

# Energy Savings from Programmable Thermostats in the C&I Sector

*Brian Eakin, Navigant Consulting*

*Bill Provencher, Navigant Consulting & University of Wisconsin-Madison*

*Julianne Meurice, Navigant Consulting*

## Abstract

Energy consumers are often told that programmable thermostats can provide energy savings of 10-30%, though the empirical evidence from a number of studies indicates that savings in the residential sector are usually lower – around 5-10%. But what are the savings in the commercial and industrial sector? The available empirical evidence is much less clear. This study examines this question using billing data for a large sample of C&I customers of two large Midwestern U.S. utilities (Detroit Edison and Consumers Energy) who received programmable thermostats. The context is an opt-in rebate program begun by both utilities in 2009 and continuing today. The analysis uses a matching-with-regression approach described in Ho et al (2007). Each participant is matched to a non-participant based on Euclidean distance in monthly energy use over a 12-month pre-program period. Data for participants and their matches are then used in a regression analysis to control for remaining, non-program differences between program customers and their matches. Over 100,000 non-program C&I customers provide the pool of feasible matches. A 4-month pre-program “test window” comparing the average energy use of program customers and their matches provides a proxy test for selection bias, which is always a concern with opt-in energy efficiency programs. Results indicate the following: (a) for small office buildings, the business type with the largest share of the sample (about 20% of the sample), average annual gas savings are 10.2%; for small retail, the next largest business type, annual gas savings are 5.0%; (c) for all other business types, annual gas savings are 5.0%; and (d) There is no evidence of electricity savings.

## Introduction

Energy consumers are often told that programmable thermostats can provide energy savings of 10-30%, though the empirical evidence from a number of studies indicates that savings in the residential sector are usually lower –in the neighborhood of 5-10%. But what are the savings in the commercial and industrial sector? The available empirical evidence is much less clear. This study examines this question using gas and electric billing data for a large sample of C&I customers of two large Midwestern U.S. utilities (Detroit Edison and Consumers Energy) who received programmable thermostats in the period 2009-2013 as part of an opt-in rebate program. The analysis attempts to differentiate savings by energy type (gas vs. electric) and building type (small office, small retail, grocery, etc.).

## Key Findings

The statistical quality of estimates of savings due to installation of a programmable thermostat varied by fuel type (gas vs. electric) and building type, and depended on the sample size for the building type and the variation in energy use over time and across customers within the sample for the building type. Reliable estimates of energy savings in percentage terms were found for the fuel/building type combinations presented in Table 1. As indicated in Table 1, for gas the evaluation team was able to generate precise estimates of energy savings for Small Office (10.2%) and Small Retail (5.0%), but not for any other of the 17 individual building types examined. We combined these other building types in a single group, “Other”, and obtained statistically significant savings for this group (5.0%). For all other fuel/building type combinations a good estimate could not be generated. In particular, even with all program customers combined it was not possible to conclude that electric savings were statistically different from zero.

**Table 1.** Estimated average annual percent savings due to participation in C&I programmable thermostat program

	Gas			Electric
	Small Office	Small Retail	Other	Overall
<b>Baseline average energy use per year (kWh)</b>	395	475	1,009	10,071
<b>Estimated average percent savings per t-stat (RPP model)</b>	10.2%	5.0%	5.0%	0.7%
<b>90% Confidence Interval:</b>	[7.9%, 12.5%]	[2.7%, 7.3%]	[3.8%, 6.3%]	[-2.2%, 0.8%]
<b>Estimated average annual energy savings (kWh)</b>	40	24	50	-67

*\*Based on average energy use during the matching period*

## Data Used in Analysis

Tracking and billing data were provided by DTE (Detroit Edison) and CE (Consumers Energy). Billing data for customers were available for the period 2008-2013. After cleaning the data, the gas analysis involved 3,783 DTE customers and 3,304 CE customers, while the electric analysis involved 2,848 CE customers and 2,376 DTE customers. DTE and CE also provided data for a large pool of C&I customers for developing the matched comparison group. DTE provided 150,000 customers for gas and 260,000 customers for electric. CE provided 46,000 customers for gas and 68,000 customers for electric.

## Statistical Method

We use a matching method for estimating savings. The particular approach is outlined in Ho et al. (2007), who essentially argue that matching a comparison group to the treatment group is a useful “pre-processing” step in a regression analysis to assure that the distributions of the covariates (i.e., the explanatory variables on which the output variable depends) for the treatment group are the same as those for the comparison group that provides the baseline measure of the output variable. This minimizes the possibility of model specification bias.<sup>1</sup> The regression model is applied only to the post-treatment period, and the matching focuses on those variables expected to have the greatest impact on the output variable. Variables affecting energy use not used for matching can be used in the regression analysis.

In the analysis, participants were categorized by 19 building types: Assembly, Big Box Retail, College/University, Fast Food, Full Service Restaurant, Grocery, Heavy Industry, Hospital, Hotel, Large Office, Light Industry, Medical, Other, School, School (K-12), Retail/Service, Small Office, Small Retail, and Warehouse. The analysis described below was applied to each building type and for each of the two fuel types (gas and electric), as sample sizes permitted. Building types were aggregated if their individual estimated savings were not statistically significant.

### Developing the matched comparison group

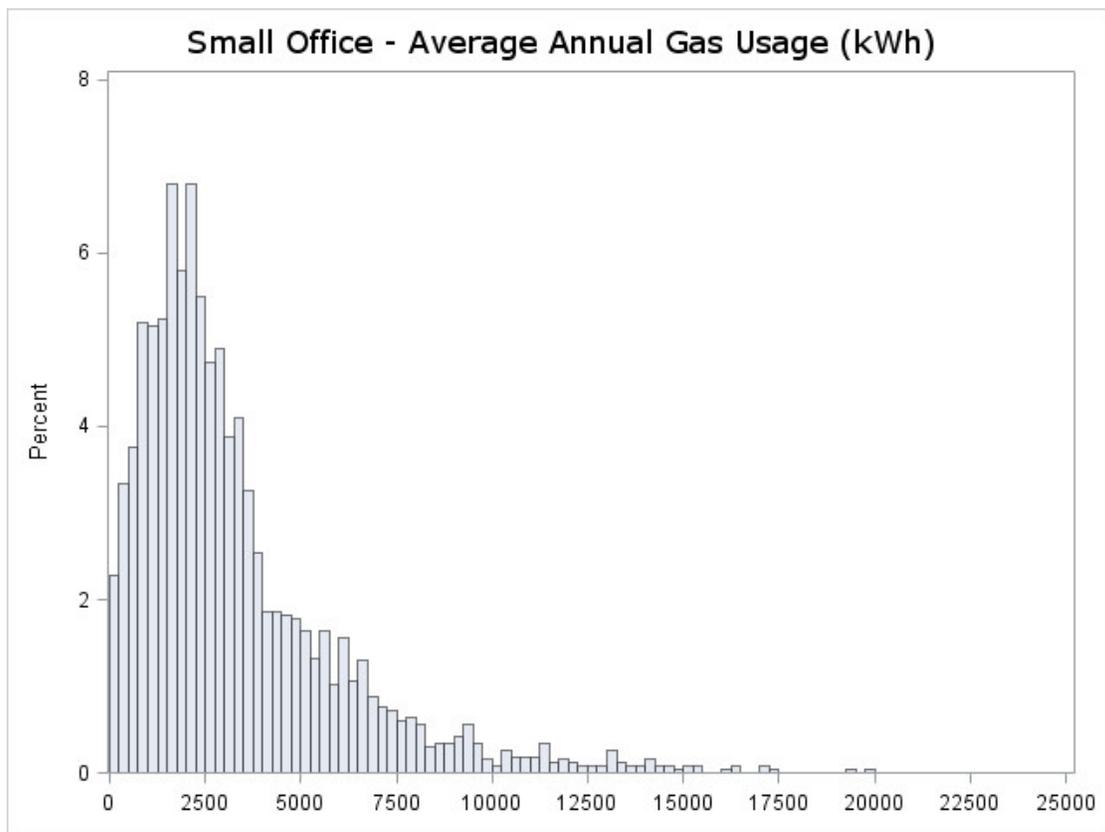
Matching is done on those covariates (explanatory variables) expected to have a high correlation with the dependent variable in the regression analysis, which is, in this study, monthly energy use after programmable thermostat installation. In a billing analysis, the covariate with by far the greatest correlation with monthly energy use after program enrollment is monthly energy use in the same calendar month before program enrollment. The logic is simple: if one finds an excellent match for a participant based on energy use over a 12-month period before program enrollment, then that match is very likely to provide an outstanding counterfactual (baseline) for the participant after program enrollment. It is feasible, of course, to find matches based on additional variables. An obvious example is matching on building type. It is intuitively sensible to match on building type, but matching on discrete variables like building type, especially if the pool of potential matches is small, can hinder the quality of the match with respect to past energy use. The alternative taken here is to limit matching to past energy use and to control for differences between participants and their matches with respect to building type in the regression analysis.

---

<sup>1</sup> Ho, Daniel E., Kosuke Imai, Gary King, and Elizabeth Stuart. 2007. Matching as nonparametric preprocessing for reducing model dependence in parametric causal inference. *Political Analysis* 15(3): 199-236.

The basis of the matching is the difference in monthly energy use between a participant and a potential match,  $D_{PM}$  (**D**ifference between **P**articipant and potential **M**atch). The quality of a match is denoted by the Euclidean distance to the participant over the 12 values of monthly  $D_{PM}$  used for matching; that is, denoting by SSD the sum of squared  $D_{PM}$  over the matching period, it is  $SSD^{1/2}$ . The non-participant customer with the shortest Euclidean distance to a participant is chosen as the matched comparison for the participant. Matching is done with replacement, meaning that a non-participant can be used as the matched comparison for more than one participant.

Because there is considerable variation in customer size in the C&I sector, a problem with using matching methods in program evaluation for the C&I sector is that, for large customers, poor matches and normal random variation in the difference in energy use between participant and match in the post-program period, can exert excessive influence on the estimated program effect. To illustrate, Figure 1 presents a histogram of annual gas use for Small Office customers. The mean energy use is 3,275 kWh and the median energy use is 2,500 kWh. Approximately 5% of participants have energy use greater than two standard deviations above the mean – about 8,800 kWh, roughly 3.5 times greater than the median energy use. If these customers have bad matches, and/or deviate substantially from their matches due to random variation, the estimate of the program effect will be affected substantially.



**Figure 1.** Histogram of annual gas use by small office participants

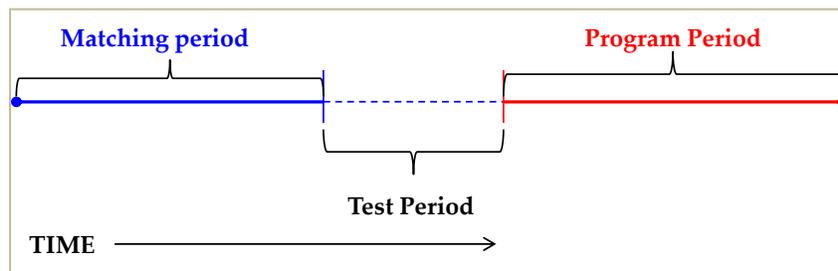
To address this problem we took two steps:

1. Within a building type, only customers within two standard deviations of the mean pre-program energy use are included in the analysis.
2. Among customers that satisfy this size criterion, we apply the analysis to the top 95% of matches.

We assume that the estimated percent savings obtained from the analysis applies as well to large customers.

In program evaluation, a concern with any non-experimental analysis is that program participants are different from non-program participants in unobservable ways that affect their energy use in the post-enrollment period. In statistical analysis, this difference is mistakenly assigned to the program effect. This is called “self-selection bias”, referring to the idea that factors motivating enrollment in the program are correlated with unobserved factors affecting energy use. It is not possible to statistically test for selection bias, but Imbens and Wooldridge (2009) present a test that is suggestive (hereafter called the “IW test”).<sup>2</sup> In the current context the logic of the test is that in the absence of selection bias there should be no difference between participants and the matched comparison group in average energy use outside of the matching period and outside of the program period. Letting  $t_k$  denote the month of program enrollment by customer  $k$ , we implemented the test by matching on energy use over the 12-month period  $t_k - 16$  to  $t_k - 5$ , and comparing average energy use for participants and their matches in the four month test window,  $t_k - 4$  to  $t_k - 1$ .

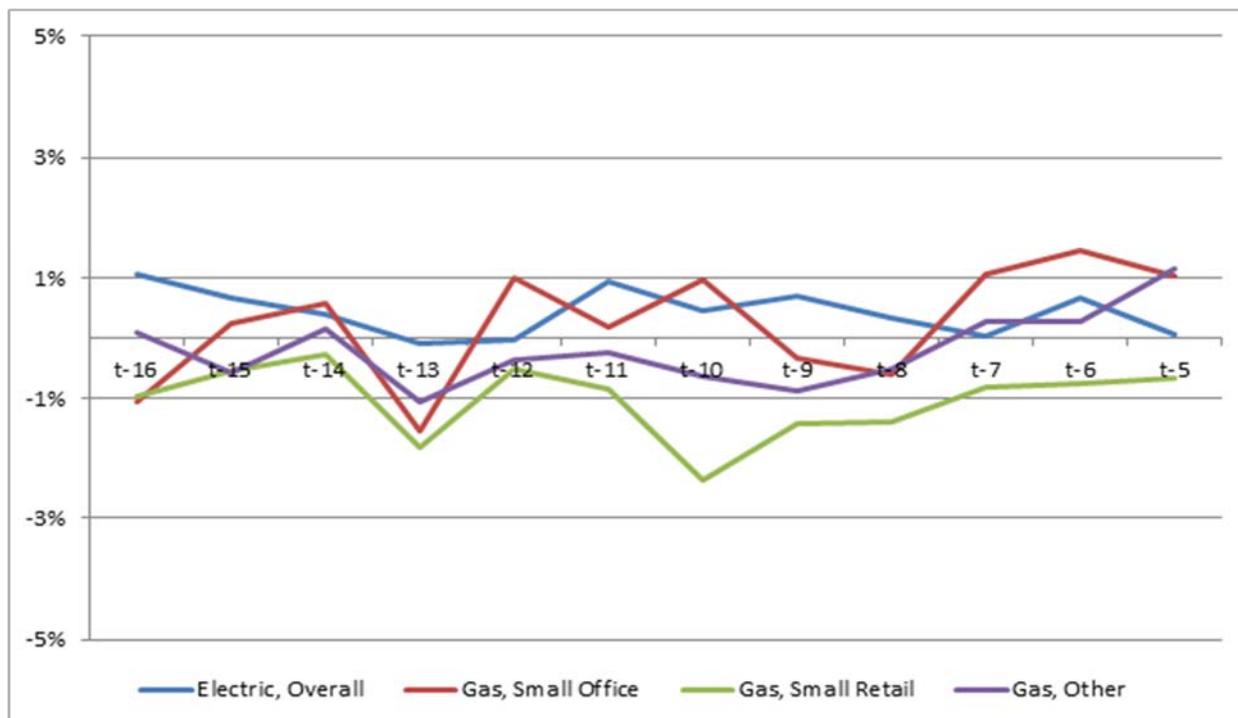
Figure 2 presents a schematic of the test. Finding that the average difference in energy use between treatment and control customers is not different from zero during the test period is consistent with (but not proof of) no selection bias.



**Figure 2.** Schematic of IW “test” for selection bias

<sup>2</sup> Imbens, Guido W., and Jeffrey M. Wooldridge. 2009. “Recent developments in the econometrics of program evaluation.” *Journal of Economic Literature*, 47(1): 5-86.

Figure 3 presents the results of matching for each of the four combinations of fuel/building types discussed in this report.<sup>3</sup> During the matching period the average energy use of participants is very similar to that of their matches – within 1%.



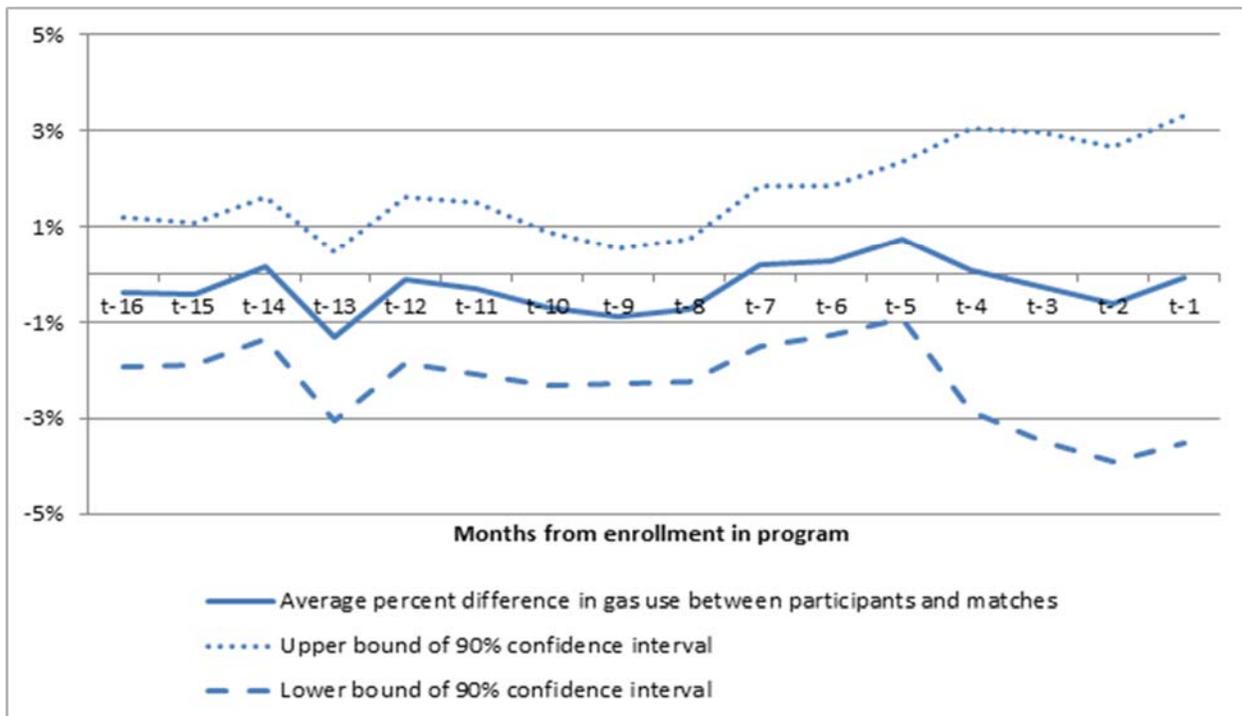
**Figure 3.** Average percent difference between treatment and control customers (matching period is t-16 to t-5, where t=0 is the month of enrollment)

Figure 4 presents the results of the test of the matching for all gas participants combined. Combining all participants reflects the assumption that selection bias is not building-type specific, and a better overall estimate is obtained by combining all participants. Results are consistent with no selection bias.

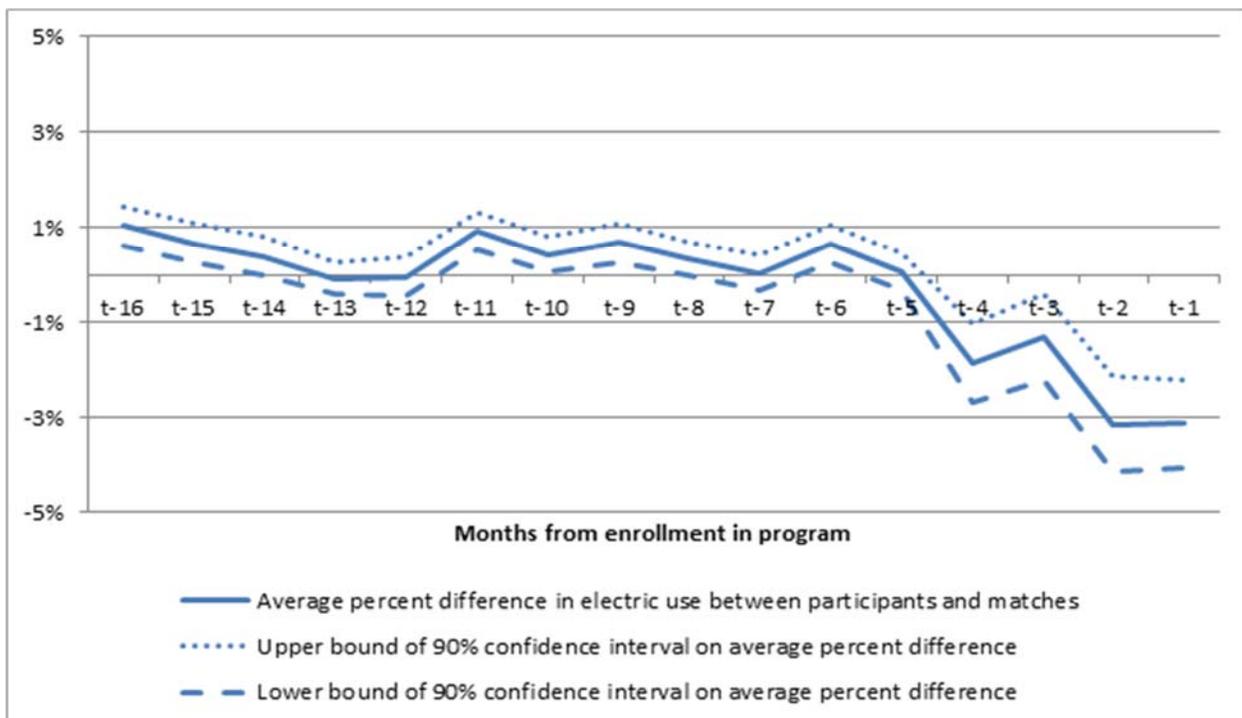
For electric customers, participants appear to use less electricity than their matches during the test period, as shown in Figure 5. This indicates the potential for selection bias, as would occur, for instance, if participants had decided to reduce their energy use just prior to enrolling in the program, and program enrollment is an effect of this decision. However, as discussed in “Results”, this narrative is not supported by the observed differences between participants and their matches in the program period.

In the regression analysis we include two matches for each participant, the best match and the next best match. This allowed us to test whether estimated savings are sensitive to the particular set of matches.

<sup>3</sup> The focus on these four fuel/building types reflects the results of the analysis. See “Key Findings” at the start of the report.



**Figure 4.** Average percent difference in energy use between gas participants and their matches in the matching and test periods (test period begins at t-4)



**Figure 5.** Average percent difference in energy use between electric participants and their matches in the matching and test periods

## Regression model

We use a log-linear specification for the regression model, in which coefficient values are interpreted as percentages. This specification expressly accounts for the fact that at the whole building level the savings from the installation of programmable thermostats increases with energy use. The model takes the specific form,

$$\ln NMU_{kt} = \delta_t + \alpha_1 Participant_k + \alpha_2 Match1_k + \alpha_3 DTE_k + \sum_{j=1}^J \beta^j PreEnergy_{kt} \cdot jSector_k + \varepsilon_{kt},$$

where:

- $\ln NMU_{kt}$  is the average daily electricity use by customer  $k$  during month  $t$ ;
- Greek characters denote coefficients to be estimated, and in particular  $\delta_t$  is a monthly fixed effect;
- $Participant_k$  is an indicator variable taking a value of 1 if customer  $k$  is in the program (as opposed to being a match) and 0 otherwise;
- $Match1_k$  is an indicator variable of whether the customer is the best match. Finding that this is not statistically different from zero indicates that there is no difference between the best match and the next best match.
- $DTE_k$  is an indicator variable for whether the customer is a DTE customer. This variable serves to correct for differences in energy use between DTE and CE customers.
- $PreEnergy_{kt}$  is the energy use by the customer in the pre-enrollment period in the same calendar month as month  $t$ , and  $jSector_k$  is an indicator variable for whether customer  $k$  is in building type (sector)  $k$ , and so  $PreEnergy_{kt} \cdot jSector_k$  is the effect of pre-enrollment energy use on energy customer  $k$ , given that customer  $k$  is in building sector  $j$ ;
- $\varepsilon_{kt}$  is the error term.

In this model  $\alpha_1$  indicates average monthly percent savings by program participants. For gas the model was estimated for the heating season, October-April. A companion model for gas savings during the cooling season revealed no statistically significant savings. For electricity the estimated model is an annual model.

## Estimation Results

Among the 12 original building types, we obtained statistically significant gas savings only for Small Office and Small Retail.<sup>4</sup> We combined all other participants in a single category, Other, and obtained statistically significant savings for this category as well. For electricity, we found no statistically significant savings for any of the building types individually, and failed to

---

<sup>4</sup> Unless otherwise noted, statistical significance refers to significance at the 90% confidence level.

find statistically significant savings even after combining all participants in a single group. The discussion below provides details of the results.

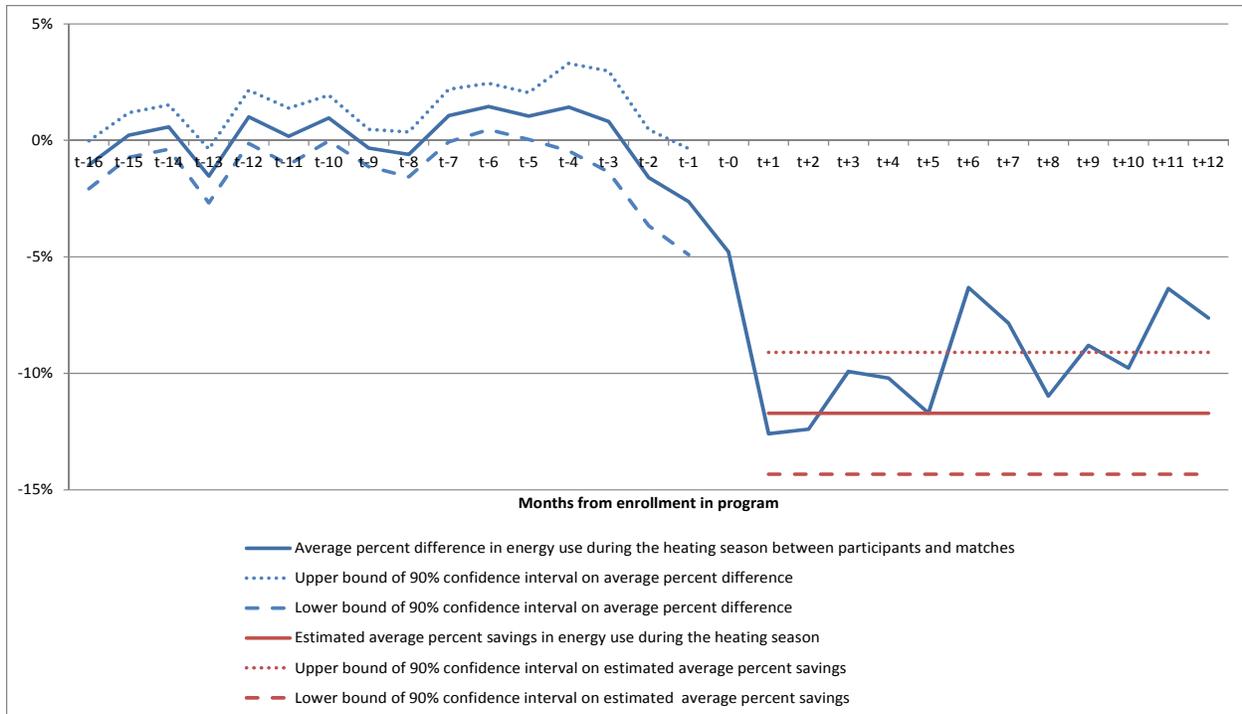
In none of the models discussed below is there a statistically significant difference in energy use between the best and next best matches, indicating that results are not sensitive to the matched sample. For all of the gas models, DTE customers use more energy than CE customers, and for the electricity model, CE customers use more energy than DTE customers.

Finally, to check the sensitivity of estimated savings to influential observations, we re-estimated the models after removing all observations for which the model residual was greater or less than two standard deviations from the sample mean of zero. In none of the models did this have a significant impact on estimated savings.

## Gas Small Office

The sample to estimate gas savings for small offices included 2,536 participants, 894 from DTE and 1,642 from CE. Results of gas savings during the heating season are shown in Figure 6. The red line is the estimated average percent savings during the heating season as estimated using the regression model. Average savings are 11.7%. Under the assumption that gas savings during the cooling season are not different from zero (a result supported by regression analysis), average annual percent savings is 10.2%.

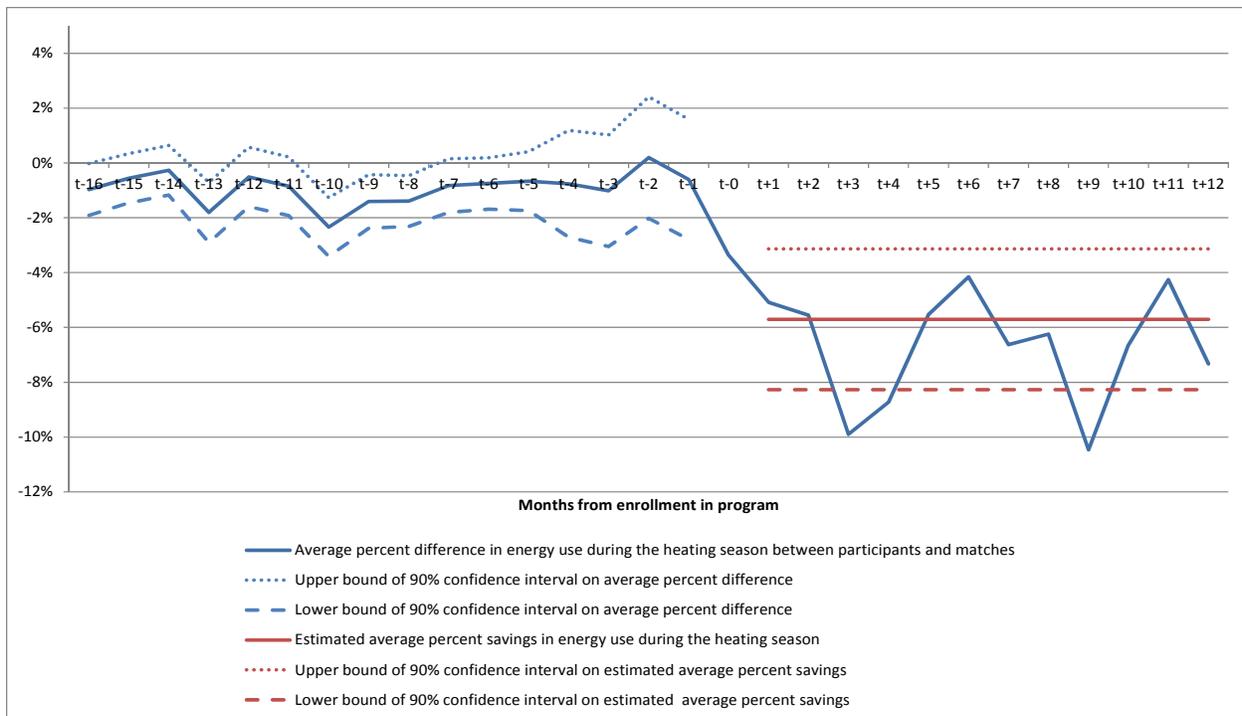
The blue line shows the month-to-month average percent difference in energy use between participants and matches during the heating season. So, for instance, average percent savings 10 months after installing a programmable thermostat ( $t+10$ ), conditional on being in the heating season, is 9.77%. Although the simple matching estimator is usually indicative of savings, it is not recommended to estimate savings. Its use is equivalent to using a regression model in which the terms accounting for energy use in the pre-program period,  $PreEnergy_{kt} \cdot jSector_k$ , are omitted. Without these terms, there is no correction for differences between participants and matches based on the small differences between them during the pre-program period, and there is no accounting for the fact that matches are not necessarily from the same sector as their participants. Nonetheless, a graph of the simple matching estimator makes clear the sharp drop in energy use upon installation of the thermostat, and hints at the possibility that savings decay over time. This is an issue that may warrant future attention.



**Figure 6.** Gas percent savings per building in the heating season, Small Office

### Gas Small Retail

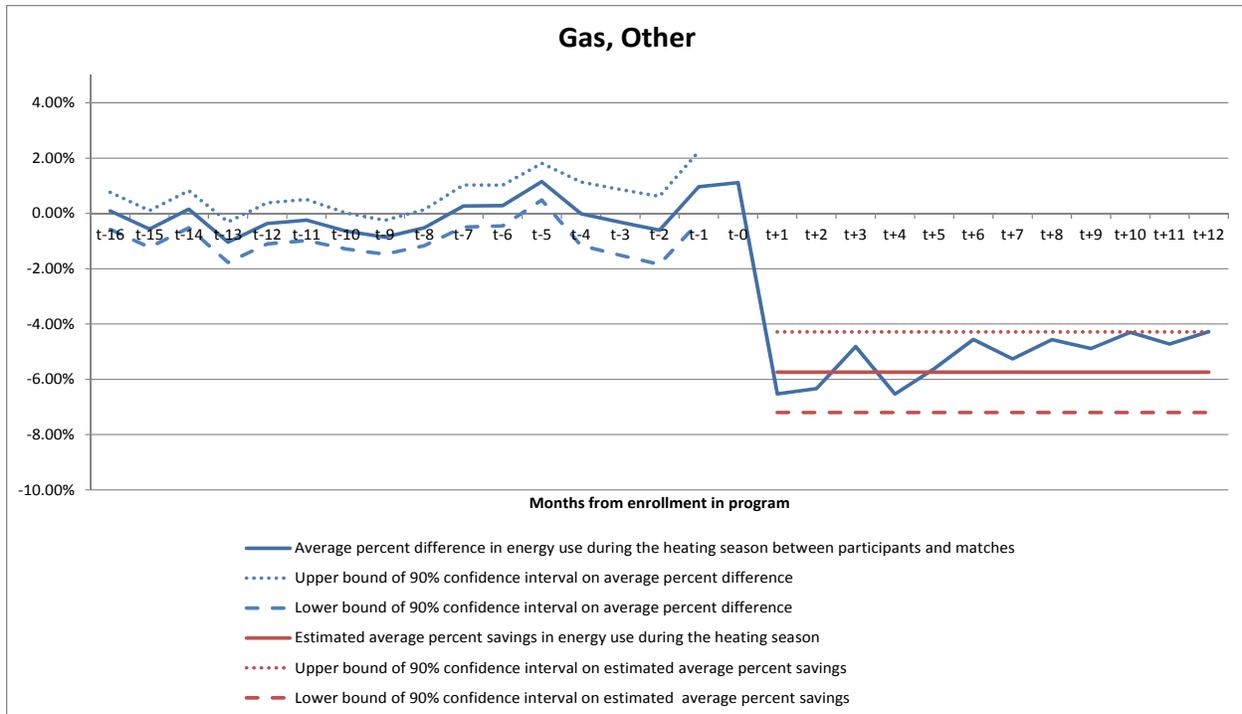
The sample to estimate gas savings for small retail customers included 2,954 participants, 1,702 from DTE and 1,252 from CE. Results of gas savings during the heating season are shown in Figure 7. The red line is the estimated average percent savings during the heating season as estimated using the regression model. Average savings are 5.7%. Under the assumption that gas savings during the cooling season are not statistically different from zero (a result supported by regression analysis), average annual percent savings is 5.0%. The pronounced cycling in the average percent difference in energy use between participants and matches in the program period is due to cohort effects (i.e., the fact that program enrollment is not uniformly distributed during the year), which are controlled for in the regression analysis.



**Figure 7.** Gas percent savings per building in the heating season, Small Retail

## Gas Other

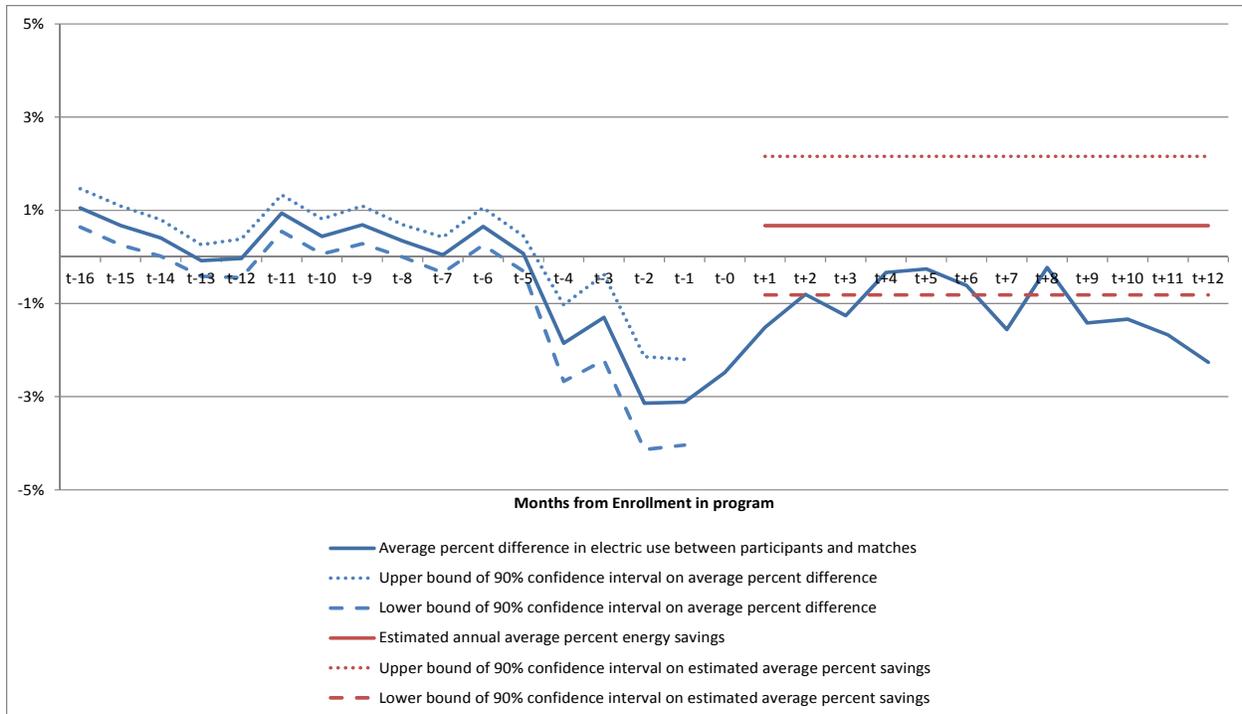
The sample to estimate gas savings for all other gas customers except small office and small retail customers included 7,066 participants, 2,433 from DTE and 4,633 from CE. Results of gas savings during the heating season are shown in Figure 8. The red line is the estimated average percent savings during the heating season as estimated using the regression model. Average savings are 5.7%. Under the assumption that gas savings during the cooling season are not statistically different from zero (a result supported by regression analysis), average annual percent savings is 5.0%. As is the case for small offices, the graph of the simple matching estimator hints at the possibility that savings decay over time.



**Figure 8.** Gas percent savings per building in the heating season, all other participants (excluding small office and small retail participants)

### Overall Electricity Savings

The sample to estimate overall electricity savings included 7,879 participants, 2,422 from DTE and 5,457 from CE. As illustrated in Figure 9, in which the 90% confidence interval on savings contains zero, it was not possible to identify statistically significant electricity savings. As noted previously, electricity use by participants is relatively low compared to their matches in the months before program enrollment, suggesting that participants start saving electricity even before entering the program. But this narrative is contradicted by the climb in energy use in the first few months after program enrollment.



**Figure 9.** Annual electricity percent savings per building, all participants

## Conclusion and Recommendations

A comprehensive statistical analysis of the bills of participants and a matched comparison group estimated annual percent energy savings due to programmable thermostats for a range of building types. Energy savings were estimated on a percentage basis per building, reflecting available data and the perspective that savings in these terms are easily communicated as reflecting typical real-world installation and operating practices, and readily applied by the industry.

For small offices, estimated annual gas savings are 10.2% per building, and for all other customers estimated annual gas savings are 5.0% per building. No electricity savings could be statistically identified.