REPORT

SoCalGas Demand Response: 2017/2018 Winter Load Impact Evaluation

August 14, 2018

Prepared for
Southern California Gas Company

Prepared by
Eric Bell, Ph.D.
Principal

Stephanie Bieler
Consultant

Nexant, Inc.

CALMAC ID: SCG0220

Table of Contents

1 Executive Summary

SoCalGas was directed by the California Public Utilities Commission (CPUC) to continue and expand the SoCalGas Thermostat Program in response to the potential need for demand reductions during the 2017-2018 winter and future winters. The Smart Thermostat Program for 2018 was an offering where two vendors (Vendor 1 and Vendor 2) recruited from their installed smart thermostat customer base, and offered incentives for customers to enroll. The program was event-based, meaning that it targeted relatively few hours on days of peak demand. Load reductions were attained on event days from temporary degree setbacks on thermostats, which led to a reduction in demand for heating. All activations took place either between the hours of 5 AM to 9 AM or 5 PM to 9 PM.

Gas load impacts (usage reductions) on event days were estimated by applying the best practices that have been developed for electric Demand Response (DR) program measurement and evaluation in California. As in the annual electric DR evaluations, the SoCalGas Smart Thermostat Program load impact estimates leverage the wide availability of interval data from advanced meters to estimate the usage reductions.

Table 1-1 provides a summary of the 2017-2018 winter SoCalGas Smart Thermostat Program hourly event impacts for each event and for the average morning and evening event by vendor. The average load reduction for a Vendor 1 morning event was 0.031 thm per participant leading to an aggregate reduction of 217.152 thm, or 16.0%. The average load reduction for a Vendor 1 evening event was 0.012 thm, leading to an aggregate reduction of 81.795 thm, or 10.7%. The average load reduction for a Vendor 2 morning event was 0.050 thm, leading to an aggregate reduction of 102.308 thm, or 25.0%. The average load reduction for a Vendor 2 evening event was 0.014 thm, leading to an aggregate reduction of 37.768 thm, or 15.6%. Vendor 2 event impacts were consistently larger than Vendor 1 event impacts, and both vendors saw morning event impacts that were larger than evening event impacts.

Table 1-1: Winter 2017-2018 Load Impact Estimates

Date	Event Window	Vendor 1				Vendor 2				Avg. Event Temp. (°F)
		Number of Participants	Average Impact (thm)	Aggregate Impact (thm)	Impact (%)	Number of Participants	Average Impact (thm)	Aggregate Impact (thm)	Impact (%)	
20-Feb	AM	6,976	0.029	201.36	12.5%	2,029	0.052	105.14	21.2%	44.8
20-Feb	PM	-	-	-	-	2,029	0.031	63.70	22.0%	51.8
21-Feb	AM	6,976	0.032	224.78	15.5%	2,029	0.052	104.96	23.7%	50.2
21-Feb	PM	-	-	-	-	2,029	0.023	46.84	18.0%	53.2
22-Feb	AM	6,976	0.031	214.12	14.1%	2,029	0.048	96.41	21.2%	48.7
22-Feb	PM	-	-	-	-	2,029	0.016	32.47	13.3%	53.5
23-Feb	AM	6,976	0.030	211.73	15.3%	-	-	-	-	48.3
26-Feb	PM	6,976	0.012	85.01	11.4%	2,029	0.017	34.29	16.4%	55.5
27-Feb	AM	6,976	0.031	214.71	16.5%	2,029	0.050	101.86	25.9%	46.8
28-Feb	PM	6,976	0.015	105.33	12.7%	2,029	0.010	20.76	9.6%	54.2
1-Mar	AM	6,976	0.032	222.01	16.4%	2,029	0.058	116.67	28.7%	51.0
1-Mar	PM	6,976	0.008	55.05	8.0%	2,029	0.014	28.55	14.3%	54.7
2-Mar	AM	6,976	0.033	231.36	21.6%	2,029	0.044	88.81	29.1%	52.3
All Events										
Avg.	**AM**	6,976	0.031	217.15	16.0%	2,029	0.050	102.31	25.0%	48.9
Avg.	**PM**	6,976	0.012	81.80	10.7%	2,029	0.019	37.77	15.6%	53.8
Common Events across both vendors										
Avg.	**AM**	6,976	0.031	218.01	16.1%	2,029	0.050	102.31	25.0%	49.0
Avg.	**PM**	6,976	0.012	81.80	10.7%	2,029	0.014	27.87	13.4%	54.8

The SoCalGas Thermostat program is one of the first, if not the first, natural gas based demand response programs in the US. It has proven that smart thermostats can be used to reduce demand for natural gas during targeted periods of time in the morning and evening. However, the thermostat setback strategy was also shown to be important, and can significantly affect the size of the load reductions and the post-event "snap back", as shown by the different vendor performance. The snap back following the event when a customer's preferred temperature settings are restored can be quite significant, and generally erases any net daily therm savings.

From a technical perspective, it's clear the program met the objectives of reducing gas consumption during specific windows of time. However, due to gas usage snap backs in the hours following events, there were no statistically significant net daily therm savings that resulted from this program. Without statistically significant net daily therm savings there is an open question regarding whether the program created value from a reliability or economic perspective. While on the electric grid blackouts can be caused by an immediate supply/demand imbalance, gas supply shortages causing low gas system pressure and deliverability issues are typically a more protracted event due to the slow speed of how gas travels. It's unclear how much of a supply shortage may exist for only a few hours in Southern California. If there aren't supply shortages lasting only a few hours, it's possible that traditional energy efficiency and behavioral conservation based programs, most notably Seasonal Energy Update energy reports, may yield greater savings over longer periods of supply shortage. These interventions have the dual benefit of providing significant gas savings on both DR event days and non-DR days throughout the winter.

2 Overview

SoCalGas was directed by the California Public Utilities Commission (CPUC) to continue and expand the SoCalGas Thermostat Program in response to the potential need for demand reductions during the 2017-2018 winter and future winters. The Smart Thermostat Program for 2018 was an offering where two vendors (Vendor 1 and Vendor 2) recruited from their installed smart thermostat customer base, and offered incentives for customers to enroll. The program was event-based, meaning that it targeted relatively few hours on days of peak demand. Load reductions were attained on event days from temporary degree setbacks on thermostats, which led to a reduction in demand for heating. Further details regarding the implementation of the pilot are contained in Section 2.1.

Gas load impacts on event days were estimated by applying the best practices that have been developed for electric Demand Response (DR) program measurement and evaluation in California. In 2008, the California Public Utilities Commission (CPUC) and joint electric Investor-Owned Utilities (IOUs) developed California's Load Impact Protocols, which required the electric utilities to conduct annual evaluations of all DR programs in the state. As in the annual electric DR evaluations, the SoCalGas Smart Thermostat Program load impact estimates leverage the wide availability of interval data from advanced meters to estimate usage reductions. The program evaluation methodology that uses a matched control group is similar to how most electric DR programs have been evaluated for several years, including Southern California Edison's (SCE's)® Save Power Days (also known as Peak Time Rebate) Program,[1] which is also a smart thermostat program.

Throughout this report, Nexant will define event, program, and load as follows:

- Event – refers to the four-hour period during which SoCalGas adjusted a customer's thermostat in order to reduce heating demand during that period (an "activation"). There can be multiple events in a single day.
- Program – refers to the SoCalGas Smart Thermostat Program, which is a combination of the Vendor 1 Program and the Vendor 2 Program
- Load – refers to customer gas usage, measured in therms (thm)

2.1 Program Design and Implementation

The SoCalGas Smart Thermostat program used the Bring Your Own Thermostat (BYOT) model to recruit existing customers with Vendor 1 and Vendor 2 thermostats into the program by offering up to $75 of incentives. Customers who enrolled in the program received a $50 enrollment incentive, as well as a $25 participation incentive after the winter season for remaining in the program. To recruit customers into the program, SoCalGas promoted the program using social media and radio advertising, and the vendors reached out to customers who had already adopted smart thermostat technologies. SoCalGas additionally sent out bill inserts to customers and had an email campaign for the program. Before the start of the program, customers were told that if an event was called, customer thermostats could be

[1] Nexant. "2017 Load Impact Evaluation of Southern California Edison's Peak Time Rebate Program." April 1, 2018. CALMAC Study ID: SCE0420.

adjusted remotely by SoCalGas by a few degrees, and there would be no "penalty of non-participation" for overriding a smart thermostat during a Natural Gas Conservation event. As shown in Table 2-1, at the end of recruitment, Vendor 1 had a little over 7,000 customers enroll in the Vendor 1 program and Vendor 2 had almost 2,000 customers enroll in the Vendor 2 program, for a total of approximately 9,000 customers enrolled in the SoCalGas Smart Thermostat Program.

Table 2-1: Vendors and Respective Pilot Program Enrollment

Contracted Vendor	Smart Thermostat Program	Enrolled Customers
Vendor 1	Vendor 1 Program	7,132
Vendor 2	Vendor 2 Program	1,842

Table 2-2 provides a summary of eligibility screens that each vendor applied to customers who had agreed to participate. Customers needed to own a thermostat from the respective vendor and needed to be a current SoCalGas residential gas service account holder. Vendor 2 additionally required that participants could not currently be enrolled in the SCE Save Power Days Program or the "SoCalGas Advanced Meter Opt-Out Program".

Table 2-2: Smart Thermostat Program Vendor Eligibility Requirements

Vendor 1 Criteria	Vendor 2 Criteria
Own Vendor 1 Thermostat with an active account	Own Vendor 2 Thermostat with active account
Have a wireless network installed at service address	Have a wireless network installed at service address
Active SoCalGas Account	Active SoCalGas Account
	Not enrolled in SCE Save Power Days
	Installed Advanced Meter at service address
	Natural gas furnace
	Not enrolled in "SoCalGas Meter Opt-Out Program"

Natural Gas Conservation events took place during periods of system constraint by adjusting thermostats to a lower temperature by no more than four degrees. Once the activations came to an end, thermostats were returned to their original set points.[2] All activations took place either between the hours of 5 AM to 9 AM or 5 PM to 9 PM, and customers who participated in the program received a notice at least two hours before the event.[3]

[2] Vendor 1 limits its thermostat adjustment to three degrees. Vendor 1 thermostats additionally will pre-adjust the temperature in the home before the event to maximize comfort. However, in the case of a morning event the thermostat will not pre-adjust the temperature unless the customer has a specific setting enabled. This is to ensure noise comfort for the customer.

[3] With the exception of Vendor 1's second event in a day, which notifies the customer at the time of the activation.

In May 2018, SoCalGas conducted a focus group in order to evaluate overall customer satisfaction with the DR program. In the focus group, customers did not report any pain points for enrollment in either program, and they found enrollment in the program to be “fast and easy”. The focus group also found that both Vendor 1 and Vendor 2 customers were very satisfied with the program, and were likely to recommend the program to a friend and participate in the program again.[4]

2.2 Program Participants

Customers who signed up to participate in the SoCalGas Smart Thermostat Program are inherently different from customers who did not sign up to participate in the program or customers who were not targeted by SoCalGas marketing or thermostat vendors. Before the evaluation, specific customer segments were examined to observe how program participants differed from the overall population. Table 2-3 compares the portion of CARE customers who enrolled in the pilot to the overall population. Program participants were less likely to be CARE customers compared to the general residential population.

Table 2-3: Comparison of Program and Participation CARE Customers

CARE	% of Program Participants	% of SoCalGas Residential Customers
Yes	9%	28%
No	91%	72%
All	100%	100%

Table 2-4 compares the breakout of SoCalGas program participant housing type to the SoCalGas residential customer population. Program participants were more likely to reside in a single family home compared to the general population.

Table 2-4: Comparison of Program and Population Housing Types

Housing Type	% of Program Participants	% of SoCalGas Residential Customers
Single Unit	84%	65%
2 or More Separate Units	2%	3%
2-4 Connected Units	4%	10%
5 or More Connected Units	10%	22%
Mobile Home Park	0%	0%
All	100%	100%

4 From Vendor 1 program and Vendor 2 progam Focus Group Report.

Figure 2-1 shows a heat map of the locations of pilot participants throughout the SoCalGas service territory. The largest concentrations of customers are in the LA Basin and Orange County areas. The next largest concentration is in the Riverside, Palm Springs and Bakersfield areas.

Figure 2-1: Heat Map of Pilot Participant Location

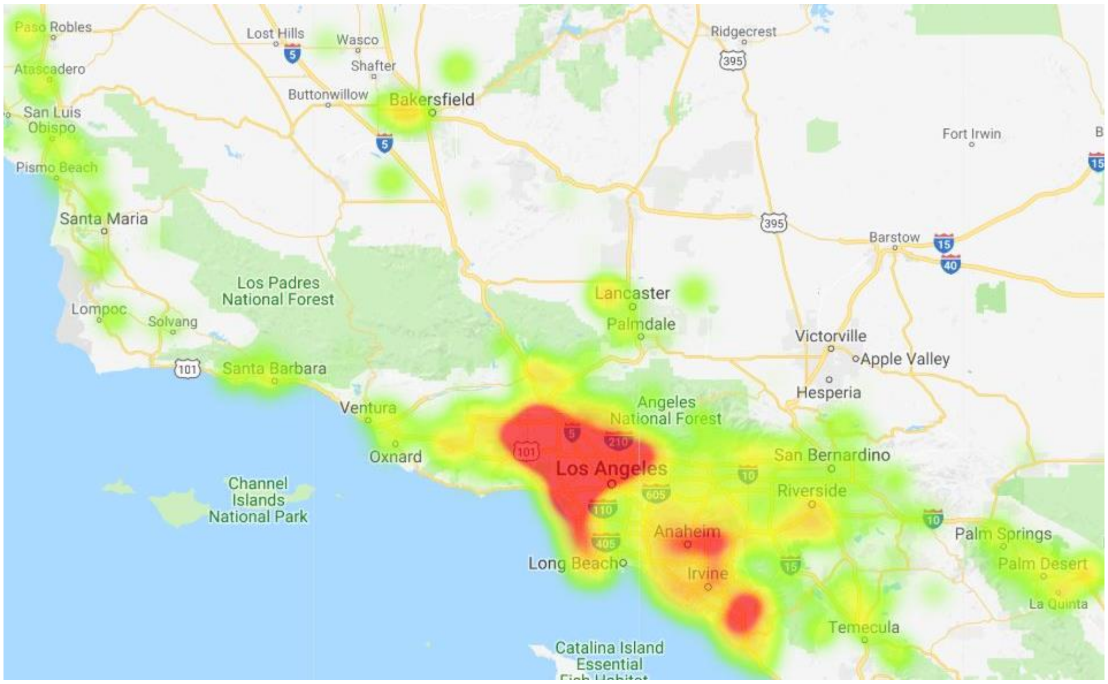
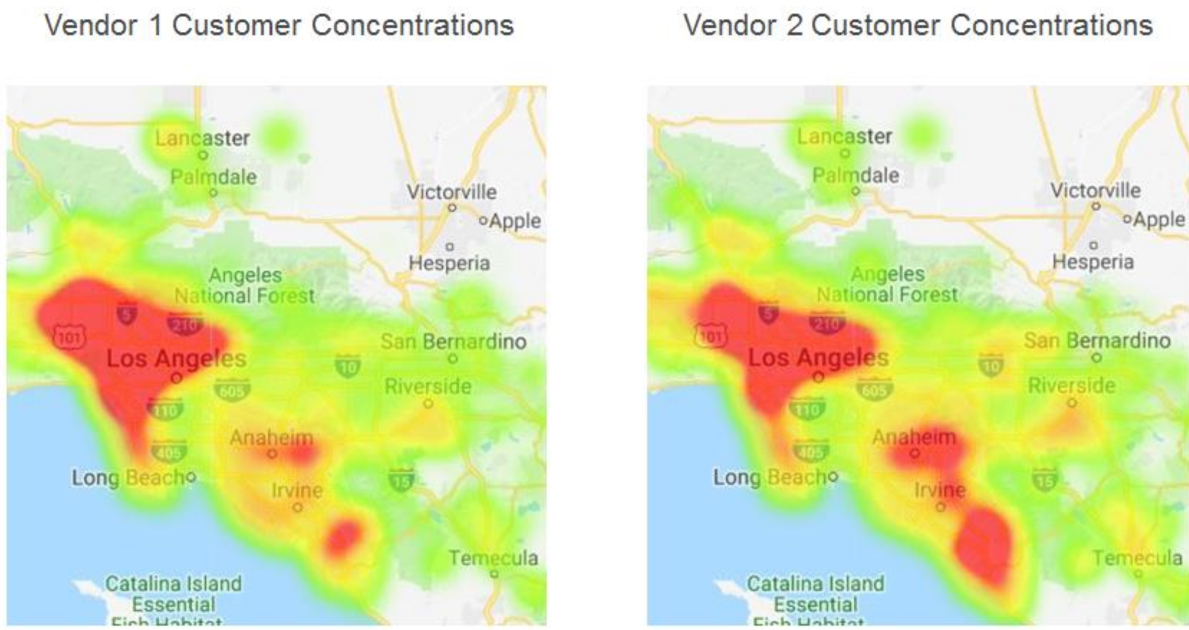


Figure 2-2 shows a heat map of pilot participants broken out by vendor. Vendor 2 has greater concentrations of customers in the Orange County region than Vendor 1, but the two vendors have similar customer concentrations in the LA Basin.

Figure 2-2: Heat Map of Pilot Participant Location By Vendor



2.3 Event Summary

Events were four hours long and took place either in the morning from 5 AM to 9 AM or in the evening from 5 PM to 9 PM. All of the events took place between February 20, 2018 and March 2, 2018. There were a total of thirteen events on nine different days, with seven morning events and six evening events. On four of the nine days, both morning and evening events were called.

Table 2-5 provides a summary of the events called during the 2017/2018 season. The thermostat vendor identifies which vendor(s) was called for each event, and the devices targeted column refers to the number of devices that were activated for an event. The last four columns record the participation status of the activated devices. Full participation refers to devices that were successfully accessed and the DR settings were in place for the entire event. An opt-out refers to customers that overrode the DR event settings. Vendor 1 kept track of which customers opted out before or during events. Vendor 2 did not, and so all opt-outs are counted as opting out before an event for Vendor 2 customers. Other refers to devices that were either “off”, in an incompatible mode, or were not accessible due to technical issues. On average, 57% of devices targeted participated in the entire event, 22% of devices targeted opted out before the event, 13% of devices targeted opted out during the event, and 8% did not participate in the event due to technical issues.

Table 2-5: Overall Event Summary

Date	Event Window	Thermostat Vendor	Devices Targeted	Full Participation	Opt-out Before	Opt-out During	Other
2/20/2018	AM	Vendor 1 and Vendor 2	9,384	51%	25%	19%	5%
2/20/2018	PM	Vendor 2 only	1,564	59%	17%	0%	24%
2/21/2018	AM	Vendor 1 and Vendor 2	9,374	55%	24%	17%	4%
2/21/2018	PM	Vendor 2 only	1,550	59%	19%	0%	22%
2/22/2018	AM	Vendor 1 and Vendor 2	9,354	55%	23%	17%	4%
2/22/2018	PM	Vendor 2 only	1,541	60%	19%	0%	21%
2/23/2018	AM	Vendor 1 only	7,801	56%	23%	20%	1%
2/26/2018	PM	Vendor 1 and Vendor 2	9,317	57%	23%	15%	5%
2/27/2018	AM	Vendor 1 and Vendor 2	9,317	55%	24%	17%	4%
2/28/2018	PM	Vendor 1 and Vendor 2	9,313	55%	23%	17%	4%
3/1/2018	AM	Vendor 1 and Vendor 2	9,575	56%	23%	17%	5%
3/1/2018	PM	Vendor 1 and Vendor 2	9,814	59%	22%	14%	5%
3/2/2018	AM	Vendor 1 and Vendor 2	9,807	59%	20%	16%	4%
Average	-	-	-	57%	22%	13%	8%

Each vendor was called for a different number of events. Vendor 1 customers were called for ten of the thirteen events and Vendor 2 customers were called for twelve of the thirteen events. Vendor 1 customers did not participate in the first three evening events due to technical difficulties, but participated in the remaining events. Vendor 2 customers were not called for a morning event on February 23, but participated in the remaining events. Both vendors were called for nine of the thirteen events, and there was one day where both vendors were called for both a morning and an evening event.

Tables 2-6 and 2-7 give the event summaries for each vendor. Vendor 2 had a higher participation rate on average than Vendor 1, with an average of 59% of Vendor 2 customers participating in events compared to 55% of Vendor 1 customers participating in events. Vendor 2 also had a higher percent of customers that did not participate due to technical issues, with 22% of customers characterized with a participation status of other, while Vendor 1 had only 1% of participants categorized as other. These differences could be due to different methods of recording participation between the two vendors, as Vendor 2 did not record different opt-out times in the same way that Vendor 1 did. Vendor 1 broke out its opt-outs into customers that opted-out before an event and customers that opted-out during an event. On average, 24% of

Vendor 1 customers opted-out before an event and 20% of Vendor 1 customers opted-out during an event. This distribution did not change significantly between morning and evening events. On average, about 19% of Vendor 2 customers opted out either before or during an event.

Table 2-6: Vendor 1 Event Summary

Date	Time	Devices Targeted	Full Participation	Opt-out Before	Opt-out During	Other
2/20/2018	AM	7,816	51%	26%	22%	1%
2/20/2018	PM					
2/21/2018	AM	7,812	55%	24%	20%	1%
2/21/2018	PM					
2/22/2018	AM	7,806	55%	24%	21%	1%
2/22/2018	PM					
2/23/2018	AM	7,801	56%	23%	20%	1%
2/26/2018	PM	7,792	56%	25%	18%	1%
2/27/2018	AM	7,792	55%	24%	20%	1%
2/28/2018	PM	7,792	54%	24%	21%	1%
3/1/2018	AM	7,793	56%	23%	21%	1%
3/1/2018	PM	8,034	58%	23%	18%	1%
3/2/2018	AM	8,029	59%	21%	19%	1%
Average	-	7,847	55%	24%	20%	1%

Table 2-7: Vendor 2 Event Summary

Date	Event Window	Devices Targeted	Full Participation	Opt-out Before	Opt-out During	Other
2/20/2018	AM	1,568	54%	22%		24%
2/20/2018	PM	1,564	59%	17%		24%
2/21/2018	AM	1,562	57%	23%		20%
2/21/2018	PM	1,550	59%	19%		22%
2/22/2018	AM	1,548	57%	22%		21%
2/22/2018	PM	1,541	60%	19%		21%
2/23/2018	AM					
2/26/2018	PM	1,525	64%	12%		24%
2/27/2018	AM	1,525	57%	23%		20%
2/28/2018	PM	1,521	61%	17%		22%
3/1/2018	AM	1,782	57%	22%		20%
3/1/2018	PM	1,780	61%	15%		23%
3/2/2018	AM	1,778	60%	20%		21%
Average	-	1,604	59%	19%		22%

3 Load Impact Estimation Methodology

The primary challenge in estimating load impacts for DR programs such as the Smart Thermostat Program is estimating how much gas participants would have used during an event in the absence of SoCalGas dispatching the program. The estimated participants' usage in the absence of the event is referred to as the counterfactual or the reference load. This was not a randomized control trial, so the primary source of data used to develop reference loads is a matched control group. Control customers were selected from a pool of non-participant customers that passed several filters that were also applied to the program participants, and were statistically matched to program participants. The fundamental idea behind the matching process is to find customers who were not subject to DR events that have similar observable characteristics to those who were subject to DR events.

Once a suitable control group was created from a group of non-participants, the next step was to use a "difference-in-differences" analysis to estimate load impacts. Difference-in-differences helps to yield more precise estimates and can correct for observable differences in load not accounted for through matching. This calculation was done using a fixed-effects regression methodology, which reduces the standard error of the estimates. The underlying approach for difference-in-differences is comprised of the following:

- Measure gas demand for both treatment and control customers on proxy (similar non-event) days;
- Measure gas demand for both treatment and control customers on event days;
- Treatment effects are calculated by taking the difference between the treatment and matched control group in the event hours and subtracting any difference between the two groups in the event period hours on proxy days.

Additional details on the load impact estimation methodology including the selection of the matched control group and difference-in-differences regression model can be found in Appendix A.

4 Load Impacts

During the 2017-2018 winter, thirteen events were called on nine different days. All thirteen events ran for four hours and were called either from 5 AM to 9 AM or from 5 PM to 9 PM. Load impacts were evaluated separately for each vendor due to differences in when vendor customers were called for events and the ways in which events were implemented for each vendor. The remainder of this section presents the load impacts for each vendor for each event the vendor participated in.

4.1 Load Impacts for Vendor 1

Table 4-1 summarizes the average and aggregate impacts for each Vendor 1 event as well as the event temperature. Vendor 1 customers participated in eight morning events and two evening events for a total of 10 events. The average hourly impact during a morning event was 0.031 thm per participant, representing a 16% load reduction from an average reference load of 0.204 thm. The average hourly aggregate impact during a morning event was a 217.152 thm load reduction from a reference load of 1,423.104 thm. The average hourly per-customer impact during an evening event was 0.012 thm, an 11% load reduction from an average reference load of 0.114 thm. The average hourly aggregate impact was a 81.795 thm load reduction from a reference load of 711.552 thm.

Time of day and corresponding levels of consumption, which are at least partially driven by temperature, were large drivers of impact differences. Morning event impacts and reference loads were consistently higher than evening event impacts and reference loads, with higher reference loads generally associated with larger event impacts. On average, there was a 5 degree temperature difference between the average morning event hour and the average evening event hour. The afternoon events also likely had reduced heating load due to the heat buildup in the home during the day as well as warmer event period temperatures.

Table 4-1: Vendor 1 Event Summary for Average Customer

Date	Event Window	Vendor 1					Avg. Event Temp. (°F)
		Average Load w/o DR (thm)	Average Load w/ DR (thm)	Average Impact (thm)	Aggregate Impact (thm)	Impact (%)	
20-Feb	AM	0.241	0.213	0.029	201.36	12.5%	44.8
21-Feb	AM	0.214	0.182	0.032	224.78	15.5%	50.2
22-Feb	AM	0.222	0.192	0.031	214.12	14.1%	48.7
23-Feb	AM	0.206	0.176	0.030	211.73	15.3%	48.3
26-Feb	PM	0.109	0.097	0.012	85.01	11.4%	55.5
27-Feb	AM	0.192	0.161	0.031	214.71	16.5%	46.8
28-Feb	PM	0.123	0.108	0.015	105.33	12.7%	54.2
1-Mar	AM	0.195	0.163	0.032	222.01	16.4%	51.0
1-Mar	PM	0.109	0.101	0.008	55.05	8.0%	54.7
2-Mar	AM	0.154	0.121	0.033	231.36	21.6%	52.3
All Events							
Avg.	**AM**	0.204	0.172	0.031	217.15	16.0%	48.9
Avg.	**PM**	0.114	0.102	0.012	81.80	10.7%	54.8

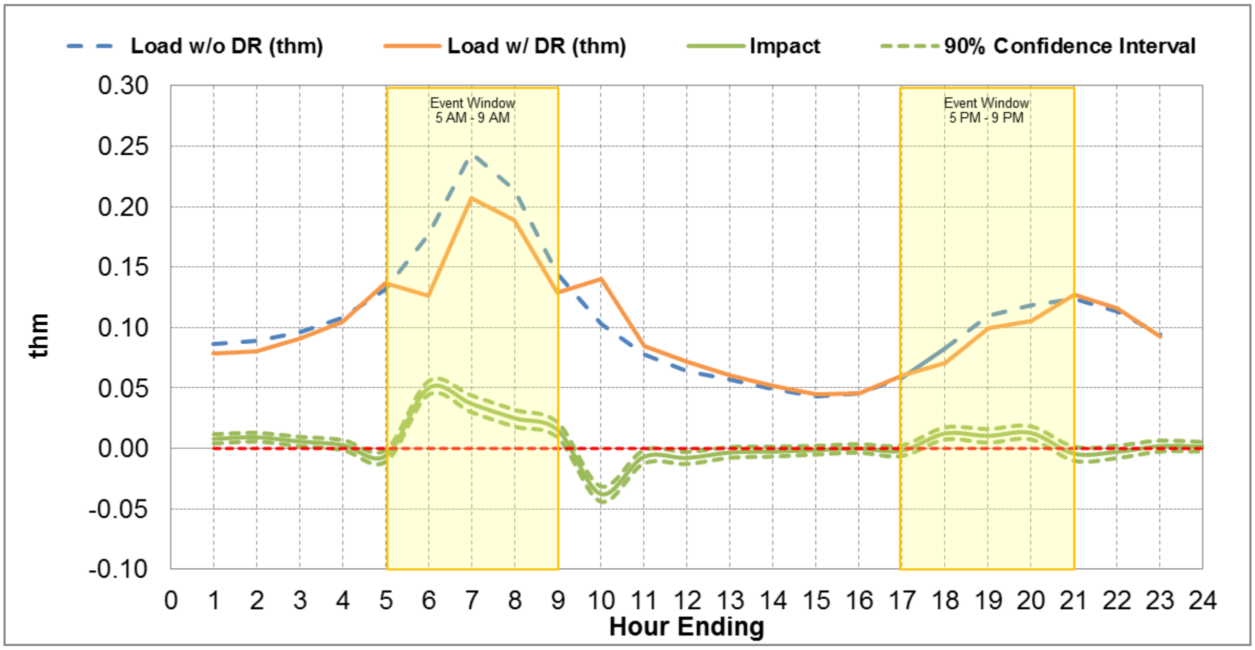
Vendor 1 customers experienced three different event day types: days with only morning activations, days with only evening activations, and days with both morning and evening activations. There was one day (March 1) where both a morning and evening event were called. Figure 4-1 provides the average per customer load with DR, load without DR (reference load), and load impact for that day. The load shape and usage patterns for the morning event window in Figure 4-1 are illustrative of customer behavior during all morning events, and the load shapes and usage patterns during the evening event window in Figure 4-1 are illustrative of the customer load shapes during all evening events.[5] Morning event windows had the highest overall reference load and highest overall impacts, with the largest impact occurring in the first hour of the morning event. Evening events had a much lower reference load and lower impacts.

In the hour following both morning and evening events, there is what is referred to as “snap back”, which is when customer gas usage is higher after an event than would be expected if an event had not taken place. This is because during an event, the Vendor 1 thermostat temperature is lowered by up to 3°F. After the event, the thermostat temperature is returned to its pre-event temperature. In order to increase the temperature in the home to the non-event temperature, the HVAC system has to run more consistently for up to the first hour following the event (or longer). This can result in increased consumption in the hours following an event compared to what would typically be expected on a similar non-event day. The average snap back for Vendor 1 customers following morning events was 0.033 thms, with the load of the average participant 26% greater than customers that did not participate in the event. The

[5] This figure does not represent average morning event impacts across all morning events or average evening event impacts across all evening events. Its purpose is to illustrate what both events looked like, and shows exact impacts only for days where both morning and evening events were called.

average snap back for Vendor 1 customers following evening events was 0.015 therms, representing a 12% load increase compared to customers that did not participate in the event.

Figure 4-1: Vendor 1 Average Hourly Load Impact per Customer on Average Event Day with both Morning and Evening Events Called



4.2 Load Impacts for Vendor 2

Table 4-2 summarizes the average and aggregate impacts for each Vendor 2 event as well as the event temperature. Vendor 2 customers participated in six morning events and six evening events for a total of twelve events. The average impact during a morning event was 0.050 thm, representing a 25% reduction from an average reference load of 0.205 thm. The average hourly aggregate impact was a 102.308 thm reduction from a reference load of 415.905 thm. The average impact during an evening event was 0.019 thm, representing a 16% load reduction from an average reference load of 0.120 thm. The average aggregate impact was a 37.768 thm reduction from a reference load of 243.48 thm.

Similar to Vendor 1, all events Vendor 2 customers participated in were within approximately 10°F of each other. Time of day and corresponding levels of consumption, which are at least partially driven by temperature, were large drivers of impact differences. Morning event impacts and reference loads were also consistently higher than evening event impacts and reference loads, with higher reference loads generally associated with larger event impacts.

Table 4-2: Vendor 2 Event Summary for Average Customer

Date	Event Window	Vendor 2					Event Temp. (°F)
		Average Load w/o DR (thm)	Average Load w/ DR (thm)	Average Impact (thm)	Aggregate Impact (thm)	Impact (%)	
20-Feb	AM	0.244	0.192	0.052	105.141	21.2%	44.78
20-Feb	PM	0.147	0.116	0.031	63.701	22.0%	51.76
21-Feb	AM	0.218	0.166	0.052	104.957	23.7%	50.23
21-Feb	PM	0.133	0.110	0.023	46.844	18.0%	53.24
22-Feb	AM	0.224	0.176	0.048	96.413	21.2%	48.70
22-Feb	PM	0.127	0.111	0.016	32.466	13.3%	53.51
26-Feb	PM	0.104	0.087	0.017	34.285	16.4%	55.49
27-Feb	AM	0.195	0.145	0.050	101.858	25.9%	46.81
28-Feb	PM	0.111	0.101	0.010	20.764	9.6%	54.23
1-Mar	AM	0.198	0.140	0.058	116.673	28.7%	50.99
1-Mar	PM	0.099	0.085	0.014	28.552	14.3%	54.73
2-Mar	AM	0.152	0.108	0.044	88.805	29.1%	52.33
All Events							
Avg.	**AM**	0.205	0.155	0.050	102.308	25.0%	48.97
Avg.	**PM**	0.120	0.101	0.019	37.768	15.6%	53.83

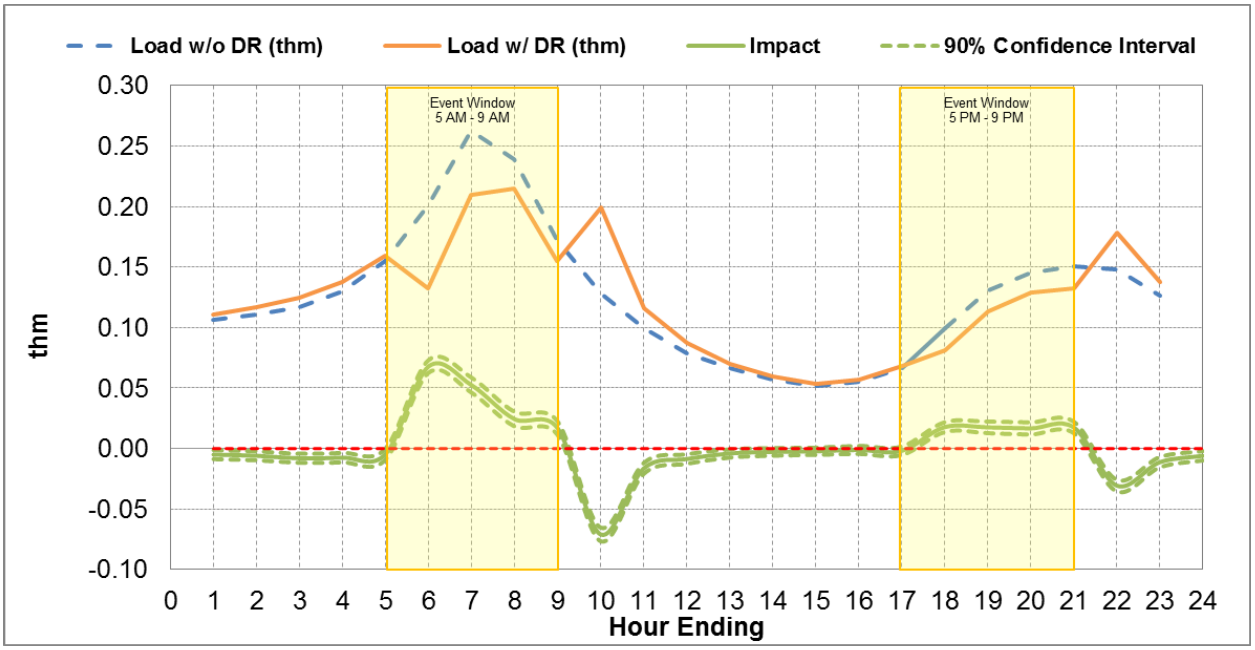
Vendor 2 customers experienced three different event day types: days with only morning events, days with only evening events, and days with both morning and evening events. There were four days where both a morning and evening event was called in the same day. Figure 4-1 provides the average per customer load with DR, load without DR (reference load), and load impact for the average event day for Vendor 2 customers where there were both morning and evening activations. The load shape and usage patterns for the morning event window in Figure 4-2 are illustrative of customer behavior during all morning events, and the load shapes and usage patterns during the evening event window in Figure 4-2 are illustrative of the customer load shapes during all evening events.[6] Morning event windows had the highest overall reference load and highest overall impacts, with the largest impact occurring in the first hour of the morning event. Evening events had a much lower reference load and lower impacts.

In the hour following the event, the snap back for the average Vendor 2 customer was larger than with Vendor 1 customers. The average snap back for Vendor 2 customers following morning events was 0.068 thm, with the load of the average participant 60% greater than customers that did not participate in the event. The average snap back for Vendor 2 customers following evening events was 0.028 thm, representing a 24% load increase compared to customers that did not participate in the event. In the evening, the post-event snap back

[6] This figure does not represent average morning event impacts across all morning events or average evening event impacts across all evening events. Its purpose is to illustrate what both events looked like, and shows exact impacts only for days where both morning and evening events were called.

increased the hourly consumption to a new higher hourly peak for Vendor 2 treatment customers between the hours of 9 PM and 10 PM.

Figure 4-2: Vendor 2 Average Hourly Load Impact per Customer on Average Event Day with both Morning and Evening Activations



4.3 Comparison of Vendor Load Impacts

Table 4-3 contains a summary of the average customer load impacts for each event for each vendor. The two vendors experienced a different mix of events during the 2017-2018 winter. Vendor 1 customers participated in seven morning events and three evening events, while Vendor 2 customers participated in six morning events and six evening events. Both vendors participated together in a total of nine events. In this section, we will use events where both vendors participated when comparing impacts since during these events customers experienced the same weather conditions. Each vendor took a different approach to the thermostat setback during the events, which is evident in the different load impacts and snap back patterns observed between the two vendors under similar weather conditions.

Vendor 2 and Vendor 1 customers both participated in a total of six morning events. During morning events, the average temperature was 48.97°F. Vendor 2 customers had a slightly higher baseline than Vendor 1 customers, with an average reference load of 0.205 thm compared to the Vendor 1 average reference load of 0.203 thm. Vendor 2 also had a much higher event impact than Vendor 1, with an average hourly impact of 0.050 thm during the event, 25% of the reference load. Vendor 1 customers had an average hourly impact of 0.031 thm, 16% of the reference load. However, as discussed above it should be noted that Vendor 2 customers also had a much larger snapback than Vendor 1 customers in the hour following an event, with Vendor 2 DR customers using 60% more load than would be expected in the absence of an event and Vendor 1 customers using 26% more load than would be expected in the absence of an event.

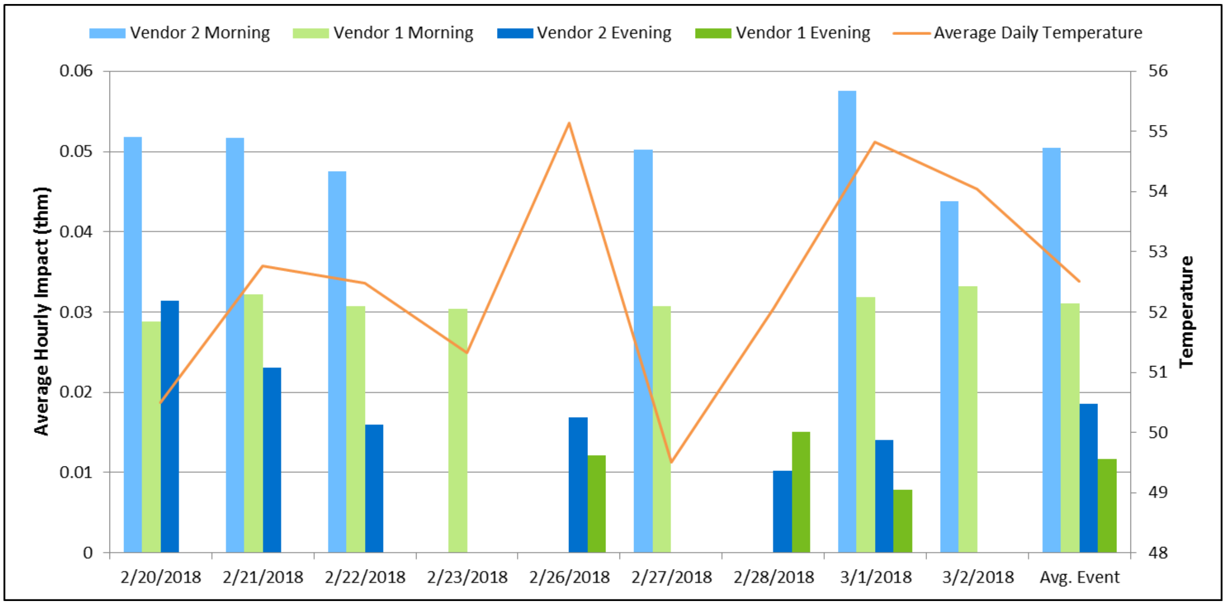
Vendor 2 and Vendor 1 customers both participated in a total of three evening events. During evening events, the average temperature was 54.82°F. Vendor 1 customers had a higher baseline than Vendor 2 customers, with an average reference load of 0.114 thm, compared to the Vendor 2 reference load of 0.104 thm. Similar to the morning impacts, Vendor 2 had a slightly higher event impact than Vendor 1, with an average hourly impact of 0.014 thm, 13% of the reference load. Vendor 1 customers had an average hourly impact of 0.012 thm, 10.7% of the reference load. Vendor 2 also again had a higher snapback after evening events than Vendor 1, seeing a 24% increase in load relative to the reference load in the hour following an event. Vendor 1 customers saw a 12% increase in load relative to the reference load in the hour following an event.

Table 4-3: Summary Load Impacts for Common Events Across Both Vendors

Date	Event Window	Vendor 1				Vendor 2				Event Temp. (°F)
		Average Load w/o DR (thm)	Average Load w/ DR (thm)	Average Impact (thm)	Impact (%)	Average Load w/o DR (thm)	Average Load w/ DR (thm)	Average Impact (thm)	Impact (%)	
20-Feb	AM	0.241	0.213	0.029	12.5%	0.244	0.192	0.052	21.2%	44.78
21-Feb	AM	0.214	0.182	0.032	15.5%	0.218	0.166	0.052	23.7%	50.23
22-Feb	AM	0.222	0.192	0.031	14.1%	0.224	0.176	0.048	21.2%	48.70
26-Feb	PM	0.109	0.097	0.012	11.4%	0.104	0.087	0.017	16.4%	55.49
27-Feb	AM	0.192	0.161	0.031	16.5%	0.195	0.145	0.050	25.9%	46.81
28-Feb	PM	0.123	0.108	0.015	12.7%	0.111	0.101	0.010	9.6%	54.23
1-Mar	AM	0.195	0.163	0.032	16.4%	0.198	0.140	0.058	28.7%	50.99
1-Mar	PM	0.109	0.101	0.008	8.0%	0.099	0.085	0.014	14.3%	54.73
2-Mar	AM	0.154	0.121	0.033	21.6%	0.152	0.108	0.044	29.1%	52.33
Common Events across both vendors										
Avg.	**AM**	0.203	0.172	0.031	16.1%	0.205	0.155	0.050	25.0%	48.97
Avg.	**PM**	0.114	0.102	0.012	10.7%	0.104	0.091	0.014	13.4%	54.82

Figure 4-3 illustrates the variation in impacts across events for each vendor for all events. Vendor 2 event impacts are blue and Vendor 1 event impacts are green. Vendor 2 consistently delivered larger impacts than Vendor 1 customers for morning events, and morning events consistently had larger impacts than evening events. Vendor 1 impacts varied very little across each event type, with all morning event impacts within 0.002 thm of the average morning event impact and all evening event impacts within 0.004 thm of the average evening event impact. Vendor 2 impacts varied more, with one morning event impact up to 0.008 thm greater than the average morning event impact and one evening event impact up to 0.011 thm greater than the average evening event impact.

Figure 4-3: Event Impact Summary by Vendor



4.4 Daily Therm Savings

Table 4-4 illustrates the average and aggregate daily savings for each event day type by vendor. It should be noted that neither vendor saw statistically significant daily savings for any event day type due to the snap-back in the hours following both morning and evening events. However, with a larger sample size it is possible that both vendors could see statistically significant daily savings in the future. Vendor 1 customers had a maximum daily saving of 4.9%[7] on March 1, when SoCalGas called both a morning and evening event. Vendor 2 customers had maximum average daily savings when only morning events were called, with an average daily impact of 2.5%. However, due to the small number of each event type, these numbers may not represent which event type would provide the largest daily savings on average.

Table 4-4: Estimated Daily Therm Savings by Vendor

Vendor	Event Day Type	Average Daily Impact (thm)	Aggregate Daily Impact (thm)	Aggregate Daily Impact (CCF)	Daily Impact (%)	Statistically Significant	Event Day Type Count
Vendor 1	AM Only	0.068	472.147	458.395	2.3%	No	6
	PM Only	0.047	328.490	318.923	1.8%	No	2
	AM & PM	0.118	826.482	802.410	4.9%	No	1
Vendor 2	AM Only	0.066	133.083	129.207	2.5%	No	2
	PM Only	0.016	31.463	30.546	0.6%	No	2
	AM & PM	0.045	91.226	88.569	1.6%	No	4

[7] Not statistically significant.

5 Conclusions and Recommendations

Table 5-1 provides a summary of the 2017-2018 winter SoCalGas Smart Thermostat Program hourly event impacts for each event and for the average morning and evening event by vendor. The average load reduction for a Vendor 1 morning event was 0.031 thm per participant leading to an aggregate reduction of 217.152 thm, or 16.0%. The average load reduction for a Vendor 1 evening event was 0.012 thm, leading to an aggregate reduction of 81.795 thm, or 10.7%. The average load reduction for a Vendor 2 morning event was 0.050 thm, leading to an aggregate reduction of 102.308 thm, or 25.0%. The average load reduction for a Vendor 2 evening event was 0.014 thm, leading to an aggregate reduction of 37.768 thm, or 15.6%. Overall, Vendor 2 customers consistently produced larger average event impacts relative to Vendor 1 customers. Across both vendors morning events provided larger impacts relative to evening events. Due to gas usage snap-backs after the event window, neither vendor had statistically significant daily therm savings, regardless of when an event was called or how many events were called.

Table 5-1: Winter 2017-2018 Load Impact Estimates

Date	Event Window	Vendor 1				Vendor 2				Event Temp. (°F)
		Number of Participants	Average Impact (thm)	Aggregate Impact (thm)	Impact (%)	Number of Participants	Average Impact (thm)	Aggregate Impact (thm)	Impact (%)	
20-Feb	AM	6,976	0.029	201.355	12.5%	2,029	0.052	105.141	21.2%	44.78
20-Feb	PM	-	-	-	-	2,029	0.031	63.701	22.0%	51.76
21-Feb	AM	6,976	0.032	224.779	15.5%	2,029	0.052	104.957	23.7%	50.23
21-Feb	PM	-	-	-	-	2,029	0.023	46.844	18.0%	53.24
22-Feb	AM	6,976	0.031	214.118	14.1%	2,029	0.048	96.413	21.2%	48.70
22-Feb	PM	-	-	-	-	2,029	0.016	32.466	13.3%	53.51
23-Feb	AM	6,976	0.030	211.733	15.3%	-	-	-	-	48.25
26-Feb	PM	6,976	0.012	85.005	11.4%	2,029	0.017	34.285	16.4%	55.49
27-Feb	AM	6,976	0.031	214.712	16.5%	2,029	0.050	101.858	25.9%	46.81
28-Feb	PM	6,976	0.015	105.334	12.7%	2,029	0.010	20.764	9.6%	54.23
1-Mar	AM	6,976	0.032	222.013	16.4%	2,029	0.058	116.673	28.7%	50.99
1-Mar	PM	6,976	0.008	55.048	8.0%	2,029	0.014	28.552	14.3%	54.73
2-Mar	AM	6,976	0.033	231.357	21.6%	2,029	0.044	88.805	29.1%	52.33
All Events										
Avg.	**AM**	6,976	0.031	217.152	16.0%	2,029	0.050	102.308	25.0%	48.87
Avg.	**PM**	6,976	0.012	81.795	10.7%	2,029	0.019	37.768	15.6%	53.83
Common Events across both vendors										
Avg.	**AM**	6,976	0.031	218.055	16.1%	2,029	0.050	102.308	25.0%	48.97
Avg.	**PM**	6,976	0.012	81.795	10.7%	2,029	0.014	27.867	13.4%	54.82

The SoCalGas Thermostat program is one of the first, if not the first, natural gas based demand response programs in the US. It has proven that smart thermostats can be used to reduce demand for natural gas during targeted periods of time in the morning and the evening.

However, the snap back following the event when a customer's preferred temperature settings are restored can be quite significant, and generally erase any net daily therm savings. Though, with larger sample sizes it may be possible to achieve statistically significant net daily therm savings. The thermostat setback strategy was also shown to be important, and can significantly affect the size of the load reductions and the post-event snap back, as shown by the different vendor performance. The performance differential actually provides a valuable data point, in that the setback strategy could be fine-tuned or adjusted to better meet a distribution system's specific need.

From a technical perspective, it's clear the program met the objectives of reducing gas consumption during specific windows of time. However, without statistically significant net daily therm savings there is an open question regarding whether the program created value from a reliability or economic perspective. While on the electric grid blackouts can be caused by an immediate supply/demand imbalance, gas supply shortages causing low gas system pressure and deliverability issues are typically a more protracted event due to the slow speed of how gas travels. It's unclear how much of a supply shortage may exist for only a few hours in Southern California. If there aren't supply shortages lasting only a few hours, it's possible that traditional energy efficiency and behavioral conservation based programs, most notably Seasonal Energy Update energy reports, may yield greater savings over longer periods of supply shortage. These interventions have the dual benefit of providing significant gas savings on both DR event days and non-DR days throughout the winter.

Appendix A Load Impact Methodology Details

A.1 Selection of Matched Control Group

Customers who signed up to participate in the Vendor 1 or Vendor 2 programs are inherently different from customers who did not sign up to participate in the SoCalGas DR programs or customers who were not targeted by the thermostat vendors. For this reason, a control group must be constructed using statistical matching. It is possible that the customers who enrolled in the SoCalGas DR programs had particular characteristics that made them more likely to enroll than customers who did not enroll or customers who were not targeted to enroll. This is particularly important when studying early adopters of a new technology such as smart thermostats who may have very different gas consumption patterns from those of the rest of the population. This type of behavior introduces selection bias because the difference in usage between the two groups caused by characteristics differences could be mistaken as the impact of treatment. A matched control group is the primary source for reference loads which are used to estimate impacts. The method used to assemble the matched control group is designed to ensure that the control group load on event days is an accurate estimate of what load would have been among SoCalGas DR customers on event days if an event hadn't taken place.

Nexant selected the control groups using propensity score matching to find residential SoCalGas customers who are non-DR program participants with load shapes most similar to those of SoCalGas DR participants. In this procedure, a probit model is used to estimate a score for each customer based on a set of observable variables that are assumed to affect the decision to join a SoCalGas DR program. A probit model is a regression model designed to estimate probabilities—in this case, the probability that a customer would enroll in a SoCalGas DR program. The score can be interpreted two different ways. First, the propensity score can be thought of as a summary variable that includes all the relevant information in the observable variables about whether a customer would choose to participate in a SoCalGas DR program. Each customer in the DR program population was matched with a customer in the non-DR population that has the closest propensity score. The second way to think of the propensity score is as the probability that a customer will join a SoCalGas DR program based on the included independent variables. Thinking of it this way, each customer in the control group was matched to a SoCalGas DR customer with a similar probability of joining a SoCalGas DR program given the observed variables. Nexant performed the match within four clusters that grouped customers based on their load shape similarity. In other words, the match was conducted separately for SoCalGas DR customers that had load shapes similar to one-another.

In order to select the probit model used to find the best match for each treatment customer, "out of sample" testing was performed to evaluate several different probit model specifications. Out of sample testing involves running each of the different model specifications using all but one of the proxy days, leaving the unused proxy day to test how well the model performed. By leaving a different proxy day out each time the matching selection is run, one is able to see how well the matches look on a day that was not used to select the match. During this process, sixteen different model specifications were tested using different observable variables including usage during event hours, average total daily usage, and usage from 12pm to 9pm. For each of the eleven models six different "calipers" were tested. Calipers set a maximum threshold of how large the difference in propensity scores can be for a matched pair. During the matching process, the treatment customers are matched to the control customer who has the most similar propensity score to them. Additionally, treatment customers can only be matched to a control customer in the same load shape cluster. If the difference between a treatment customer and control customer's propensity score is higher than the set caliper, the treatment customer will not be matched. Therefore, a caliper sets the standard for how close the matched pairs need to be. In order to find the closest control customer matches, the SoCalGas DR customers were split out by vendor to find the optimal probit model for each vendor. This provided much closer matches for each of the two thermostat vendor customers.

Figure A-1 and Figure A-2 show the results of the matched control group for the two thermostat vendors. The Vendor 1 customers match very well to their matched control group on proxy days. Vendor 2 also matches very well, although not quite as well as Vendor 1 customers do. This is in part due to the difference in sample size between the two vendors. Vendor 1 has over 7,000 customers while Vendor 2 has less than 2,000 customers. Both vendors also do not match perfectly during the daytime hours. This is because when selecting the model matching during event hours was given priority over non-event hours, since non-event hours are not as crucial for estimating event impacts.

Figure 5-1: Hourly Average Demand for Vendor 1 Customers on Proxy Days

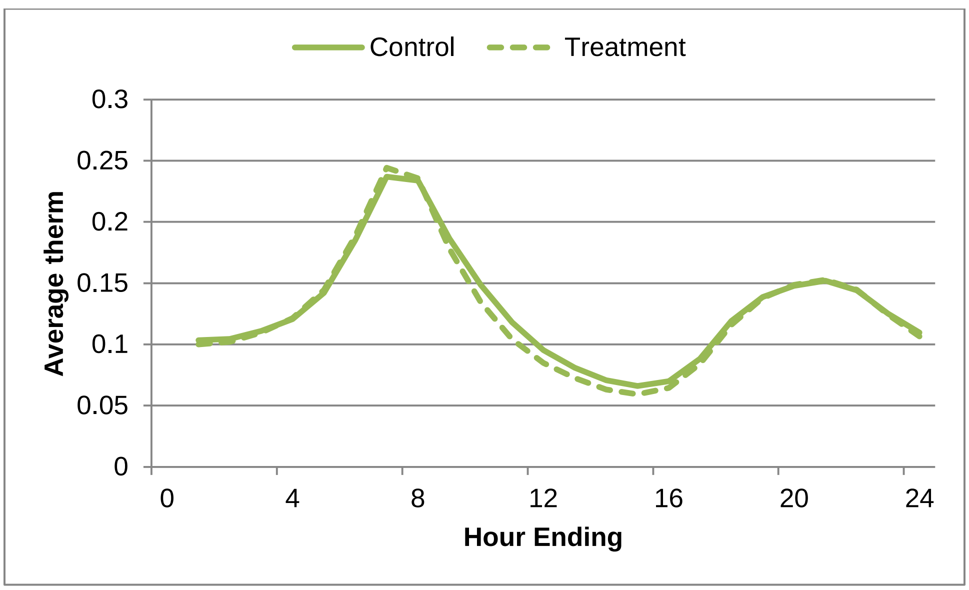
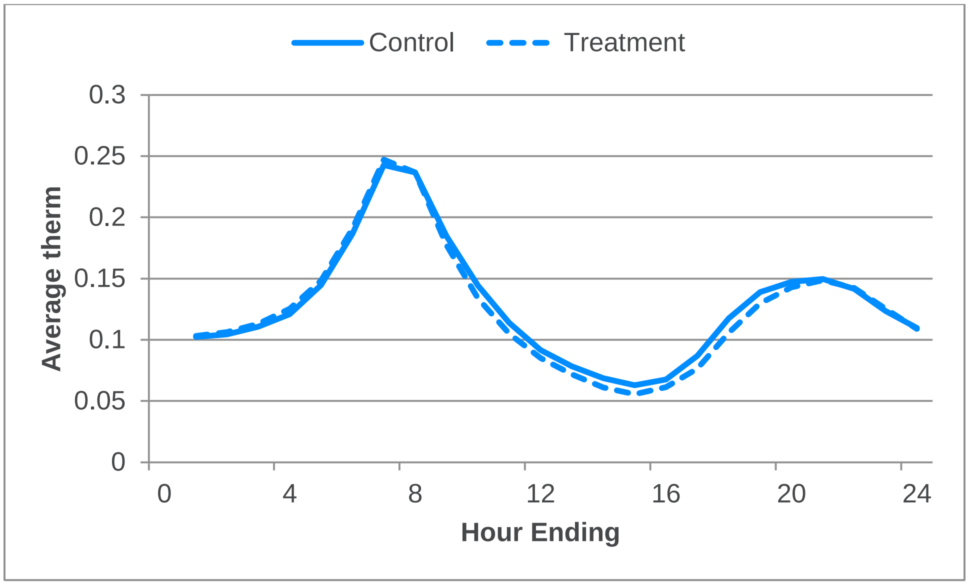


Figure 5-2: Hourly Average Demand for Vendor 2 Customers on Proxy Days



A.2 Difference-in-Differences Regression Models

After a matched control group was created, program impacts were estimated using a difference-in-differences regression model. This methodology is based on the assumption that the program impact is equal to the difference in usage between the treatment and the control groups during the event period, minus any pre-existing difference between the two groups. When using difference-in-differences, the matched control group does not need to perfectly match the treatment group on the proxy days. Any differences that may be due to observable differences in load not accounted for through matching will be netted out by the differencing. It is a reasonable assumption that any unobservable differences between the treatment and the control groups during the event period hours on proxy days stay the same during the DR event hours. Therefore any further difference between the groups in the DR event hours is assumed to be the impact of treatment. This regression model is shown in Equation A-1 below:

Equation A-1: Difference-in-Differences Models

$$thm_{i,t} = a + b \cdot Treatment_i + c \cdot Event_t + d \cdot (Treatment_i \cdot Event_t) + u_t + \varepsilon_{i,t} \text{ for } i \in \{1, \ldots, n_i\} \text{ and } t \in \{1, \ldots, n_t\}$$

Variable	Definition
i, t, n	Indicate observations for each individual i, date t and event number n
a	The model constant
b	Pre-existing difference between treatment and control customers
c	The difference between event and proxy days common to both treatment and control group members[8]
d	The net difference between treatment and control group customers during event days–this parameter represents the difference-in-differences
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
v	Customer fixed effects that control for unobserved factors that are time-invariant and unique to each customer; fixed effects do not control for fixed characteristics such as air conditioning that interact with time varying factors like weather
E	The error for each individual customer and time period
$Treatment$	A binary indicator or whether or not the customer is part of the treatment or control group
$Event$	A binary indicator of whether an event occurred that day–impacts are only observed if the customer is enrolled in DR ($Treatment$ = 1) and it was an event day

The model was estimated using both event days and proxy days, which are nonevent days with similar weather conditions and system load usage as days when events are called. The difference in loads between treatment and control customers for the event period hours on proxy days is subtracted from the differences on DR event hours to adjust for any differences between the treatment and control groups due to random chance.

As an extra validation, the simple difference in loads between treatment and control customers during event hours on event and proxy days was calculated to ensure that the regression model produces a similar output. The regression model also reduces the standard errors of the impact estimates compared to those that can be calculated from a simple difference in loads.

8 In practice, this term is absorbed by the time effects, but it is useful for representing the model logic.