR-only price plans paid the standard retail rate during non-event hours.[20] Rebates were calculated by a third-party that implemented customer billing. Reference levels for calculating the rebates were based on customer-specific algorithms that combined historical usage patterns on non-holiday weekdays with weather data.

The three TOU-only pricing treatments differed both with respect to time windows and prices. TOU-1 was a two-period program with a long and modestly priced peak. TOU-2 was also a two-tiered program, but had a shorter peak with a higher price. TOU-3 used the same peak time window and a similar peak price-level as TOU-2, but split the non-peak period into two periods: mid-peak and off-peak. The specific prices for the TOU treatments were as follows: 1) TOU-1: 7.5 cents/kWh off-peak (10pm-6am) and 13.6 cents/kWh on-peak (6am-10pm); 2) TOU-2: 8.3 cents/kWh off-peak (8pm-3pm) and 17.6 cents/kWh on-peak (3pm-8pm); and 3) TOU-3: 6.9 cents/kWh off-peak (10pm-11am), 11.9 cents/kWh mid-peak (11am-3pm and 8pm-10pm), and 18 cents/kWh on-peak (3pm-8pm).[21] The three hybrid pricing programs, CPR-2 + TOU-1, CPR-2 + TOU-2, and CPR-2 + TOU-3, combined CPR-2 with each TOU pricing scheme.[22] Events were called at identical times across hybrid and

[18]Households that opted-in, but were designated as control households, were informed that their participation could not be accommodated at this time and thanked for their interest in the program.

[19]Standard residential customers were on a two-tier block-rate pricing system. Marginal prices increased by about one cent once monthly consumption exceeded 1,000 kWh.

[20]While not embedded in the experimental design, broad deployment of rebates would require the utility to collect additional revenue by raising the standard rate or increasing fixed charges.

[21]All pricing treatments involving TOU pricing were charged off-peak prices on weekends.

[22]Note that the hybrid pricing programs create a slightly *elevated* marginal incentive to conserve during events, relative to event-only pricing, because the marginal price is the rebate incentive plus the TOU peak rate, whereas for event-only pricing the marginal price is the rebate incentive plus the standard rate. For example, during events, the marginal price for the CPR-2-only plan is 166 cents (155 + 11) and the marginal price for the CPR-2 + TOU-2 plan is 172.6 cents (155 + 17.6). This difference should, if anything, lead to a stronger response to hybrid pricing during events absent behavioral considerations related to price complexity, which is the opposite of what we find.

CPR-only treatment groups, typically for three-hour blocks on unusually hot days.

We depict the general price structure across types of time-variant pricing in Figure 1. In these figures, the marginal price represents the cost of consuming a kWh plus, for plans that include event-based pricing, the foregone rebate. Each row in the figure corresponds to a general class of pricing: standard, TOU-only, CPR-only, or TOU + CPR.[23] The left panels represent prices across hours on non-event days and the right panels represent prices across hours on event days. Standard pricing is the simplest pricing plan as there is never any variation in the price across hours of the day. TOU-only and CPR-only are more complex than standard pricing because they have a single period when the marginal price is elevated. For TOU, this period is afternoon hours on both event and non-event days. For CPR-only, the period is critical hours on critical-event days. The hybrid pricing plan is the most complex, as it involves the most pricing components and the most changes in the marginal price of electricity. Note, for example, that in the depiction of hybrid pricing, customers experience four different changes in the marginal price of electricity on an event day, but CPR-only customers experience only two changes. For the TOU-3 hybrid plan, which is not depicted in the figure, the complexity is even more elevated relative to the CPR-only price plans, as there are six changes in the marginal price of electricity on event days.

The utility began marketing the program in February 2016 through email and postal mail. Marketing materials emphasized the opportunity for reducing bills and improving sustainability by enrolling in time-variant pricing. In order to obtain the desired sample sizes, recruitment efforts were dispatched in waves whereby marketing materials were sent to a new randomly selected group of treatment and control individuals in each treatment arm. The CPR-2, CPR-2 + TOU-2, and TOU-2 experimental groups were viewed as priority price plans by the utility that were likely to perform the best in terms of energy savings and customer satisfaction. These groups, especially TOU-2, were over-represented in some of the recruitment waves.

[23]We do not report a scale on the left axis because these figure are meant to represent the general price structure for each type of plan as opposed to representing any of the precise price plans described in the previous paragraph. For both the TOU-only and hybrid price schedules, we depict the rate schedules based on a two-tier TOU component with an afternoon on-peak period (i.e. TOU-2). In these depictions, we assume that the household's consumption level is at or below the household's reference usage level, thereby allowing households that are enrolled in plans that include event-based rebates to earn a rebate by conserving energy.

All households that enrolled by December of 2016 are included in the experimental data. Opt-in rates, treatment rates, drop-out rates, and total enrollees are displayed in Table 1. There are about 250-400 households per experimental group, with the exception of the TOU-2 experimental group which has about 900 enrollees because it was the most substantially over-represented in recruitment. Across experimental arms, roughly half of households were placed into treatment. Opt-in rates for CPR-2 and CPR-2 + TOU-2 were lower than for other experimental versions of rebate-only or hybrid pricing. Based on discussions with the utility and the third-party marketer, our understanding is that this occurred because these groups were overrepresented in earlier marketing waves that predominantly used emails as opposed to postal mail and were less successful at recruiting participants.

While each household in the experimental population was randomly assigned to a specific treatment and to the treatment or control group, potential for differential selection into treatments potentially complicates interpretation of the experimental results. Differences in the timing of the marketing across experimental groups amplifies these concerns. While the randomization of treatment and control ensures that we can estimate a causal local average treatment effect (LATE) for each experimental arm, differences in LATEs across experimental arms could be driven by either 1) differences in the effect of pricing program or 2) differences in who selects into each group. We discuss this issue further in Section 6, but to preview our conclusion, we believe the collective evidence suggests that the differential effects we document across stand-alone versus hybrid pricing schemes are *not* driven by selection-based differences. As mentioned above, the primary reason for this is that rebate-only pricing has a much stronger effect than hybrid pricing even within experimental groups that do not have significant differences in observable pre-experiment characteristics (energy consumption, age, income, household size, enrollment in auto-pay billing, etc.).

Beginning in the late spring of 2016, TOU pricing was in place and the utility held the option of calling critical events. Critical events were called six times during the summer of 2016 (twice in July and four times in August). All events were called for the time period between 4pm and 7pm. During the summer of 2017, seven events were called (one in July, five in August, and one in September). There was some slight variation in the timing of these events depending on projected system conditions, but all events started between 3pm and 5pm and lasted for three hours. The mean high temperature on event days was 92°F. For context, on non-event days during summer months over the experimental window, the mean high temperature was 78°F. For all events, customers on price schedules that included rebate incentives were provided notification of the event on the day prior to the event and day of the event via email, text message, or voice mail.

The key experimental data are hourly consumption data from households for the summer months during 2015 (pre-period), 2016, and 2017. Across all experimental groups, there are 3,422 total households. One caveat regarding the meter consumption data is that, due to a technical issue, the pre-period data are often measured as integer values. The meters compute hourly usage levels as the difference between the aggregate usage (across the life of the meter) recorded on the meter at the end of the current hour and the aggregate usage recorded on the meter at the end of the previous hour. For many of the pre-period observations, this difference was done by the metering software after rounding the current and previous meter readings down to the largest preceding whole number. As a result, usage for these observations is reported in integers and contains classical measurement error. To address this issue, we aggregate the pre-period data into means on any occasion when we employ it. In particular, we use the pre-period data to 1) graphically evaluate trends in mean consumption between treatment and control groups across hours of the day during the pre-period; and 2) control for each household's mean hourly pre-period consumption levels (e.g., mean consumption during 1pm-2pm, 2pm-3pm, etc.). Because both of these applications are based on mean levels computed over many observations for each household, the measurement error embedded in the data is vastly mitigated.[24]

For the primary analysis, we use observations from the June-September of 2016 and 2017, which are the two summers after the program was initiated. We arrange the data such that each observation includes a variable recording the household's consumption dur-

[24]We use the post-period data, which were not rounded, to confirm that aggregating the data reduces issues related to measurement error in the pre-period data. In particular, we convert the post-period data into what it would have been had it gone through the same rounding procedure as the pre-period data. We then collapse this rounded post-period data into means for each household for each of the twenty-four hours of the day. We compare these means to the means we get if we directly collapse the un-rounded post-period data into means for each household for each of the twenty-four hours of the day. The correlation between the two measures is .996. Similarly, if we regress the means from the un-rounded data on the means from the rounded data, we get a coefficient of .991 (standard error of .0003).

ing the corresponding hour-of-the-sample as well as a variable recording the mean level of consumption during the summer of 2015 for the corresponding hour-of-the-day. Mean usage levels per hour during the post-intervention period were 1.22 kWh. We exclude outlier values with usage levels greater than 8 kWh (0.2% of the data). We also exclude observations from weekends and holidays because time-variant pricing was not in effect on these days. Households only enter the post-intervention data once they opt into time-variant pricing regardless of whether they were then placed into treatment or control. Altogether, there are about 10 million observations in the data used for the primary analysis. Because households were signing up over the course of the summer and fall of 2016, there are more households in the data for the summer of 2017 (3,355) than for the summer of 2016 (2,167).

4 Analysis

Our general approach to the analysis is to evaluate the effects of time-variant pricing by comparing usage patterns for treatment households to usage patterns for control households. Due to the randomized assignment of households to treatment and control status, this comparison should lead to unbiased estimates of the local average treatment effect for households that opt into each type of pricing program. We initially compare treatment and control households by evaluating patterns in means during the pre- and post-intervention periods and then proceed for more formal estimates from regression models based on the post-period data that employ controls computed based on pre-intervention consumption.[25]

4.1 Comparison of Means

We begin the analysis by presenting mean usage levels during the summer of 2015 for all experimental categories during off-peak, mid-peak, and on-peak hours as defined by the most detailed TOU rate schedule, TOU-3. The means are reported in Table 2. As expected, mean consumption levels are lowest during the off-peak hours and highest during on-peak hours. For each period, we test for pre-period differences in mean consumption levels across all ex-

[25]Because they deal well with the periodic nature of energy consumption, "Post-only" models of energy consumption with controls computed based on pre-intervention data, as opposed to fixed effects models, are arguably the state-of-the-art for RCT evaluations of household energy consumption (see Allcott and Rogers, 2014 and Stewart and Todd, 2017).

perimental categories using an ANOVA. In all three cases, the results indicate statistically insignificant differences in mean consumption levels across categories with large p-values that exceed 0.5.

To compare mean levels of consumption in greater detail, we next present graphs of mean consumption levels by hour-of-day for the treatment and control groups. Due to space constraints, in these graphs, we aggregate the data into coarse experimental groups: CPR-only, TOU-only, and CPR + TOU. The first two experimental groups experienced relatively simple pricing treatments with between two and four marginal price changes on event days and the latter group experienced more complex pricing treatments with between four and six marginal price changes on event days. Graphs of means for each detailed experimental group (e.g., CPR-1) are available in the Supplementary Material (SM).

Pre-experiment mean usage patterns from 2015 are presented in Figure 2. Treatment observations are presented in triangles and control observations in circles. The difference in mean levels between the treatment and control households and the associated 95% confidence interval are presented in diamonds. The vertical dashed lines mark the period between 4pm and 7pm, which is when events are most frequently called. As expected during summer, consumption tends to peak during afternoon hours between 4pm and 10pm. Across all hours, treatment and control households have consumption levels that are similar, although not identical to each other. Due in part to the large number of comparisons, there are some cases, such as the morning hours for the CPR-only experimental groups, in which the difference in means between the treatment and control groups is statistically significant. To control for differences that do exist due to random chance, portions of the analysis presented below will control for pre-period differences in mean consumption.

We next move to an examination of means during the post-period. In these graphs, in cases when the price treatments were effective at changing consumption, we expect to observe differences between treatment and control means. We first display mean usage levels by hour-of-the-day on *non-event* days during the post-intervention period. These graphs are presented in the left panels of Figure 3. For all treatment groups and across hours of the day, the differences between control and treatment means are very small in magnitude and almost never differ by a statistically significant margin. The implication of these graphs is that none of the pricing treatments were effective on non-event days. This is a notable result because, for any treatment involving TOU rates, consumers have a financial incentive on non-event days to shift their consumption from on-peak to mid-peak or off-peak time windows.

For *event* days, means are presented in the right panels of Figure 3. Figure 3.2 provides clear evidence that the relatively simple CPR-only intervention led to a decrease in consumption. Usage fell sharply when events began and rebounded thereafter. In contrast, focusing on Figure 3.4, there is at best modest evidence that the more complex hybrid pricing plan led to a decrease in consumption and, to the extent that it did, it was a much smaller decrease than the response seen under CPR-only pricing.[26] Lastly, Figure 3.6 provides no evidence that event day incentives led to a change in consumption in the TOU-only interventions, as might be expected.

Collectively, Figures 2 and 3 present evidence that households were not responsive to TOU pricing for either TOU-only or hybrid interventions and were primarily responsive to event-based CPR pricing when it was offered as a stand-alone intervention as opposed to when it was bundled with TOU pricing. We next more formally evaluate the effects of the program using regression models. A benefit of these models is that they aggregate temporal periods and can directly control for pre-existing differences in consumption patterns.

4.2 Estimates

We estimate treatment effects using regression models based on the following specification,

$$\text{kWh}_{it} = \beta_k T_{ik} + \lambda \text{ Pre-Period Mean Cons.}_{ij} + \gamma_{kt} + \epsilon_{it} \tag{1}$$

where kWh_{it} is hourly electricity consumption (i indexes households and t indexes hours of the sample), T_{ik} represents a vector of treatment indicators for each treatment group (k

[26]The modest visual evidence in Figure 3.4 of some effectiveness during events for households on hybrid pricing stems from consumption decreasing during the 4pm-7pm window relative to 1) the hours just before or just after the event, in which treatment consumption tended to be greater than control consumption by about 0.1 kWh and 2) the pre-program difference in mean consumption between 4pm and 7pm, when treatment households tended to have usage levels that were about 0.1 kWh greater than control households, albeit at a statistically insignificant level (see Figure 2.2).

indexes experimental groups), Pre-Period Mean Cons.$_{ij}$ is a household's mean consumption during the pre-period (the summer of 2015) during the corresponding hour of the day (j indexes 1 through 24),[27] and γ_{kt} represents a separate vector of fixed effects for each hour of the sample for each experimental group.[28] We also present models that do not include pre-period consumption as a control variable. Due to the randomized nature of the treatment, these models are expected to produce similar, but noisier estimates of treatment effects.[29] In all models, we cluster standard errors by household.

We begin by focusing on effects on event days, which were most readily apparent based on the graphical comparisons of means. Estimates from hours when critical events were called by the utility are reported in Table 3. The first two columns report estimates in which only three treatment indicators are used: CPR-only, CPR + TOU, and TOU-only.[30] Focusing on column 2, which reports the more precise estimates because pre-intervention household consumption is included as a control variable, the relatively simple CPR-only pricing plan led to a large and statistically significant decrease in consumption of about 0.37 kWh, or about a 19% decrease relative to the mean. The more complex hybrid pricing plan also led to a decrease in consumption of 0.10 kWh (5% relative to the mean) during event hours, but the coefficient is only about a fourth of the size as the CPR-only effect, significantly different than the CPR-only coefficient, and only significant in models that include a control for pre-period consumption. There is little evidence that the TOU-only pricing treatments decreased energy usage during events as the TOU coefficient is insignificant and small in

[27]While pre-period mean consumption is a single variable, it provides a detailed level of control because it varies for each household by hour. For example, for an observation recorded for 7-8pm during the post-period, pre-period mean consumption is measured as the household's mean consumption during 7-8pm in the pre-period. When computing these means, we exclude observations from weekends and holidays because, as mentioned above, time-variant prices were not in effect on those days.

[28]Note that the vectors of fixed effects for each hour of the sample for each experimental group are not collinear with the treatment indicators because each experimental group (e.g., CPR-2, TOU-2) is defined to include both the treatment and control households assigned to that experimental group.

[29]In our primary models, the experimental indicators are not time-varying within households and therefore our estimates are effectively intent-to-treat (ITT) estimates conditional on opting into a time-variant pricing program. Because some customers dropped out after enrolling, we can also estimate treatment-on-treated (TOT) effects by instrumenting for current enrollment with an indicator for ever having opted-in. These models are essentially identical to the ITT models because less than 2% of observations correspond to households that were part of a treatment group but not actively enrolled. We report TOT estimates in the Supplementary Material (Tables SM.A and SM.E).

[30]The variables are simple aggregations of the more detailed treatment variables. For example, CPR-only is a summation of the indicators for CPR-1, CPR-2, and CPR-3.

magnitude.

In columns 3 and 4 of Table 3, treatment effects are estimated for each detailed variety of time-variant pricing. As can be seen by inspecting the coefficients on CPR-1, CPR-2, and CPR-3, each of the relatively simple event-only interventions led to large energy savings during event hours, with point estimates indicating a decrease of between 0.27 and 0.48 kWh (13.5% to 24%).[31] In contrast to the clear evidence that each CPR-only intervention was effective at reducing consumption, the coefficients on two of the three hybrid pricing indicators are statistically insignificant. Additionally, all of the coefficients on the hybrid pricing variants are smaller than the coefficients on the indicators for the CPR-only variants, with estimates ranging from -0.01 to -0.16 kWh (a 0.5% to 8% decrease). It is notable that the smallest coefficient, which is extremely close to zero, occurs for the CPR-2 + TOU-3 pricing plan. The CPR-2 + TOU-3 price plan was the most complex pricing plan embedded in the experiment. It involved both hybrid pricing and a three-tier TOU schedule, leading to six changes in the marginal price of electricity on event days.

Continuing with the discussion of Table 3, in terms of evaluating whether CPR-only pricing produces different effects on event days than hybrid pricing, it may be helpful to focus only on terms including CPR-2, which is the type of CPR-pricing embedded in all of the hybrid price plans. Focusing on these terms, the point estimates indicate that CPR-2-only pricing led to a larger decrease (36%) than the CPR-2 + TOU-1 hybrid (16%), the CPR-2 + TOU-2 hybrid (13%), and the CPR-2 + TOU-3 hybrid (1%). The CPR-2-only coefficient is statistically different than both the CPR-2 + TOU-2 coefficient and the CPR-2 + TOU-3 coefficient ($p < .05$) and almost statistically different than the coefficient on the CPR-2 + TOU-1 coefficient ($p = .12$).[32]

[31]Note that the estimated magnitude of the effect of the CPR-only interventions increases as the size of the rebate increases, suggesting that larger financial incentives may produce large responses. While none of the differences between the estimated coefficients on the CPR-only interventions are statistically significant, the CPR-1 and CPR-3 coefficients are on the cusp of being significantly different. A Wald test for the equality of the two coefficients produces a p-value of .13. Nonetheless, the evidence that the size of the rebate influences household responsiveness is at best modest, especially given the large differences in the size of the incentives, which is consistent with Gillan (2018), who finds that consumers are mostly insensitive to the size of event-based incentives.

[32]The coefficient on TOU-1 is *positive* and statistically significant, although we suspect this is spurious correlation. Treatment households in TOU-1 tended to experience relative increases in afternoon consumption in both the pre-period and on non-event days (see Figures SM.A.7 and SM.B.7). If this tendency was amplified on event days because they tend to be hotter, then it could drive the results observed in Table 3.

Detailed estimates from non-event days are presented in Table 4. Models were run separately for each group of hours embedded in the TOU-3 rate schedule, which is the most detailed TOU rate schedule and includes off-peak, mid-peak, and on-peak periods. Consistent with the graphical comparison of means, almost all estimates are small and statistically insignificant. The few that are significant may be spurious correlation due to the large number of comparisons.[33]

Taken together, the results support the patterns seen in the comparison of means. The relatively simpler CPR-only interventions produced large and significant decreases in consumption during event hours. The more complex hybrid approach led to smaller decreases in consumption in some cases and no decreases in others. The effect of hybrid pricing was especially weak when the underlying TOU-rate schedule was relatively more complex (i.e. three-tiered instead of two-tiered).

5 Discussion: Are Hybrid Prices Harder to Understand?

We posit that our results are driven by the enhanced complexity of hybrid pricing. A recent on-line experiment found that utility customers perceive electricity tariffs with more pricing components to be more complex (Layer et al., 2017). In the experiment, researchers presented 664 utility customers with electricity pricing tariffs and assessed customer perceptions of price complexity and depth of information processing. The authors found that study participants perceived hybrid price plans to be significantly more complex than the standard rate or the two-tiered TOU rate and that price complexity was associated with less information processing.

One way to investigate whether hybrid prices are meaningfully more complex in our setting is to investigate whether consumers found them more difficult to understand, which we do in this section using data from an online survey distributed by the utility. The survey was designed to investigate whether customers could identify their rate schedule and understood the prices they faced. Treatment households that were enrolled in any time-variant pricing program prior to the first event that was called during the summer of 2016

[33]Through chance alone, about one coefficient in each specification might be expected to be significant at the 10% level.

were invited to take the survey via email in November 2016. About one in five households completed the survey and the total number of respondents was 304. The key questions from the survey were: 1) Identify the rate you pay for electricity based on TOU pricing; 2) Yes/No: I can save money on my energy bill by using energy during off-peak hours instead of during peak hours; and 3) Yes/No: On event days, I can earn a rebate on my energy use if I shift my energy use from the event hours to other periods.

We focus the analysis on whether respondents were able to respond correctly to questions about their prices. For questions 1 and 2, which focus on TOU-pricing, the analysis is based only on customers on TOU-only or hybrid pricing. The correct answer to question 1 for each respondent was the TOU pricing schedule that they were enrolled in and, for question 2, the correct answer for all respondents was "yes." For question 3, the analysis is based only on customers on CPR or hybrid pricing. The correct answer for question 3 is "yes."

To analyze the survey data, we use a simple linear probability model where the dependent variable equals one if the respondent answered correctly for the corresponding survey question and the independent variable is an indicator equaling one if the respondent was on hybrid pricing and zero if they were on stand-alone pricing. The results are reported in Table 5. As can be inferred from the constants in the models, respondents on stand-alone pricing tended to answer correctly, with rates of correctness ranging from 66% to 85%. For respondents on hybrid pricing, the probability of answering correctly was significantly lower. In particularly, respondents on hybrid pricing were 7.9 percentage points less likely to identify their TOU-rate correctly, 17.9 percentage points less likely to identify they could save money on TOU pricing by shifting their consumption from on-peak to off-peak periods, and 11.0 percentage points less likely to understand that they could earn rebates by shifting their consumption to non-event hours on event days. While the results should be interpreted with some caution due to the modest response rate, potential for survey response bias, and the small sample size, the results from the survey data indicate that customers on hybrid pricing did not understand the features of their time-variant pricing plans as well as customers on stand-alone pricing treatments.

6 Limitations

The key limitation of the study, as with many experimental studies, is that it takes place in a very specific context. We study time-variant pricing for a single utility, based on an opt-in enrollment procedure, and without pairing the pricing with complementary technology. Invariably, changing the context would likely change the results. For example, it seems reasonable to think that consumers might be less affected by complexity in a setting where they have access to smart thermostats that can be programmed to automatically respond to prices. If consumers are, in fact, more responsive to complex price signals when the prices are implemented in a setting with enabling technology, then our results suggest two paths for utilities. They can adopt a simplified version of time-variant pricing and not bother pairing the new pricing scheme with a simultaneous effort to increase adoption of enabling technology, or they can adopt a complex pricing schedule and simultaneously pair the pricing initiative with incentives for the installation of new varieties of technology.[34]

It is also unclear the extent to which findings from our setting might extend to other utilities. We can only speculate on this matter and it seems likely that there will be substantial variation in the effectiveness of time-variant pricing across utilities based on the traits of the utility's programs and the features of its customer base. However, regarding our main finding—that event-based pricing is more effective when offered in isolation—there is a non-peer-reviewed study that supports it. In particular, a pilot from Xcel Energy in Colorado tested both CPP-only and TOU + CPP pricing interventions.[35] For households without central air-conditioning, the reduction in peak usage for the CPP-only intervention (31.9%) was over twice as large as the reduction achieved by the TOU + CPP intervention (15.1%). For households with central air-conditioning, there was a similar, albeit less-dramatic pattern (38.4% vs 28.8%). While it is not clear if the differences in the estimated treatment effects were significantly different, they provide evidence that our findings regarding diminished effectiveness from hybrid interventions may hold in other settings.

A final limitation of our study is that it is based on an opt-in experimental design. One

[34] See Jacobsen (2019) for a discussion of common types of incentive programs offered in the energy sector.

[35] This study is not publicly available but is described in a review by Faruqui and Sergici (2010).

issue posed by the opt-in experimental design is that, because households were invited to enroll in a specific form of time-variant pricing based on their randomized pre-assignment, it is possible that households that selected into each type of time-variant pricing systematically differed across experimental groups in the manner in which they respond to time-variant pricing. For example, it is possible that households that opted into rebate-only pricing happened to be more responsive to time-variant pricing, of any form, than households that opted into hybrid pricing. While we do not observe significant differences in mean on-peak, mid-peak, and off-peak electricity consumption across experimental categories in the pre-period (see Table 2), which we might expect if there were large selection-based differences across experimental groups in their energy-related behavioral dispositions, it is possible that the experimental groups might differ along other dimensions that could influence their response to time-variant pricing.

To investigate whether there are differences across households in experimental categories in non-energy variables, we obtained information from the utility's central information system database that were recorded during June 2017. These variables include age, income, household size, whether the household is living in a single or multi-family housing unit, whether the household is living in rented or owner-occupied housing, an indicator for donating toward renewable energy through their energy bill, an indicator for enrollment in auto-pay billing, an indicator for paperless billing, and an indicator for participation in other programs offered by the program related to energy efficiency.[36] We test for differences in these variables by regressing each variable on indicators for each experimental category. In these models, the omitted group is the CPR-2 control group. We chose this group as a point of comparison because it is the control group for the CPR price plan that was offered both as a stand-alone rebate and embedded in all hybrid price plans.

The results of the regressions described above are reported in Table SM.B. Examining the overall model p-values reported at the bottom of the table, there are six variables in which the experimental variables do not do a significantly better job than an intercept-only model (income, household size, renter, renewable energy, auto-pay, and energy efficiency)

[36]Income was recorded using a 13-category response variable corresponding to a range of values. For the analysis, it was converted to a continuous variable based on the midpoints for each range.

and four variables where significant differences are observed (age, home age, multi-family, and paperless). Looking at the coefficients, the CPR-2 only control group tended to be younger, live in newer houses, more likely to live in multifamily units, and more likely to be enrolled in paperless billing. These differences likely stem from the fact that the CPR-2 experimental arm was over-represented in earlier marketing efforts, as described in Section 3, which used email marketing (as opposed to direct mailing) more intensively than later marketing efforts.

Are the differences in some of the covariates evidence that selection into each experimental arm are driving the differential effectiveness we observe across the price plans? While we cannot rule this possibility out entirely, we think the answer is that it is not, especially for our key finding that more complex price schedules are less effective. The reason is that, with the exception of a single coefficient that is only significant at the 10-percent level (in the other energy efficiency program model), there are no significant differences documented between the CPR-2-only and TOU-2 + CPR-2 hybrid experimental groups. This is not surprising because both groups were prioritized in early marketing efforts. Even in this group, for which we observe no evidence of selection-driven differences in the covariates, we still observe large difference between the estimated effect of the relatively simple rebate-only pricing plan and the more complex hybrid pricing plan (see Table 3).

A second issue related to the opt-in experimental design is that households that chose to opt-in may be systematically different than households that did not opt-in. This is a limitation in that it makes it unclear how much our results would generalize to time-variant pricing programs implemented as mandatory or opt-out programs. We think it is reasonable to assume that households that opt into an experimental time-variant pricing program might tend to be more sophisticated and better at absorbing complex prices that the average consumer, which, if anything, would make our estimates conservative with respect to our central finding of reduced responsiveness to complex prices. Nonetheless, we view the extent to which our findings are indicative of what might be expected in a mandatory or opt-out setting as an important question and something that we hope other researchers will shed light on in the future. Effects based on both an opt-in and opt-out design are of significant policy relevance because time-variant pricing programs are implemented as both opt-in and opt-out programs. Opt-in programs are common, in part, because certain state utility commissions require time-variant pricing to be implemented as voluntary opt-in programs.

7 Conclusion

Due to technological advancement, new types of prices are being used in a variety of settings where they have not historically been applied. One setting where the type of prices that consumers face is rapidly changing is the power sector, where time-variant prices are increasingly replacing standard flat-rate or block-rate pricing. Time-variant electricity prices offer a potential avenue by which to align consumer demand with the conditions facing electricity producers, which is becoming an important objective as intermittent sources of electric generation, such as wind and solar, become more common. While time-variant pricing holds significant promise, appropriately implementing time-variant prices will require a detailed understanding of how consumers respond to them.

This paper presents experimental evidence on the effects of the two most common types of time-variant pricing, time-of-use pricing and event-based pricing (in our case, critical peak rebates), when offered both in isolation as well as in tandem. We find that event-based pricing leads to conservation during critical events, when conservation is especially valuable. While event-based prices are effective regardless of whether they are offered as a stand-alone measure or paired with TOU prices, the magnitude of their effectiveness is strikingly larger when they are employed in isolation. Specifically, the decrease in consumption during critical events, relative to a control groups on standard pricing, is almost four times larger for households on rebate-only pricing (a 19% decrease) than it is for households on hybrid pricing (a 5% decrease).

The results support the idea that more complex pricing systems can overwhelm consumers and result in unexpected outcomes, which is consistent with other findings from the behavioral economics literature that show that complexity plays an important role in consumer decision-making. While it is uncertain how much our findings will translate to other settings where new types of prices are being considered, such as the transportation or health care sectors, we believe there may be some generality to the notion that complex pricing structures make consumers less responsive to price signals. We look forward to future research that further investigates how price complexity plays a role in consumer decision-making, especially in the power sector, where enhancing understanding of consumer behaviors is likely to remain of high value as the electricity system evolves.

8 References

Allcott, H. 2011. "Rethinking real-time pricing," *Resource and Energy Economics* 26(1) 3-28.

Allcott, H. and Greenstone, M. 2012. "Is There an Energy Efficiency Gap?," *Journal of Economic Perspectives* 26(1) 3-28.

Allcott, H. and Rogers, T. 2014. "The Short-Run and Long-Run Effects of Behavioral Interventions: Experimental Evidence from Energy Conservation," *American Economic Review* 104(10) 3003-3037.

Auffhammer, M. and E. Rubin. 2018. "Natural Gas Price Elasticities and Optimal Cost Recovery Under Consumer Heterogeneity: Evidence from 300 million natural gas bills," NBER Working Paper No. 24295.

Bhargava, S. and Manoli, D. 2015. "Psychological Frictions and the Incomplete Take-Up of Social Benefits: Evidence from an IRS Field Experiment," *American Economic Review* 105(11) 3480-3529.

Bettinger, E. P., Long, B. T., Oreopoulos, P., and Sanbonmatsu, L. 2012. "The Role of Application Assistance and Information in College Decisions: Results from the H&R Block Fafsa Experiment," *Quarterly Journal of Economics* 127(3) 1205-1242.

Bollinger, B. K. and Hartman, W. R. 2019. "Information versus Automation and Implications for Dynamic Pricing," *Management Science*, in press.

Borenstein, S. 2005. "The Long-Run Efficiency of Real-Time Electricity Pricing," *The Energy Journal* 26(3) 93-116.

Brown, J. R., Kapteyn, A., Luttmer, E. F. P., Mitchell, O. S., and Samek, A. 2020. "Behavioral Impediments to Valuing Annuities: Complexity and Choice Bracketing," *Review of Economics and Statistics*, in press.

Burkhardt, J., Gillingham, K., Kopalle, P. K. 2019. "Experimental Evidence on the Effect of Informationand Pricing on Residential Electricity Consumption," Working Paper.

Cambridge Dictionary. 2020. Definition of Complex. Available at: https://dictionary.cambridge.org/us/dictionary/english/complex.

Chetty, R., Looney, A., and Kroft, K. 2009. "Salience and Taxation: Theory and Evidence," *American Economic Review* 99(4) 1145-1177.

Clerides, S. and Courty, P. 2017. "Sales, Quantity Surcharge, and Consumer Inattention," *Review of Economics and Statistics* 99(2): 357-370.

DellaVigna, S. 2009. "Psychology and Economics: Evidence from the Field," *Journal of Economic Literature* 47(2) 315-372.

Dynarski, S. and Scott-Clayton, J. 2006. "The Cost of Complexity in Federal Student Aid: Lessons from Optimal Tax Theory and Behavioral Economics," *National Tax Journal* 59(2) 319-356.

Earl, P.E. 1990. "Economics and Psychology: A Survey," *The Economic Journal* 100(402) 718-755.

Deryugina, T., MacKay, A., and Reif, J. 2019. "The Long-Run Dynamics of Electricity Demand: Evidence from Municipal Aggregation," *American Economic Journal: Applied Economics* in press.

Faruqui, A. 2018. "Modernizing Distribution Tariffs for Households," presented to Energy Consumers Australia (Nov. 9). The Brattle Group. Online at: https://energyconsumersaustralia.com.au/wp-content/uploads/Ahmad-Faruqui-Modernizing-distribution-tariffs-for-households.pdf

Faruqui, A. and Sergici, S. 2010. "Household response to dynamic pricing of electricity: a survey of 15 experiments," *Journal of Regulatory Economics* 38 193-225.

Faruqui, A., Sergici, S., and Warner, C. 2017. "Arcturus 2.0: A meta-analysis of time-varying rates for electricity," *The Electricity Journal* 30 64-72.

Feldman, B. 2018. "The Rise of Dynamic Power Pricing," Navigant Insights. Online at: https://www.navigant.com/insights/energy/2018/the-rise-of-dynamic-power-pricing

Ferraro, P. J. and Price, M. K. 2013. "Using Nonpecuniary Strategies to Influence Behavior: Evidence from a Large-Scale Field Experiment," *Review of Economics and Statistics* 95(1) 64-73.

Fowlie, M., Greenstone, M., and Wolfram, C. 2015. "Are the Non-Monetary Costs of Energy Efficiency Investments Large? Understanding Low Take-Up of a Free Energy Efficiency Program," *American Economic Review* 105(5) 201-204.

Fowlie, M., Wolfram, C., Spurlock, C. A., Todd, A., Baylis, P., and Cappers, P. 2021. "Default Effects and Follow-On Behavior: Evidence from an Electricity Pricing Program," *Review of Economics Studies*, in press.

Gandhi, R., Knittel, C. R., Pedro, P., and Wolfram, C. 2016. "Running Randomized Field Experiments for Energy Efficiency Programs," *Economics of Energy & Environmental Policy* 5(2) 7-26.

Gillan, J. M. 2018. "Dynamic Pricing, Attention, and Automation: Evidence from a Field Experiment in Electricity Consumption," working paper.

Harding, M. and Lamarce, C. 2016. "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* 35(4) 906-931.

Hobman, E. V., Frederiks, E. R., Stenner, K., and Meikle, S. 2016. "Uptake and usage of cost-reflective electricity pricing: Insights from psychology and behavioural economics," *Renewable and Sustainable Energy Reviews* 57 455-467.

Holland, S. P. and Mansur, E. T. 2008. "Is Real-Time Pricing Green? The Environmental Impacts of Electricity Demand Variance," *Review of Economics and Statistics* 90(3) 550-561.

Hossain, T. and Morgan, J. 2006. "...Plus Shipping and Handling: Revenue (Non) Equivalence in Field Experiments on eBay," *Advances in Economic Analysis & Policy* 6(2) 1-26.

Huck, S., Lünser, G., and Tyran, J. 2016. "Price competition and reputation in markets for experience goods: an experimental study," *RAND Journal of Economics* 47(1) 99-117.

Huck, S. and Wallace, B. 2015. "The impact of price frames on consumer decision making: Experimental evidence," working paper.

Jacobsen, G. D. 2015. "Do Energy Prices Influence Investment in Energy Efficiency? Evidence from Energy Star Appliances," *Journal of Environmental Economics and Management* 74 94-106.

Jacobsen, G. D. 2019. "An Examination of How Energy Efficiency Incentives Are Distributed Across Income Groups," *Energy Journal* 40(6) 171-198.

Jessoe, K. and Rapson, D. 2014. “Knowledge is (Less) Power: Experimental Evidence from Residential Energy Use,” *American Economic Review* 104(4) 1417-1438.

Kahneman, D., Knetsch, J.L., Thaler, R.H. 1991. “Anomalies: the endowment effect, loss aversion, and status quo bias,” *Journal of Economic Perspectives* 5 193-206.

Kalayci, K. and Potter, J. 2020. “Buyer Confusion and Market Prices,” *International Journal of Industrial Organization* 29 14-22.

Layer, P., Feurer, S., and Jochem, P. 2017. “Perceived price complexity of dynamic energy tariffs: An investigation of antecedents and consequences,” *Energy Policy* 106: 244-254.

Matisoff, D.C., Beppler, R., Chan, G., and Carley, S. 2020. “A review of barriers in implementing dynamic electricity pricing toachieve cost-causality,” *Environmental Research Letters* 15 093006.

Matjasko, J. L., Cawley, J. H., Baker-Goering, M. M., Yokum, D. V. 2016. “Applying Behavioral Economics to Public Health Policy Illustrative Examples and Promising Directions,” *American Journal of Preventive Medicine* 50(5S1) S13-S19.

Samuelson, W. and Zeckhauser, R. 1988. “Status quo bias in decision making,” *Journal of Risk and Uncertainty* 1:7-59.

Schneider, I. and Sunstein, C. R. 2017. “ Behavioral considerations for effective time-varying electricity prices,” *Behavioural Public Policy* 1(2) 219-251.

Sonsino, D., Benzion, U., and Mador, G. 2002. “The Complexity Effect on Choice with Uncertainty - Experimental Evidence,” *The Economic Journal* 112 936-965.

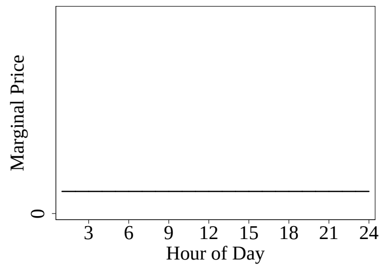
Stewart, J. and Todd, A. 2017. “Chapter 17: Residential Behavior Protocol,” The Uniform Method Project: Methods for Determining Energy Efficiency Savings for Specific Measures. National Renewable Energy Laboratory, United States Department of Energy. Report NREL/SR-7A40-62497.

9 Tables and Figures

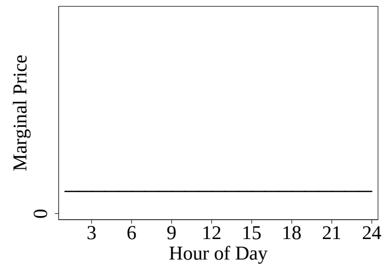
Table 1: Experimental Overview

	Opt-in Rate	Share in Treatment	Drop-out Rate	Total Enrollees
CPR-1	4.62	47.06	0.78	272
CPR-2	2.27	54.02	6.05	398
CPR-3	6.10	50.62	2.96	401
CPR-2 + TOU-1	4.18	53.69	4.58	244
CPR-2 + TOU-2	1.87	53.65	5.63	397
CPR-2 + TOU-3	4.73	50.72	5.67	278
TOU-1	3.01	53.66	2.27	246
TOU-2	3.04	55.42	4.89	922
TOU-3	3.23	52.27	4.35	264
Total	3.12	52.95	4.42	3,422

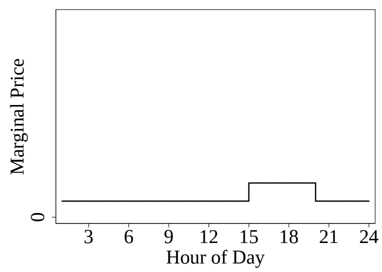
Notes: Opt-in rate, share in treatment, and drop-out rate are reported as percentages.



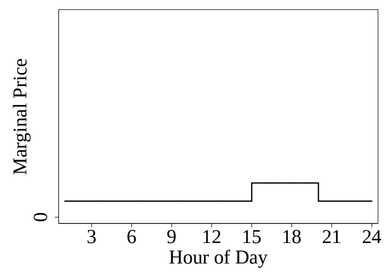
(1) Standard: Non-Event Day



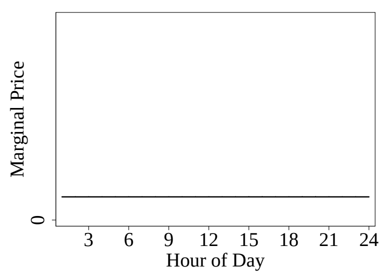
(2) Standard: Event Day



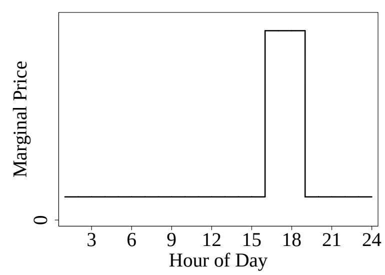
(3) TOU-only: Non-Event Day



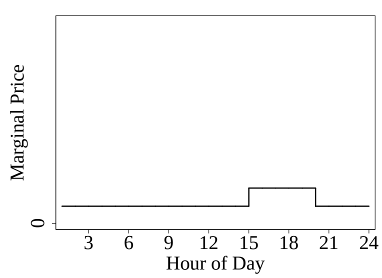
(4) TOU-only: Event Day



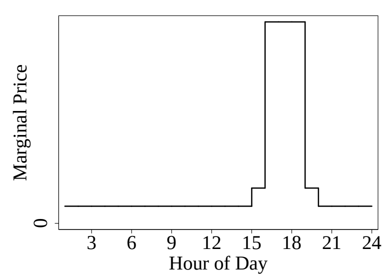
(5) CPR-only: Non-Event Day



(6) CPR-only: Event Day



(7) TOU + CPR: Non-Event Day



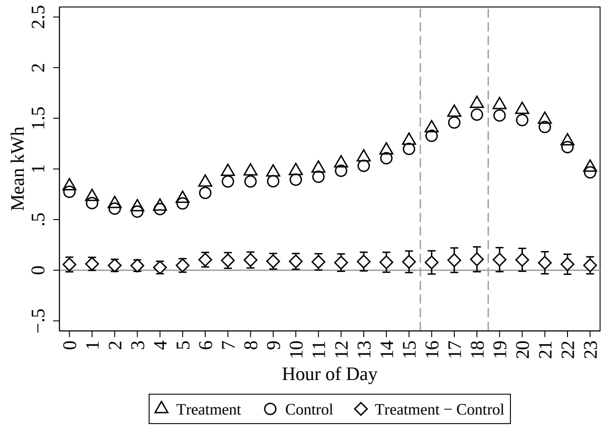
(8) TOU + CPR: Event Day

Figure 1: **Price Depiction for Each Type of Time-Invariant Pricing**. The panels on the left represent the prices on a non-event day and the panels on the right represent the prices on an event day. TOU-pricing elements are depicted based on a two-tier TOU component with an afternoon on-peak period (i.e. TOU-2).

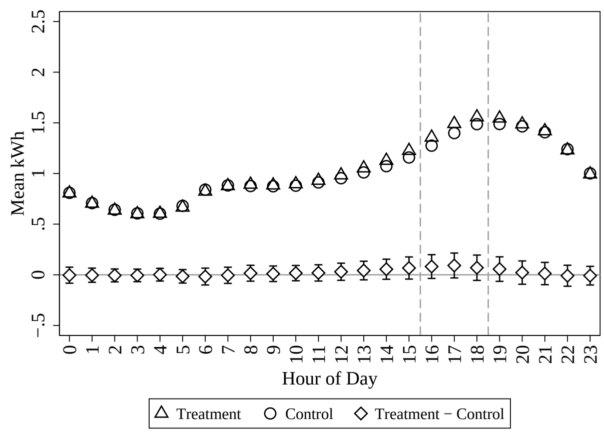
Table 2: Comparison of Pre-Program Mean Consumption Levels

	Off-Peak	Mid-Peak	On-Peak	
	10pm-11am	11am-3pm 8pm-10pm	3pm-8pm	Households
CPR-1 - Treatment	0.87	1.28	1.56	128
	(0.59)	(0.85)	(1.02)	
CPR-1 - Control	0.78	1.15	1.40	144
	(0.39)	(0.65)	(0.84)	
CPR-2 - Treatment	0.84	1.21	1.45	215
	(0.57)	(0.81)	(1.01)	
CPR-2 - Control	0.80	1.14	1.39	183
	(0.54)	(0.80)	(0.97)	
CPR-3 - Treatment	0.87	1.21	1.49	203
	(0.60)	(0.77)	(0.98)	
CPR-3 - Control	0.83	1.22	1.49	198
	(0.47)	(0.72)	(0.96)	
CPR-2 + TOU-1 - Treatment	0.82	1.17	1.44	131
	(0.56)	(0.67)	(0.83)	
CPR-2 + TOU-1 - Control	0.82	1.18	1.43	113
	(0.55)	(0.75)	(0.90)	
CPR-2 + TOU-2 - Treatment	0.80	1.14	1.40	213
	(0.46)	(0.65)	(0.88)	
CPR-2 + TOU-2 - Control	0.82	1.13	1.36	184
	(0.54)	(0.76)	(0.97)	
CPR-2 + TOU-3 - Treatment	0.83	1.19	1.46	141
	(0.60)	(0.89)	(1.11)	
CPR-2 + TOU-3 - Control	0.84	1.15	1.37	137
	(0.51)	(0.60)	(0.78)	
TOU-1 - Treatment	0.87	1.19	1.44	132
	(0.53)	(0.72)	(0.94)	
TOU-1 - Control	0.91	1.23	1.42	114
	(0.70)	(0.95)	(1.14)	
TOU-2 - Treatment	0.86	1.18	1.40	511
	(0.55)	(0.77)	(0.95)	
TOU-2 - Control	0.91	1.23	1.41	411
	(0.63)	(0.78)	(0.96)	
TOU-3 - Treatment	0.93	1.26	1.48	138
	(0.62)	(0.80)	(0.98)	
TOU-3 - Control	0.83	1.10	1.27	126
	(0.55)	(0.65)	(0.77)	
Overall Mean	0.85	1.19	1.42	
p-value (ANOVA)	0.55	0.87	0.84	

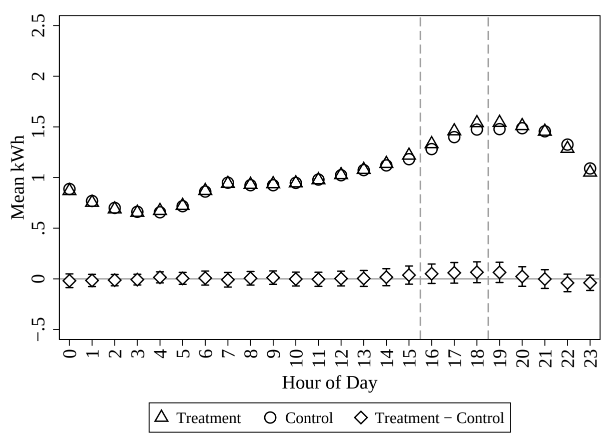
Notes: Standard errors are reported in parentheses.



(1) CPR-only

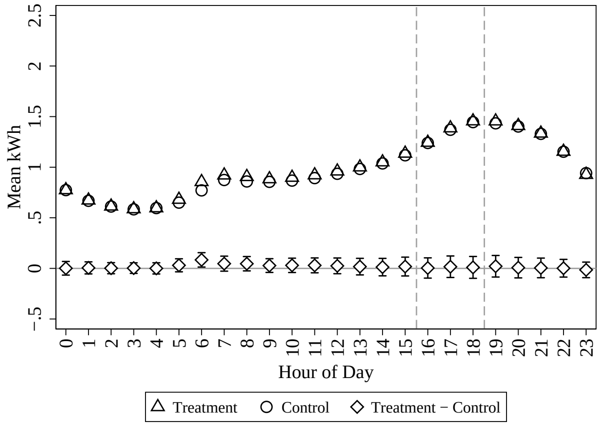


(2) CPR + TOU

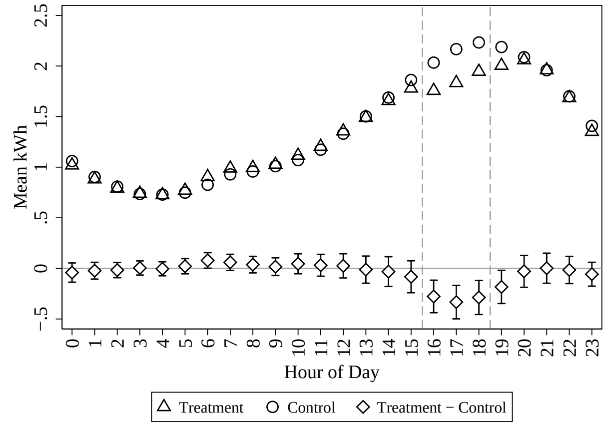


(3) TOU-only

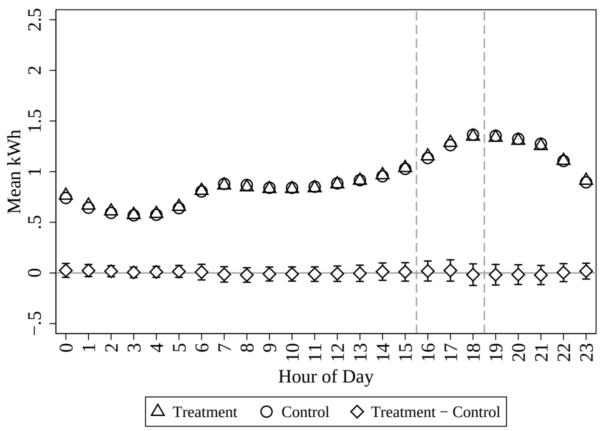
Figure 2: **Comparison of Pre-Program Means**. The vertical dashed lines denote 4pm-7pm, which is the hourly window when events are most frequently called.



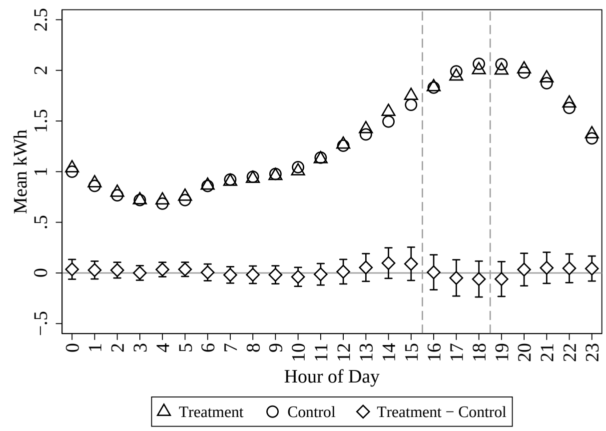
(1) CPR-only: Non-Event Day



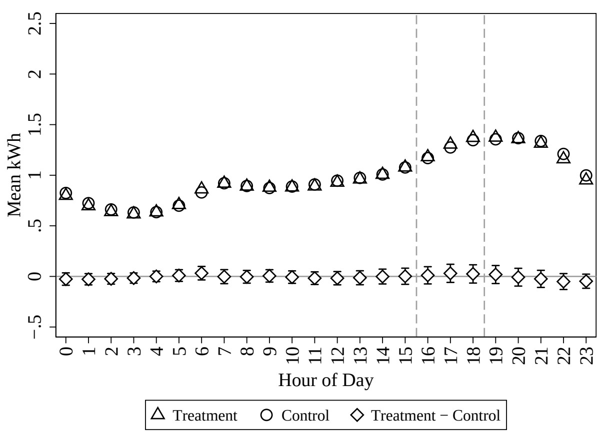
(2) CPR-only: Event Day



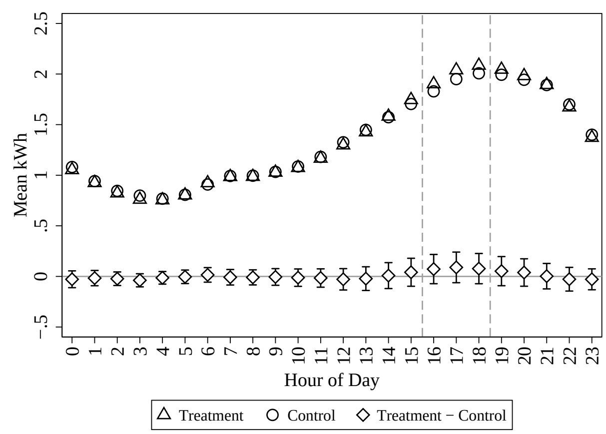
(3) CPR + TOU: Non-Event Day



(4) CPR + TOU: Event Day



(5) TOU-only: Non-Event Day



(6) TOU-only: Event Day

Figure 3: **Comparison of Treatment Period Means**. The vertical dashed lines denote 4pm-7pm, which is the hourly window when events are most frequently called.

Table 3: Treatment Effects During Event Hours

	(1)	(2)	(3)	(4)
CPR Only	-0.29***	-0.38***		
	(0.08)	(0.05)		
CPR + TOU	-0.04	-0.10**		
	(0.09)	(0.05)		
TOU Only	0.08	0.05		
	(0.07)	(0.04)		
CPR-1 Pricing			-0.05	-0.27***
			(0.17)	(0.10)
CPR-2 Pricing			-0.30**	-0.36***
			(0.13)	(0.07)
CPR-3 Pricing			-0.45***	-0.48***
			(0.14)	(0.09)
CPR-2 + TOU-1 Pricing			-0.09	-0.16
			(0.15)	(0.11)
CPR-2 + TOU-2 Pricing			-0.07	-0.13*
			(0.13)	(0.07)
CPR-2 + TOU-3 Pricing			0.06	-0.01
			(0.17)	(0.10)
TOU-1 Pricing			0.25	0.18*
			(0.18)	(0.10)
TOU-2 Pricing			-0.02	0.04
			(0.09)	(0.05)
TOU-3 Pricing			0.25	-0.05
			(0.16)	(0.09)
Control for Pre-Cons.	No	Yes	No	Yes
Control Mean	2.0	2.0	2.0	2.0
R-squared	0.03	0.44	0.03	0.44
Observations	117,027	117,027	117,027	117,027

Notes: The unit of analysis is a household-hour. The dependent variable is electricity consumption (kWh). All models are linear regression models with standard errors clustered by household. All models include hour-of-sample-by-experimental-group fixed effects. The sample is limited to hours when critical events were called. One, two, and three stars indicate 10 percent, 5 percent, and 1 percent significance, respectively.

Table 4: Non-Event Day Treatment Effects by Time of Day

	Off-Peak 10pm-11am	Mid-Peak 11am-3pm 8pm-10pm	On-Peak 3pm-8pm	Off-Peak 10pm-11am	Mid-Peak 11am-3pm 8pm-10pm	On-Peak 3pm-8pm
	(1)	(2)	(3)	(4)	(5)	(6)
CPR-1 Pricing	0.04	0.06	0.11	-0.01	-0.02	-0.00
	(0.05)	(0.08)	(0.10)	(0.03)	(0.04)	(0.05)
CPR-2 Pricing	0.01	0.01	-0.01	-0.03	-0.05	-0.06
	(0.05)	(0.07)	(0.09)	(0.03)	(0.03)	(0.04)
CPR-3 Pricing	0.00	-0.04	-0.09	-0.03	-0.04	-0.09**
	(0.05)	(0.06)	(0.08)	(0.02)	(0.04)	(0.05)
CPR-2 + TOU-1 Pricing	-0.03	-0.06	-0.02	-0.03	-0.05	-0.04
	(0.05)	(0.07)	(0.09)	(0.04)	(0.04)	(0.05)
CPR-2 + TOU-2 Pricing	0.01	0.01	0.01	0.02	0.00	-0.02
	(0.05)	(0.06)	(0.08)	(0.02)	(0.03)	(0.04)
CPR-2 + TOU-3 Pricing	0.04	0.01	0.02	0.05	0.01	-0.01
	(0.06)	(0.08)	(0.09)	(0.04)	(0.05)	(0.06)
TOU-1 Pricing	-0.02	0.05	0.10	0.02	0.07*	0.08
	(0.07)	(0.09)	(0.11)	(0.03)	(0.04)	(0.05)
TOU-2 Pricing	-0.04	-0.06	-0.05	-0.00	-0.01	-0.02
	(0.04)	(0.04)	(0.05)	(0.02)	(0.02)	(0.03)
TOU-3 Pricing	0.09*	0.10	0.16*	0.01	-0.05	-0.03
	(0.05)	(0.07)	(0.09)	(0.03)	(0.04)	(0.05)
Control for Pre-Cons.	No	No	No	Yes	Yes	Yes
Control Mean	0.8	1.1	1.3	0.8	1.1	1.3
R-squared	0.06	0.07	0.07	0.33	0.32	0.35
Observations	5,407,589	2,496,016	2,077,484	5,407,589	2,496,016	2,077,484

Notes: The unit of analysis is a household-hour. The dependent variable is electricity consumption (kWh). All models are linear regression models with standard errors clustered by household. All models include hour-of-sample-by-experimental-group fixed effects. All models are based on the hours listed in the column headings on days when critical events were not called. One, two, and three stars indicate 10 percent, 5 percent, and 1 percent significance, respectively.

Table 5: Regression of Correct Response on Hybrid Pricing Indicator

	Identify TOU Rate (1)	TOU Shift & Save (2)	Earn Rebates (3)
Hybrid Indicator	-0.079	-0.179***	-0.110*
	(0.068)	(0.059)	(0.057)
Constant	0.660***	0.845***	0.833***
	(0.048)	(0.037)	(0.037)
R-squared	0.007	0.043	0.017
Obs.	202	202	207

Notes: The unit of analysis is a household. The dependent variable is a binary variable equaling one if the survey respondent correctly answered the corresponding question. For column 1, the dependent variable is based on a question asking respondents to correctly identify their TOU-based rate schedule. For column 2, the dependent variable is based on a question asking respondents whether they would save money by moving consumption from on-peak to off-peak periods. For column 3, the dependent variable is based on a question asking respondents whether they could save money by conserving energy during critical events. Column 1 and 2 include households on TOU-only pricing or hybrid pricing. Column 3 includes households on CPR-only pricing or hybrid pricing. All models are linear regression models with robust standard errors. One, two, and three stars indicate 10 percent, 5 percent, and 1 percent significance, respectively.

SM Supplementary Material

1. Additional Figures (Summer)

2. Treatment on Treated IV Analysis (Summer)

3. Covariate Balance

4. Winter Discussion

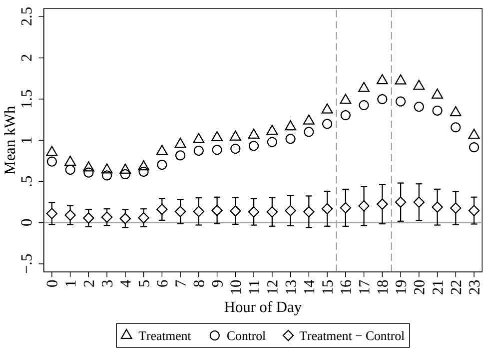
5. Primary Tables and Figures for Winter Analysis

6. Additional Figures (Winter)

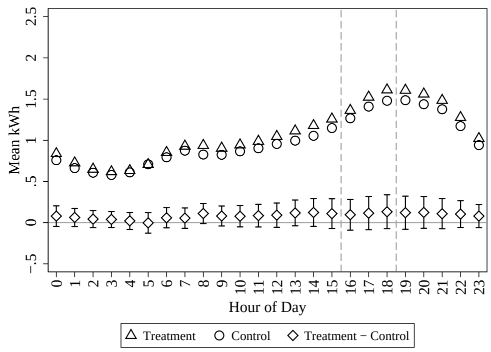
7. Treatment on Treated IV Analysis (Winter)

8. Event Effects - Robustness (Winter and Summer)

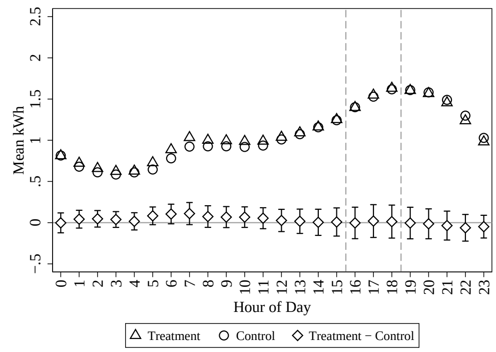
SM.1 Additional Figures (Summer)



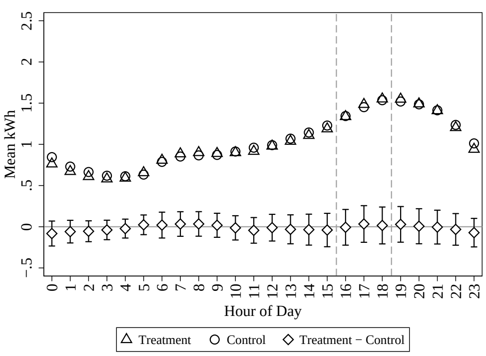
(1) CPR-1



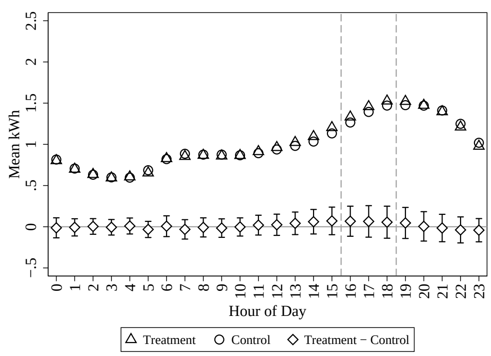
(2) CPR-2



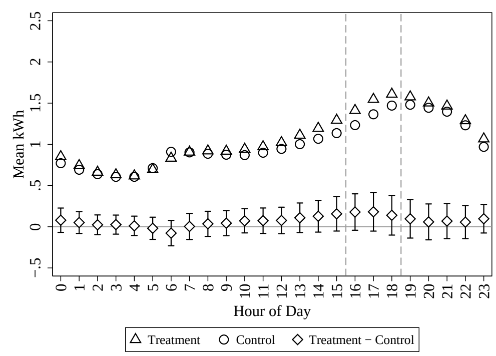
(3) CPR-3



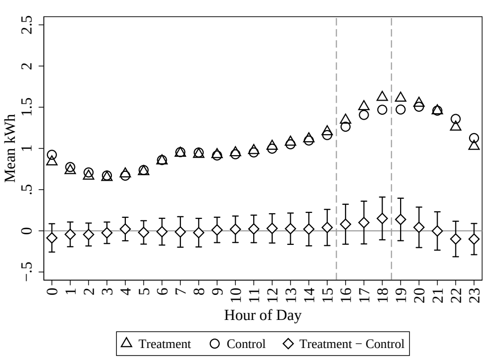
(4) CPR-2 + TOU-1



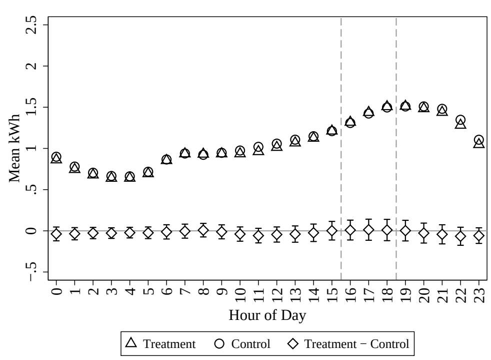
(5) CPR-2 + TOU-2



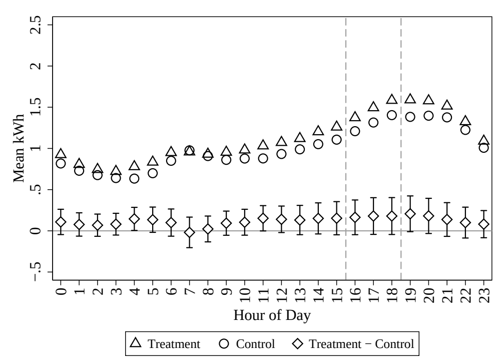
(6) CPR-2 + TOU-3



(7) TOU-1

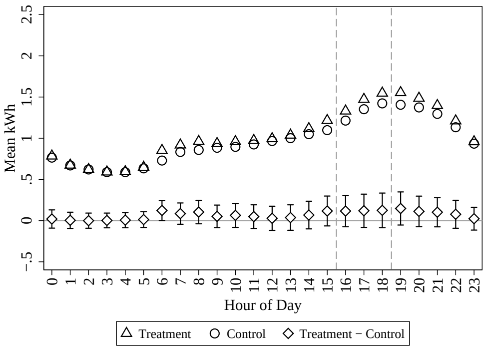


(8) TOU-2

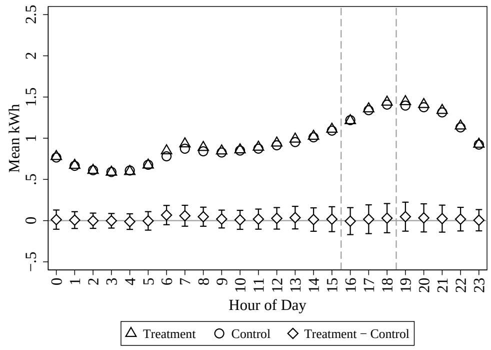


(9) TOU-3

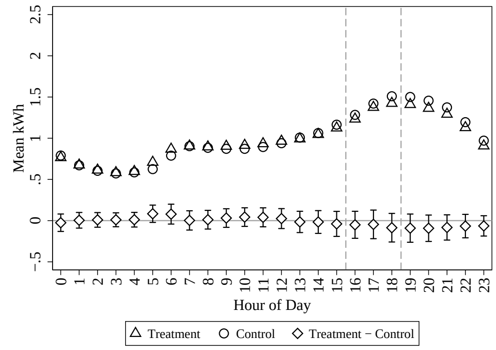
Figure SM.A: **Comparison of Summer Pre-Program Means: Graphs for Each Experiment Group**. The vertical dashed lines denote 4pm-7pm, which is the hourly window when events are most frequently called.



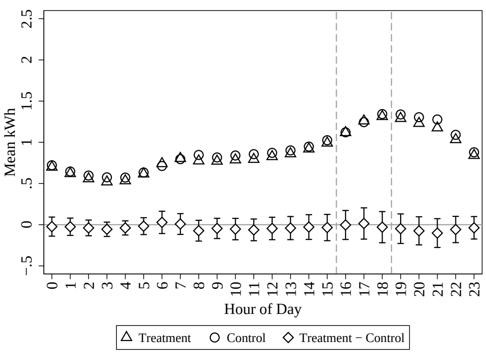
(1) CPR-1



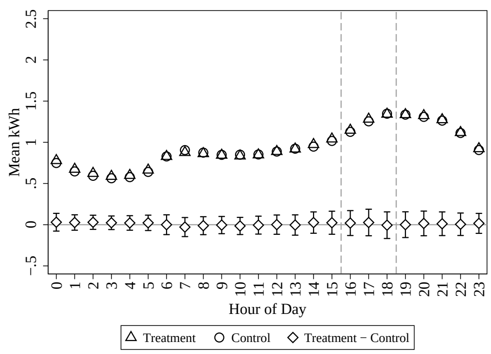
(2) CPR-2



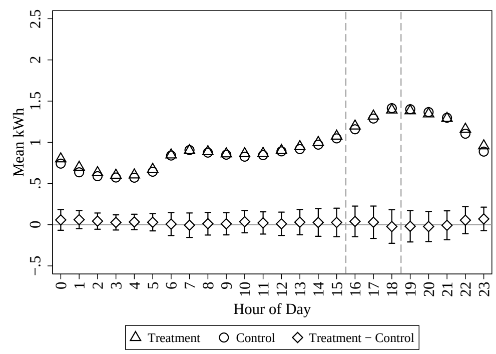
(3) CPR-3



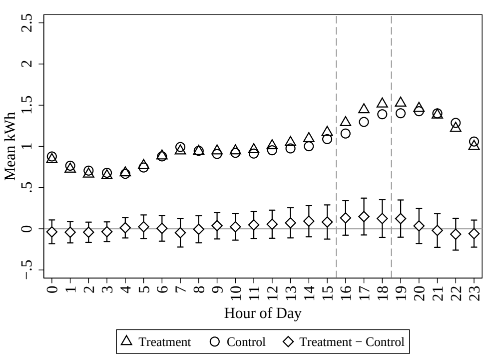
(4) CPR-2 + TOU-1



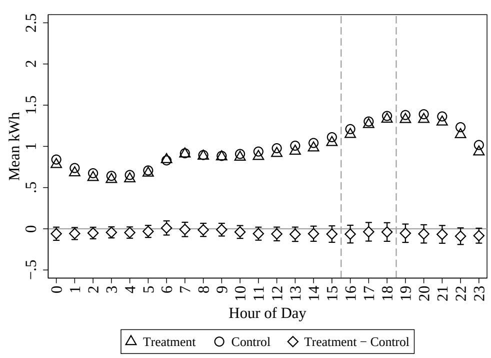
(5) CPR-2 + TOU-2



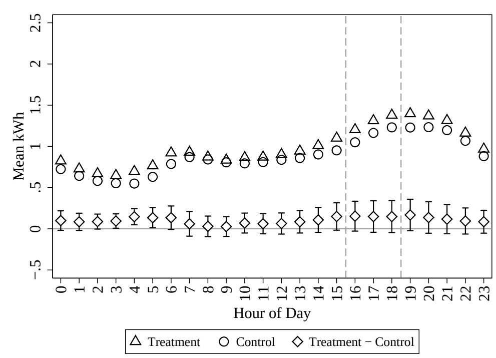
(6) CPR-2 + TOU-3



(7) TOU-1

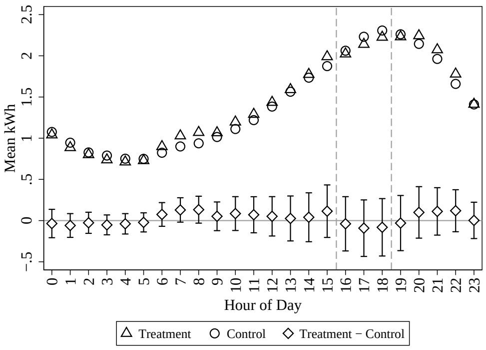


(8) TOU-2

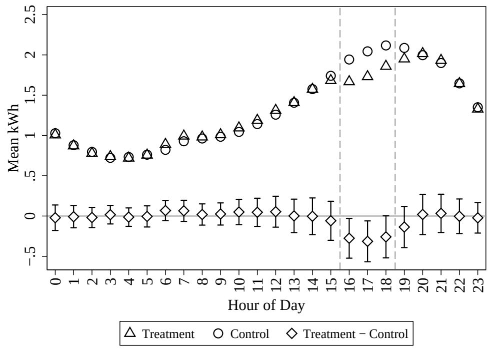


(9) TOU-3

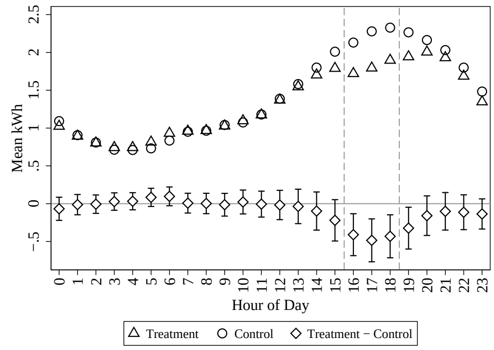
Figure SM.B: **Comparison of Treatment Period Means on <u>Non-Event</u> Days: Graphs for Each Experiment Group**. The vertical dashed lines denote 4pm-7pm, which is the hourly window when events are most frequently called.



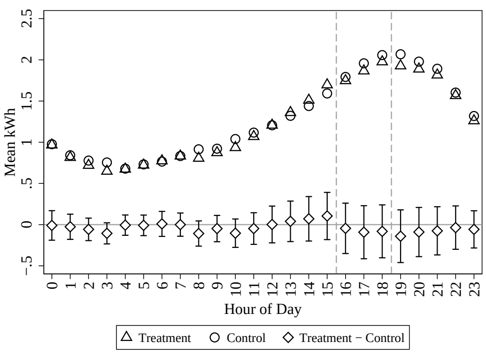
(1) CPR-1



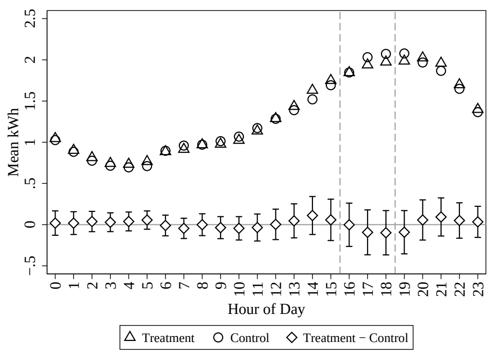
(2) CPR-2



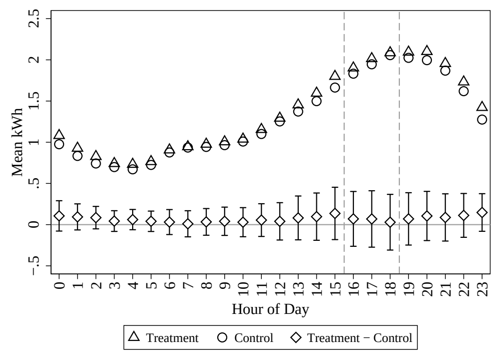
(3) CPR-3



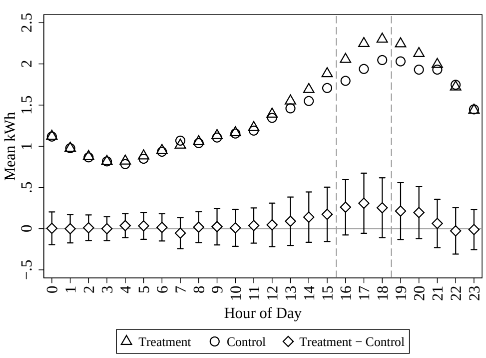
(4) CPR-2 + TOU-1



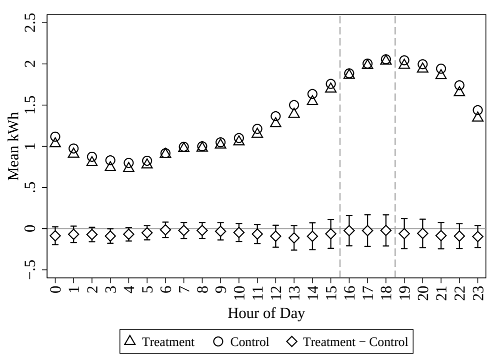
(5) CPR-2 + TOU-2



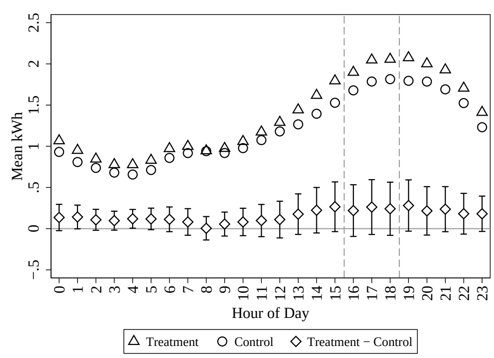
(6) CPR-2 + TOU-3



(7) TOU-1



(8) TOU-2



(9) TOU-3

Figure SM.C: **Comparison of Treatment Period Means on Event Days: Graphs for Each Experiment Group**. The vertical dashed lines denote 4pm-7pm, which is the hourly window when events are most frequently called.

SM.2 Treatment on Treated IV Analysis (Summer)

Table SM.A: Effects During Event Hours - Treatment on Treated IV Analysis

	(1)	(2)	(3)	(4)
CPR Only (Active Enrollment)	-0.30***	-0.39***		
	(0.08)	(0.05)		
CPR + TOU (Active Enrollment)	-0.04	-0.11*		
	(0.09)	(0.05)		
TOU Only (Active Enrollment)	0.08	0.05		
	(0.08)	(0.04)		
CPR-1 Pricing (Active Enrollment)			-0.05	-0.27**
			(0.17)	(0.10)
CPR-2 Pricing (Active Enrollment)			-0.30**	-0.37***
			(0.13)	(0.07)
CPR-3 Pricing (Active Enrollment)			-0.46***	-0.49***
			(0.14)	(0.09)
CPR-2 + TOU-1 Pricing (Active Enrollment)			-0.09	-0.16
			(0.16)	(0.12)
CPR-2 + TOU-2 Pricing (Active Enrollment)			-0.08	-0.13*
			(0.14)	(0.08)
CPR-2 + TOU-3 Pricing (Active Enrollment)			0.06	-0.01
			(0.17)	(0.10)
TOU-1 Pricing (Active Enrollment)			0.25	0.19*
			(0.18)	(0.10)
TOU-2 Pricing (Active Enrollment)			-0.02	0.04
			(0.10)	(0.05)
TOU-3 Pricing (Active Enrollment)			0.25	-0.05
			(0.16)	(0.09)
Control for Pre-Cons.	No	Yes	No	Yes
Control Mean	2.0	2.0	2.0	2.0
R-squared	0.00	0.43	0.00	0.43
Observations	117,027	117,027	117,027	117,027

Notes: The unit of analysis is a household-hour. The dependent variable is electricity consumption (kWh). All models are instrumental variables models, where current enrollment in the program is instrumented for based on whether the household was ever enrolled in the program. Standard errors are clustered by household. All models include hour-of-sample-by-experimental-group fixed effects. The sample is limited to hours when critical events were called. One, two, and three stars indicate 10 percent, 5 percent, and 1 percent significance, respectively.

SM.3 Covariate Balance

Table SM.B: Investigating Covariate Balance

	Age	Income	Household Size	Age of Home	Multifamily Unit	Renter	Renewable	Auto-Pay	Paperles	Other EE Prog.
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
CPR-1 - Control	4.29**	3215.98	0.36**	1.04	-0.05	-0.08	-0.01	0.02	-0.15***	0.11*
	(1.81)	(5313.77)	(0.17)	(1.66)	(0.05)	(0.05)	(0.05)	(0.05)	(0.06)	(0.06)
CPR-1 - Treatment	2.74	3105.28	0.11	0.34	-0.12***	-0.03	0.05	0.02	-0.21***	0.10*
	(1.77)	(5211.03)	(0.18)	(1.76)	(0.04)	(0.06)	(0.06)	(0.06)	(0.06)	(0.06)
CPR-2 - Treatment	-0.87	-1795.00	0.04	-2.00	-0.05	0.03	0.02	0.02	0.04	0.05
	(1.54)	(4538.35)	(0.16)	(1.55)	(0.04)	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)
CPR-3 - Control	4.90***	2327.02	0.22	-0.11	-0.02	-0.04	-0.03	-0.01	-0.17***	0.06
	(1.63)	(4684.78)	(0.16)	(1.51)	(0.04)	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)
CPR-3 - Treatment	2.97*	-4678.83	0.20	0.55	-0.10**	-0.04	-0.02	0.01	-0.18***	-0.00
	(1.64)	(4626.73)	(0.16)	(1.53)	(0.04)	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)
TOU-1 + CPR-2 - Control	6.72***	-354.26	0.05	2.97*	-0.09*	-0.02	-0.10*	0.08	-0.09	0.03
	(1.88)	(5614.50)	(0.18)	(1.71)	(0.05)	(0.06)	(0.05)	(0.06)	(0.06)	(0.06)
TOU-1 + CPR-2 - Treatment	5.51***	-3518.87	0.20	1.43	-0.14***	-0.08	-0.02	0.01	-0.17***	0.10*
	(1.78)	(5308.47)	(0.18)	(1.73)	(0.04)	(0.05)	(0.05)	(0.06)	(0.06)	(0.06)
TOU-2 + CPR-2 - Control	2.22	823.00	-0.06	-0.63	-0.00	-0.03	0.05	0.07	0.07	0.09*
	(1.63)	(4869.30)	(0.16)	(1.60)	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)
TOU-2 + CPR-2 - Treatment	1.96	-4749.31	0.01	-0.04	-0.04	0.00	0.01	-0.00	-0.01	0.05
	(1.60)	(4685.27)	(0.15)	(1.54)	(0.04)	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)
TOU-3 + CPR-2 - Control	3.60**	-1276.46	0.26	2.79*	-0.07	-0.06	0.02	0.00	-0.08	0.05
	(1.79)	(5171.78)	(0.19)	(1.68)	(0.05)	(0.05)	(0.05)	(0.05)	(0.06)	(0.06)
TOU-3 + CPR-2 - Treatment	7.23***	-2690.49	-0.00	1.04	-0.07	0.05	-0.07	0.10*	-0.08	0.02
	(1.77)	(4981.40)	(0.17)	(1.70)	(0.05)	(0.06)	(0.05)	(0.06)	(0.06)	(0.06)
TOU-1 - Control	2.22	888.43	0.20	2.25	-0.01	-0.03	0.05	0.04	-0.10*	0.01
	(1.88)	(5632.68)	(0.20)	(1.74)	(0.05)	(0.06)	(0.06)	(0.06)	(0.06)	(0.06)
TOU-1 - Treatment	3.73**	-7987.16	0.06	3.92**	-0.11***	-0.06	-0.04	0.03	-0.15***	0.07
	(1.79)	(5332.58)	(0.17)	(1.61)	(0.04)	(0.05)	(0.05)	(0.06)	(0.06)	(0.06)
TOU-2 - Control	3.94***	2735.73	0.08	0.51	-0.03	-0.05	-0.02	0.00	-0.15***	0.06
	(1.40)	(4116.89)	(0.14)	(1.35)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)
TOU-2 - Treatment	5.78***	-2308.56	0.09	1.01	-0.07**	-0.06	-0.03	0.01	-0.16***	0.02
	(1.36)	(3978.18)	(0.13)	(1.31)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)
TOU-3 - Control	3.15*	6053.21	0.08	2.38	-0.06	-0.06	-0.04	-0.02	-0.23***	-0.01
	(1.88)	(5283.27)	(0.18)	(1.69)	(0.05)	(0.06)	(0.05)	(0.06)	(0.06)	(0.06)
TOU-3 - Treatment	4.78***	-3151.90	0.09	3.09*	-0.07	-0.04	-0.04	0.02	-0.18***	0.03
	(1.79)	(5348.49)	(0.18)	(1.65)	(0.05)	(0.05)	(0.05)	(0.06)	(0.06)	(0.06)
Constant	50.58***	81661.07***	2.59***	36.00***	0.24***	0.37***	0.34***	0.36***	0.63***	0.44***
	(1.14)	(3368.74)	(0.12)	(1.14)	(0.03)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)
Model p-value	0.00	0.53	0.68	0.02	0.01	0.53	0.34	0.89	0.00	0.67
R-squared	0.018	0.006	0.004	0.009	0.009	0.005	0.005	0.003	0.027	0.004
Observations	2,939	2,834	3,263	3,379	3,350	3,369	3,379	3,379	3,379	3,379

Notes: The unit of analysis is a household. Dependent variables are indicated in the column headings. “Multifamily Unit” is a binary variable where one indicates residing in a multifamily unit. “Renter” is a binary variable where one indicates living in a rental unit. “Renewable” is a binary variable where one indicates participating in a green power program. “Auto-Pay” is a binary variable where one indicates enrollment in auto-pay billing. “Paperless” is a binary variable where one corresponds to enrollment in paperless billing. “Other EE Prog.” is a binary variable where one indicates enrollment in another program related to energy efficiency offered by the utility. All models were estimated using least squares. Heteroskedasticity-robust standard errors are reported in parentheses. One, two, and three stars indicate 10 percent, 5 percent, and 1 percent significance, respectively.

SM.4 Winter Analysis

In this section, we effectively replicate the summer analysis (June-September) for the winter months (December-February). As mentioned in footnote 9, the results from the winter analysis are less pronounced and more sensitive to modeling assumptions than are the results from the summer analysis. That said, the winter results also provide evidence that event-based pricing is more effective when offered in isolation than it is when paired with TOU pricing. We discuss some reasons why the results may be less pronounced during winter at the end of this section.

The time-variant pricing programs operated slightly differently during winter months than they did during summer months. First, for both TOU-2 and TOU-3 prices, winter prices included two peak time windows: 7am-11am and 3pm-8pm (the off, mid, and on-peak time periods were otherwise identical to summer). Additionally, unlike the summer events that were always called in the evening, the utility sometimes called critical events in the morning hours during the winter. Across the study period, the utility called fifteen winter events. Four were in the morning, typically between 7am and 10am. The remainder were in the afternoon, typically starting at 4pm or 5pm and lasting for three hours.

We replicate the main graphical analysis for the winter months in Figure SM.D and SM.E. Figure SM.D presents pre-period means and generally indicates statistically insignificant differences between the treatment and control group in the pre-period, although the CPR-only treatment group did experience some increases relative to the control group in the early morning hours. Consistent with the summer analysis, the left panels of Figure SM.E shows few significant differences on non-event days after the new prices were initiated, which suggests that TOU-prices were ineffective at changing consumption. The right panels, which present trends on event days, do not show nearly as dramatic visual evidence of event effects as was seen in the summer analysis.

To more precisely identify event effects, we again turn to regression models. Table SM.C displays estimates from event days. There is some evidence that both CPR-only and CPR + TOU pricing led to event savings.[37] As with the summer analysis, the CPR-only coefficient is larger than the CPR + TOU coefficient, but the difference is minor and not statistically

[37]Estimates showing evidence of event savings may seem surprising given the insignificant differences during event hours observed in the graphical analysis. However, the graphs also show that the treatment groups for CPR and CPR + TOU tended to have relatively elevated consumption during peak hours during the pre-period yet have relatively depressed consumption during peak hours on event days during the post-period.

significant. Both coefficients are less than half the magnitude of the CPR-only coefficient from the summer analysis.

Non-event day effects are presented for the winter months in Table SM.D. As with the summer analysis, most of the coefficients are insignificant. The main coefficients that are significant, are those for CPR-2 + TOU-1, but those coefficients indicate reductions across all periods, as opposed to effects primarily during peak time periods, as might be expected if the effect was driven by TOU pricing. It is possible that the negative coefficients on CPR-2 + TOU-1 are driven by a general decline in consumption for the CPR-2 + TOU-1 treatment group that occurred through random chance and was not caused by the shift to time-variant pricing.

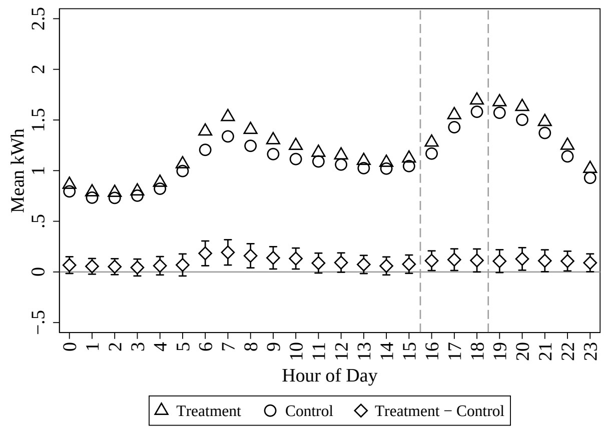
Generally, the summer-time estimates provide much stronger evidence of event treatment effects. This can be seen graphically as, during the summer, the CPR treatment group experienced a dramatic decline in consumption during the exact hours when events were called and not elsewhere (see Figure 3). Within the context of a regression framework, one way to assess whether the treatments created noticeable intra-day changes in energy consumption, as might be expected if households are truly responding to critical events, is to control for consumption on event days that occurred outside of event hours.[38] We apply this control to both the summer and winter models and report the results in Table SM.F. For the summer event estimates, adding the control produces nearly identical results. This can be seen, for example, by comparing column 1, which does not control for consumption during non-event hours on event days, and column 2, which does control for it. The winter estimates, however, are dramatically affected. Focusing on columns 5 and 6, the estimated CPR-only effect falls by nearly half, to -.09, and the CPR + TOU coefficient falls by even more, to -.04. Also, as with the summer analysis, the CPR + TOU coefficient is now half the size as the CPR-only coefficient, which is consistent with layering the programs dampening consumer responsiveness.

Collectively, the analysis from the winter months presents much weaker evidence that the pricing interventions had a substantial influence on consumption patterns. One possible explanation for this is that customers may have less tolerance for being cold than hot, which would reduce their willingness to modify their thermostats during winter months.

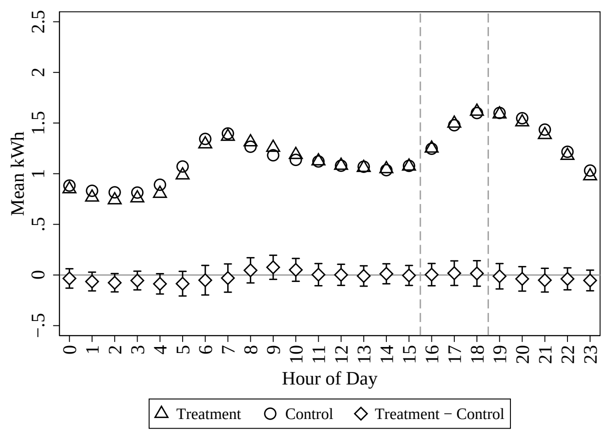
[38]This is not the optimal way to assess event effects, because events can create load-shifting to non-event hours (which makes non-event-hour consumption a "bad" control), but it is helpful for assessing the credibility of the winter event estimates.

Secondly, many homes in the study area use gas heating, whereas cooling must occur by using electricity. Households that have gas heating cannot achieve savings through thermostat adjustments and therefore may be less sensitive to price fluctuations.

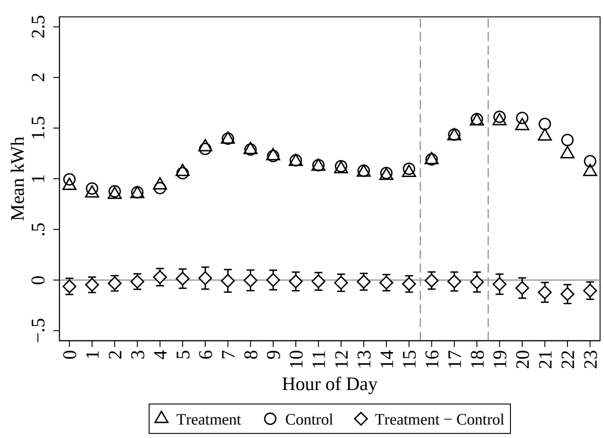
SM.5 Primary Tables and Figures for Winter Analysis



(1) CPR-only

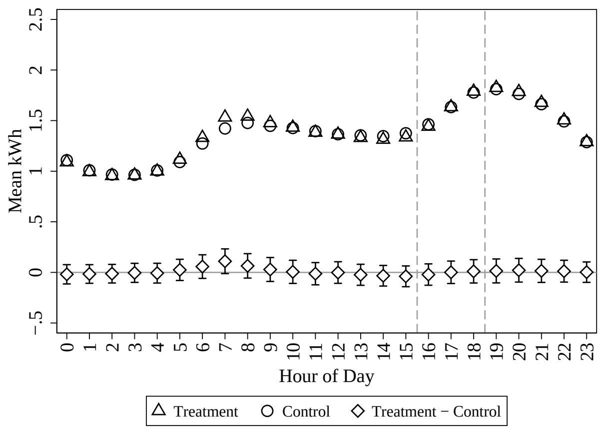


(2) CPR + TOU

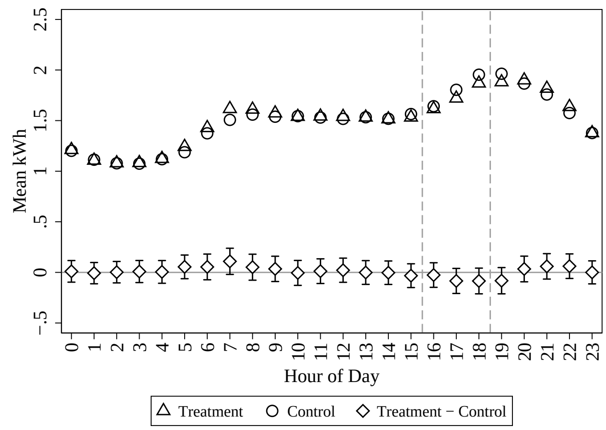


(3) TOU-only

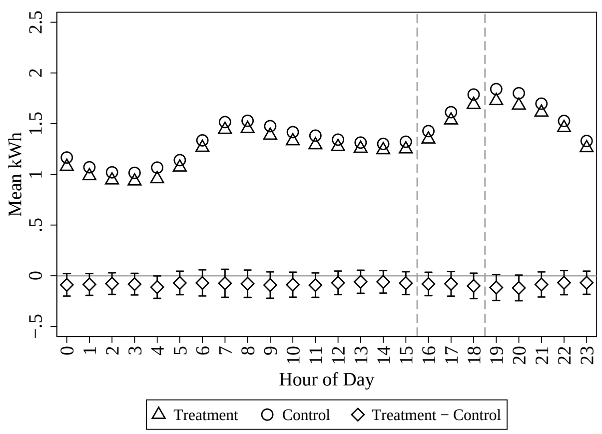
Figure SM.D: **Winter: Comparison of Pre-Program Means.** The vertical dashed lines denote 4pm-7pm, which is the hourly window when events are most frequently called.



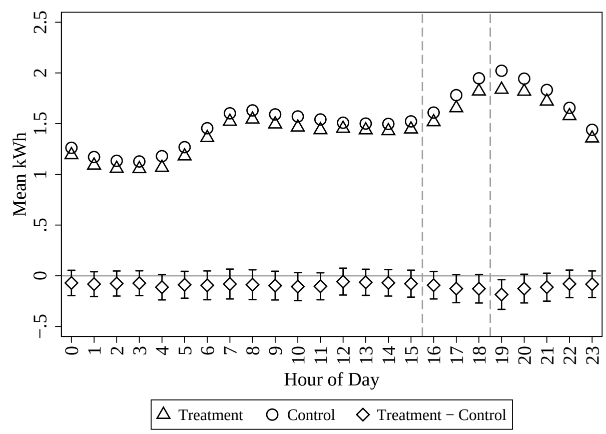
(1) CPR-only: Non-Event Day



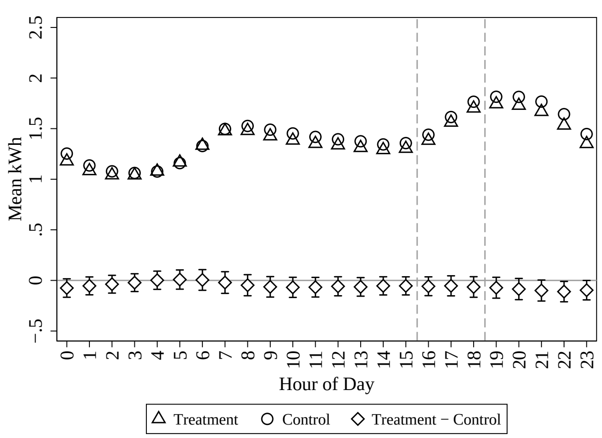
(2) CPR-only: Event Day



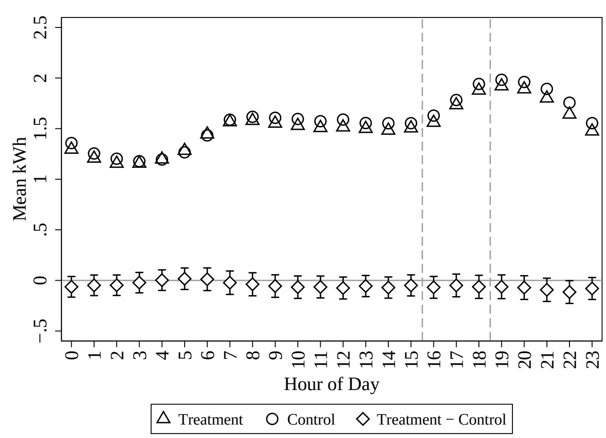
(3) CPR + TOU: Non-Event Day



(4) CPR + TOU: Event Day



(5) TOU-only: Non-Event Day



(6) TOU-only: Event Day

Figure SM.E: **Winter: Comparison of Treatment Period Means**. The vertical dashed lines denote 4pm-7pm, which is the hourly window when events are most frequently called.

Table SM.C: Winter: Treatment Effects During Event Hours

	(1)	(2)	(3)	(4)
CPR Only	-0.07	-0.17***		
	(0.06)	(0.04)		
CPR + TOU	-0.14**	-0.14***		
	(0.07)	(0.05)		
TOU Only	-0.06	-0.03		
	(0.05)	(0.04)		
CPR-1 Pricing			-0.05	-0.14*
			(0.11)	(0.08)
CPR-2 Pricing			0.02	-0.10
			(0.11)	(0.07)
CPR-3 Pricing			-0.18*	-0.26***
			(0.09)	(0.07)
CPR-2 + TOU-1 Pricing			-0.17	-0.24***
			(0.13)	(0.09)
CPR-2 + TOU-2 Pricing			-0.18*	-0.14**
			(0.10)	(0.07)
CPR-2 + TOU-3 Pricing			-0.04	-0.04
			(0.12)	(0.09)
TOU-1 Pricing			-0.06	-0.02
			(0.13)	(0.09)
TOU-2 Pricing			-0.12*	-0.04
			(0.07)	(0.05)
TOU-3 Pricing			0.18	0.03
			(0.13)	(0.08)
Control for Pre-Cons.	No	Yes	No	Yes
Control Mean	1.8	1.8	1.8	1.8
R-squared	0.03	0.29	0.03	0.29
Observations	140,270	140,270	140,270	140,270

Notes: The unit of analysis is a household-hour. The dependent variable is electricity consumption (kWh). All models are linear regression models with standard errors clustered by household. All models include hour-of-sample-by-experimental-group fixed effects. The sample is limited to hours when critical events were called. One, two, and three stars indicate 10 percent, 5 percent, and 1 percent significance, respectively.

Table SM.D: Winter: Non-Event Day Treatment Effects by Time of Day

	Off-Peak 10pm-11am	Mid-Peak 11am-3pm 8pm-10pm	On-Peak 3pm-8pm	Off-Peak 10pm-11am	Mid-Peak 11am-3pm 8pm-10pm	On-Peak 3pm-8pm
	(1)	(2)	(3)	(4)	(5)	(6)
CPR-1 Pricing	0.02	0.00	0.02	-0.03	-0.06	-0.05
	(0.09)	(0.10)	(0.10)	(0.06)	(0.07)	(0.06)
CPR-2 Pricing	0.02	0.01	0.06	-0.05	-0.09	-0.06
	(0.08)	(0.09)	(0.09)	(0.06)	(0.06)	(0.06)
CPR-3 Pricing	-0.03	-0.03	-0.02	-0.08*	-0.06	-0.09
	(0.08)	(0.08)	(0.09)	(0.04)	(0.05)	(0.06)
CPR-2 + TOU-1 Pricing	-0.12	-0.13	-0.14	-0.10*	-0.13*	-0.19***
	(0.11)	(0.11)	(0.12)	(0.06)	(0.07)	(0.07)
CPR-2 + TOU-2 Pricing	-0.10	-0.08	-0.12	0.01	-0.04	-0.08
	(0.08)	(0.09)	(0.09)	(0.05)	(0.06)	(0.06)
CPR-2 + TOU-3 Pricing	-0.00	-0.03	0.01	0.01	-0.04	-0.01
	(0.10)	(0.10)	(0.11)	(0.08)	(0.07)	(0.07)
TOU-1 Pricing	-0.04	-0.04	-0.02	0.02	0.01	0.01
	(0.10)	(0.11)	(0.11)	(0.07)	(0.07)	(0.07)
TOU-2 Pricing	-0.10*	-0.13**	-0.12*	-0.02	-0.03	-0.05
	(0.06)	(0.06)	(0.06)	(0.04)	(0.04)	(0.04)
TOU-3 Pricing	0.14	0.10	0.12	0.03	-0.05	-0.01
	(0.11)	(0.11)	(0.11)	(0.06)	(0.06)	(0.06)
Control for Pre-Cons.	No	No	No	Yes	Yes	Yes
Control Mean	1.2	1.5	1.5	1.2	1.5	1.5
R-squared	0.05	0.04	0.03	0.34	0.29	0.30
Observations	2,997,275	2,000,981	2,996,015	2,996,175	2,000,228	2,995,441

Notes: The unit of analysis is a household-hour. The dependent variable is electricity consumption (kWh). All models are linear regression models with standard errors clustered by household. All models include hour-of-sample-by-experimental-group fixed effects. All models are based on the hours listed in the column headings on days when critical events were not called. One, two, and three stars indicate 10 percent, 5 percent, and 1 percent significance, respectively.

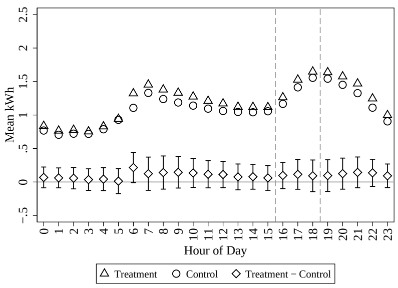
SM.6 Treatment on Treated IV Analysis (Summer)

Table SM.E: Winter: Effects During Event Hours - Treatment on Treated IV Analysis

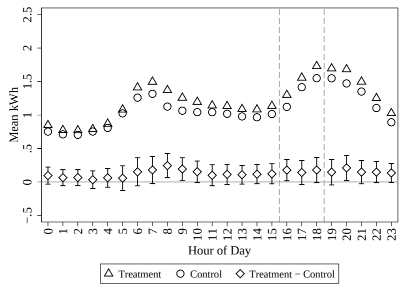
	(1)	(2)	(3)	(4)
CPR Only (Active Enrollment)	-0.07	-0.17***		
	(0.06)	(0.04)		
CPR + TOU (Active Enrollment)	-0.14*	-0.14***		
	(0.07)	(0.05)		
TOU Only (Active Enrollment)	-0.06	-0.03		
	(0.06)	(0.04)		
CPR-1 Pricing (Active Enrollment)			-0.05	-0.14*
			(0.11)	(0.08)
CPR-2 Pricing (Active Enrollment)			0.03	-0.10
			(0.11)	(0.07)
CPR-3 Pricing (Active Enrollment)			-0.19*	-0.27***
			(0.10)	(0.07)
CPR-2 + TOU-1 Pricing (Active Enrollment)			-0.18	-0.25**
			(0.14)	(0.10)
CPR-2 + TOU-2 Pricing (Active Enrollment)			-0.19*	-0.15*
			(0.11)	(0.07)
CPR-2 + TOU-3 Pricing (Active Enrollment)			-0.04	-0.04
			(0.13)	(0.09)
TOU-1 Pricing (Active Enrollment)			-0.06	-0.02
			(0.13)	(0.09)
TOU-2 Pricing (Active Enrollment)			-0.12*	-0.04
			(0.07)	(0.05)
TOU-3 Pricing (Active Enrollment)			0.18	0.03
			(0.13)	(0.08)
Control for Pre-Cons.	No	Yes	No	Yes
Control Mean	1.8	1.8	1.8	1.8
R-squared	0.00	0.27	0.00	0.27
Observations	140,270	140,270	140,270	140,270

Notes: The unit of analysis is a household-hour. The dependent variable is electricity consumption (kWh). All models are instrumental variables models, where current enrollment in the program is instrumented for based on whether the household was ever enrolled in the program. Standard errors are clustered by household. All models include hour-of-sample-by-experimental-group fixed effects. The sample is limited to hours when critical events were called. One, two, and three stars indicate 10 percent, 5 percent, and 1 percent significance, respectively.

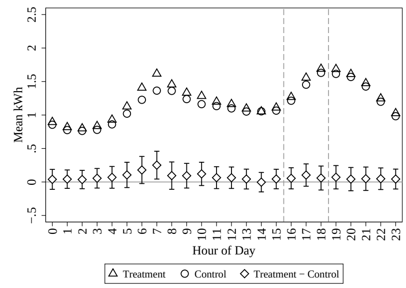
SM.7 Additional Figures (Winter)



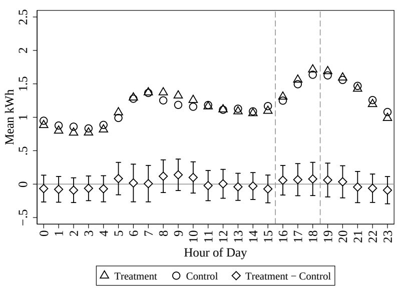
(1) CPR-1



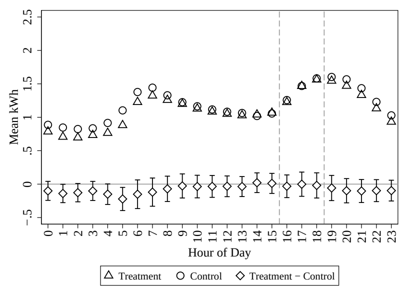
(2) CPR-2



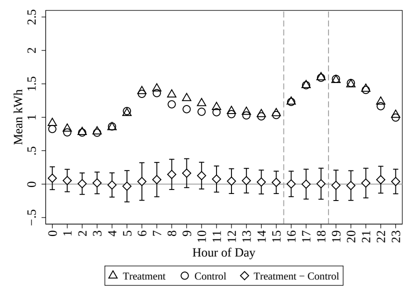
(3) CPR-3



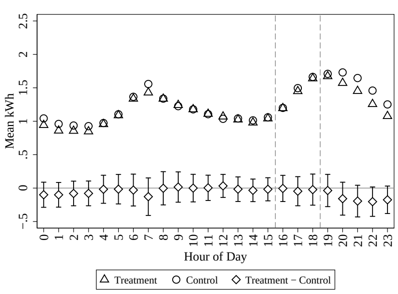
(4) CPR-2 + TOU-1



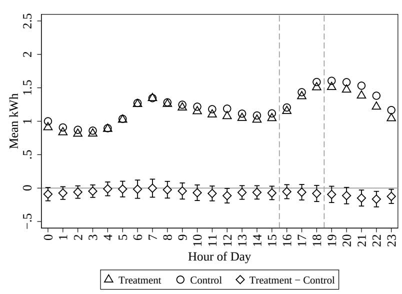
(5) CPR-2 + TOU-2



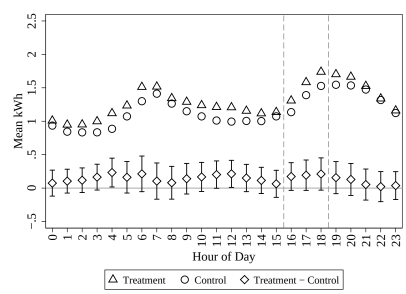
(6) CPR-2 + TOU-3



(7) TOU-1

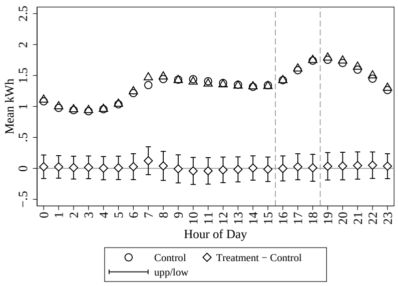


(8) TOU-2

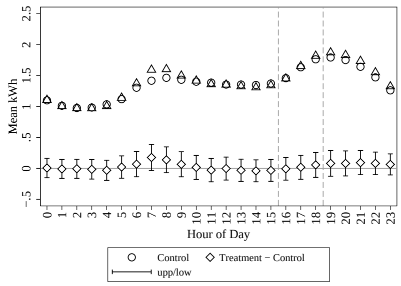


(9) TOU-3

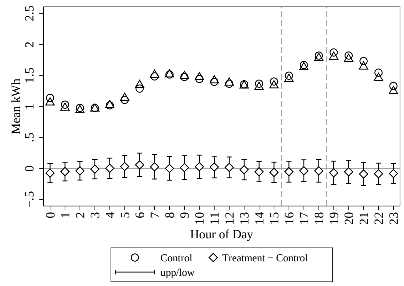
Figure SM.F: **Winter: Comparison of <u>Pre-Program</u> Means**. The vertical dashed lines denote 4pm-7pm, which is the hourly window when events are most frequently called.



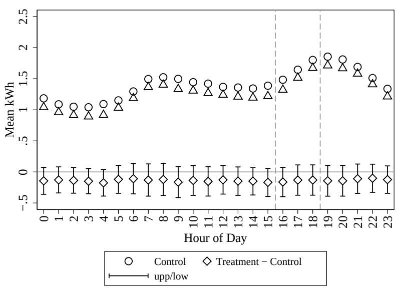
(1) CPR-1



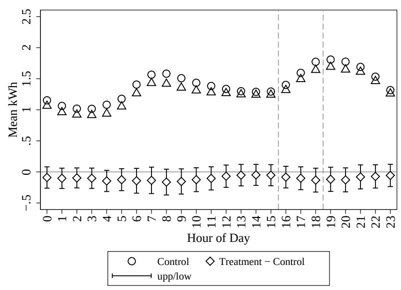
(2) CPR-2



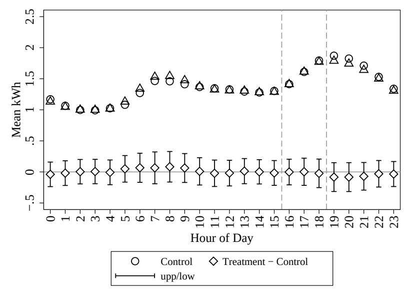
(3) CPR-3



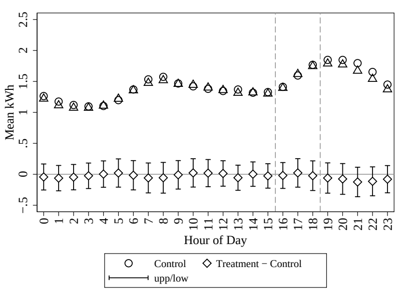
(4) CPR-2 + TOU-1



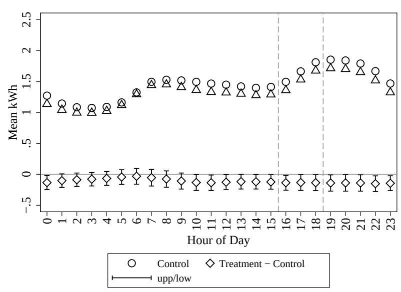
(5) CPR-2 + TOU-2



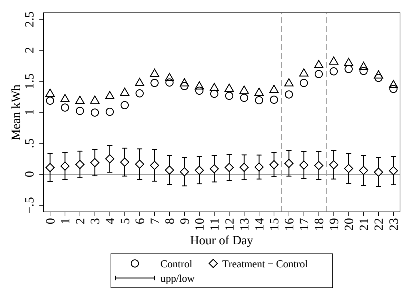
(6) CPR-2 + TOU-3



(7) TOU-1

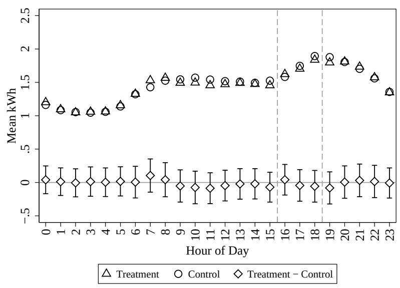


(8) TOU-2

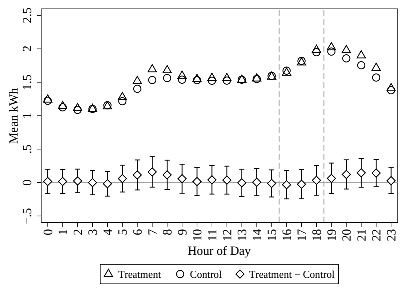


(9) TOU-3

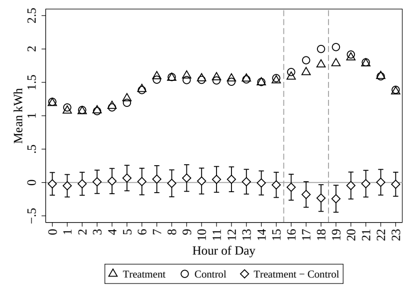
Figure SM.G: **Winter: Comparison of Treatment Period Means on <u>Non-Event</u> Days**. The vertical dashed lines denote 4pm-7pm, which is the hourly window when events are most frequently called.



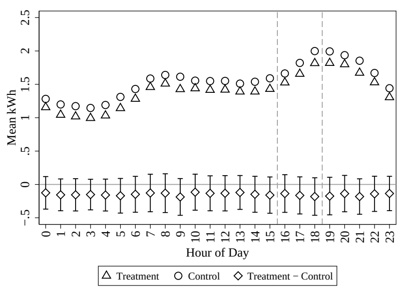
(1) CPR-1



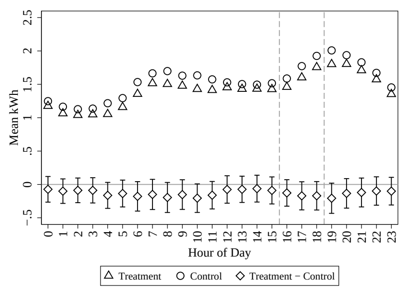
(2) CPR-2



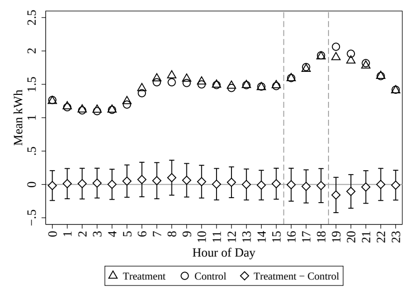
(3) CPR-3



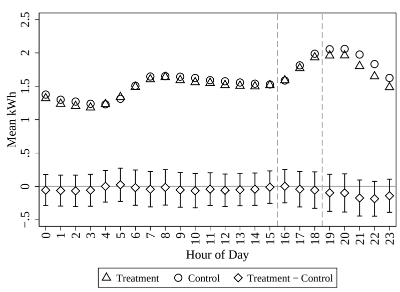
(4) CPR-2 + TOU-1



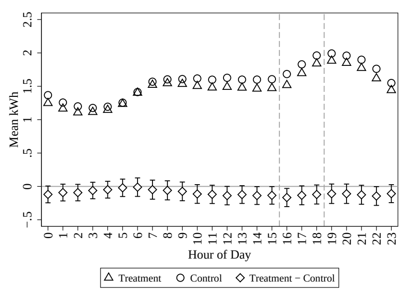
(5) CPR-2 + TOU-2



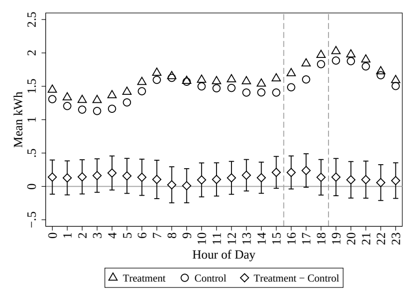
(6) CPR-2 + TOU-3



(7) TOU-1



(8) TOU-2



(9) TOU-3

Figure SM.H: **Winter: Comparison of Treatment Period Means on Event Days**. The vertical dashed lines denote 4pm-7pm, which is the hourly window when events are most frequently called.

SM.8 Event Effects -Add Non-Event Day Consumption Controls (Winter and Summer)

Table SM.F: Event-Day Estimates Based on Models Including Non-Event Day Consumption Controls

	Summer				Winter			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
CPR Only	-0.38***	-0.33***			-0.17***	-0.09***		
	(0.05)	(0.04)			(0.04)	(0.02)		
CPR + TOU	-0.10**	-0.08**			-0.14***	-0.04**		
	(0.05)	(0.04)			(0.05)	(0.02)		
TOU Only	0.05	0.05*			-0.03	0.01		
	(0.04)	(0.03)			(0.04)	(0.01)		
CPR-1 Pricing			-0.27***	-0.27***			-0.14*	-0.08**
			(0.10)	(0.08)			(0.08)	(0.03)
CPR-2 Pricing			-0.36***	-0.33***			-0.10	-0.04
			(0.07)	(0.06)			(0.07)	(0.03)
CPR-3 Pricing			-0.48***	-0.38***			-0.26***	-0.15***
			(0.09)	(0.07)			(0.07)	(0.03)
CPR-2 + TOU-1 Pricing			-0.16	-0.12			-0.24***	-0.04
			(0.11)	(0.08)			(0.09)	(0.03)
CPR-2 + TOU-2 Pricing			-0.13*	-0.11**			-0.14**	-0.05*
			(0.07)	(0.05)			(0.07)	(0.03)
CPR-2 + TOU-3 Pricing			-0.01	0.01			-0.04	-0.03
			(0.10)	(0.07)			(0.09)	(0.03)
TOU-1 Pricing			0.18*	0.08			-0.02	-0.03
			(0.10)	(0.08)			(0.09)	(0.03)
TOU-2 Pricing			0.04	0.06			-0.04	0.01
			(0.05)	(0.04)			(0.05)	(0.02)
TOU-3 Pricing			-0.05	-0.00			0.03	0.04
			(0.09)	(0.07)			(0.08)	(0.03)
Control for Pre-Cons.	No	No	No	No	Yes	Yes	Yes	Yes
Control for Non-Event Cons.	No	Yes	No	Yes	No	Yes	No	Yes
Control Mean	2.0	2.0	2.0	2.0	1.8	1.8	1.8	1.8
R-squared	0.44	0.54	0.44	0.54	0.29	0.53	0.29	0.53
Observations	117,027	117,027	117,027	117,027	140,270	140,270	140,270	140,270

Notes: The unit of analysis is a household-hour. The dependent variable is electricity consumption (kWh). All models are linear regression models with standard errors clustered by household. All models include hour-of-sample-by-experimental-group fixed effects. One, two, and three stars indicate 10 percent, 5 percent, and 1 percent significance, respectively.