Demand Side Analytics
DATA DRIVEN RESEARCH AND INSIGHTS

CADMUS

REPORT

Program Year 2021 Impact Evaluation of Central Hudson's Energy Insights Behavioral Program

Prepared for Central Hudson
March 2022

ACKNOWLEDGEMENTS

Research Team, Demand Side Analytics

- Josh Bode
- Savannah Horner
- Kate Bleakley

Central Hudson Team

- Michael Lauchaire
- Cory Scofield

ABSTRACT

This report summarizes the program year 2020 (PY2021) energy and gas consumption saving due to the Central Hudson's Energy Insights Behavioral program. In total, Central Hudson sent reports to over 96,000 electric accounts and over 25,000 gas accounts, tracking their historical usage, comparing it to similar households, and providing customized energy saving tips. The program was evaluated using a randomized control trial and compares energy use before and after the intervention for both the participants and the control group (difference-in-differences). While home energy reports deliver small percentage changes in energy use, because they reach a large numbers of customers and do not require rebates or installations, they typically yield large aggregate savings. In 2021, the program delivered 14,804 MWh of electric savings and 35,131 MMBtu of gas savings.

TABLE OF CONTENTS

1 Introduction 4
1.1 Key research questions 4
1.2 Program Description 4
1.3 Program Background 6
1.4 Report structure 7
2 Methodology 8
2.1 Approach Overview 8
2.2 Simple Difference-in-Differences 9
2.3 Calendarizing of Bi-monthly Billing Data 10
2.4 Opt Outs and Attrition 11
2.5 Scaling of Gas Impacts 12
3 Electric Impacts 13
3.1 Comparison of Control Group and Participant Electricity Use 13
3.2 Electric Savings Ex Post Analysis 14
3.3 Monthly Impacts 18
4 Gas Impacts 21
4.1 Comparison of Control Group and Participant Gas Use 21
4.2 Gas Savings Ex-Post Analysis 22
4.3 Monthly Impacts 25
5 Conclusions and Recommendations 28

1 INTRODUCTION

Central Hudson's Energy Insights Behavioral program delivers paper and electronic home energy reports (HERs) to residential customers. While home energy reports deliver small percent changes in energy use, they typically yield large aggregate savings because they reach a large number of customers and do not require applications, rebates, or installation. The primary challenge of this evaluation is the need to accurately detect small changes in energy consumption while systematically eliminating plausible alternative explanations for those changes, including random chance. This report summarizes the electric and gas savings realized in 2021.

The Central Hudson behavioral program underwent a substantial shift in 2021. In prior years, Central Hudson worked with a contractor that sent customers multiple paper or electronic reports per year (typically six per year), which included energy use comparisons to similar, neighboring households. For 2021, Central Hudson brought the behavioral program operations in-house and shifted it into maintenance mode to reduce costs and improve cost-effectiveness.

1.1 KEY RESEARCH QUESTIONS

The study has four main research questions:

- Was the participant and control group energy use similar when neither group had yet received home energy reports?
- What is the magnitude of annual electricity savings?
- What is the magnitude of annual gas savings?
- Is there decay in the per customer energy savings? How do the results compare to savings before the shift to maintenance mode?
- What steps can be undertaken to improve delivery and performance?

1.2 PROGRAM DESCRIPTION

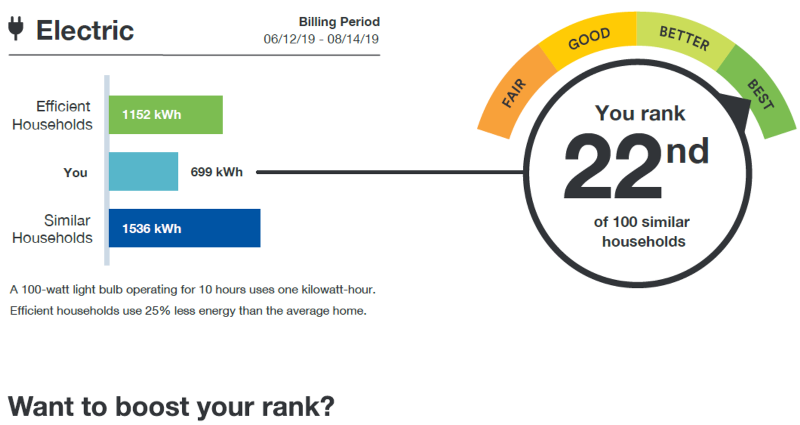
Home energy reports are behavioral interventions designed to encourage energy conservation of either gas or electricity. In Central Hudson, these reports are coupled with an online tool for customers to track energy consumption.

In previous program years, customers received roughly six paper or electronic reports per year comparing their energy consumption to similar neighboring households. The reports leveraged behavioral psychology and social norms to lower residential energy usage. These reports compared customers' historical use to see how it changed over time and compared their usage to similar, neighboring homes. Based upon these metrics, the reports included a personalized assessment of energy use and customized recommendations with concrete steps on how to save.

By 2021, Central Hudson took over the administration of the behavioral program and shifted it into maintenance mode to reduce costs and improve cost-effectiveness. Rather than send six reports with household energy usage comparisons, Central Hudson sent two communications with energy-saving tips to customers. Central Hudson maintained the randomly assigned control groups and ensured the communications to the customer used the same mode – email or print – as in prior years. An example of the difference in reports is provided in Figure 1.

Figure 1: Example of Change in Reports

PRIOR YEARS



MAINTENANCE MODE

1.3 PROGRAM BACKGROUND

The initial implementation of this energy savings program occurred in April 2011. Since then, there have been several new waves – in 2013, 2014, 2016, and 2019 – where Central Hudson added to the population receiving home energy reports. With each successive wave, Central Hudson randomly assigned customers to receive home energy reports or to act as the control group and provide a baseline of energy use in the absence of home energy reports.

Table 1 summarizes the waves of participants, with over 96,000 electric accounts and 25,000 gas accounts receiving home energy reports in 2021. Several of the electric and gas waves are too small for evaluation and lack statistical power on their own – based on the population and control group sizes, statistically significant impacts cannot be expected for many individual waves. Thus, the focus of this study is on the overall program savings rather than on the savings delivered by a specific wave.

Table 1: Wave Start Dates and Actively Enrolled Accounts

Wave	Experiment Start	Electric		Gas	
		Treatment	Control	Treatment	Control
D1	4/1/2011	13,317	4,147	13,680	3,494
E1	4/1/2011	36,324	13,658	672	.
E2O	1/1/2013	10,543	1,677	665	.
G3	9/1/2014	3,045	2,954	3,910	.
E2T	1/1/2016	9,156	1,357	114	.
G3 top up 1	1/1/2016	1,888	1,862	2,175	.
D1 top up 1	3/1/2019	2,468	756	2,463	755
E2O top up 1	3/1/2019	10,545	1,715	73	.
E2T top up 1	3/1/2019	7,703	1,164	93	.
G3 top up 2	3/1/2019	1,545	1,557	1,875	1,871
TOTAL		**96,534**	**30,847**	**25,720**	**6,120**

Three features of the program administration are relevant to the following evaluation. First, different companies have administered the program at different times – Opower, Simple Energy, Uplight, and directly by Central Hudson. Second, Central Hudson's billing system only retains the last 24 bills for each customer. As a result, the tracking of the control and treatment group's usage history has undergone multiple handoffs. Finally, Central Hudson's billing system underwent restructuring in September 2021. Thus, a key element of the study was assessing if the control and treatment groups had similar usage before either group received home energy reports and reviewing how well the historical billing data was tracked.

1.4 REPORT STRUCTURE

Section 2 describes the program implementation methodology and the evaluation methodology. Section 3 presents the electric impacts. Section 4 presents the gas impacts. Section 5 summarize the key findings and provides recommendations.

2 METHODOLOGY

The primary challenge of impact evaluation is the need to accurately detect changes in energy consumption while systematically eliminating plausible alternative explanations for those changes, including random chance. Did the introduction of HERs cause a decrease in customer energy consumption? Or can the differences be explained by other factors? To estimate energy reductions, it is necessary to estimate what these patterns would have been in the absence of treatment—this is called the counterfactual or reference value. At a fundamental level, the ability to measure energy reductions accurately depends on four key components:

- **The effect or signal size:** The effect size is most easily understood as the percent change. It is easier to detect large changes than it is to detect small ones. For most HER programs, the expected impact is between 1% and 2%, a relatively small impact.
- **Inherent data volatility or background noise:** The more volatile a customer's billing data are from month to month (or bimonthly billing period), the more difficult it is to detect small changes.
- **The ability to filter out noise or control for volatility:** At a fundamental level, statistical models, baseline techniques, and control groups—no matter how simple or complex—are tools to filter out noise (or explain variation) and allow the effect or impact to be more easily detected.
- **Population or sample size:** It is easier to precisely estimate average impacts for a large population than a small one because individual customer behavior patterns smooth out and offset across large populations. When control groups are employed, the size of the control group and the ratio between the control and participant groups determine the ability to detect different effect or signal sizes. As a rule of thumb, a control group of 10,000 is needed for sufficient statistical power to detect a 1% effect.

2.1 APPROACH OVERVIEW

Because the expected percent reduction from HERs is typically small (i.e., less than 3%), we followed the below principles to ensure accurate results:

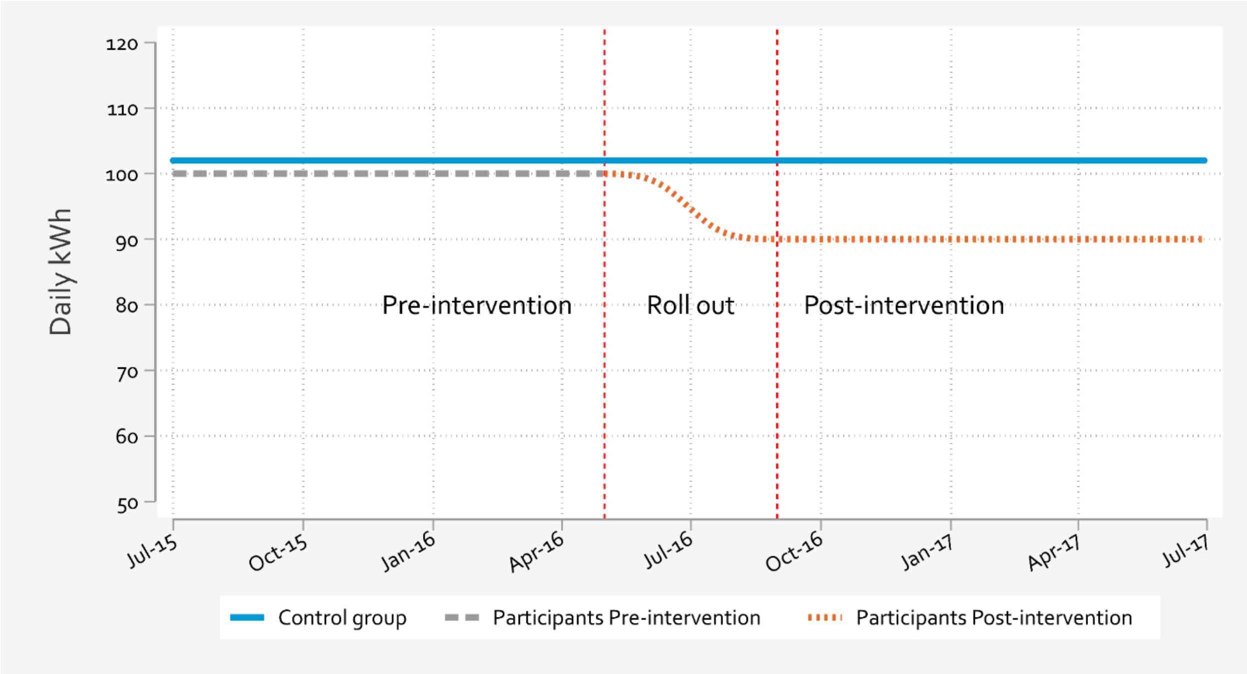
1. **Verify that participant and control customers had similar usage before the introduction of HERs.** By design, randomized control trials ensure that the only systematic difference between the two groups is that one receives the HER and one does not. However, random assignment is sometimes not implemented correctly or maintained. Thus, we compare the treatment and control groups across a host of characteristics—electricity use, location, etc.— to ensure the randomization was correctly implemented and tracked.

2. **Include at least one year of pre-treatment data and post-treatment data for both HER and control groups.** The pre-treatment data is useful for assessing if energy consumption changed and allows the use of more powerful statistical techniques such as difference-in-differences and lagged dependent variable models. If HERs lead to reductions in consumption, we should observe a change in consumption for customers who received the HER treatment but no similar change for the control group. Thus, the analysis did not include participant and control customers that lacked pre-intervention data.
3. **Ensure sample sizes large enough to detect meaningful differences**. If sample sizes are too small, it is not possible to distinguish meaningful differences from random noise. Several groups are too small to evaluate on their own (i.e., they lack statistical power). Thus, the focus of this study is on the overall program savings rather than on the savings delivered by specific waves.
4. **Apply the same data management procedures to both the HER and control groups.** Because of random assignment, data management decisions should have the same effect on the treatment and control group.
5. **Pre-specify the analysis method, econometric models, and segmentation in advance of the study.** The goal was to engage in science and avoid after-the-fact analysis where there is a temptation to modify models to find the desired results. This required documenting the hypothesis, specifying the intervention, randomly assigning customers to treatment and control conditions, establishing the sample size and the ability to detect meaningful effects, identifying the data that will be collected and analyzed, identifying the outcomes that will be analyzed, and documenting in advance the statistical techniques and models that will be used to estimate energy savings. The goal was to leave no ambiguity regarding how the data would be analyzed.
6. **Use difference-in-differences to estimate program impacts.** This approach for estimating impacts is conceptually simple and transparent. Compare energy usage before and after the intervention for both the participant and control groups and net out any pre-existing differences. The approach can be implemented through a comparison of means or via a difference-in-difference panel regression with fixed effects. In the evaluation, we estimate both.

2.2 SIMPLE DIFFERENCE-IN-DIFFERENCES

Figure 2 illustrates the difference-in-differences approach. The treatment and control groups for a randomized control trial should be nearly identical. If the behavioral intervention leads to reductions in consumption, we should observe a change in consumption for customers who received the energy reports and no similar change for the control group. We would also expect the timing of the change to align with the implementation of the HER. There may be small pre-existing differences between the two groups due to random chance, which are removed in the modeling. Randomization allows for clear attribution of energy impacts since the approach produces net savings due to the home energy report program.

Figure 2: Conceptual Example of Difference-in-Differences



The difference-in-difference calculation can be implemented in a variety of ways. The simplest approach is to use t-tests to estimate the impacts.[1] We used the t-test approach to estimate the impacts, but the difference-in-difference technique also can be implemented using a panel regression model. Only customers who have data before and after the intervention were included in the analysis.

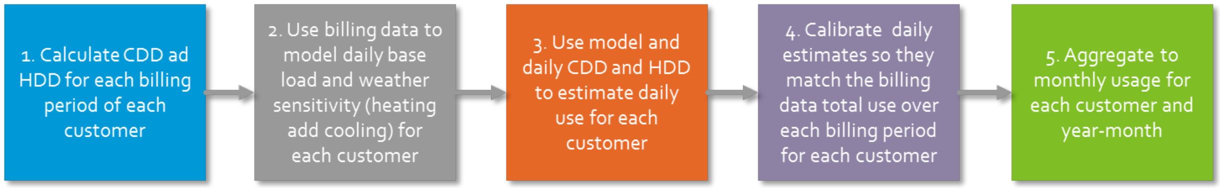
$$Impact = (Usage\ Post_{Treat} - Usage\ Pre_{Treat}) - (Usage\ Post_{Control} - Usage\ Pre_{Control})$$

2.3 CALENDARIZING OF BI-MONTHLY BILLING DATA

Central Hudson collects information about residential customer consumption on a bi-monthly basis. During months between bill reads, Central Hudson sends customers estimated bills. The timing of the meters read and the number of days between reads varies by customer. We prorate bimonthly billing data into a common calendar month basis in order to estimate the energy savings in 2021. The process of converting bills to usage is known as calendarization. Figure 3 summarizes the process employed to calendarize the data.

[1] Since each customer has before and after data, a paired t-test is used to estimate the change in use (and the standard errors) for the treatment and control group. The second step is to net of the difference observed in the control group.

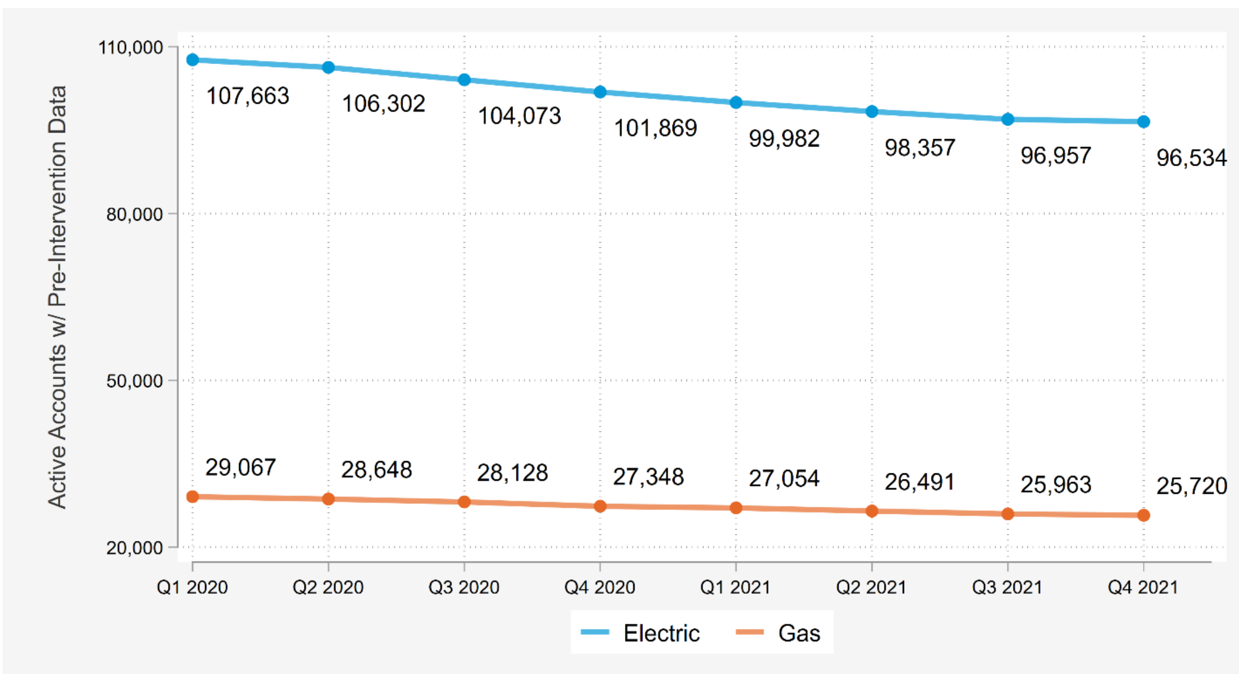
Figure 3: Calendarization of Bi-monthly Billing Data



2.4 OPT OUTS AND ATTRITION

Over time, some homes assigned to the HER program will close their accounts with Central Hudson. The most common reason for this is that the occupant is moving, but other possibilities exist. This account churn happens at a predictable rate and can be forecasted with some degree of certainty. Over the last two years, we have seen that electric account growth is decaying by about 5.3% each year and gas account growth is decaying by about 5.9% each year (Figure 4).

Figure 4: Account Attrition



This attrition is completely external to the program, so there is no reason to suspect that it happens differently in the treatment and control groups, if randomization is implemented properly. Once an account closes, there will no longer be consumption records in the billing data, so the home is removed naturally from the analysis without requiring any special steps.

Additionally, treatment group participants are allowed to opt out of receiving HER mailings if they choose. Typically, only a small proportion of the treatment group exercises this option. It is important that those who opt out are not removed from the analysis because doing so could compromise the randomization (control group homes do not the option to opt out).

2.5 SCALING OF GAS IMPACTS

Central Hudson lacks pre-intervention data for a subset of gas participants due, in part, to the transitions of program administration over time. Thus, impacts are estimated for the population with pre-intervention data – 100% of electric and 60% of gas customers – and the percent savings are applied to the remaining customers receiving reports. This is similar to the field of medicine where randomized control trials are performed on a subset of the population, and those results are applied to a broader population. Because we have usage information during PY2021, we apply the percent savings to the actual energy use of customers who lack pre-intervention data.

3 ELECTRIC IMPACTS

Electricity use is characterized by a wide range of end uses and technologies, including lighting, cooking and cleaning appliances, entertainment, and more. But the primary driver of energy loads is the heating and cooling systems. Electric usage peaks in the summer as air conditioning systems are running and in the winter for electrically heated homes. Because of this, energy use is highly dependent on weather.

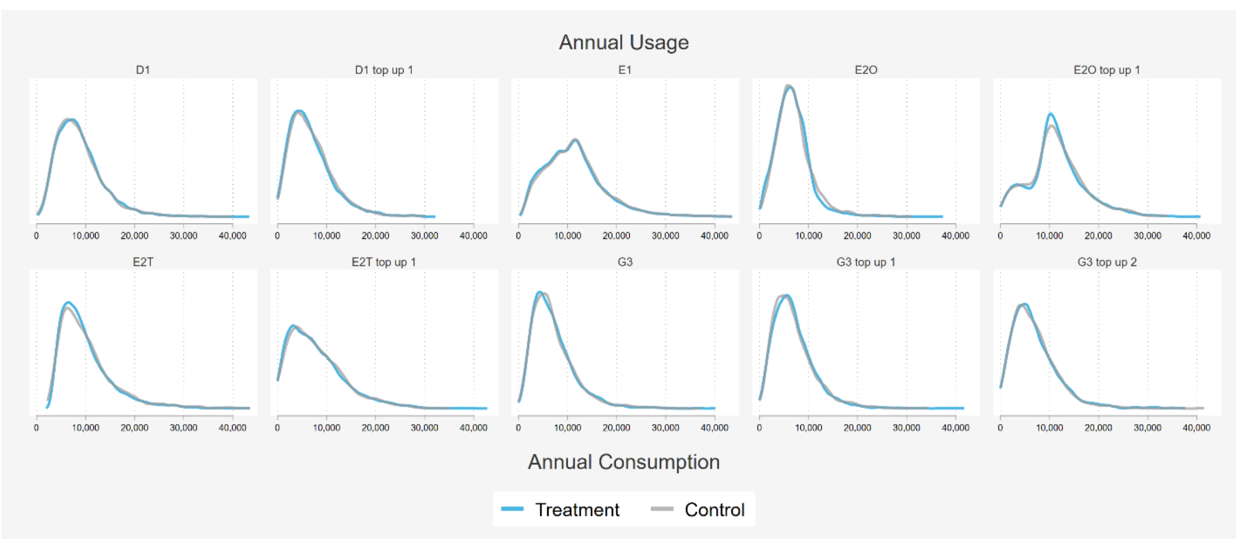
The main objectives of this section are to:

- Compare electricity use of the control group and participants prior to HER distribution
- Present summary statistics and figures of the energy use before and after the intervention
- Summarize electric energy savings over the evaluation period

3.1 COMPARISON OF CONTROL GROUP AND PARTICIPANT ELECTRICITY USE

For each wave of the HER program, pre-treatment energy consumption should be identical. A good control group should use energy in a similar manner as the participants before receiving HERs. Figure 5 shows the consumption distribution by wave for the treatment and control groups prior to the HER intervention. Treatment and control groups are comparable, but the average customer size varies between groups.

Figure 5: Pre-Treatment Annual Electric Consumption by Wave



While the pre-treatment analysis was similarly implemented in 2020, there are constant changes in the active participants over time. As customers move (or unenroll), they are removed from this analysis.

Between 2020 and 2021, there about 6,000 electric participants were deactivated. The fluctuation in participants alters the treatment and control groups.

Table 2 shows the average annual usage between treatment and control groups by wave. There are minor differences between the two groups for each wave. On average, the pre-intervention annual usage is 0.19% different between the groups, as compared to the 0.67% difference in the 2020 analysis. The customers who remain active in the program seem to be more similar than last year's active participants. Two waves also show statistically significant differences – E1 and E2O. In both cases, the average use in the treatment group is less than the control group. These pre-existing differences were removed in the evaluation by using the difference-in-differences technique to estimate energy savings.

Table 2: Pre-Treatment Electric Differences

Wave	Start Date	Control (n)	Treated (n)	Annual Use Control	Annual Use Treated	Diff	% Diff	S.E.	t	Lower Bound	Upper Bound
D1	4/1/2011	4,147	13,317	8,696.2	8,821.0	124.8	1.43%	84.6	1.47	-41.1	290.6
E1	4/1/2011	13,658	36,324	11,595.7	11,386.5	-209.2	-1.80%	63.8	-3.28	-334.2	-84.2
E2O	1/1/2013	1,677	10,543	7,070.7	6,728.8	-341.9	-4.84%	93.2	-3.67	-524.6	-159.2
G3	9/1/2014	2,954	3,045	7,099.3	7,104.7	5.3	0.07%	118.9	0.04	-227.8	238.4
E2T	1/1/2016	1,357	9,156	10,571.1	10,275.2	-295.9	-2.80%	190.2	-1.56	-668.7	76.9
G3 top up 1	1/1/2016	1,862	1,888	7,099.8	7,076.8	-23.0	-0.32%	162.4	-0.14	-341.4	295.4
D1 top up 1	3/1/2019	756	2,468	6,937.0	6,721.9	-215.2	-3.10%	185.1	-1.16	-577.9	147.6
E2O top up 1	3/1/2019	1,715	10,545	11,345.7	11,334.3	-11.4	-0.10%	148.7	-0.08	-302.8	280.0
E2T top up 1	3/1/2019	1,164	7,703	7,945.0	7,943.4	-1.6	-0.02%	184.0	-0.01	-362.3	359.1
G3 top up 2	3/1/2019	1,557	1,545	6,947.9	7,095.4	147.5	2.12%	183.0	0.81	-211.1	506.1
TOTAL		**30,847**	**96,534**	**9,712.4**	**9,730.7**	**18.3**	**0.19%**	**109.3**	**0.17**	**-195.9**	**232.6**

3.2 ELECTRIC SAVINGS EX POST ANALYSIS

Table 3 includes the results for the simple difference-in-differences analysis. The results are presented as energy impacts per day and can be scaled to annual energy savings per participant or program aggregate savings. Due to small sample sizes, some of the waves don't have much statistical power are not expected to produce statistically significant results. The focus should be on the program level impacts, which synthesizes the results for the over 96,000 participants and just under 31,000 control customers. The energy impact is highly statistically significant at the program level with a t-statistic of 5.40. On average, customers reduced electricity use by 1.76% in 2021.

By multiplying the impacts from Table 3 by the number of post-HER billing days, we can calculate the annual energy savings for the average customer. The number days per site varies by customers and wave due to differences in the billing cycles. Table 4 shows the annual savings per customer and the aggregate program energy savings. Annual savings were 153.4 kWh per customer. On an aggregate basis, Central Hudson's behavioral program reduced electricity use by 14.8 GWh.

Table 3: 2021 Difference-In-Difference Electric Analysis

Wave	Start Date	Days	Treatment				Control Group				Energy Impact per Customer per Day (kWh/day-account)					
			Accts (n)	Daily Use Post	Daily Use Pre	Diff	Accts (n)	Daily Use Post	Daily Use Pre	Diff	Impact (Diff-in-diff)	% Impact	S.E.	t	Lower Bound	Upper Bound
D1	4/1/2011	341	13,317	22.40	24.17	-1.77	4,147	22.60	23.83	-1.22	-0.54	-2.49%	0.18	-3.04	-0.89	-0.19
E1	4/1/2011	343	36,324	28.85	31.20	-2.35	13,658	29.75	31.77	-2.02	-0.33	-1.17%	0.13	-2.62	-0.58	-0.08
E2O	1/1/2013	343	10,543	20.12	18.44	1.69	1,677	22.40	19.37	3.03	-1.34	-7.14%	0.28	-4.76	-1.89	-0.79
G3	9/1/2014	340	3,045	20.51	19.46	1.04	2,954	21.24	19.45	1.79	-0.75	-3.81%	0.24	-3.12	-1.23	-0.28
E2T	1/1/2016	342	9,156	29.83	28.15	1.68	1,357	30.57	28.96	1.61	0.07	0.25%	0.40	0.19	-0.70	0.85
G3_top_up_1	1/1/2016	340	1,888	20.76	19.39	1.37	1,862	21.36	19.45	1.90	-0.54	-2.65%	0.33	-1.63	-1.18	0.11
D1_top_up_1	3/1/2019	339	2,468	19.12	18.42	0.71	756	19.68	19.01	0.67	0.03	0.16%	0.31	0.10	-0.57	0.63
E2O_top_up_1	3/1/2019	342	10,545	30.55	31.05	-0.50	1,715	31.11	31.08	0.03	-0.53	-1.75%	0.27	-1.95	-1.06	0.00
E2T_top_up_1	3/1/2019	340	7,703	22.84	21.76	1.07	1,164	23.02	21.77	1.25	-0.18	-0.77%	0.29	-0.60	-0.75	0.40
G3_top_up_2	3/1/2019	340	1,545	20.03	19.44	0.59	1,557	19.81	19.04	0.78	-0.19	-0.94%	0.34	-0.55	-0.86	0.48
TOTAL		**341**	**96,534**				**30,847**				**-0.45**	**-1.76%**	**0.08**	**-5.40**	**-0.61**	**-0.29**

[1] Only participant and control customers active at the end of 2021 are included.

[2] The Post period includes the relevant portion of 2021 when customers received reports.

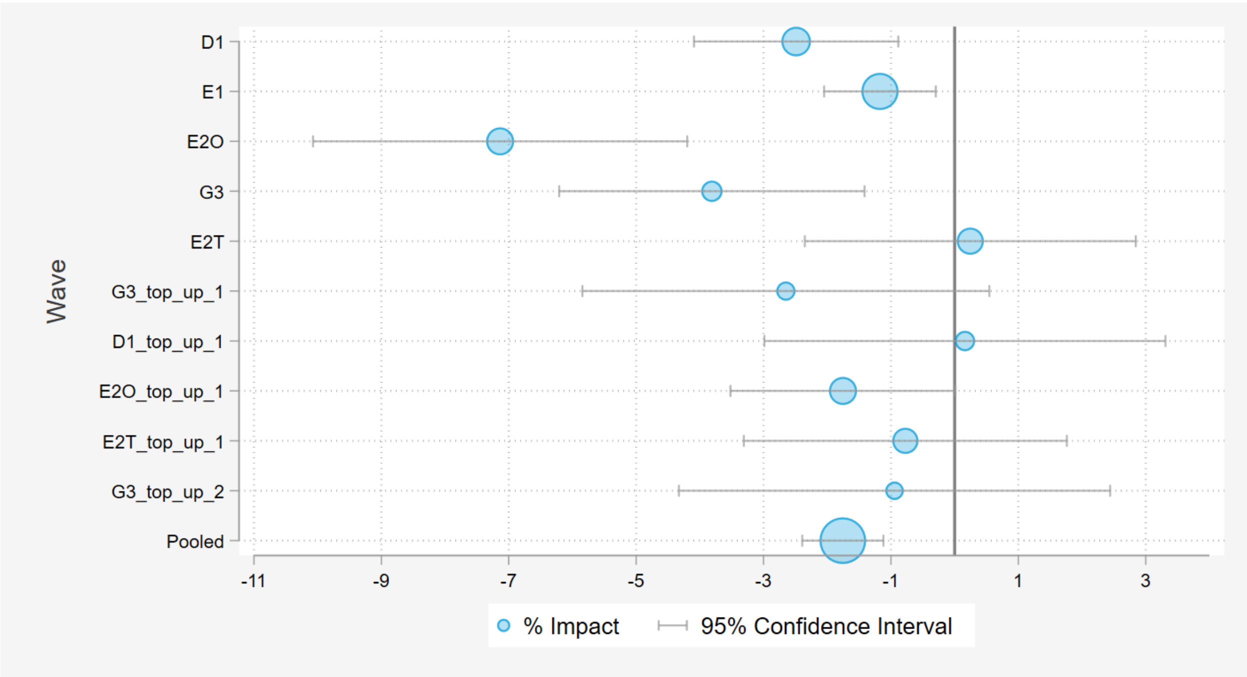
[3] The Pre period is based on data during the year immediately before the experiment launch.

Table 4: Per Customer Annual Savings (kWh) and Program Aggregate Impacts (GWh)

Wave	Experiment Start	Days	Participant Accounts (n)	% Impact	Impact per Customer per Day (kWh)	Annual Impact per Customer (kWh)	Aggregate Impact (GWh)	Aggregate 95% Confidence Interval	t-statistic
D1	4/1/2011	341	13,317	-2.49%	-0.54	-185.20	-2.47	(-4.06, -0.88)	-3.04
E1	4/1/2011	343	36,324	-1.17%	-0.33	-114.61	-4.16	(-7.27, -1.05)	-2.62
E2O	1/1/2013	343	10,543	-7.14%	-1.34	-460.14	-4.85	(-6.85, -2.85)	-4.76
G3	9/1/2014	340	3,045	-3.81%	-0.75	-255.91	-0.78	(-1.27, -0.29)	-3.12
E2T	1/1/2016	342	9,156	0.25%	0.07	25.12	0.23	(-2.20, 2.66)	0.19
G3_top_up_1	1/1/2016	340	1,888	-2.65%	-0.54	-181.94	-0.34	(-0.76, 0.07)	-1.63
D1_top_up_1	3/1/2019	339	2,468	0.16%	0.03	10.48	0.03	(-0.48, 0.53)	0.10
E2O_top_up_1	3/1/2019	342	10,545	-1.75%	-0.53	-180.13	-1.90	(-3.81, 0.01)	-1.95
E2T_top_up_1	3/1/2019	340	7,703	-0.77%	-0.18	-59.58	-0.46	(-1.96, 1.05)	-0.60
G3_top_up_2	3/1/2019	340	1,545	-0.94%	-0.19	-63.61	-0.10	(-0.45, 0.25)	-0.55
TOTAL		**341**	**96,534**	**-1.76%**	**-0.45**	**-153.36**	**-14.80**	**(-20.17, -9.44)**	**-5.40**

Figure 6 shows the percent impacts by wave and the percent impact for all waves pooled. When average usage varies significantly across waves, percent impacts provide meaningful insight by putting the values on a similar scale. The size of the marker indicates the relative participant population size for each wave. There is a savings of 1.76% for the pooled analysis and an aggregate savings of 14,804 MWh across the 96,534 participants.

Figure 6: Electric Percent Impacts by Wave



3.3 MONTHLY IMPACTS

The magnitude of potential savings is a function of how much energy is typically used. Customers have a greater opportunity to reduce use on days when their usage is typically higher. For electric power, peak usage tends to occur during summer cooling and winter heating. Figure 7 shows the savings calculated for each individual month.

Figure 7: Monthly Electric Savings

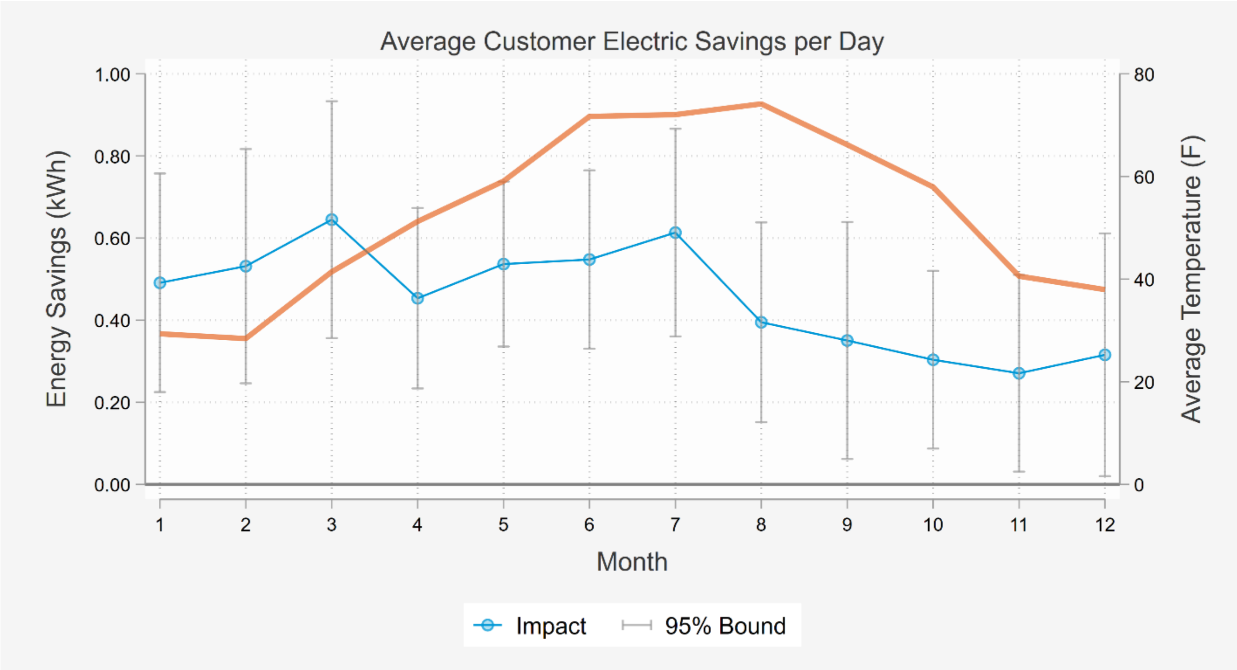
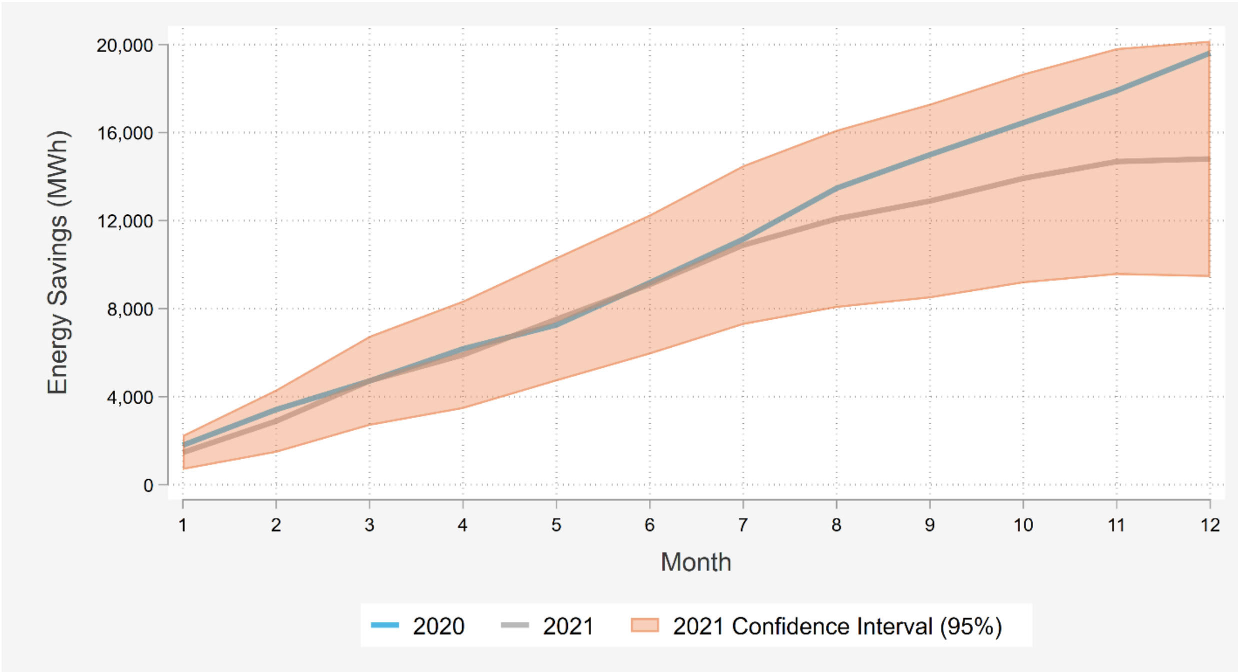


Figure 8 combines the aggregate monthly savings to show the cumulative savings over the course of 2021 and compares it to 2020. The trend starts to deviate between the two years, which is a function of:

1. The reduction of participants (-5.3%) due to natural attrition.
2. The weather in 2021, which had fewer cooling hours and more heating hours than in 2020.

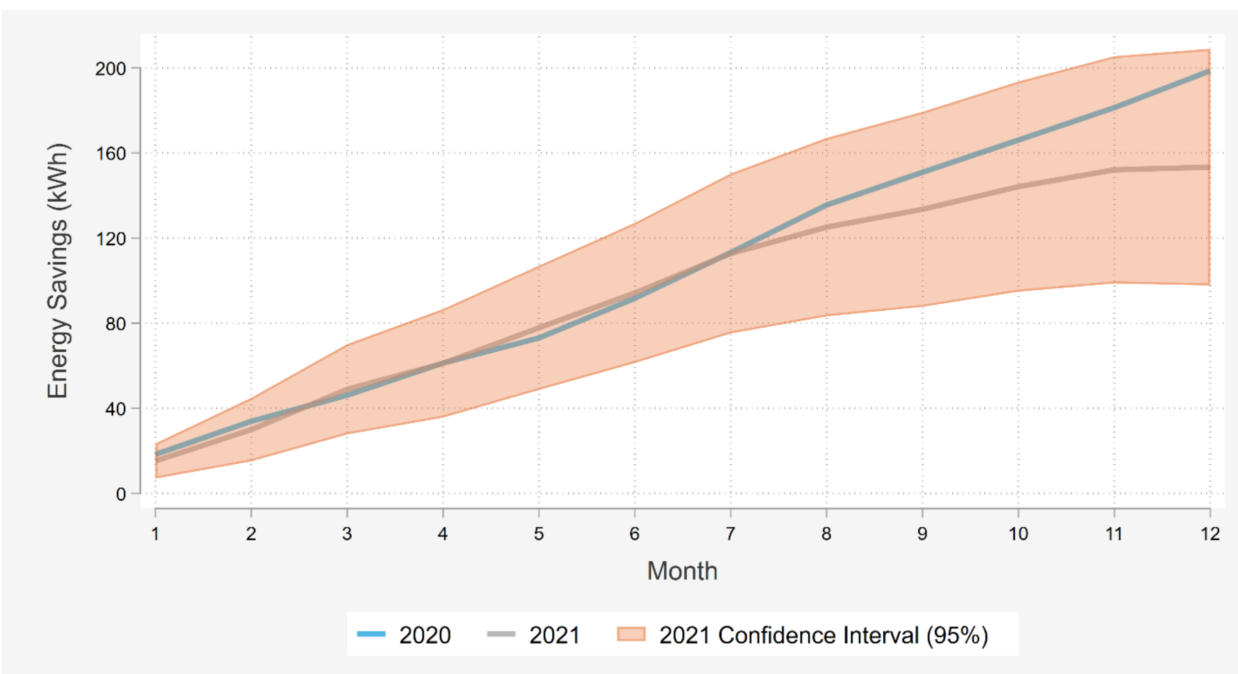
Despite the differences, the 95% confidence interval produced this year still includes the point-estimates reported last year.

Figure 8: Cumulative Aggregate Monthly Electric Savings



The trend in Figure 9, which shows the cumulative per customer savings, controls for customer attrition and depicts a similar trend to that of the aggregate savings. This indicates that the difference between years is more likely a function of weather.

Figure 9: Cumulative Per Customer Monthly Electric Savings



4 GAS IMPACTS

Gas is primarily used for heating in the winter. A few other end-uses use gas as a fuel source, but peak demand for gas is characterized by heating. Because of this, gas use is highly dependent on weather.

The main objectives of this section are to:

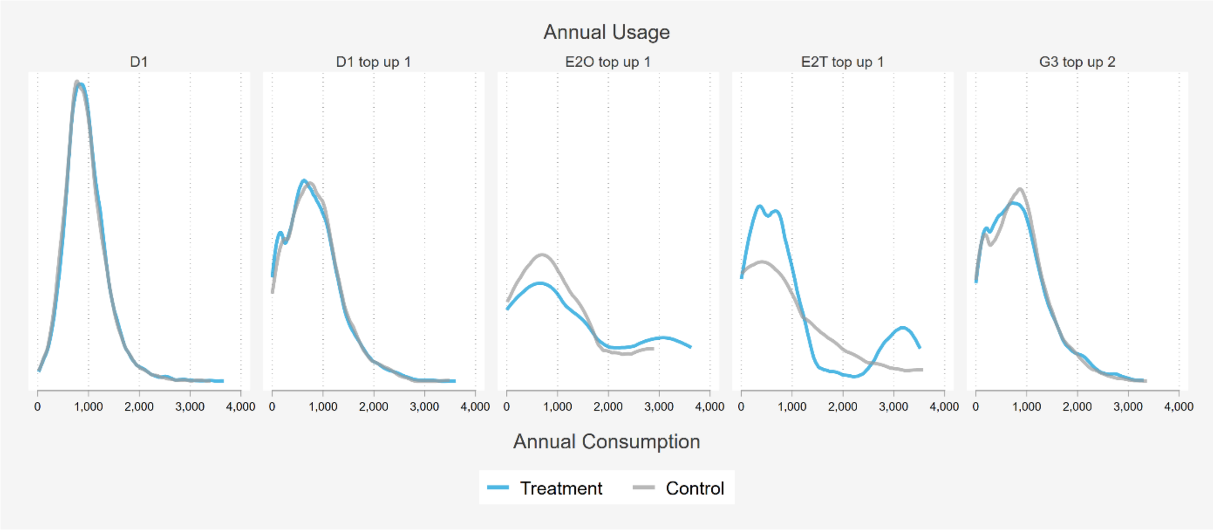
- Present summary statistics and figures of the energy use by participants and the control group before and after the intervention
- Summarize gas energy savings over the evaluation period

Central Hudson lacks pre-intervention data for a subset of gas participants due in part to the transitions of program administration over time. Thus, the impacts are estimated for the participants and controls with pre-intervention data – over 15,000 participants and over 6,000 control customers. The percent energy savings are then applied from the estimating sample to the full number of participants who received home energy reports. The approach aligns with evidence-based practices in the field of medicine where randomized control trials are performed on a subset of the population, and those results are applied to the broader population.

4.1 COMPARISON OF CONTROL GROUP AND PARTICIPANT GAS USE

For each wave of the HER program, pre-treatment energy consumption should be identical. A good control group should use energy in a similar manner as the participants before either group receives the intervention. Figure 10 shows the consumption distribution by wave for the treatment and control groups prior to HER intervention. Treatment and control groups are comparable for the three largest waves – D1, D1 top up 1, and G3 top up 2 – but vary for waves with less than 100 customers in the sample.

Figure 10: Pre-Treatment Annual Gas Consumption by Wave



Similar to the electric participant fluctuations, gas customers may also move and become inactive in the program. Between 2020 and 2021, about 2,000 gas participants deactivated. On average, the annual usage is 3.47% different between the groups, as compared to the 3.25% difference in the 2020 analysis. The only statistically significant difference occurs in the D1 wave, which is the largest of all gas waves.

Table 5: Pre-Treatment Differences

Wave	Start Date	Control (n)	Treated (n)	Annual Use Control	Annual Use Treated	Diff	% Diff	S.E.	t	Lower Bound	Upper Bound
D1	4/1/2011	3,470	11,206	936.6	957.2	20.6	2.20%	8.2	2.50	4.5	36.8
D1 top up 1	3/1/2019	751	2,446	808.3	798.6	-9.7	-1.20%	22.4	-0.43	-53.7	34.2
E2O top up 1	3/1/2019	12	47	1,733.1	1,449.9	-283.2	-16.34%	434.1	-0.65	-1134.1	567.6
E2T top up 1	3/1/2019	12	66	1,010.8	1,363.1	352.3	34.85%	422.6	0.83	-476.1	1,180.6
G3 top up 2	3/1/2019	1,857	1,856	830.9	836.4	5.6	0.67%	18.9	0.29	-31.6	42.7
TOTAL		**6,102**	**15,621**	**890.4**	**921.2**	**30.9**	**3.47%**	**32.0**	**0.96**	**-31.9**	**93.7**

4.2 GAS SAVINGS EX-POST ANALYSIS

The same approach is used for gas impacts as was used in the electric analysis, a simple difference-in-difference model to estimate impacts at the wave level. Table 6 presents the results in energy savings per day. For estimation, we analyzed waves that included at least 100 active customers and had a full year of pre-intervention data. In total, savings estimates were based on 15,500 treatment and 6,100 control customers. Across all estimation waves in the experiment, the energy reduction of 1.76% is statistically significant with a t-statistic of 3.28. We apply the average customer percent reductions to the actual 2021 energy usage for the waves without pre-treatment data.

Table 6: Difference-In-Difference Gas Analysis (Estimating Sample)

Wave	Start Date	Days	Treatment				Control Group				Energy Impact Per Customer Per Day (Ccf/day-account)					
			Estimate Accts	Daily Use Post	Daily Use Pre	Diff	Estimate Accts	Daily Use Post	Daily Use Pre	Diff	Impact (Diff-in-diff)	% Impact	S.E.	t	Lower Bound	Upper Bound
D1	4/1/2011	341	11,283	2.42	2.62	-0.20	3,494	2.42	2.57	-0.15	-0.05	-2.25%	0.02	-3.45	-0.08	-0.02
D1_top_up_1	3/1/2019	339	2,463	1.90	2.19	-0.29	755	1.93	2.21	-0.28	-0.00	-0.11%	0.02	-0.09	-0.05	0.04
G3_top_up_2	3/1/2019	340	1,875	1.99	2.28	-0.30	1,871	1.99	2.28	-0.29	-0.01	-0.26%	0.02	-0.24	-0.05	0.04
TOTAL		**340**	**15,621**				**6,120**				**-0.04**	**-1.76%**	**0.01**	**-3.28**	**-0.06**	**-0.02**

[1] Only participant and control customers active at the end of 2021 are included.

[2] The Post period includes the relevant portion of 2021 when customer received reports.

[3] The Pre period is based on data during the year immediately prior to the experiment launch.

Table 7 shows the annual savings per customer and the aggregate program savings. To arrive at program impacts, for waves included in the estimating sample, we apply the respective percent reductions. For waves without pre-treatment data, we apply the average percent reductions from the waves shown in Table 6. Annual savings were 13.45 Ccf per customer. On an aggregate basis, Central Hudson's behavioral program reduced gas usage by 35,131 MMBtu.

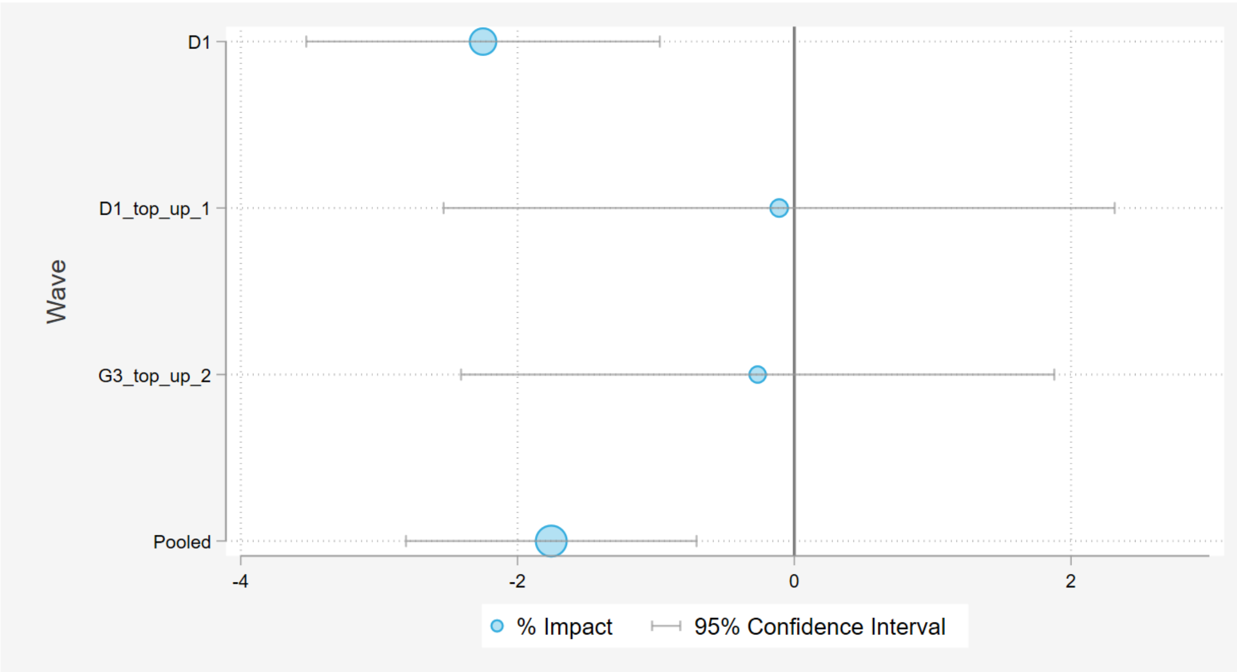
Table 7: Per Customer Annual Savings (Ccf) and Program Aggregate Impacts (MMBtu)

Wave	Experiment Start	Days	Participant Accounts	% Impact	Impact per Customer per Day (CCf)	Annual Impact per Customer (CCf)	Aggregate Impact (MMBtu)	Aggregate 95% Confidence Interval	t-statistic
D1	4/1/2011	341	13,680	-2.25%	-0.05	-18.17	-26,362	(-41,335, -11,389)	-3.45
E1	4/1/2011	.	672	-1.76%	-0.04	-13.45	-947	(-1,513, -381)	.
E2O	1/1/2013	.	665	-1.76%	-0.04	-13.45	-584	(-933, -235)	.
G3	9/1/2014	.	3,910	-1.76%	-0.04	-13.45	-4,713	(-7,530, -1,896)	.
E2T	1/1/2016	.	114	-1.76%	-0.04	-13.45	-131	(-210, -53)	.
G3_top_up_1	1/1/2016	.	2,175	-1.76%	-0.04	-13.45	-2,684	(-4,288, -1,080)	.
D1_top_up_1	3/1/2019	339	2,463	-0.11%	-0.00	-0.71	-180	(-4,168, 3,807)	-0.09
E2O_top_up_1	3/1/2019	.	73	-1.76%	-0.04	-13.45	-131	(-210, -53)	.
E2T_top_up_1	3/1/2019	.	93	-1.76%	-0.04	-13.45	-163	(-260, -66)	.
G3_top_up_2	3/1/2019	340	1,875	-0.26%	-0.01	-1.78	-348	(-3,163, 2,467)	-0.24
TOTAL		**340**	**25,720**	**-1.76%**	**-0.04**	**-13.45**	**-35,131**	**(-56,129, -14,134)**	**-3.28**

[1] Impacts scaled based on D1, D1 top up 1, and G3 top up 2 waves. The waves were employed because they included greater than 100 active participants with the required one-year of pre-period data.

Figure 11 shows the percent impacts by wave and the percent impact for all waves pooled. When average usage varies significantly across waves, percent impacts provide meaningful insight by putting the values on a similar scale. The size of the marker indicates the relative participant population size for each wave. There is a savings of 1.76% for the pooled analysis and an aggregate savings of 35,131 MMBtu across the 25,720 participants.

Figure 11: Gas Percent Impacts by Wave



4.3 MONTHLY IMPACTS

Potential savings is a function of how much energy is being used. Days with higher usage have a greater opportunity to save. For gas, peak usage tends to occur during winter heating. Figure 12 shows the savings calculated for each individual month. Summer months exhibit low or no gas savings. The 95% confidence interval for each month is shown by the gray bars. The results for individual months have wider confidence bands than the results for annual savings, which are more precise.

Figure 12: Monthly Gas Savings

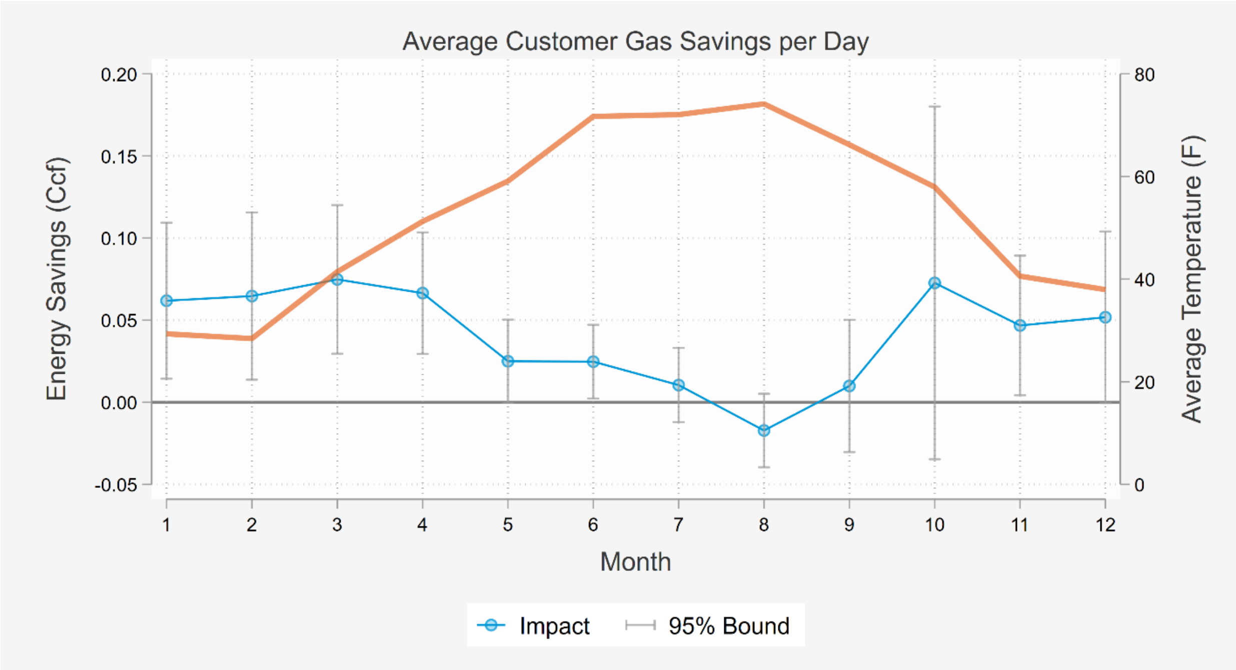
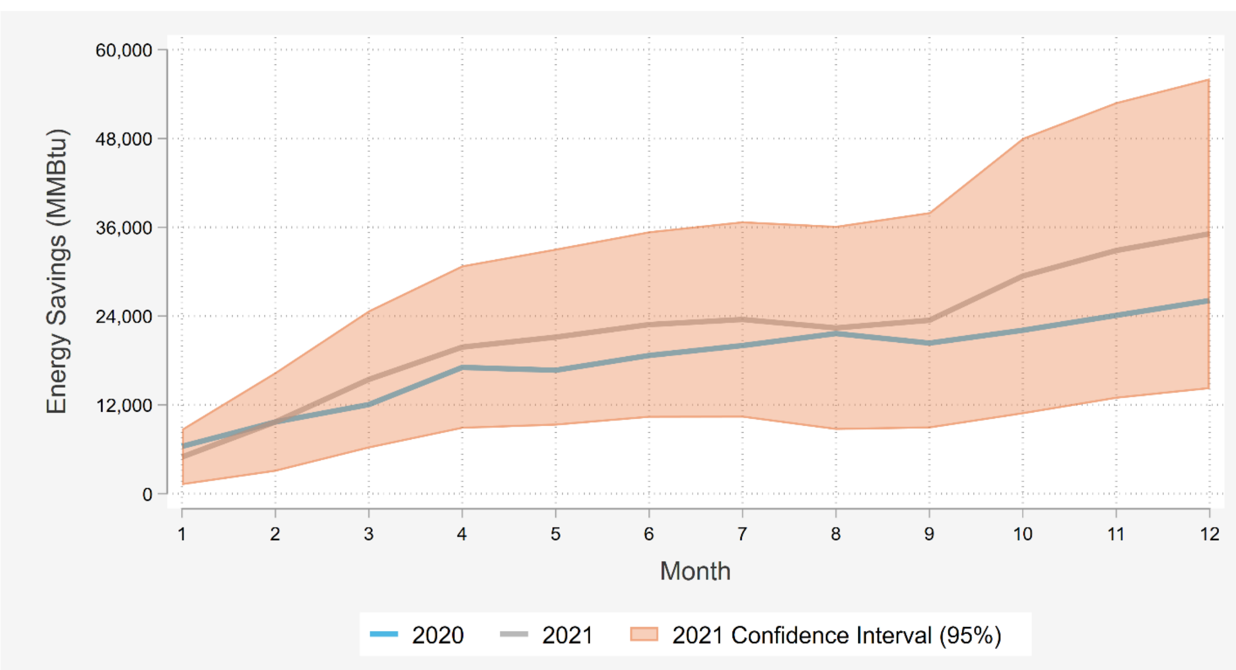


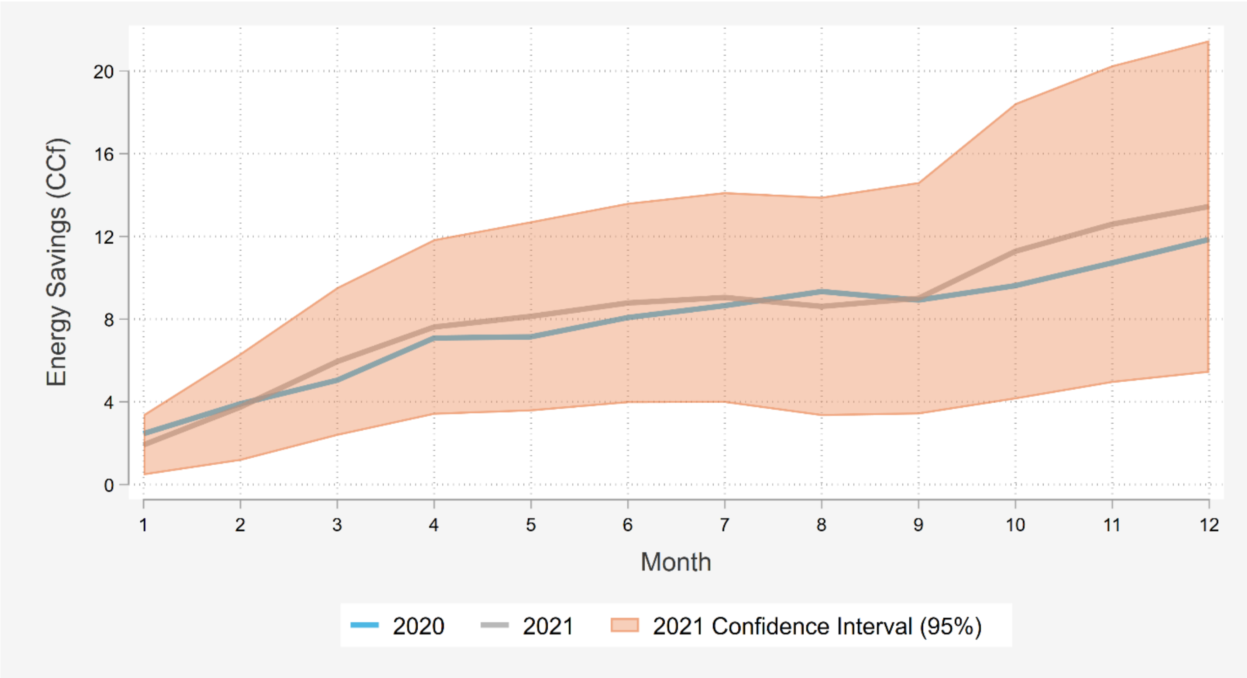
Figure 13 combines the aggregate monthly savings to show the cumulative savings over the course of 2021 and compares it to 2020. The trend starts to deviate between the two years, which is a function of the weather in 2021 having fewer cooling hours and more heating hours than in 2020. Despite the difference, the 95% confidence interval produced this year still includes the point-estimates reported last year.

Figure 13: Cumulative Aggregate Monthly Gas Savings



The trend in Figure 14, which shows the cumulative per customer savings, controls for the attrition of customers and depicts a more similar trend than the aggregate savings.

Figure 14: Cumulative Per Customer Monthly Gas Savings



5 CONCLUSIONS AND RECOMMENDATIONS

Central Hudson's Energy Insights Behavioral program continues to deliver a significant portion of Central Hudson's energy efficiency savings. It currently reaches over 96,000 electric accounts and over 25,000 gas accounts. While home energy reports deliver small percent changes in energy use, they typically yield large aggregate savings because they reach a large number of customers and do not require applications, rebates, or installations. In PY2021, the program yielded, 14,804 MWh of electric savings and 35,131 MMBtu of gas savings. However, while electric savings estimates were precise, gas savings estimates had wide confidence intervals, even though the results were statistically significant.

Starting in 2021, Central Hudson brought the program administration in house and shifted it maintenance mode in order to reduce costs and improve cost-effectiveness. The number of reports send decreases from six to two per year, and the content shifted from comparison of energy usage plus energy tips to energy tips alone. Despite the changes and natural attrition due to household churn, the program delivered energy savings in 2021 comparable to those attained in 2020. Over time, we anticipate the per customer energy savings will decay.

Based on the 2021 impact evaluation, our recommendations include:

- **Revise the earning adjustment mechanism rules.** Notably, the shift in strategy was partly in response to the recent emphasis on lifetime savings in New York policy. Under current rules, Central Hudson can be penalized for delivering energy savings via the behavioral program because it lowers the average expected useful life of the portfolio measures.
- **Decide whether or not to reintroduce energy use comparisons within the next two years.** The behavioral program was initially designed to include energy comparisons and, thus, leverage behavioral psychology and social norms to lower residential energy usage. The program no longer includes the energy comparisons and it is effectively in maintenance mode. It can continue to accrue savings, but at some point, the savings and the ability to detect them will decay.
- **Increase the number of reports.** In 2021, the program administrator switched from Uplight to Central Hudson. With this switch, the amount of reports has decreased from six to two. To continue to produce program savings and mitigate decay, we recommend increasing the number of reports sent to customers.
- **Monitor gas savings closely.** With time, the population of customers receiving the behavioral intervention and the control group is expected to dwindle. As the control group size decreases, the ability to detect statistically significant savings will decay.
- **Continue to analyze the impacts using a randomized control trial.** While the approach requires withholding a subset of customers to serve as controls and provide a baseline, it is necessary because the signal, or the percent savings, is small.

- **Continue to conduct quarterly updates of program impacts.** The cadence allows Central to assess if and how energy savings are changing with the modifications to the behavioral intervention.