Impact Evaluation of a Peak Time Rebate Program with Universal Enrollment – Estimation With and Without Control Groups

Dr. Steven D. Braithwait, CA Energy Consulting, Madison, WI

Marlies Hilbrink, CA Energy Consulting, Madison, WI

Abstract

This paper provides an update to a previous impact evaluation of one of the first peak time rebate (PTR) programs in the United States in which all residential customers were enrolled by default and eligible to receive bill credits. The primary objective of the original evaluation was to estimate hourly PTR event-day load impacts at the program level and for various subsets of customers of interest. The evaluation approach involved designing samples of two large components of the San Diego Gas and Electric (SDG&E) residential population and conducting customer-level regression analysis of hourly load data for each sampled customer. The primary finding from that study was that, on average, only customers who opted to receive electronic notifications, or alerts, of PTR events reduced their electricity usage during PTR event-hours by statistically significant amounts. The update reported in this study uses the finding of no response by non-notified customers to select a matched control group and apply a traditional treatment and control group evaluation approach to produce updated estimates of program impacts.

Introduction

This paper summarizes the results of an update to a previous impact evaluation of one of the first peak time rebate (PTR) programs in the United States in which all residential customers were enrolled by default and eligible to receive bill credits. In 2012, SDG&E enrolled nearly all of its residential customers (those who had received Smart Meters) in PTR. Day-ahead announcements of PTR events were provided through public media (*e.g.*, through radio and TV news, and weather features). Customers were also encouraged to sign up to receive electronic notification, or alerts, of events through email or text messages (or both). About four percent of customers opted to receive notification.

The objective of this update was to examine the sensitivity of the estimated load impacts in the original evaluation to the method of estimation. In particular, the original evaluation[1] applied participant-only customer-specific regression analysis to billing-based hourly load data to estimate *ex-post* load impacts for both opt-in alert and non-alert customers.[2] One of the key findings was that only opt-in alert customers on average reduced usage during PTR event windows by statistically significant amounts. The analysis reported in this study leveraged on the finding of the lack of response of non-notified customers to select a matched control group sample of the non-notified customers, and to re-estimate load impacts of opt-in alert customers using a treatment and control group method. The resulting estimates are compared to the participant-only results.

PTR Program Features

SDG&E's PTR program includes the following features:

---

[1] See S. Braithwait and M. Hilbrink, "Impact Evaluation of a Peak Time Rebate Program with Universal Enrollment," IEPEC Chicago, 2013.

[2] The original study also involved a group of customers located in the city of San Diego who enrolled in the San Diego Energy Challenge (SDEC), a separate effort within PTR that involved a competition among middle schools in the San Diego Unified School District. Results for those customers are not reported in this paper, though they were included in the follow-on study.

2014 International Energy Policy & Programme Evaluation Conference, Berlin

- Two rebate levels are available—a basic level of $0.75/kWh and a premium level of $1.25/kWh for customers who use automated enabling technology installed through a SDG&E program. For 2012, only those customers who were enrolled in SDG&E's Summer Saver air conditioner direct load control program were eligible for the premium level.
- There is no maximum number of events, though rebate levels were designed assuming nine events each year. Seven events were called in 2012. The event window is 11 am to 6 pm.
- Load reductions for rebate purposes are measured relative to a customer-specific reference level (CRL) based on an average of the highest three out of the most recent five similar non-event days.[3]
- Customers who register for My Account online have access to information on their consumption history, CRL, event performance, and online rebate calculation, as well as online bill paying.

Customer Characteristics

Table 1 provides an overview of the customer base that was enrolled in PTR, and indicates the number of customers in certain subgroups that were used in the previous analysis. As indicated in the table, participants enrolled in SDG&E's Summer Saver (SS) program were excluded from that evaluation, but were analyzed in an evaluation of the SS program. Relatively large samples of the opt-in alert customers and the remaining population (after excluding all of the other subgroups) were selected and analyzed.[4] The two rows in bold are the subject of this study update.[5]

Table 1: PTR Subgroup Populations and Sample Sizes

PTR Subgroup	Population	Analysis Samples
Summer Saver (excluded)	23,998	-
SDEC (excluding SS)	4,633	4,631
Opt-in Alert	**41,243**	**13,745**
Remaining Population	**1,154,144**	**29,692**
Total (Excluding SS)	1,200,020	48,068

Previous Analysis Approach and Findings

Our original *ex-post* load impact evaluation applied customer-level regression analysis to hourly load data for samples of participating customers, using observations on customers' own loads on non-event days to estimate what their load would have been on event days had the events not been called.[6] This "participant-only" evaluation approach is commonly used for event-based demand response programs, in which relatively few events are called during a season, and a number of non-event days are available for estimating counterfactual loads on event days.[7] The regression analysis

[3] The "highest" days are those with the highest total consumption between the event window hours of 11 am to 6 pm. For events called on weekend or holiday days, the CRL is total consumption during the above hours on the highest of the three preceding weekend days.

[4] The samples were stratified by climate zone (Coastal and Inland) and customer size (Small, Medium, and Large).

[5] The number of customers who opted to receive event notification in this first year of default PTR may be considered relatively small. However, the percentage of customers requesting notification expanded considerably in 2013.

[6] In the terminology of California demand response evaluations, *ex post* load impacts represent the measured load impacts in the historical period. The evaluations typically also forecast *ex ante* load impacts based on *ex post* results and enrollment forecasts.

[7] Some evaluations of pilot programs, such as our evaluation of the SDG&E PTR pilot for 2011, compare PTR participant usage to that of a control group in order to estimate load impacts. However, this approach was ruled out in the

2014 International Energy Policy & Programme Evaluation Conference, Berlin

controls for factors other than PTR events that influence customers' load profiles, including hour of day, day of week, and weather conditions, and also includes hourly variables that indicate event days. The coefficients on the event-day variables allow direct estimation of hourly PTR load impacts for each event day.

An important overall finding from the original study was that only customers who opted to receive electronic notifications, or alerts, of PTR events reduced their electricity usage during PTR event hours by statistically significant amounts. Specifically, opt-in alert customers in the Coastal and Inland regions were estimated to have *reduced* usage on average by small but statistically significant amounts of 0.064 to 0.067 kWh per hour, or 6 and 4 percent of their respective reference loads.[8] However, the estimated load impacts varied considerably across events, were only statistically significant for some events, and did not always appear consistent with observed load data on event days and comparable non-event days.[9] In contrast, estimates of PTR load impacts for customers who did *not* request notification were *positive* on average, suggesting small usage *increases* during PTR event-hours.[10] However, the estimates were not statistically significant.

Analysis Approach in Current Study

For this update to the original evaluation, SDG&E was interested in learning whether estimation using a treatment and control group approach would produce substantially different results from the original study based on the participant-only analysis. It was still the case that no true non-participant population was available from which to select a classical non-treated control group. However, the original evaluation's finding that the average non-alert customer did not reduce usage by a statistically significant amount on PTR event days suggested using that population as the source for a control group. The ready availability of a sample of those customers from the original evaluation meant that no additional data collection was required.

The objective in selecting a control group was to match each customer in the opt-in alert (treatment) group sample to the most similar customer in the full sample of potential control group customers. The only available demographic information on the customers was their ZIP code location and prior-year usage data. Thus, matching was conducted using these two factors.[11] The process involved constructing usage metrics for each opt-in alert customer and potential control group customer, and then comparing them to determine the best matches. Specifically, for each opt-in alert customer, we constructed normalized indexes that compared average usage measures for each of the 24 hours and the average across hours for the pre-treatment months of August and September of 2011 to each customer in the non-alert population sample within the same ZIP code. For each opt-in alert customer, we selected the customer for which the average normalized difference was smallest

original evaluation by the universal nature of the program and public announcements of events.

[8] These relatively small usage reductions are due largely to the fact that average customer loads in the mild-climate SDG&E area are relatively small compared to those in other regions. For example, event-day reference loads for the Coastal and Inland areas averaged only 1.06 kW and 1.6 kW respectively during event-window afternoon hours. However, the *percentage* reductions are not that much smaller than those found in other PTR programs.

[9] Close examination of customer-level results in the original study indicated that approximately 25 to 35 percent of the opt-in alert customers, differentiated by climate zone and size, reduced usage by consistent and statistically significant amounts that were on the order of five to six times the magnitude of those for the average opt-in alert customer.

[10] The estimated load impacts are based on the values of regression coefficients on variables representing PTR event hours. Negative values represent *reductions* in usage during the event window relative to a reference load that is adjusted for event-day weather conditions and other factors included in the regression. However, estimates for the average non-alert customer were positive, indicating that their usage during the event window was slightly *greater* than the implied reference load, although the increase was not statistically significantly different from zero.

[11] A number of control group selections in recent pilot evaluations have used a "propensity score matching" procedure, in which control group customers are matched on the basis of a number of demographic characteristics. The process that we used is analogous to that approach, except that with no demographic characteristics other than location, we focused on the usage profiles on non-event days, since the objective of the control group was to provide suitable reference loads for comparison with participant loads on actual event days.

2014 International Energy Policy & Programme Evaluation Conference, Berlin

from among the available control group customers. The final samples included approximately 13,000 opt-in alert customers and 10,000 matched control group customers.[12] Comparisons of average non-event day load profiles for the opt-in alert and matched control group samples, by size category, indicated close correspondence between usage patterns for the two groups.

Once the matched control group was selected, the analysis approach involved constructing *daily* observations on average hourly usage within the event-window hours (hours-ending 12 – 18) for each treatment and control group customer, pooling the observations within a climate zone and usage size category, and estimating a *fixed-effects* regression model. These models allow customer-specific constant terms, but estimate the same coefficients for variables that do not vary across customers (*e.g.*, occurrence of PTR event days). These essentially represent an average effect across customers. The approach represents a classic *difference-in-differences* evaluation approach, in that systematic differences in average usage between treatment and control group customers on *non-event* days are used to adjust the estimated differences in average *event-day* usage.

Study Findings

Estimated Load Impacts for Average Event

Table 2 reports estimated average hourly PTR load impacts for the average event for the opt-in alert customers, by climate zone and overall, and for three alternative analysis approaches.[13] The lower portion of the table presents estimated load impacts as percentages of the indicated reference loads. The first row in the table shows estimated average event-day kW load impacts for the selected fixed-effects treatment and control group regression models.[14] The estimates are statistically significant for both climate zones. To enable comparison to the estimated load impacts that were reported in the original study, the third line in the table reproduces estimated average hourly load impacts based on the individual customer-level regressions. That analysis differs in several aspects from the analysis approach used here. Specifically, separate equations were estimated for each opt-in alert customer, hourly observations were used rather than daily, and load impacts were estimated for each event hour, rather than for the average event-hour.

Table 2: Estimated Per-Customer PTR Usage Reductions, by Analysis Type

Model	Type	Coastal	Inland	All
Fixed Effects (kW)	Treatment & Control	0.073	0.108	0.088
	Treatment only	0.049	0.067	0.057
Customer Level (kW)	Treatment only	0.062	0.067	0.064
	Reference load (kW)	1.06	1.58	1.29
% Load impacts	Treatment/Control (FE)	6.8%	6.8%	6.8%
	Treatment only (FE)	4.7%	4.2%	4.4%
	Treatment only (CL)	5.9%	4.2%	5.0%

Comparing the third row to the first row indicates that the estimated load impacts from the two methods are reasonably similar for the coastal region, but differ by a fairly large amount for the inland region. The fixed-effects estimates of usage reductions for both climate zones and overall exceed those from the treatment-only customer-level regression approach. The overall estimated load impact using the treatment and control group approach, shown in the last column of the table, is 0.088 kW per

[12] The matching was conducted with replacement, such that the same customer could be selected multiple times for inclusion in the matched control group. In those cases, the customers were weighted appropriately in the regression analysis and in the construction of the average customer loads shown in the figures below.
[13] Sample weights are applied to size category results to produce average load impacts by climate zone and overall.
[14] The estimated model coefficients are actually negative, reflecting lower event-window usage on event days. By convention, these values are converted to positive numbers in the table to represent load *reductions*.

2014 International Energy Policy & Programme Evaluation Conference, Berlin

customer, or 6.8 percent of the average customer reference load. These values are in contrast to the overall average of 0.064 kW, representing a 5 percent load reduction that were reported in the original evaluation.

To explore potential reasons for the differences across methods, we also applied the fixed effects estimation approach to data for the opt-in alert customers alone, without the control group customers. This approach differs from the customer-level regression approach in three ways: 1) daily rather than hourly observations are used, 2) a single average load impact coefficient is estimated rather than averaging across separately estimated load impacts, and 3) coefficients on variables other than event-day indicators are constrained to be the same within a customer group rather than varying by customer. Comparing these results in the second row to the average load impacts from the individual customer regression results in the third row indicates that the estimates are quite close, particularly for the Inland climate zone. Thus, it appears that including the control group is the source of the apparent differences between the treatment/control and treatment-only approaches. This, rather than the fixed-effects estimation approach itself, applied to daily observations.

Estimated Load Impacts by Event

We also estimated fixed-effects models in which load impacts are estimated separately for each event. These are designed to explore the effect of the treatment and control group approach on the magnitude and variability of estimated load impacts across events. Results are shown in Figure 1 for the two primary methods of estimation – fixed-effects treatment and control; and customer-level treatment-only, in which estimated load impacts for each event are averaged (using appropriate sample weights) across customers in the relevant group. Within each method of estimation and climate zone category, eight bars are shown, one for each of the seven 2012 PTR events and one for the average event, where the height of the bars represents estimated average event-hour load impacts as a percent of the reference load.

Looking across events for each grouping, it is apparent that the treatment and control group approach generally produces larger and more consistent percentage load impacts than does the treatment-only approach. This is particularly the case for the third, fourth, and fifth events for the Inland climate zone, for which non-significant load *increases* were estimated in the original analysis.[15] It should be noted that the fourth and fifth events were characterized by certain unusual features. The fourth event, on August 11, was called on a Saturday, following two consecutive event days. The fifth event, on August 14, was called on a day of moderate temperatures following several hot days. It is thus likely that the problematic load impact estimates for those events in the original study were caused by a lack of comparable non-event days to provide sufficient information to control for non-event factors on those event days. In contrast, including control group loads for the same days as the event days serves to better isolate and estimate PTR load impacts for the opt-in alert groups.

[15] These estimated load increases are based on the estimated regression coefficients, as described in a previous footnote, where the estimated coefficients were positive but not statistically significantly different from zero.

2014 International Energy Policy & Programme Evaluation Conference, Berlin

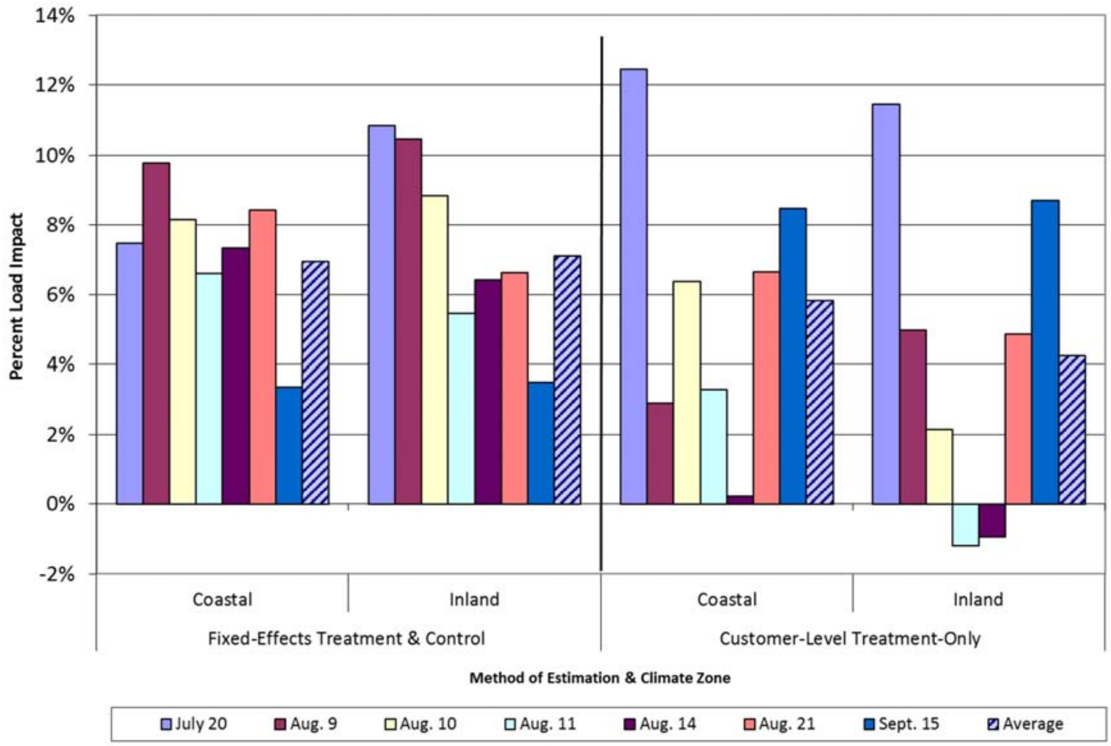

Figure 1: Estimated PTR Load Impacts by Event and Analysis Approach

An analogous issue exists with the last event, on September 15. This was also a Saturday, and was the hottest day of the summer by some margin. As a result, no comparable non-event day was available to inform the regression about usage on a comparable non-event day. In this case, the load impact estimates in the original study were *higher* than suggested by the observed load data, for both climate zones, at nearly 9 percent. In contrast, the fixed-effects treatment and control approach produces estimated load impacts of 3 to 3.5 percent.

To quantify the difference in variability of the estimated load impacts across events for the two methods, we can compare the standard deviations of the percentage load impacts relative to the average values. For the treatment and control group approach, the standard deviation is approximately *one-third* of the average 6.8 percent load impact. In contrast, the standard deviation for the treatment-only approach is nearly *90 percent* of the average 5 percent load impact. It is logical to expect that customers' response to the seven PTR events would be reasonably similar.

To illustrate the nature of the observed data that is used in the fixed-effects estimation, Figure 2 shows the average opt-in alert treatment group load profile for the August 14 event, for the Inland climate zone, along with the average control group load profile for that day, where the control group load has been adjusted for the average difference between the two load types on non-event days. Since it is difficult to find a comparable non-event day for the August 14 event to serve as a comparison day, due to the unique weather conditions described above, the control group load for that day serves a valuable role in suggesting the likely nature of the counterfactual load for that event. The average difference between the two load curves during the event hours (shown by the lower line) implies an event-hour load impact of 0.12 kW, or 7 percent, which is quite close to the fixed-effects estimate of 6.4 percent shown in Figure 1 above. Analogous load comparisons for the other events show comparable similarity of observed load differences and estimated load impacts from the treatment and control group fixed-effects approach.

2014 International Energy Policy & Programme Evaluation Conference, Berlin

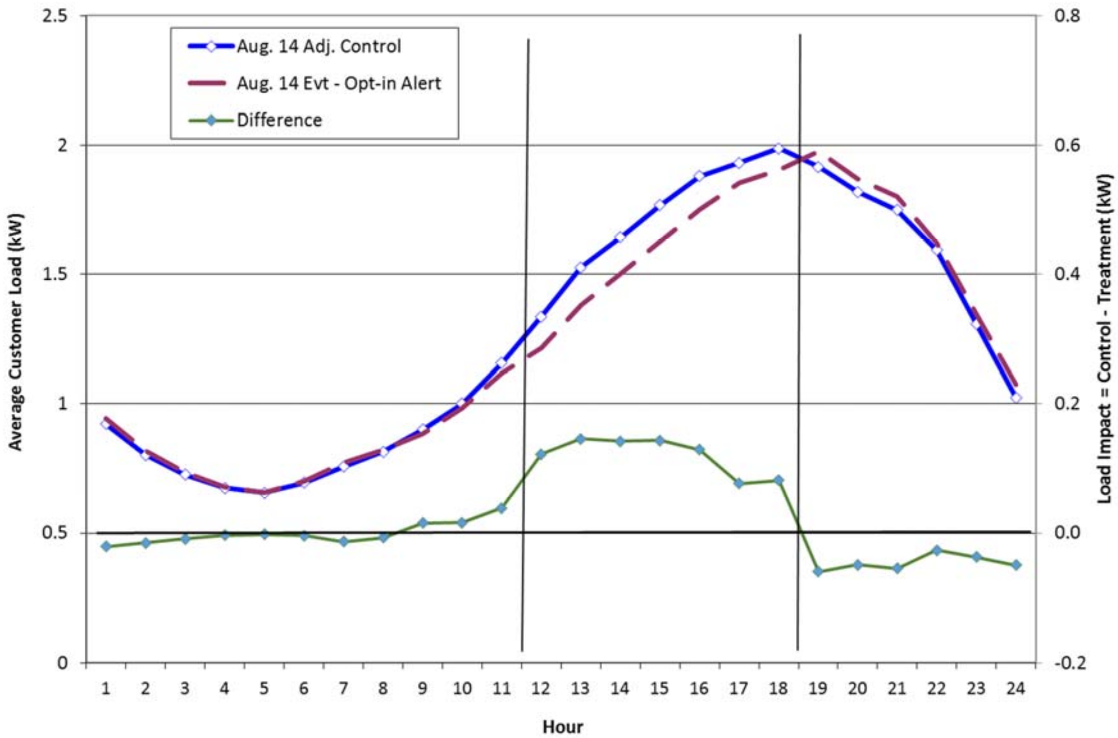

Figure 2: Average Adjusted Control Group and Opt-in Alert Treatment Group Loads for August 14 Event – *Inland Climate Zone*

Conclusions and Observations

The findings of this study demonstrate the potential value of incorporating control-group data into the estimation of treatment-group load impacts for a residential demand response program like PTR. Most importantly, the treatment and control group approach generally produced more consistent (*i.e.*, less variable) load impact estimates across events than did the treatment-only approach, particularly for the Inland climate zone, and particularly for events that were characterized by unusual weather conditions.[16] At this point, it is not possible to determine the extent to which this result is driven by unique weather conditions in the Inland area in 2012, or is a more general result. However, it is the case that demand response program events tend to be called on the most extreme days, which suggests that unavailability of comparable non-event days may be a common problem in attempting to estimate event-day load impacts using treatment customers' data alone. In fact, the evaluation of SDG&E's 2011 pilot PTR faced a similar challenge of events being called on all of the hottest days of the summer, leaving no event-like days to assist in estimating treatment customers' counterfactual reference load using only their load data. Fortunately, a control group had been selected as part of the pilot and could be used in estimating event-day load impacts.

Going forward, two countervailing factors suggest how evaluations of PTR programs may be conducted in the future. One the one hand, if PTR remains a default element of the standard residential tariff, then the availability of potential control group customers will remain an issue. However,

[16] As noted above, the selected control groups are somewhat tainted in that customers in those groups were exposed to the same PTR events as were the opt-in alert treatment customers, and thus may have reduced usage to some degree, even though on average they were found to have not done so. Thus it could be argued that load impacts measured relative to these control group loads may be *understated*. On the other hand, it appears that including the control group in the analysis helps to control for weather effects and potential unobserved event-day effects that might otherwise affect the estimated load impacts.

2014 International Energy Policy & Programme Evaluation Conference, Berlin

SDG&E has proposed to modify the program so that only those customers who opt to receive electronic notification of events are eligible to receive bill credits for load reductions. This proposed change has been driven by two primary factors. One is the finding that non-notified customers appear to not reduce usage on average. The second factor is that related research has found that the customer-specific baseline methods used to measure PTR usage reductions are relatively inaccurate due to the inherent variability of residential customer loads. This inaccuracy leads to substantial numbers of either *over-payments* for "phantom" load reductions, or *under-payments* to customers who tried to reduce usage but whose baseline loads understated those reductions.

Limiting eligibility for bill credits to only notified customers reduces the potential scale of over-payments, and also implies that a substantial number of non-notified customers will likely remain available for selection into a control group for purposes of evaluation.

References

[CAEC 2013] Steven D. Braithwait, Daniel G. Hansen and Marlies Hilbrink, Christensen Associates Energy Consulting, *2012 Evaluation of San Diego Gas & Electric's Peak Time Rebate Program*, April 1, 2013.

2014 International Energy Policy & Programme Evaluation Conference, Berlin