Is Behavioral Energy Efficiency and Demand Response Really Better Together?

David Thayer, Wendy Brummer, Brian Arthur Smith, Rick Aslin, Pacific Gas and Electric Company and Jonathan Cook, Nexant

ABSTRACT

After a number of successful experiments of Home Energy Reports (HER) at Pacific Gas and Electric Company (PG&E), a field study of Behavioral Demand Response (BDR) was launched to households in HER treatment and control conditions in summer 2015. The field study was limited to residential customers being served by electric substations that are capacity-constrained so that the effectiveness of BDR as a means to curtail peak demand could be tested. Unlike HER reports which present general energy efficiency (EE) messaging with a goal of reducing overall consumption, the BDR messaging seeks to reduce residential consumption on specific days. This paper presents results of a layered study whereby HER treatment and control customers were assigned to BDR treatment or control conditions so that the interactive effects of HER and BDR messaging can be studied.

The experiment found statistically significant savings for both energy efficiency (EE) and demand response (DR). The analysis found that the layered program delivers incremental off-peak EE from participants in households treated with both HER and BDR. Households treated with both HER and BDR messaging tended to deliver less demand savings than did customers receiving BDR treatment only, though both groups experienced peak demand savings over their associated BDR control group (1.8% vs 2.4% per event day on average).

Although the results were favorable, conducting a layered trial involving multiple stimuli has challenges. Careful consideration must be made when layering EE and DR messaging to avoid contaminating ongoing HER experiments or shifting control group baselines significantly.

Introduction

Home Energy Reports (HER) has proven to be a successful Energy Efficiency (EE) program at Pacific Gas and Electric Company (PG&E), as it has elsewhere. PG&E's HER initiative began its first major randomized control trial in August 2011. Subsequently, ten additional HER experiments have been launched. In conjunction with the ongoing in-field HER programs, PG&E conducted a Behavioral Demand Response (BDR) study during the summer of 2015 to assess the impact of BDR on peak electricity usage during four designated event days.

Unlike many traditional residential demand response interventions, BDR does not offer financial incentives to participants to reduce their usage, nor does it require the installation of technology at a customer's residence to directly control air conditioning load. Instead, BDR provides customers with pre/post-event communications and social comparisons specifically aimed at reducing usage on peak demand (event) days. The fundamental concepts of BDR are similar to those in the well-established HER program. BDR messaging delivers social comparisons through direct marketing to motivate energy conservation behaviors (Karlin, Zinger, and Ford 2015). A key difference between HER and BDR is that BDR has the objective of reducing customer load for a few hours (5-8 PM) on days when electricity demand is high.

The BDR study was implemented as a randomized control trial (RCT) using the treatment and control households of an HER experiment. Specifically, the existing treatment/control design of HER is layered with a treatment/control BDR experiment, resulting in a 2 X 2 experimental design (HER treatment or control, BDR treatment or control). This design is shown in Table 1 and has a total of four different customer groups, which are randomly assigned

- Customers receiving both HER and BDR treatments
- Customers receiving only BDR treatment
- Customers receiving only HER treatment, and
- Customers receiving neither treatment

The 2 X 2 factorial experimental design allows for BDR impacts to be estimated separately within each HER group (treatment and control) and then compared in order to assess possible interactive effects of BDR treatment on HER recipients and HER control customers.

Table 1. Design of the BDR Experiment

	HER Recipients	HER Control Customers	Total
BDR Treatment	30,200	9,800	**40,000**
BDR Control	26,400	8,500	**34,900**
Total	**56,600**	**18,300**	**74,900**

Behavioral Demand Response

Behavioral Demand Response (BDR) builds on the fundamental concept of using neighbor comparisons–as pioneered in Home Energy Reports (HER)–to encourage residential customers to reduce peak usage. Both BDR and HER reports were provided by Opower, Inc. BDR relies on pre/post-event communications and social comparisons to provide customers with information and motivation to reduce their electricity usage on days of high peak usage, referred to as "Summer Saving Days." Both BDR and HER are "opt-out" programs: customers are not asked to participate; rather those households that are randomly assigned to treatment conditions are simply enrolled to receive the messaging but are given an opportunity to remove themselves from treatment ("opt out"). Importantly, BDR does not offer any financial incentives for customers to reduce their usage, nor does it require the installation of technology at a customer's residence.

Whereas the HER program is designed to motivate energy conservation every day, BDR is designed to target efficiency behaviors for only a few hours during days when electricity demand is high. For this study, four event days were called to test the performance of BDR on different weekdays in July through September of 2015 between the hours of 5 and 8 PM.[1]

A few weeks before the first event day, BDR treatment households were introduced to the trial through a welcome letter[2] that described the concept of Summer Saving Days and explained how customers could participate. The PG&E-branded welcome letter was sent by surface mail to the address on file for each household.

Calling an event required coordination between the PG&E and Opower operations teams in order to deliver personalized customer communications via email or phone. During the summer, PG&E monitored weather and system conditions and identified Summer Saving Days on a day-ahead basis. Upon deciding that an event would occur on the following day, PG&E informed Opower of the event, which initiated pre-event communications with customers. The pre-event messaging consisted of either an email or an automated phone call to announce the event, provide each household with a customized peak event normative comparison of a customer's peak usage to their neighbors' usage for the most recent event, and suggest actionable behavioral tips for reducing peak usage during the event.

After the event, Opower imported hourly interval electric consumption data for all participating customers on the day of the event and developed personalized rankings for each participating household. These rankings, plus additional energy saving tips, were distributed to customers 24-48 hours after the event via another email or automated phone call.

Methodology

The load impact evaluation for the BDR study is based on a randomized control trial (RCT) design involving approximately 75,000 HER-eligible customers spread across 31 capacity-constrained substations in the PG&E territory. This section outlines the details of the RCT design and the methods used to estimate the load impacts of the program.

In an RCT design, load impacts are measured as the difference in average usage between BDR treatment and control customers on event days. To implement this approach, PG&E's evaluation firm Nexant built a linear regression model specified to include dummy variables indicating the experimental group to which the customer was assigned and the event date. The linear model was estimated using ordinary least squares (OLS) regression so that the average load of the treatment group was subtracted from the average load of the control group on each event day. The key assumption of the model is that the usage of the control group accurately predicts what usage would have been for treatment customers if they had not experienced the treatment. With large samples, this assumption is satisfied by virtue of the random assignment to the BDR treatment condition from the population of HER treatment and control conditions.

[1] The four event days were July 29, August 27, September 9 and September 11.
[2] Welcome letters were mailed on June 29.

To reduce the standard error of the impact estimates, date fixed effects were included to control for weather variations and other factors that change with time. Clustered standard errors at the customer level were used to account for likely serial correlation that exists between hourly observations for the same customer.

The regression analysis employed a simple model that relies on no explanatory variables other than the treatment and time effects. This model does not rely on modeling the relationship between customers' electricity usage and other factors such as weather; it is informed by control group customers that experience the event day weather, but do not experience the BDR treatment. Nexant did estimate load impacts using a model that included weather control variables, but the added control variables did not notably improve precision. In this case the simple model has sufficient explanatory power to justify its use over a less transparent model that provides only a marginally better fit to the data.

Separate specifications were used to estimate average impacts for event days and for each individual event. These specifications are shown as Equations 1 and 2 in Figure 1; the definitions of the variables and parameters are shown in Table 2. In all equations, the dependent variable was the average load (kW) during each hour and the regression errors were specified to be robust to the serial correlation.

Avg. Event Equation:

$$kW_{i,t} = a + b \cdot Treatment_i + u_t + \varepsilon_{i,t} \text{ for } i \in \{1, \dots, n_i\} \text{ and } t \in \{1, \dots, n_t\} \quad (1)$$

Individual Event Equation:

$$kW_{i,t} = a + \sum_{t=1}^{max} b_t \cdot (Treatment_i \cdot Date_t) + u_t + \varepsilon_{i,t} \text{ for } i \in \{1, \dots, n_i\} \text{ and } t \in \{1, \dots, n_t\} \quad (2)$$

Figure 1. Equations used to estimate impacts of the experiment

Table 2. Definitions of model variables and parameters

Variable	Definition
i, t	Indicate observations for each individual *i*, date *t* and event number *n*, where the number of events varies by utility and is denoted *max*
a	The model constant
b	The difference between BDR and control group customer during event days – this is the coefficient of interest
u	Time effects for each date that control for unobserved factors that are common to all treatment and control customers but unique to the time period
ε	The error for each individual customer and time period
Treatment	A binary indicator or whether or not the customer is part of the treatment (BDR) or control group
Date	A binary indicator of date

Results

In an RCT design with large sample sizes, load impacts can be estimated accurately by calculating the difference in average peak period (5-8 PM) usage between customers in the treatment and control groups on each event day. For both customers in HER treatment and HER control conditions, any differences between the BDR treatment and control groups during the pre-treatment period are nearly zero. The lack of differences prior to treatment is expected with random assignment into the treatment group and serves as a randomization check.

The results for energy reduction during the peak hours on event days are shown in Table 3. Customers that received BDR treatment saw incremental demand savings in both HER treatment and HER control groups. For HER control customers, percent impacts of demand savings resulting from exposure to BDR messaging range from 1.7% to 3.1% while absolute impacts range from 0.05 kW to 0.09 kW. For HER recipients, percent impacts of demand savings resulting from exposure to BDR messaging range from 1.2% to 2.2% while absolute impacts range from 0.03 kW to 0.07 kW. The larger number of HER treatment customers assigned to the BDR treatment group (N=30,200) compared to the number of HER control customers assigned to the BDR treatment group (N=9,800) results in more precise estimates for the treatment group, which is reflected by the tighter confidence intervals as shown in the table. These results support the use of BDR messaging to curtail residential peak demand in capacity-constrained substations.

Table 3. Peak demand reductions estimated for each BDR experimental condition

Category	Event Date	Control Load (kW)	Treatment Load (kW)	Impact (kW)	95% CI (kW)	Impact (%)
HER Control Customers	July 29	3.35	3.27	0.08	(0.01; 0.14)	2.4%
	August 27	2.89	2.80	0.09	(0.03; 0.15)	3.1%
	September 9	3.11	3.05	0.05	(-0.01; 0.12)	1.7%
	September 11	2.86	2.79	0.07	(0.01; 0.13)	2.4%
	Avg. Event Day	*3.05*	*2.98*	*0.07*	*(0.02, 0.12)*	*2.4%*
HER Recipients	July 29	3.21	3.14	0.07	(0.04, 0.11)	2.2%
	August 27	2.74	2.69	0.05	(0.02, 0.09)	2.0%
	September 9	2.93	2.87	0.05	(0.02, 0.09)	1.8%
	September 11	2.65	2.62	0.03	(0.00, 0.06)	1.2%
	Avg. Event Day	*2.88*	*2.83*	*0.05*	*(0.02, 0.08)*	*1.8%*

There is insufficient evidence to suggest that BDR impacts are significant at the 95% level on the September 9 event for HER control customers, and the September 11 event for HER recipients. The impacts on these days, highlighted in grey, are the smallest of each group's estimated impacts across event days in both absolute and percentage terms. On average, BDR produced an average 2.4% reduction in peak usage for HER control customers across the four event days while BDR produced an average 1.8% reduction for HER recipients. We take this result to indicate that HER recipients, who on average do show energy efficiency savings compared to HER control group customers, have reduced capacity to make additional peak demand-specific behavioral changes. The reference load of HER recipients, 2.88 kW on the average event day, is lower on average than that of HER control customers, which is 3.05 kW on the average event day. The lower reference load of HER-treated households strengthens our interpretation that they already have realized the energy savings that HER control customers achieved during the BDR event as a consequence of ongoing HER treatment. The lower percent reductions and lower reference loads of HER recipient households therefore results in lower absolute impacts for the BDR intervention.

To formally test for differences in both absolute and percent impacts between HER recipients and HER control customers in the BDR study, all customers were pooled into one dataset and an interaction term for HER recipient and BDR treatment dummy variables was added to the regression specification. The coefficient of the interaction estimates the difference in the treatment effect (reductions in load due to BDR) for the average customer in both the HER recipient and BDR treatment group, relative to a customer in the BDR treatment group only. The large p-values in Table 4 indicate that there is insufficient evidence to suggest either absolute or percent load reductions were different for BDR customers in the HER recipient group compared to HER control customers. This result holds for both individual event days and for the average of all event days. Such a result does not mean that there is no difference between the impacts for the two groups (the difference is consistently present across the four events), but instead suggests that the observed difference in impact between customers receiving both HER and BDR messaging and customers receiving only BDR messaging is too small to rule out chance as a possible explanation given the sample sizes.

Table 4. Tests for differences in BDR impacts between HER treatment and control households.

Impact Type	Event Date	BDR x HER Interaction Coefficient	SE	t-statistic	p-value
Absolute (kW)	July 29	0.007	0.038	0.17	0.86
	August 27	0.033	0.035	0.96	0.34
	September 9	0.000	0.037	0.01	1.00
	September 11	0.036	0.034	1.05	0.29
	Avg. Event Day	0.019	0.018	1.06	0.29
Percent (%)	July 29	-0.004	0.017	-0.25	0.80
	August 27	0.011	0.017	0.67	0.51
	September 9	-0.008	0.017	-0.49	0.62
	September 11	0.008	0.017	0.48	0.63
	Avg. Event Day	0.002	0.008	0.20	0.84

In addition to encouraging customers to reduce usage on Summer Saving Days, there is also evidence that the impacts of BDR spill over into non-event days. Figure 2 shows the estimated difference in peak period usage for BDR-treated households compared to households untreated by BDR (all within the HER control group) for every day of the summer. Starting from the date welcome letters were sent, households treated with BDR consistently have lower peak usage compared to customers in the BDR control group.[3] This reduction in peak usage could be a result of some customers responding to the information and energy saving tips that were included as part of the welcome letter, a conservation effect caused by BDR-treated household members being aware that other customers are being asked to reduce their usage (social comparison), or possibly a small Hawthorne effect (Landsberger 1958) that results from customers adjusting their behavior in response to knowing that their usage might be studied as part of the trial.

[3] A formal hypothesis test of the difference between the BDR treatment and control groups for HER controls before and after the welcome letters were sent (May 1-July 28) shows the difference to be borderline statistically significant (p=0.12) at the 90% confidence level.

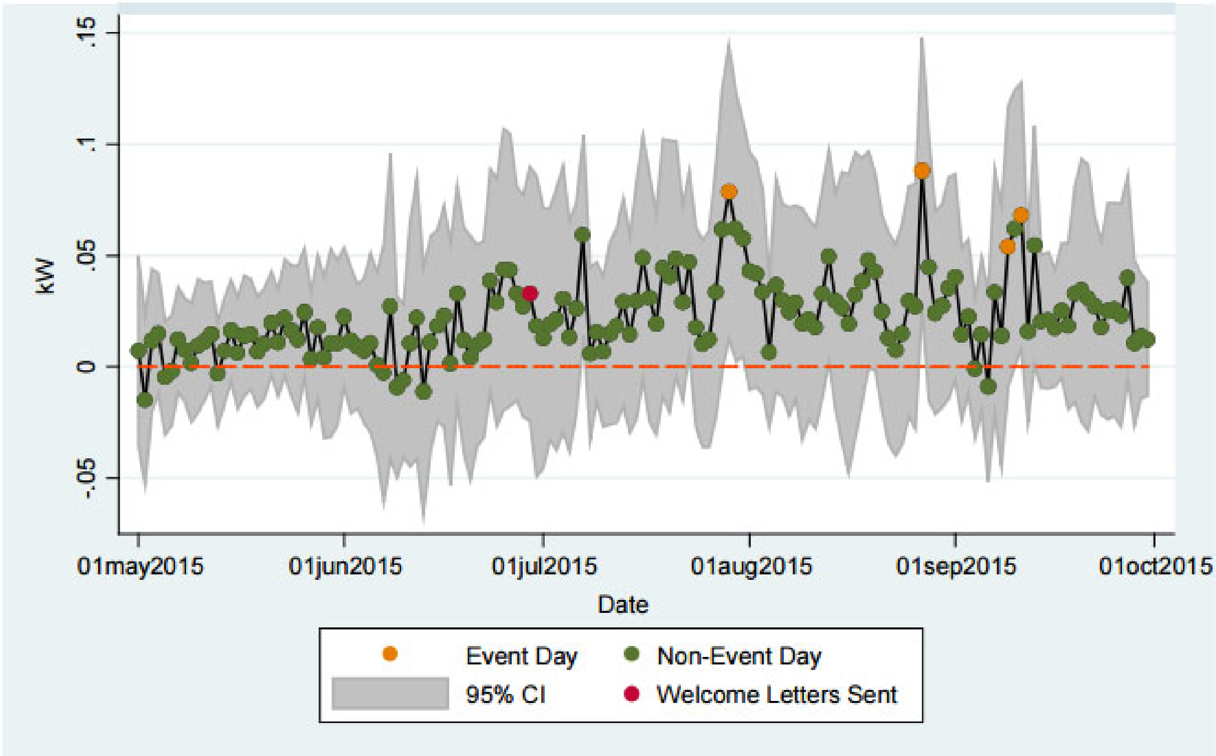

Figure 2. Average BDR-only impact (as compared to customers not receiving BDR) on summer cooling days (all customers are in HER control conditions and therefore were not exposed to HER messaging). Numbers above the horizontal axis represent a decline in energy use.

Figure 3 shows the average customer's hourly load shapes and impacts for both BDR study groups. The reference load, shown in blue, is the load that was observed in the absence of the BDR treatment. The observed load, shown in red, is the load observed in the treatment group on the average event day. The raw load shapes are associated with the left-most y axis and event hours are denoted by filled in circles. The impact (orange) is the difference between observed load and reference load, which is measured on the right y axis. The 95% confidence interval for the impact is shown in grey.

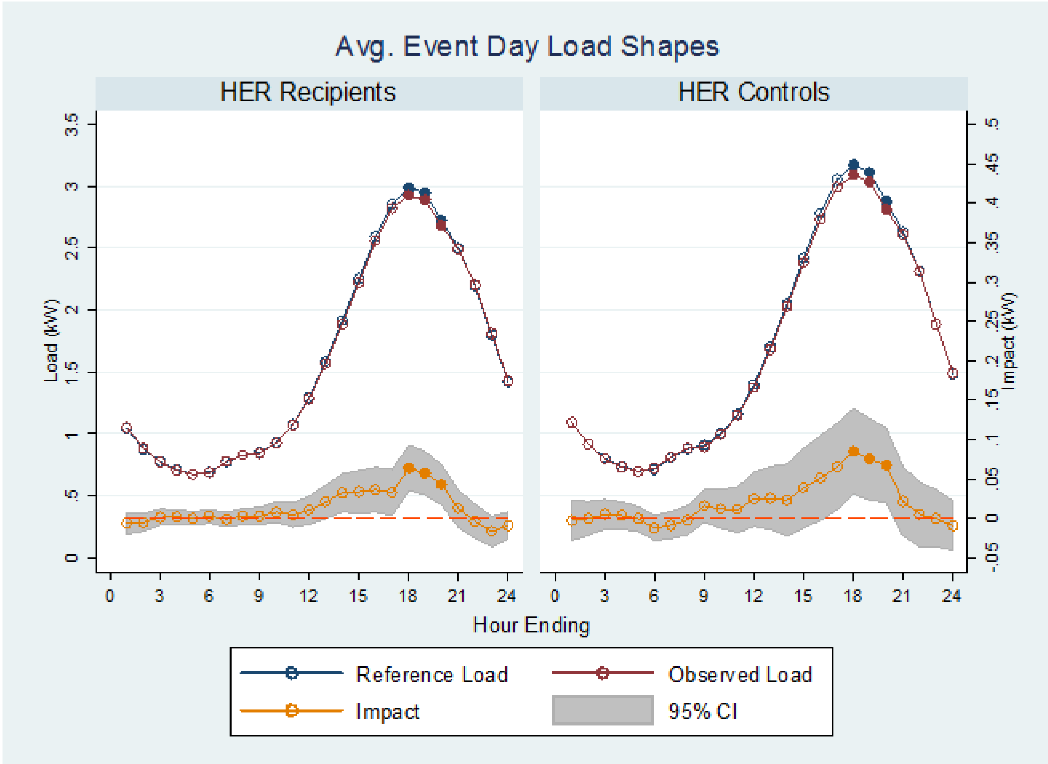

Figure 3. Load shapes for the average of four event days.

Conclusions and Recommendations

Load impact estimates from the BDR study show consistent reductions in peak usage of about 2-3% for BDR participants relative to the control group on event days. Multiplied by a large number of participants, a large-scale BDR program has the potential to provide significant load reductions—particularly for capacity-constrained substations—without the costs and time associated with launching traditional residential demand response programs. While the production costs and timeline of a full scale BDR program is not yet known, it is reasonable to assume that both costs and implementation time for launching BDR would be less than the cost of traditional load control programs requiring incentives or load control technologies that must be installed at customers' homes.

It is also evident that the effect of BDR messaging on energy consumption is not confined to the hours of the peak period for event days. In fact, peak demand reductions observed in the BDR condition persisted from day to day throughout the summer season. This persistence of peak demand reductions is observed in both households treated with HER and those that are not.

BDR appears to be an effective mechanism for reducing residential energy consumption during critical time periods. However, peak demand reductions are higher in the BDR-only condition compared to the BDR and HER condition. This finding suggests that layering on BDR messaging onto households already receiving HERs will result in less incremental demand savings than households not exposed to HERs. Customers treated with HERs may lack the capacity to affect as wide a range of behavioral changes to curtail energy use at peak demand periods as BDR-only treated households, most likely due to the behavioral changes already made in response to the messaging in HERs. The incremental benefits of BDR in addition to the energy reduction achieved through HER treatment should be carefully considered by portfolio managers in assessing which populations to include for participation in BDR.

There remain several important unanswered questions about the potential impacts of BDR that could be addressed in future research:

- This study was confined to 31 substations where PG&E is attempting to limit load growth. Future research would be required to determine whether load reductions in other locations would be comparable to what was observed in the study.
- The persistence of BDR impacts for HER recipients is not clear. The peak load demand decline from about 2.2% on the first event (July 27) to about 1.2% on the fourth event (September 11). This decline over time is approximately linear and might result from a coincidence of other temporal factors with the performance of the program (e.g., weather) or it might result from the fact that the impact of BDR messaging weakens with repeated exposure. Alternatively, it is also possible that impacts on the September 11 event were lower due to it occurring soon after the September 9 event and interrupting the normal BDR messaging cycle. The BDR messaging cycle may need to be adjusted to allow for multiple events in a short period of time. Future research could examine whether the observed decline in impact over successive event days is a characteristic of BDR overall.
- Since this study demonstrates that BDR has the potential to complement and increase the energy savings resulting from exposure to the HER, presenting BDR messaging to customers in the HER control group has the potential to affect their overall energy use. Further research should compare the differences in baseline energy use between customers in the HER control group that are exposed to BDR messaging and those that are not.

References

Karlin, B., J. Zinger, and R. Ford. 2015. *The Effects of Feedback on Energy Conservation: A Meta-Analysis.* Washington, DC: American Psychological Association.

Landsberger, H. 1958. *Hawthorne Revisited: Management and the Worker, Its Critics, and Developments in Human Relations in Industry*. Ithaca, NY: Cornell University.