Memo to:

Massachusetts Gas Program Administrators	**From:**	Jeremiah Robinson, DNV GL
Massachusetts Energy Efficiency Advisory Council	**Date:**	March 6, 2017

MA45 Prescriptive Programmable Thermostats

This memo summarizes the findings of Project 45, the Prescriptive Commercial & Industrial Programmable Thermostat Phase 2 study. The purpose of Project 45 was to determine an accurate natural gas savings estimate associated with programmable thermostats (PTs) installed in commercial and industrial (C&I) buildings in Massachusetts.

The results of this study were inconclusive. We were unable to estimate a statistically significant value for programmable thermostat natural gas savings using the methods detailed below. However, the results of Project 45 do suggest that PTs installed in C&I buildings save some amount of natural gas. For this reason, we recommend that the program continue to use the conservative deemed value of 32 therms per prescriptive programmable thermostat.

The following section provides some background on programmable thermostat evaluations to put this study and its implications in context.

History of Programmable Thermostat Evaluations

Programmable thermostats save natural gas in residential and commercial buildings by setting back the heating temperature to a lower temperature in the winter when areas are unoccupied. Before the existence of PTs, occupants had to remember to turn their thermostat down when leaving a space. With PTs, they can program the thermostat once for their regular schedule and forget about it, overriding it when they happen to be either arriving early and/or work late when the heat would otherwise be set back.

It is generally agreed upon that PTs save energy because they provide additional functionality to the occupants while not taking any functionality away. Occupants can still set back the thermostat manually. However, when programmed correctly, PTs set back the temperature for them and thus reduce the risk that they will forget or choose not to do so on any given day. There has been some speculation that a small percentage of occupants who were extremely diligent about setting back their old-style thermostats actually use more energy when PTs are installed because they pre-heat their spaces before they arrive or wake up, thus adding an extra 30-60 minutes of non-set back time daily. However, there is no evidence to suggest that a large percentage of occupants fall into this diligent user category.

For these reasons, debates around energy savings associated with PTs have centered around how much energy they save, not whether they save energy. Doubts about the magnitude of energy savings revolve around two difficult-to-overcome sources of uncertainty, one behavioral and one engineering-based:

1. The percentage of PTs that are actually programmed, and whether new occupants learn to program the PTs left behind when the old occupants leave.
2. The amount of energy saved by setting back temperatures in buildings with different heating systems and thermal characteristics.

Because field data collection of natural gas usage at the appliance or household level is generally cost-prohibitive (Success Factor #1), and because it is challenging to find participants who will submit to a study prior to the installation of PTs which would allow for the collection of pre-installation data (#2), analyses of PT energy savings have generally relied on customer bills and Billing Analysis.

Because residential buildings are much more numerous than commercial (#4), are similar to one-another in their usage patterns thus making it easier to find a matched control group (#3), and because common residential usage patterns are well-understood (#5), most studies of PTs have been performed with residential homes.

Despite these advantages, results from residential billing analysis studies have been inconsistent and often inconclusive.

The following six studies all attempted to estimate natural gas savings from residential PTs over the past sixteen years using billing analysis. The second and sixth studies found statistically significant savings estimates of 75 therms and 16.5 therms, respectively. The other studies were inconclusive and did not find statistically significant savings estimates.

1. *Energy and Housing in Wisconsin: a Study of Single-Family Owner-Occupied Homes*. Energy Center of Wisconsin, 2000.
2. *Validating the Impact of Programmable Thermostats*. Prepared for the GasNetworks by the RLW Analytics. January 2007.
3. *2004/2005 Statewide Residential Retrofit Single-Family Energy Efficiency Rebate Evaluation*. Prepared for California's Investor-Owned Utilities by Itron, Inc., 2007.

SUCCESS FACTORS

For years, energy efficiency program evaluators have attempted to estimate natural gas savings associated with programmable thermostats in residential and commercial buildings using a variety of methods.

The quality and statistical significance of PT savings estimates have been found to be improved by each of the following factors:

1. Direct measurements of natural gas usage, as opposed to indirect measurements of temperatures or setpoints.
2. Comparison of pre-installation and post-installation data from the same premise.
3. Comparison of the population participating in a PT installation program to a matched control group of nonparticipants.
4. Large sample sizes.
5. A minimum of variability in usage caused by factors other than the PT, such as other natural gas using appliances.
6. Additional data to suggest whether participants made other changes to their space during the analysis period, such as installing a new furnace or boiler.

***Project 45: 2013 Massachusetts Prescriptive Gas Thermostat Evaluation Study.* Prepared by KEMA, Inc. August 2015.**

This report summarized the results of Project 45 Round 1, which included a telephone survey and field data collection effort to support a final billing analysis of 91 accounts. It found savings estimates of 131 ± 108 therms and 5.3 ± 3.8 therms, depending on whether or not the analysis included one large multifamily building project with 120 PTs. Neither of these estimates was statistically significant enough to use going forwards.

MA45 Prescriptive Programmable Thermostats. Prepared by DNV GL. September 14, 2015

This memo summarized the conclusions of a discussion following the submission of the report described above, which expressed technical concerns about some of the modifications applied to the billing analysis, as well as the large standard errors. A decision was made to reduce the deemed savings from 77 therms to 32 therms for the time being, to be consistent with the current residential PT savings value.[1] The evaluation team, including PAs and EEAC consultants, also agreed that additional billing analysis, including additional billing data, would be beneficial to the evaluation. This led to the Phase 2 study.

4. *Measuring the Usability of Programmable Thermostats.* Lawrence Berkeley National Laboratories. From CS Week: April 25, 2011.
5. *NYSEG/RG&E Residential Gas Process and Impact Evaluation.* Prepared by KEMA, Inc. April 2013.
6. *Impact Evaluation of 2014 EnergyWise Single Family Program*. Prepared by DNV GL for National Grid Rhode Island. July 19, 2016

BACKGROUND TO PROJECT 45

This study represents the first serious attempt that we're aware of to perform a field data collection-assisted, billing analysis-based evaluation on C&I PTs. For the reasons listed above, we expected that this effort would be challenging and may result in an inconclusive result. However, we also approached the study with confidence (and we retain this confidence) that C&I PTs save natural gas at levels at least as high as residential for the following reasons:

- C&I spaces usually have a more consistent schedule
- Businesses often employ at least one person with the technical interest and expertise to program a PT – often the owner or a maintenance person

This study consisted of two phases which used a combination of surveys, billing analysis, and field data collection to attempt to determine natural gas savings associated with the installation of PTs into commercial and industrial buildings.

Phase 1 results are summarized in the left sidebar.

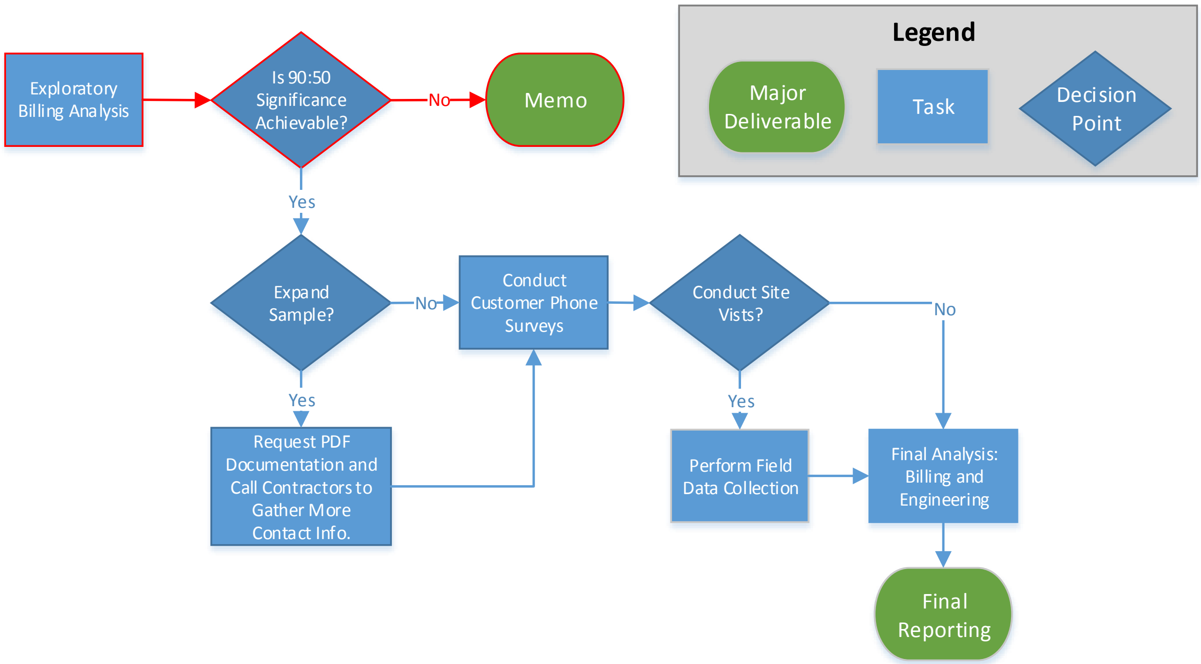
Phase 2 expanded the billing analysis sample for Phase 1 to include all program participants for all years where we have pre-post billing data (2011-2015). The goal was to reduce uncertainty by increasing the sample size.

The project followed the flow chart as laid out in the Phase 2 workplan[1] shown in Figure 1. In this chart, the diamond-shapes represent decision points or stage gates at which the evaluation team would present results to the PAs and EEAC and decide together which series of tasks to pursue next. The flowchart path actually taken is indicated in red outline. Because the Exploratory Billing Analysis found that 90:50 statistical significance was not achievable, the PAs and EEAC consultants on 11/10/2016 chose to conclude Project 45 by publishing this memo and not to pursue further study for this measure. This analysis and the choices made are described in

[1] Draft Phase 2 Scope of Work. Massachusetts Commercial and Industrial Impact Evaluation of Prescriptive Programmable Thermostats. Prepared by DNV GL: July 28, 2016.

more detail in the following sections.

Figure 1. Project 45 Phase 2 Flowchart



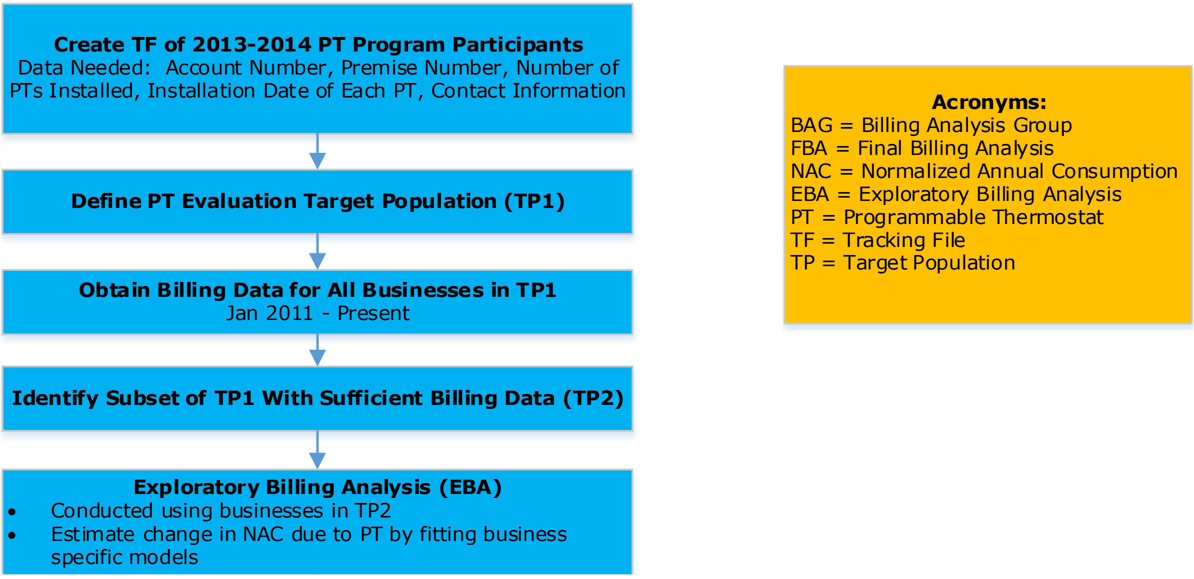
METHODOLOGY

Because the technical portion of Project 45 Phase 2 included only the Exploratory Billing Analysis, this section describes just that portion of the methodology.

The exploratory billing analysis was intended to determine if a larger-scale study including field data collection and phone surveys would be likely to produce statistically significant savings estimates at a 90:50 level of precision.

Figure 2 shows the data analysis approach taken under Phase 2.

Figure 2. Exploratory Billing Analysis Methodology Flowchart



As shown above, the exploratory billing analysis attempted to use all available tracking and billing data for all participants going back to 2011.

Sample Selection

DNV GL selected a sample of 38 program participants who each fulfilled the following conditions:

- Installed at least one programmable thermostat during the 2012-2015 period.
- Had at least 120 days of cold season (November-April) billing data from a single cold season during both the 12 month pre-installation period and the 12 month post-installation period.
    - For participants without an installation date, we excluded the full calendar year of installation as a blackout period.

This process broke up the sample into five categories depending on their data characteristics, as shown below. Sites in group 20 are those which were included in the exploratory billing analysis sample.

Figure 3. Participant Sites by Sample Group

Sample Group	Count
2. Ineligible, No Billing Records	440
3. Ineligible, Participated in Another Program During Pre or Post Period	163
5. Ineligible, Need data before 1/1/2012 AND after 8/1/2016[2]	1
6. Ineligible, Need data before 1/1/2012	336
7. Ineligible, Need data after 8/1/2016	34
10. Non-respondent, Need Data	399
20. Respondent (Eligible for Billing Analysis)	375
Total	**1,753**

Figure 4 presents four examples of sites to show how they fit into the categories described in Figure 3. The "Pre/Blackout/Post" column includes sparklines[3] showing 30-day periods of applicable pre-installation data (black squares on left), 30-day periods of blackout data (red squares) and 30-day

[2] This participant installed programmable thermostats in consecutive years and so has no pre- or post-installation billing data,

[3] Sparklines are tiny charts developed by Edward Tufte intended to display many similar datasets in small spaces for comparison.

periods of applicable post-installation data (black, on right). The columns labelled "Pre" and "Post" show color-coded assessments of whether the participant has sufficient data, with green for yes and red for no.

Figure 4. Examples of Participant Selection Process

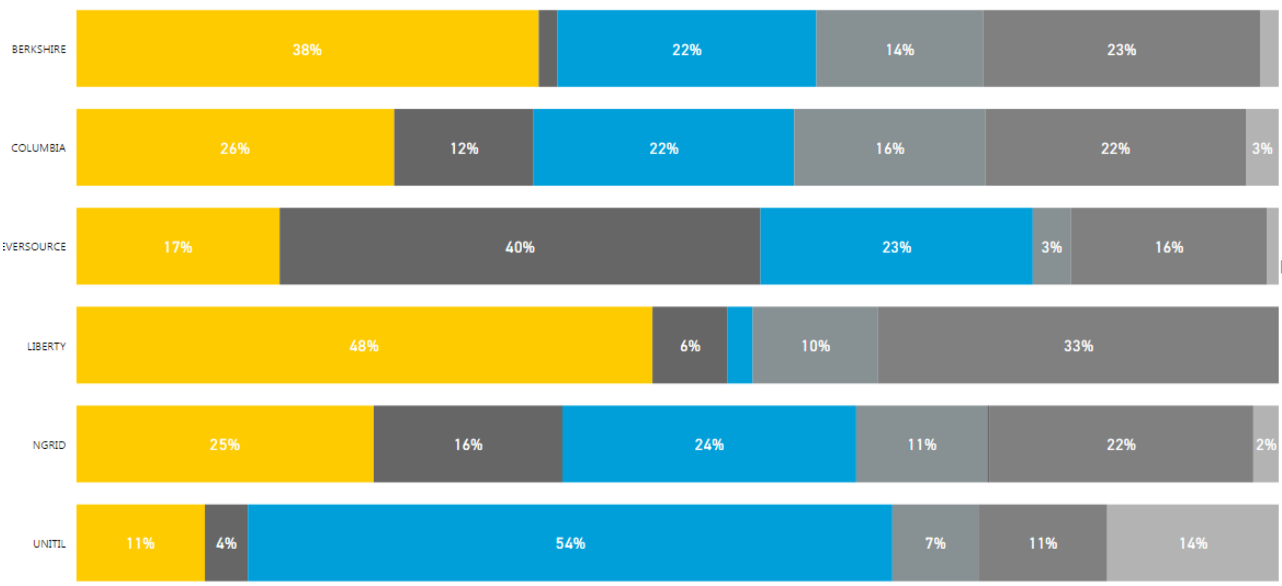
Status Code	PA	DNVID	Pre/Blackout/Post (30 day periods)	Pre	Post
20. Respondent (Eligible for Billing Analysis)	BERKSHIRE	1			
10. Nonrespondent, Need Data	BERKSHIRE	9			
7. Ineligible, Need data after 8/1/2016	BERKSHIRE	66179			
6. Ineligible, Need data before 1/1/2012	BERKSHIRE	10			

The rows shown are described below:

- The first row shows a participant with a 12-month blackout period because of a missing installation date, but who still has sufficient data for analysis.
- The second row shows a participant with a 1-month blackout period, but only 1 month of post-installation data
- The third row shows a participant without an installation date, for whom the 12-month blackout period extends right up to the end of our available billing data in 2015
- The last row shows a participant with an installation too close to the beginning of 2012 for us to have sufficient pre-installation billing data.

Figure 5 shows the portion of sites which fit into each sample group by Program Administrator (PA). Blue sites are those which are in the usable sample (group 20) and yellow are sites in the sample with insufficient billing data (group 10) during the 2011-2015 period.

Figure 5. Participant Sites by Sample Group by PA



Legend: Status codes, values shown from left to right in Figure 4.

- 10. Non-respondent, Need Data
- 2. Ineligible, No Billing Records
- 20. Respondent (Eligible for Billing Analysis)
- 3. Ineligible, Participated in Another Program During Pre or Post Period
- 5. Ineligible, Need data before 1/1/2012 AND after 8/1/2016
- 6. Ineligible, Need data before 1/1/2012
- 7. Ineligible, Need data after 8/1/2016

As shown here, the large PAs (National Grid, Columbia Gas, and Eversource) are relatively similar in the portion of participants with sufficient billing data (all between 20-25%), while the small PAs vary greatly with Unitil having the largest percentage of usable data. Eversource also shows a large number of participants without any billing data at all.

Billing Analysis

Using the sample of 375 participants with sufficient billing data, DNV GL conducted an exploratory billing analysis using a fixed effects model to estimate change in weather normalized gross therms consumption between the pre- and post- installation periods for each participant.

The primary purpose of this exploratory billing analysis was to estimate the variability in the change in consumption between the participants' pre- and post-installation periods. This variability enabled us to estimate the likely precision we would see in the final estimate of program savings that would be produced using a full, field data collection supported billing analysis.

We then used the results from this exploratory billing analysis to partition the participant target population into groups, described below.

a. By PA.

b. By number of thermostats installed.

c. We explored the possibility of grouping based on 4-digit North American Industry Classification System (NAICS) code, but after some discussion, and based on experience, determined that this would probably not be helpful because the 4-digit NAICS codes were unlikely to break the sample into groups which were similar-enough in building type and usage patterns to be meaningful.

d. Based on input from the PAs, we also created four groups based on overall consumption to determine if large participants were fundamentally different than small ones.

Mathematically, this exploratory billing analysis used a fixed effect billing model methodology that involves estimating models very similar to the pooled model discussed in Jayaweera and Haeri, 2013.[4] The various models produced during the billing analyses used linear regression techniques and variations of the commonly used PRISM®[5] model. PRISM® has been well-documented for effectiveness and accuracy, and is widely used to estimate the impact of programs on energy usage. One of the earliest references of the PRISM model can be found in Fels (1986).[6]

The important feature of the PRISM model that makes it both unique and applicable for measuring energy savings is its use of weather data as predictors. Weather predictors will be included in the models by constructing heating degree day values for each participant and each time period. Cooling degree day values may be used as well.

In summary, the basic PRISM linear model that was used in this billing analysis is the following:

---

[4] Jayaweera, T. and Haeri, H. (2013). "The Uniform Methods Project: Methods for Determining Energy Efficiency Savings for Specific Measures." Report prepared for the National Renewable Energy Laboratory (NREL) and funded by the U.S. Department of Energy's Office of Energy Efficiency and Renewable Energy, and the Permitting, Siting and Analysis Division of the Office of Electricity Delivery and Energy Reliability under National Renewable Energy Contract No. DE-AC36-08GO28308. Available electronically at http://www.osti.gov/bridge. Jan 2012 — Mar 2013.

[5] PRISM® (PRInceton Scorekeeping Method) is copyright protected. Copyright 1995, Princeton University. All rights reserved.

[6] Fels, M. (1986) "PRISM: An Introduction." Energy and Buildings 9, #1-2, pp. 5-18

Equation 1

$$E_{ki} = \mathbf{z}_{ki}\boldsymbol{\gamma} + \mathbf{x}_{ki}\boldsymbol{\beta} + \varepsilon_{ki} \qquad (1)$$

Where the subscript i denotes participant, k is time period and

E_{ki} is the energy use for participant *i* and time period *k*. This equals the therms consumption as noted in the billing data.

$\mathbf{z}_{ki}$ is a vector of model explanatory variables that are not a function of any program-related variable(s). For this evaluation, this vector includes an assortment of variables, including weather data (degree days), year/month indicators and (for the size grouping analysis) participant consumption group.

$\mathbf{x}_{ki}$ is a set of model explanatory variables that are a function of program-related variable(s). Elements in this vector are equal to zero for time period *k* in the pre-installation time frame for each participant and are generally something other than zero for time periods in the post-installation time frame. Often some or all of the components of $\mathbf{x}_{ki}$ are interaction terms between a 0/1 program indicator for *(k,i)* and the variables in $\mathbf{z}_{ki}$

$\boldsymbol{\gamma}$, $\boldsymbol{\beta}$ are the model coefficients that are estimated in a weighted least squares, regression estimation.

ε_{ki} is the model random error term.

Returning to Equation 1, assume the estimated $\boldsymbol{\gamma}$ and $\boldsymbol{\beta}$ are $\hat{\boldsymbol{\gamma}}$ and $\hat{\boldsymbol{\beta}}$ respectively, and note that for any particular $\mathbf{z}_{ki} = \tilde{\mathbf{z}}_i$ and $\mathbf{x}_{ki} = \tilde{\mathbf{x}}_i$, the model-predicted amount of therms use before program participation for participant i is the following:

Equation 2

$$\hat{E}_{i,before} = \tilde{\mathbf{z}}_i\hat{\boldsymbol{\gamma}}$$

And the predicted amount of therms use after program participation is the following:

Equation 3

$$\hat{E}_{i,after} = \tilde{\mathbf{z}}_i\hat{\boldsymbol{\gamma}} + \tilde{\mathbf{x}}_i\hat{\boldsymbol{\beta}}$$

So the gross difference in energy use that may be attributed to the program is found by subtracting Equation 2 from Equation 3, which results in the following:

Equation 4

$$\Delta\hat{E}_i = \hat{E}_{i,after} - \hat{E}_{i,before} = \left(\tilde{\mathbf{z}}_i\hat{\boldsymbol{\gamma}} + \tilde{\mathbf{x}}_i\hat{\boldsymbol{\beta}}\right) - \left(\tilde{\mathbf{z}}_i\hat{\boldsymbol{\gamma}}\right) = \tilde{\mathbf{x}}_i\hat{\boldsymbol{\beta}}$$

When $\Delta\hat{E}_i$ is negative this indicates the model prediction suggests some energy *savings* can be attributed to the program.

In the above discussion, the vectors $\tilde{\mathbf{z}}_i$ and $\tilde{\mathbf{x}}_i$ would include weather data (heating degree days) in a typical meteorological year (TMY). These "typical" TMY temperatures were derived using 3 years of historical data, and represent the outside temperature per hour, for every day in a "typical" calendar year that one would expect at any given weather station in Massachusetts. Using these outside air temperatures, we define indoor temperature to determine a value for heating degree days (base) which shows the highest-possible R-squared value. The use of the TMY data to create heating degree days enables us to predict, using Equation 4, the gross savings attributed to the program in a "typical" year.

As noted earlier, we would have varied Equation 1 slightly between the exploratory and final billing analysis, as follows:

- *The billing model described in Equation 1 was estimated for each individual business in the exploratory billing analysis. In other words, for N participants in the target population then N models of the form described in Equation 1 were estimated. A pooled version of the model described in this equation, was proposed for the final model.*
- *The explanatory variables (* $\mathbf{z}_{ki}$ *) included only weather data and an intercept term for the exploratory analysis. This vector did not include the total cooling degree days in a billing period because this tends to have little predictive power in Equation 1 when estimating therms consumption.*
- *The model coefficients would have been estimated using a weighted linear regression technique for the final billing analysis; however, weights were not used for the exploratory billing analysis since models are being fit for each individual business. The sample weights used in the final billing analysis would have been designed to account for businesses in the target population that would not have been included in the final pooled billing analysis for various reasons, such as insufficient pre/post billing data or results from the participant survey which suggested the business underwent some change that significantly affects their difference in consumption between the pre and post periods and is unrelated to the program itself.*

RESULTS

Our analysis of the available data for Massachusetts C&I programmable thermostat program participants under Project 45 does not support a confident estimate for programmable thermostat natural gas savings in C&I buildings

In other words, given the variability of C&I energy use, the sample size, and the available tracking and billing data, the work performed under Project 45 suggests that billing analysis as an evaluation approach—either alone or supplemented by field data collection—is unlikely to be able to make a statistically significant determination of the amount of energy saved by C&I PTs either at the program level or for any subgrouping.

Sample Characterization

This section provides a visual representation for the characteristics and breakdown of the overall population and the billing analysis sample.[7]

Figure 6 shows a distribution of claimed savings versus number of PTs installed for each PA during our analysis period in the overall population. As shown here, most PTs were installed during the early part of the analysis period when 77 therms was the deemed savings value. Installations dropped off after 32 therms became the deemed savings value. Columbia Gas claimed a different amount of savings for each thermostat project, usually less than the deemed value.

[7] These charts were produced in Microsoft PowerBI, and can be explored and filtered in a visual way through this software. Please contact Jeremiah Robinson if you would like a copy of the results in PowerBI for your PA, or the overall results with any personally identifiable information (PII) removed.

Figure 6. Tracking Savings per Thermostat By PA

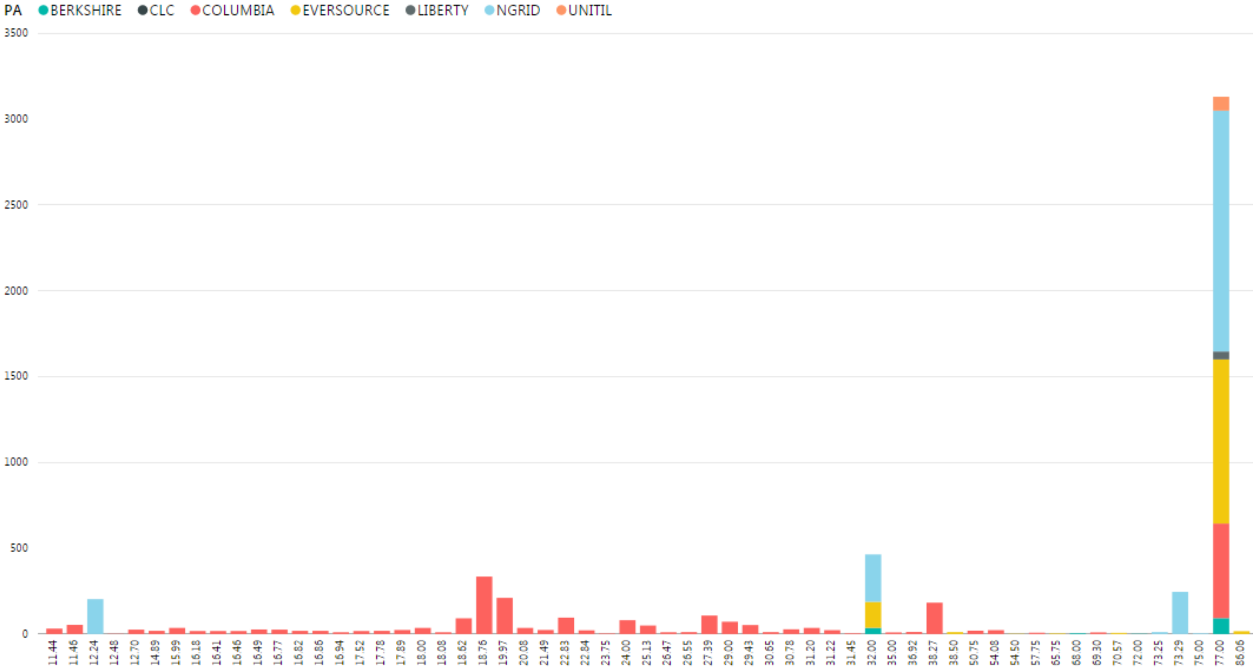


Figure 7 shows a map of the 18 sites with more than 50 PTs installed.

Figure 7. Map of Projects with More Than 50 Thermostats

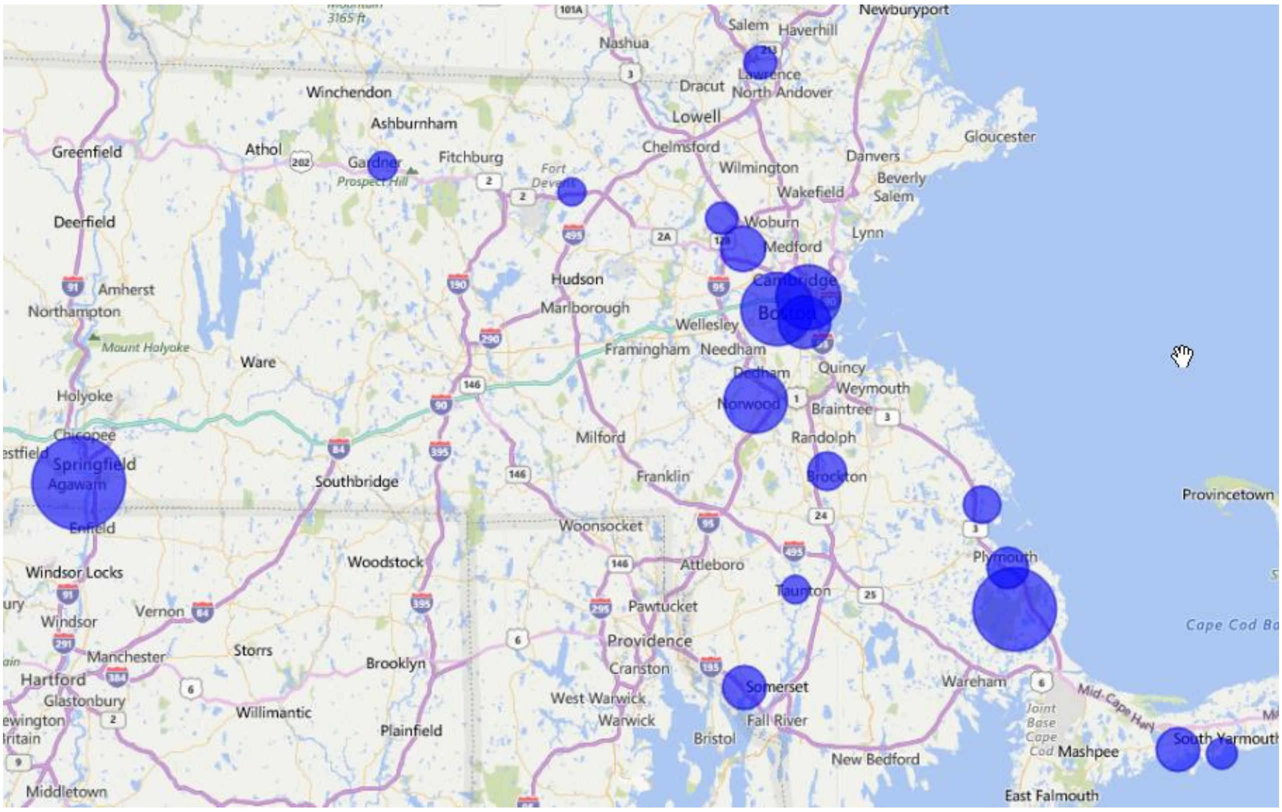
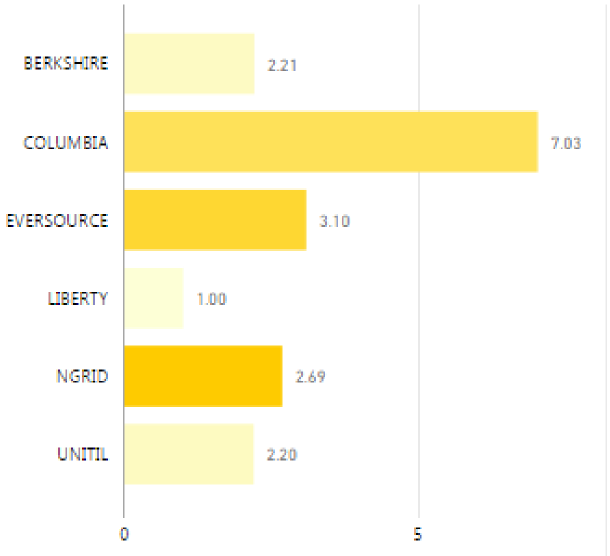


Figure 8 shows the average number of thermostats installed per site by PA in the analysis sample. The color saturation represents the sample size, with darker yellow representing a larger number of sites and, thus, the effect of that PA's data on the overall result.

Figure 8. Average Number of Thermostats Per Site in Sample



Overall Results

Following are the summary statistics from the exploratory billing analysis. As discussed above, these results are not statistically significant.

- 375 sites (1,411 PTs) used in model
- 4,112 billing periods used in model
- 112 therms saved per site, with 235 therms standard error
- 30 therms saved per thermostat, with 62 therms standard error
- R^2 value of 0.901

Figure 9 shows the amount of therms saved per programmable thermostat (chart on left) and per site (chart on right) by number of thermostats installed per site, with darker color saturation representing the total number of thermostats (left) and sites (right) used in the exploratory billing analysis sample.

Figure 9. Savings Per Thermostat (left) and Per Site (right) by Number of Thermostats Installed

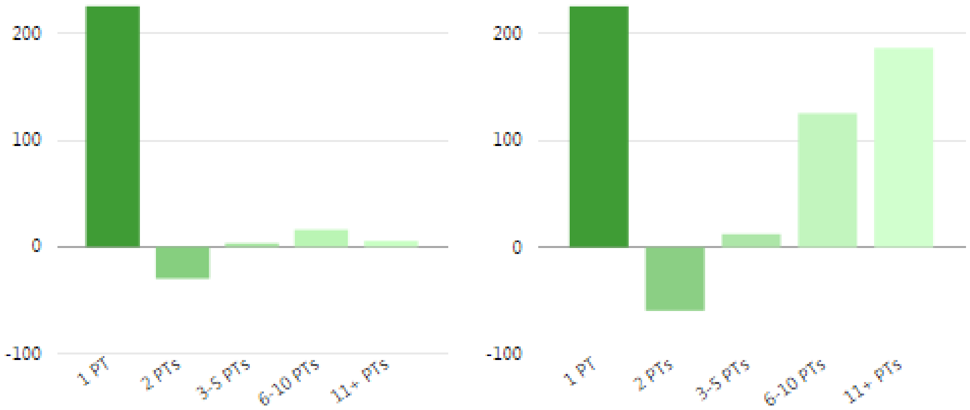


Figure 10 shows the (not statistically significant) amount of therms saved per programmable thermostat (chart of left) and per site (chart on right) by PA, with darker color saturation representing the total number of thermostats (left) and sites (right) rebated by that PA and used in the exploratory billing analysis sample.

Figure 10. Savings Per Thermostat (left) and Per Site (right) by PA

SUMMARY AND CONCLUSIONS

Our analysis of the available data for Massachusetts C&I programmable thermostat program participants under Project 45 does not support a confident estimate for programmable thermostat natural gas savings in C&I buildings

In other words, given the variability of C&I energy use, the sample size, and the available tracking and billing data, the work performed under Project 45 suggests that billing analysis as an evaluation approach—either alone or supplemented by field data collection—is unlikely to be able to make a statistically significant determination as the amount of energy saved by C&I PTs either at the program level or for any subgrouping.

The results of Phase 1 of this evaluation, while not statistically significant at the 90:50 confidence level, do suggest that PTs save some amount of natural gas in C&I buildings when replacing existing non-PTs or older PTs which the occupants no longer know how to program. For this reason, we recommend that the program continue to rebate PTs in these situations at a value of 32 therms per thermostat.

Given the lack of a persuasive result for this measure, we recommend that the program pursue the following options:

1. Continue to offer this measure at the current deemed value of 32 therms per thermostat.
2. Shift resources towards smart web-enable thermostat or mini-EMS systems, both for energy efficiency and electrical demand response.

We also include the following suggestions for future program consideration.

1. Consider the cost-effectiveness of this offering using the current 32 therms per thermostat savings value.
2. Revisit savings potential if gas interval data becomes available or gas non-intrusive load monitoring (NILM) is pursued as part of a larger evaluation.