Guidehouse

Memorandum

To: Erin Daughton; ComEd
Cc: Jennifer Morris; ICC
From: Paul Higgins, Dustin Kunkel; Guidehouse
Date: May 26, 2020
Re: Small Commercial Thermostats TRM Research

INTRODUCTION

At the request of ComEd and ICC Staff, Guidehouse undertook research in CY2019 to support a future update to the Illinois Technical Reference Manual's (TRM's) current Small Commercial Thermostat (SCT) measure.[1] The current provisional SCT measure in the TRM includes a cooling reduction parameter of 8 percent, representing the average reduction in total building cooling energy consumption due to installation of the measure; its value is based on residential advanced thermostats research.[2] This memo summarizes our research method, the data utilized, key results, and recommendations.

Overall, we found statistically significant savings of 10.4 percent of whole building electric energy consumption from programmable thermostats, equivalent to 29.7 percent of cooling energy savings, although the precision of the estimate was low due to the small size of the final sample we analyzed.[3] However, the point estimate of 10.4 percent savings is not robust to minor tweaks to our modeling assumptions. When we constrained the algorithm we used to select the baseline ("matched") comparison group in our analysis so that it required a minimum of 60 percent overlap of the available days of usable data during the matching period between a given treatment customer and each potential match, the estimated whole-building savings dropped to 2.0 percent whole building, or 5.6 percent cooling load, savings. Given the relatively low precision of these estimates, and their sensitivity to minor variations in the modeling assumptions, we recommend replacing the provisional cooling reduction value in the TRM with the midpoint of the range spanned by these values – 6.2 percent of overall savings (midpoint of 2.0 and 10.4) or, equivalently, 17.7 (midpoint of 5.6 and 29.7) percent of cooling savings – in view of the uncertainty surrounding each value.

---

[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). See IL TRM v8.0, vol. 2, pp. 426-429.
[2] As noted in footnote 760 on p. 428, this value is "Assumed equal to [the] assumption for Residential Advanced Thermostats."
[3] The 90 percent confidence interval (90% CI) on this estimate is 10.4% ± 8.8%, or [1.6%, 19.2%].

230 Horizon Drive
Suite 101-B | Verona, WI 53593
Main No.: 608.616.5970
guidehouse.com

To convert our whole-building results to the same basis as the current TRM measure cooling energy savings of 8 percent, we assumed an average 35 percent share of summer energy usage for cooling.[4]

Note that our research was on customers who installed programmable thermostats, not advanced ("smart") thermostats. There was insufficient data available to include advanced thermostats in this study.

Research Method

The goal of this research was to develop an estimated rate of cooling energy savings comparable to the cooling reduction parameter in the provisional Small Commercial Thermostats measure (measure 4.4.48) of TRM v8.0 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 daily electric energy use data of this group of customers before and after the installation of the SCT measure, together with corresponding data from a group of matched controls, during the summer months during which the predominant impact of installing this measure on energy usage would result from reduced cooling demand.[5]

Data

To maximize the odds that our analysis would be able to distinguish the impact of an installed thermostat measure from those of other EE measures, we restricted the data we used for the analysis to ComEd small business customers that had installed *only* an SCT measure through a ComEd EE program during the sample period; had usable AMI usage data sufficient to cover at least one complete cooling season, from both the pre-install and post-install cooling periods; and did not have net metering.[6]

We used data from the small business customers who installed an SCT (and no other measures) through the AirCare Plus (ACP) Program in PY8 and PY9[7] for whom sufficient AMI data were available.[8] Note that these customers all installed programmable thermostats, not advanced ("smart") thermostats. There was insufficient data available to include advanced thermostats in this study. While the initial sample included 443 ACP participants and a pool of 5,416 potential matched controls,[9] after data cleaning the final analysis data set consisted of 89 small business participant customers and 5,336 potential matches (52 of whom were selected as matches). The cleaning steps we took, and their impacts on the sample, are shown in Table 1.

[4] To estimate the cooling share of whole building summer energy usage, we used the eQuest models on which the previous programmable thermostat algorithm was based, performing a model run for each relevant building type in each of Illinois' 5 climate zones (see IL TRM v7.0, Volume 2, p. 232). The baseline run in each model was used to generate the "Monthly Energy Consumption by Enduse". The modeled cooling shares by building type in climate zones 1 and 2 (which comprise virtually all of ComEd's service territory) ranged from 28% to 53%. Since our final regression sample contained too few participant customers to support separate estimates for each building type-climate zone combination, we took the mean over all building types in climate zones 1 and 2, which was 35%.

[5] We assumed the summer cooling season to run from June 1 through August 31. As noted in the recommendations section below, additional research will be required to update the impact of the SCT measure on natural gas heating usage.

[6] In the final data set used for the regression analysis, 83 participants were listed in the tracking file as receiving a thermostat replacement only, another 5 as receiving a replacement thermostat plus an adjustment, and one as receiving a thermostat adjustment only. We chose to include all participants in our analysis to maximize the sample size. Dropping the adjustment-only customer changed the savings from 10.4 percent to 10.6 percent.

[7] PY8 began June 1, 2015 and ended May 31, 2016; PY9 began June 1, 2016 and ended December 31, 2017.

[8] The ACP participants we used all installed a programmable thermostat, and no other EE measure, during PY8 or PY9. We were unable to use customers who received SCTs through ComEd's Small Business Offering (SBO) because in that program the measure was always installed in combination with other energy efficiency measures, which would have swamped the effect of the SCT and made it less likely that we would have achieved a statistically meaningful savings estimate.

[9] The pool of potential matched controls consisted of ComEd small business customers who had not received an SCT through either the ACP or SBO program.

Table 1. Effects of Data Cleaning Steps on Sample Size

Step Name	Counts		Percent Change	
	Participant	Non-Participant	Participant	Non-Participant
Raw data	443	5,416	NA	NA
Drop participants w/o valid tracking*	270	5,416	39.1%	0.0%
Drop customers w/ net metering	270	5,341	0.0%	1.4%
Drop participants w/ install outside pre/post window†	156	5,341	42.2%	0.0%
Drop participants w/ only pre or post data‡	119	5,341	23.7%	0.0%
Drop outliers§	117	5,336	1.7%	0.1%
Drop participants w/o matching period data	89	5,336	23.9%	0.0%
Matched data	89	52	0.0%	99.0%

NA = Not applicable
* Removes participants who did not appear in the ACP tracking file after filtering to only thermostat install/adjustment.
† Measure install must have occurred between 2016-09-01 and 2018-05-31 to ensure data available for both a pre- and post-install summer.
‡ Participants must have cooling season data both before and after their measure install date.
§ Outliers are defined as any customer with daily usage over 1,500 kWh.
Source: ComEd data and Navigant analysis.

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;[10] however, the opt-in nature of the ACP Program 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.[11]

The matching method we used relied on energy usage data at 30-minute intervals, rolled up to daily totals, obtained from the meters of program participants as well as from a set of 5,416 non-participant small business customers from which the matches were drawn. For each participant, we compared the daily energy consumption during the cooling season in the pre-install year to that of all customers in the pool of potential matches over the same period. For each comparison, we calculated the difference in daily energy use, D_{PM} (**D**ifference between **P**articipant and potential **M**atch). The quality of the potential match was indicated by the Euclidean distance between their usage and that of the participant calculated over the matching period. Denoting the sum of squared D_{PM} over the matching period by SSD (**S**um of **S**quared **D**istance), the match quality was defined as $\sqrt{SSD}$. The non-

[10] In an RCT customers from the population of interest, in this case ComEd small business customers, 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.

[11] 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.

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.[12]

We ran the matching algorithm in two different ways. First, we ran it with no constraints on how many days of usable data a given participant and each potential match had to have overlapping during the match period. Second, we ran a constrained version where we required a minimum of 60 percent overlap of the available days of usable data during the matching period between a given participant and each potential match.[13]

Once the matches were selected, we applied the regression model shown in Equation 1 to the post-install[14] usage values of participants and their matched controls from each participant's install date through the end of CY2018.

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

$$DailykWh_{kt} = \beta_1 Treatment_k + \beta_2 DailykWh\_lag_{kt} + \sum_j \beta_{3j} Year\_Month_{jt} + \sum_m DOW_{mt} + \varepsilon_{kt}$$

where:

$DailykWh_{kt}$	is the daily kWh used by customer k on day t of the post-install period
$Treatment_k$	denotes whether customer k is a participant (=1) or a matched control (=0)
$DailykWh\_lag_{kt}$	is customer k's average daily kWh used on the corresponding day of the week in the pre-install year
$Year\_Month_{jt}$	comprises a set of binary variables indicating the year-month combination j into which the current observation (indexed by t) falls
DOW_{mt}	comprises a set of binary variables indicating which day of the week m into which the current observation (indexed by t) falls
ε_{kt}	is a cluster-robust disturbance term for customer k

In the above model, β_1, the regression coefficient on the $Treatment_k$ variable, estimates the average difference in daily energy 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 cooling season plus the estimated daily kWh savings. This produces the percentage of whole home savings.

Key Results

The key results of our analysis are shown in Table 2. The average daily cooling season energy savings from installing an SCT was 19.7 kWh per day, or 10.4 percent of baseline cooling season energy usage, using the minimally-constrained matching algorithm, and 3.4 kWh per day, or 2.0 percent of cooling season energy usage, using the matching algorithm that requires a potential match to have at least 60 percent of the same days of the matching period as the participant to which it is being compared. The minimally-constrained result is statistically different from zero (p-value=0.053), while the result using the 60 percent matching overlap floor is not (p-value=0.662). Both point estimates have wide 90 percent confidence intervals owing to the small size of the final sample size.

[12] 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.
[13] More details of this constraint are provided in the next section.
[14] Note that customers' pre-install usage data does enter the analysis via the $DailykWh\_lag_{kt}$ term on the right-hand side of Equation 1.

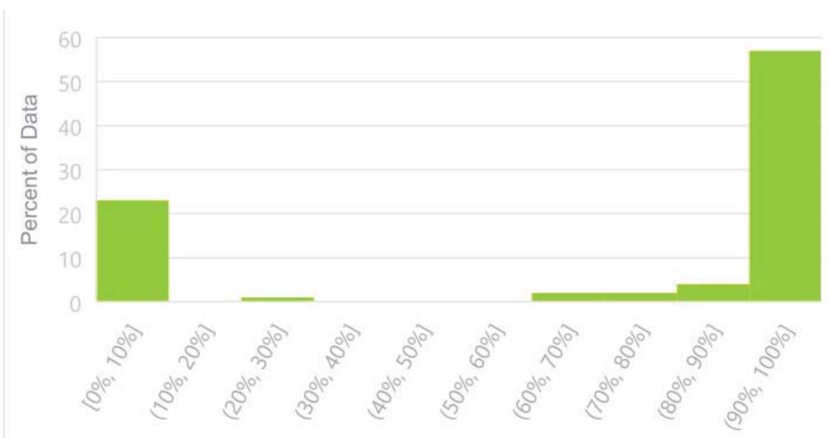
Table 2. Key Regression Results

Model	Savings		P-Value	90% C.I.		Analysis Sample Size	
	kWh/Day	Percent		Lower	Upper	# Trts	# Ctrls
Matched LDV	19.7	10.4%	0.053	1.6%	19.2%	89	52
Matched LDV w/60% Match Floor	3.4	2.0%	0.662	-5.4%	9.4%	89	84

Source: ComEd data and Navigant analysis.

We did not arrive at the 60 percent matching period overlap cutoff arbitrarily. Figure 1 shows the distribution of matches chosen using the minimally-constrained matching algorithm by the percent of days missing from the matching period. It indicates that there is a natural break in the data at a 60 percent overlap – that is, most of the matches chosen using the matching algorithm described above were selected based on a common set of days comprising at least 60 percent of the available days during the participant's pre-install year.[15] However, a significant minority of matches were based on significantly less than a 60 percent overlap. The second savings estimate ("Matched LDV w/60% Match Floor") requires matches for each participant to have at least a 60 percent overlap of days during the participant's match period.

Figure 1. Distribution of Matches by Percent of Days of Overlap in the Match Period



Source: ComEd data and Navigant analysis.

To make a valid comparison of our research value and the provisional value found in the TRM v8.0, we needed to convert the values shown in Table *2*, which represent the reduction in *whole building* energy usage, to the equivalent reduction in *cooling* usage. This was done by dividing each estimate by a representative cooling load percentage for small business customers in ComEd's service territory.

To estimate the average cooling share of total cooling season energy consumption for small business customers, we relied on the eQuest models that were used to develop the programmable thermostat algorithm used in the TRM v7.0.[16] Separate models were run for each relevant building type in each of the five climate zones in Illinois. For each model run, the time period was restricted to June 1 through August 31. The resulting cooling shares of total summer energy use are shown in Table 3.

---

[15] We do not describe this constraint as requiring that a potential match must have a specific number of days of usable data during the match period because the absolute number of available match days differs by participant. The requirement is that any potential match must have at least 60 percent of the number of days in the match period that the participant has available.
[16] IL TRM v7.0, op. cit.

Table 3. Modeled Summer Cooling Shares of Whole-Building Energy Consumption

	Climate Zone				
Building Type	CZ1	CZ2	CZ3	CZ4	CZ5
Assembly	53%	53%	59%	63%	63%
Convenience Store	31%	31%	34%	38%	38%
Office - Low Rise	32%	32%	35%	37%	37%
Fast Food Restaurant	32%	33%	39%	42%	42%
Full Service Restaurant	32%	33%	36%	39%	39%
Religious Facility*	NA	NA	NA	NA	NA
Retail - Strip Mall	33%	33%	40%	44%	45%
Retail - Department Store	28%	29%	33%	36%	37%
Average	**35%**	**35%**	**39%**	**43%**	**43%**

* We were unable to obtain results for the Religious Facility building type because the available version did not have any cooling equipment built into the model.
Source: eQuest model runs

The final regression sample included participants in a range of building types, as shown in Table 4. However, the counts of sample participants in each category are too small to support estimating separate savings values for each building type, so we chose the average value of 35 percent cooling share across all available building types in climate zones 1 and 2 as representative of the small business customers in ComEd's service territory.[17]

Table 4. Counts of Building Types in Final Regression Sample

Building Type	N
1 Convenience Store	5
2 Office - Low Rise	22
3 Religious Worship / Church	27
4 Restaurant - Fast Food	3
5 Restaurant - Full Service	14
6 Retail / Service - Department Store	4
7 Retail / Service - Strip Mall	16
Total*	**91**

* Counts sum to more than 89 because some participants in the final sample listed multiple building types.
Source: Sample data.

Dividing the average whole-building summer energy savings by the representative summer cooling load share yields the equivalent cooling energy savings, as shown in Equation 2.

[17] The final regression sample contained 50 unique ZIP Codes, 2 in climate zone 1 and 48 in climate zone 2.

Equation 2. Derivation of Cooling Energy Savings Estimate

$$\frac{\%\Delta E_{WB}}{Cooling_{Share}} = \begin{cases} \frac{10.4\%}{0.35} = 29.7\% \text{ for the minimally constrained matching algorithm} \\ \frac{2.0\%}{0.35} = 5.6\% \text{ for the matching algorithm with a 60\% overlap floor} \end{cases}$$

where:

$\%\Delta E_{WB}$	is the research whole-building energy savings estimate
$Cooling_{Share}$	is the representative small business summer cooling load share

To translate the standard error associated with the whole-building savings estimate to the equivalent cooling energy savings estimate, we relied on a standard theorem in statistics for obtaining the variance of a linear function of a random variable,[18] $V(aX) = a^2V(X)$, where a is a constant and X is a random variable, as shown in Equation 3.

Equation 3. Derivation of Cooling Energy Savings Standard Error

$$SE\left(\frac{\widehat{\beta_1}}{.35}\right) = \sqrt{\left(\frac{1}{.35}\right)^2 \cdot V(\widehat{\beta_1})}$$

$$= \frac{1}{.35} \cdot SE(\widehat{\beta_1})$$

$$= \begin{cases} 15.3\% \text{ for the minimally constrained matching algorithm} \\ 12.9\% \text{ for the matching algorithm with a 60\% overlap floor} \end{cases}$$

where $\widehat{\beta_1}$ is the estimate of the coefficient on $Treatment_k$ from Equation 1 and $SE(\widehat{\beta_1})$ is the associated standard error. Thus, the 90% CI on the estimated cooling energy savings are as shown in Equation 4.

Equation 4. Derivation of 90 Percent Confidence Interval on Cooling Energy Savings

$$90\%CI = Estimate \pm 1.645 \cdot SE(Estimate)$$

$$= \begin{cases} 29.7\% \pm 25.2\% \text{ for the minimally constrained matching algorithm} \\ 5.6\% \pm 21.2\% \text{ for the matching algorithm with a 60\% overlap floor} \end{cases}$$

where $Estimate$ is the estimated cooling energy savings and 1.645 is the Z-score for a 90 percent confidence interval.[19]

RECOMMENDATIONS

We obtained two substantially different researched cooling energy savings rates, 29.7 percent and 5.6 percent, by varying how much overlap we required between a participant and a potential match during the matching period. Our initial result, which placed minimal constraints on the matching algorithm, is much higher than the current provisional cooling reduction parameter for the SCT

[18] See, e.g., Arthur S. Goldberger, *A Course in Econometrics,* Harvard University Press 1991, p. 28.
[19] Goldberger, op. cit., p. 381 (Table A-1).

measure of 8 percent in TRM v8.0, while the result obtained with a somewhat more constrained matching algorithm was lower. Rather than select one estimate or the other, we recommend replacing the current value of cooling reduction in the TRM v8.0 with the midpoint between the two. We believe this gives due consideration to the importance of replacing the current value, which is based on residential thermostat research, with a cooling reduction parameter based on the experience of Illinois small business program participants, while also taking into account the sensitivity of the replacement values to small differences in the maintained modeling assumptions, as well as the relative uncertainty of both values as indicated by their confidence bounds.

Improving the statistical precision of the savings estimate would require a significantly larger research sample, ideally on the order of hundreds rather than tens of participants. This could be most effectively achieved by the having the utilities undertake pilot randomized trials in which representative samples of small business customers are given efficient thermostats in two waves, one group in the first year and the other group in the second. This would allow the utilities to control the size of the samples available for research.

Finally, our research did not address the provisional heating reduction parameter for this measure, which is currently 7 percent in the TRM v8.0. To do so, we would need to perform a similar analysis using natural gas and electric usage data from a comparable set of small business customers who installed thermostats at their premises, along with corresponding usage data from a set of matched controls, during the heating season.