|  |  |
| --- | --- |
| To: | Randy Opdyke, Bruce Liu, Rohith Mannam, Nicor Gas; Ted Weaver, First Tracks Consulting; Katie Parkinson, Jane Colby, Scott Dimetrosky, Apex Analytics |
| Cc: | Jennifer Morris; ICC Staff, Illinois TRM Administrator; Celia Johnson, Illinois SAG  Kevin Grabner, Stu Slote, Rick Berry, Charles Ampong; Guidehouse |
| From: | Paul Higgins, Jackson Lines; Guidehouse |
| Date: | April 21, 2021 |
| Re: | Small Commercial Thermostats TRM Research |

# Introduction

At the request of Nicor Gas and ICC Staff, Guidehouse undertook research in CY2020 to support a future update to the Illinois Technical Reference Manual’s (TRM’s) current Small Commercial Thermostat (SCT) measure.[[1]](#footnote-1) The current provisional SCT measure in TRM v9.0 includes a heating reduction parameter of 8.8 percent, representing the average reduction in total building heating energy consumption due to installation of the measure; its value is based on residential advanced thermostats research.[[2]](#footnote-2) Our basic research approach was to apply regression analysis to the monthly gas usage data of this group of customers, obtained from their billing records. This memo summarizes our research method, the data utilized, key results, and recommendations.

Overall, we found savings of either 4 percent or 4.5 percent of all-uses therm consumption due to installation of the programmable thermostat measure, depending on the precise sample used; this corresponds to an average savings of either 0.70 or 0.72 therms per day during the heating season. However, neither result was statistically significant.[[3]](#footnote-3) Converting our all-uses results to the same basis as the provisional heating reduction parameter in the TRM (i.e., expressed as the percent reduction in *heating* energy usage) would require deriving the heating share of total natural gas usage by building type and climate zone during the heating season, and using the heating shares to convert all-use savings to heating energy savings.[[4]](#footnote-4) While these results are insufficiently reliable to justify replacing the current provisional heating reduction parameter in the provisional TRM measure, they are broadly consistent with it:

* Given a 212-day heating season (October 1 – April 30), our results imply an average savings of roughly 150 therms per year; the current provisional TRM v9.0 measure yields savings of 148 therms under typical assumptions for the TRM algorithm below.[[5]](#footnote-5)

*ΔTherms = (EFLHheat \* Capacity \* 1/AFUE \* Heating\_Reduction)/ 100,000Btu/Therm*

*Where:*

*Capacity = Nominal Heating Input Capacity (Btu/hr) of heating system*

*AFUE = Annual Fuel Utilization Efficiency Rating*

*EFLHheat = Heating mode equivalent full load hours in Existing Buildings are provided in section 4.4 HVAC End Use.*

*Heating\_Reduction = Assumed percentage reduction in total building heating energy consumption due to thermostat*

* Assuming an average 80 percent heating share of total gas usage during the heating season, our results imply a heating reduction of 5 to 5.6 percent.

Note that virtually all of the customers in our research sample had installed programmable thermostats, not advanced (“smart”) thermostats. There was insufficient data available to estimate a separate impact parameter for advanced thermostats in this study.

# Research Method

The goal of this research was to estimate the average natural gas heating savings parameter in the provisional SCT measure (measure 4.4.48) of TRM v9.0 based on the observed experience of Illinois small commercial customers who installed a programmable or advanced thermostat. Our basic approach was to apply regression analysis to the monthly gas usage data of this group of customers, obtained from their billing records, before and after the installation of the SCT measure, together with corresponding data from a group of matched controls, during the months during which the predominant impact of installing this measure on energy usage would result from reduced heating demand.[[6]](#footnote-6)

# Data

For this research we identified a potential sample of 275 SCT projects from three sources: the Nicor Gas Small Business Energy Savings (SBES) Program, the Nicor Gas Business Energy Efficiency Rebates (BEER) Program, and the ComEd AirCare Plus (ACP) Program. For the Nicor Gas programs, we selected participants that installed commercial programmable thermostats as the only measure in GPY4, GPY5, GPY6, 2018, or 2019 (prior to September 1, 2019 to permit the use of the 2019/20 heating season); we included the BEER program because we deemed the size of the sample from the SBES program to be too small by itself, and the eligibility of the BEER thermostat rebate is targeted to small commercial type heating systems. The ACP sites consisted of Nicor Gas customers that had installed a programmable or advanced thermostat (and no other measure) through ComEd’s ACP Program.[[7]](#footnote-7) To be included, customers had to have billing data sufficient to cover at least one complete heating season from both the pre-install and post-install heating periods. The research sample also included a pool of 6,576 potential matched controls – small business customers that had not participated in a Nicor Gas rebate program. After data cleaning and matching, the final analysis data set consisted of 152 treatment customers and 6,559 potential matches (146 of whom were selected as matches). The cleaning steps we took, and their impacts on the sample, are shown in Table 1.

Table 1. Effects of Data Cleaning Steps on Sample Size



*Source: Nicor Gas data and Guidehouse analysis.*

The final research sample comprised a variety of building types, as shown in Table 2.

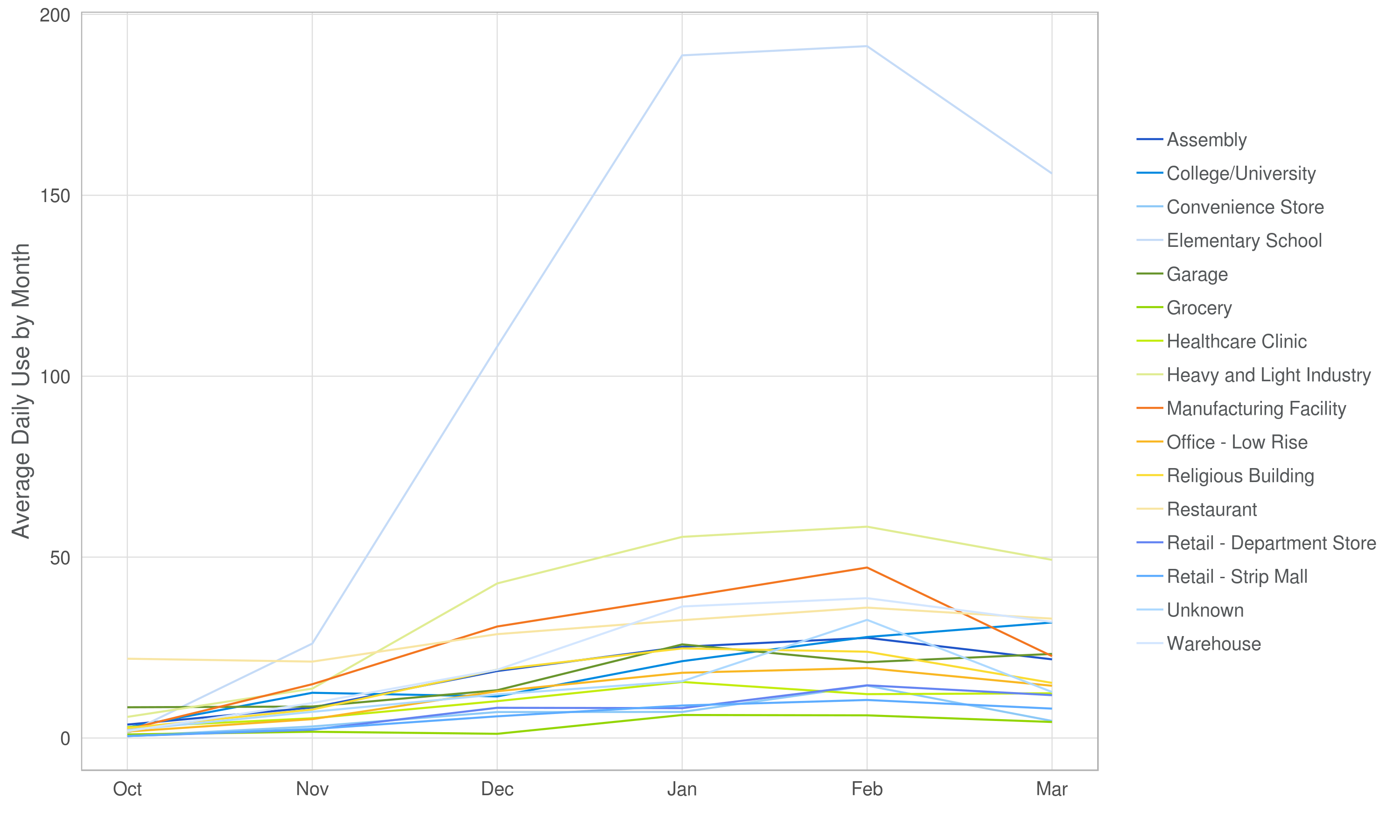
Table 2. Distribution of Treatment Customers by Building Type



*Source: Nicor Gas data and Guidehouse analysis.*

Out of concern that some building types might have usage patterns that were different enough to justify doing a separate analysis, we examined the average daily use by month for each building type. With the exception of Elementary School the usage patterns were broadly similar. Schools exhibited a much more peaked pattern, as shown in Figure 1.

Figure 1. Average Daily Use by Building Type



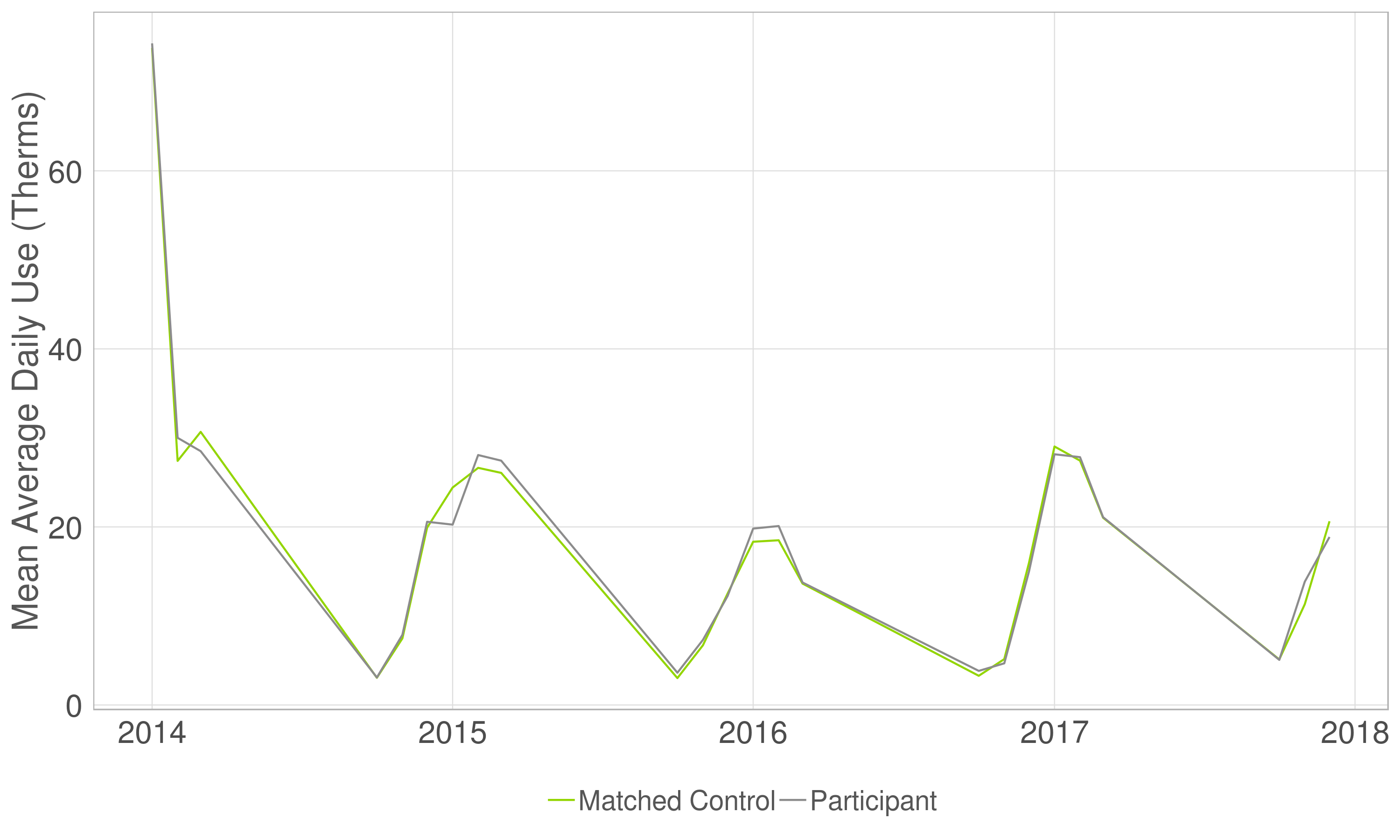
For this reason, in addition to analyzing the full sample we also ran a subsidiary analysis consisting of applying the same model to a sample from which the Elementary School data were dropped, to check whether the results were substantially different.

# Analytical Approach

Ideally, a randomized control trial (RCT) is preferred when using individual customer usage data to measure the impacts of an energy efficiency program;[[8]](#footnote-8) however, the opt-in nature of the SBES and BEER Programs made that approach infeasible. Instead, we employed a quasi-experimental design that compares the pre- to post-install changes in energy consumption of program participants to those of a set of matched non-participants using regression analysis. This method, known as regression with pre-program matching (RPPM), is described in Ho, Imai, King, and Stuart.[[9]](#footnote-9)

The matching method we used relies on gas usage data pulled from each customer’s monthly heating-season bills, divided by the number of days in the billing period to put them on a common (therms per day) basis. We compared the monthly average daily therm consumption values of each participant during the heating season in their pre-install (“match period”) year to that of all customers in the pool of potential matches over the same period. For each comparison, we calculated the difference in average daily therm use in each month, *DPM* (**D**ifference between **P**articipant and potential **M**atch). The quality of the potential match was indicated by the Euclidean distance, or sum of squared differences, between the participant’s monthly usage values and those of the potential match calculated over the matching period. Denoting the sum of squared *DPM* over the matching period by SSD (**S**um of **S**quared **D**istance), the match quality was defined as . The non-participant whose energy usage minimized this distance during the participant’s pre-install period was chosen as the match for that participant. Matching was done with replacement.[[10]](#footnote-10) In selecting the matched controls, we required them to have at least an 80 percent overlap of usable observations in the matching period with those of the relevant treatment customer. Mean average daily therm use is shown in Figure 2 for the treatment and control groups.

Figure . Mean Average Daily Use, Treatment and Control



Once the matches were selected, we applied the regression model shown in Equation 1 to the post-install usage values of participants and their matched controls from each participant’s install date through the end of the 2020 heating season.

Equation 1. Lagged-Dependent Variable (LDV) Regression Model

where:

is the average daily use, in therms, of customer *k* during billing period *t* of the post-install period

denotes whether customer *k* is a participant (=1) or a matched control (=0)

is customer *k*’s average daily use in the same period during the pre-install year

comprises a set of binary variables indicating the year-month combination *j* into which the current observation (indexed by *t*) falls

is a cluster-robust disturbance term for customer *k*

In the above model, , the regression coefficient on the variable, estimates the average difference in average daily therm use between the treatment and control groups in the post-install period. To convert to percent energy savings this value was divided by average participant daily usage during the post-install heating season plus the estimated daily therm savings.

# Key Results

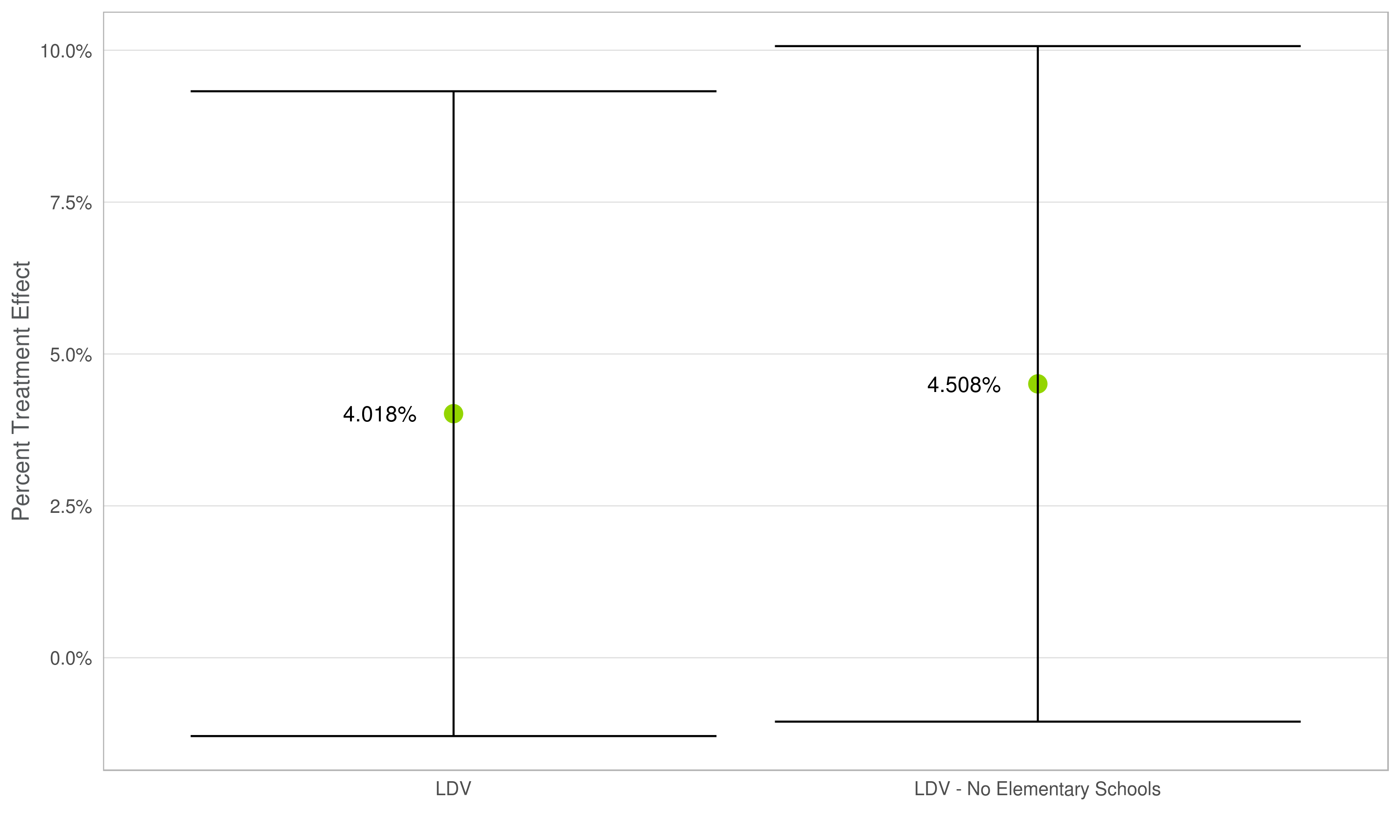
The key results of our analysis are shown in Table 3 and Figure 3. The average daily heating season gas savings from installing an SCT was 0.7 therms per day, or 4 percent of baseline heating season gas usage using the full sample. Dropping schools caused only minimal changes. Both point estimates have wide 90 percent confidence intervals owing to the small size of the final sample size, which indicates that they are not statistically different from one another.

Table 3. Key Regression Results



Source: Nicor Gas data and Guidehouse analysis.

Figure 3. Savings Percentages with Confidence Intervals



# Recommendations

We found savings of 4 percent whole-building, all-uses natural gas consumption using the full sample, or 4.5 percent when elementary schools were dropped. Neither result was statistically significant. Given the broad, overlapping confidence intervals of the two estimates there is no basis for reporting a separate estimate for Elementary Schools, or indeed for changing the current TRM parameter value.

While the present results do not provide a sufficient basis for changing the current provisional heating\_reduction parameter in the TRM v9.0, as we noted at the beginning of this memo, they are broadly consistent with the current parameter value. For this reason, we recommend retaining the current heating\_reduction parameter value.

To continue the SCT heating savings TRM research, we recommend that you consider performing a small pilot experiment focused on measuring the impact of *advanced* thermostats installed in small commercial buildings. To improve the statistical rigor of the resulting savings estimates, we recommend designing the pilot as a randomized experiment in which representative samples of small business customers are offered an efficient thermostat in two waves, one group receiving it in the first year and the other group receiving it in the second. This would obviate the need for drawing a set of matched controls while delivering a result that limits the possibility of bias.

1. The IL TRM v8.0 retired the previous Small Commercial Programmable Thermostats, Small Commercial Programmable Thermostat Adjustments, and Advanced Thermostats for Small Commercial measures (measures 4.4.18, 4.4.25, and 4.4.42, respectively) and replaced them with a single provisional Small Commercial Thermostats measure (measure 4.4.48). The measure remains provisional in v9.0 (see IL TRM v9.0, vol. 2, pp. 442-444). Guidehouse estimated an analogous cooling reduction parameter in research performed for ComEd, which was incorporated into the IL TRM v9.0. [↑](#footnote-ref-1)
2. As noted in footnote 811 on p. 442, this value is “Assumed equal to [the] assumption for Residential Advanced Thermostats.” [↑](#footnote-ref-2)
3. The 4 percent savings estimate, which corresponds to an average usage reduction of 0.70 therms/day, was based on the full sample, while the 4.5 percent estimate, corresponding to an average reduction of 0.72 therms/day, excluded school buildings from the analysis due to their qualitatively different usage pattern. (See below for more discussion of the research samples.) The 90 percent confidence intervals of both savings estimates encompass positive, negative and zero savings, indicating that they are not significantly different from zero, or one another, at a 90 percent level of confidence. [↑](#footnote-ref-3)
4. To estimate the heating share of whole building heating-season gas usage, one could use eQuest models for each relevant building type in each of Illinois’ 5 climate zones (see IL TRM v9.0, Volume 2, p. 182 ). The Nicor Gas technical potential model may also be a source of reasonable heating shares. [↑](#footnote-ref-4)
5. If we assume the heating mode equivalent full load hours for the “Unknown” building type (1,678 hours), which is the average over the known building types, a furnace capacity of 80,000 Btu/hr, and an AFUE of 0.8, the 8.8 percent heating\_reduction parameter value in the current provisional measure yields savings over the heating season of 147.7 therms. Using 2020 tracking data with TRM v9.0 percent heating savings, Nicor Gas input parameters produced a range of gross therms saved for this measure of 40-311 therms, with an average of 198 therms for 22 installations. The average was skewed by several large manufacturing facility units; dropping those lowered the average to 171 therms. Peoples Gas and North Shore Gas are using 196 therms and 200 therms, respectively, for 2021 default tracking savings. [↑](#footnote-ref-5)
6. We assumed the heating season to run from October 1 through April 30. As noted in the Introduction section, and in the Recommendations section below, additional research would be required to convert our key findings to a percent-heating-usage reduction basis. [↑](#footnote-ref-6)
7. We restricted the sample to customers that had installed SCTs as the sole measure to avoid the problem of distinguishing the effect of the thermostat from those of other measures. Nicor Gas was not able to match any of the ACP customers to their accounts. [↑](#footnote-ref-7)
8. In an RCT customers from the population of interest would be randomly assigned either to the treatment group, and thus receive the treatment, or to the control group and thus not receive it. When this is done, the randomization ensures that all potentially confounding factors that could bias the causal analysis are controlled for. [↑](#footnote-ref-8)
9. Daniel Ho, Kosuke Imai, Gary King, Elizabeth A. Stuart, “Matching as Nonparametric Preprocessing for Reducing Model Dependence in Parametric Causal Inference,” *Political Analysis* (2007) 15: 199-236. Downloadable at *https://gking.harvard.edu/files/matchp.pdf*. As the title of the article suggests, using a matched control group reduces the possibility that the results will depend on the particular model specification used, but it does not fully eliminate the possibility of bias. See also Guido W. Imbens and Donald B. Rubin, *Causal Inference for Statistics,* *Social and Biomedical Sciences: An Introduction,* Cambridge University Press 2015; Paul J. Gertler et al., *Impact Evaluation in Practice,* International Bank for Reconstruction and Development 2011; and Joshua D. Angrist and Jörn-Steffen Pischke, *Mostly Harmless Econometrics: An Empiricist’s Companion,* Princeton University Press 2009. [↑](#footnote-ref-9)
10. Matching with replacement means that the same matched control customer may be matched to more than one participant, and thus that there may be fewer (unique) matched controls than participants, as was the case here. If a match is selected twice as a result of matching with replacement, it is weighted as such. [↑](#footnote-ref-10)