Bundle of M&V Joy? M&V of Bundled DR Smart Thermostats, Direct Install, and Assessments Programs

Noah Fraser, Alexandra Horne, and Lissa Callahan, ADM Associates Inc., Reno, NV

ABSTRACT

As utilities get creative to meet higher savings goals with smaller budgets, dispatching a single install visit for three programs ("bundling") is an attractive cost saving option, but can create complications for impact evaluation. The bundling approach is successful at increasing customer participation in multiple programs but means evaluators can no longer simply remove a small number of cross-participants to have a 'clean" data set. When most program participants are cross-participants, can we still provide accurate savings estimates for bundled in-home energy assessments (IHEA), smart thermostats in residential demand response (DR), and direct install (DI) programs? Of course, we can. The analysis must separate the annual energy optimization savings from the smart thermostats in the DR program, while also disaggregating DI program savings, and IHEA program savings. Direct Install measures generally have deemed savings values that are closely tracked within the program. Smart thermostat optimization reduces A/C runtimes which provides energy savings throughout the year, which is an added benefit to their use in residential DR programs. In-Home Assessments provide a list of potential energy saving tips and must be modeled each program year to calculate savings. To have a large enough sample for the IHEA regression model to be statistically valid, we had to keep bundled participants in the analysis sample. The ordered process described in this paper allowed us to accurately estimate the annual energy savings for the smart thermostats in the DR program, the DI measure savings, and IHEA savings without double counting.

Introduction

During the 2019 program year, one of our long-term utility clients decided to bundle three residential energy efficiency (EE) programs together so that only one roll of the truck was required to complete an IHEA assessment, installation of DI measures, and a free smart thermostat which provided year-round optimized energy savings while facilitating participation in the DR program. When a customer requested to participate in any of the three programs, they were typically offered participation in the two other programs. Participants were targeted via television and online advertisements, social media, the online dashboard available through their account, educational community outreach events, direct mailing, and emailing. Both the internal and external contact centers offered the IHEA when customers contacted them regarding their billing amount, along with smart thermostats from the residential DR program and measures from the DI program. The push of the IHEA program by the contact centers resulted in that program having a high percentage of cross participants; 87 percent bundled with DI and 46 percent with DR.

The IHEA program provided energy efficiency assessments of participants' homes that resulted in a written list of improvements such as appliances, insulation, lighting, and behaviors that could be implemented to save energy in the home. Because the program did not track what energy saving measures the IHEA participant installed due to the assessment, the evaluated (ex-post) energy savings was estimated using a mixed effect panel regression model. The DI program provided installation of free measures such as LED lights, air filters, and air conditioner line insulation. Evaluation of the DI program involved desk reviews of the program tracking data and verification of the engineering algorithm inputs to calculate energy savings. Each DR program participant received a free smart thermostat that also provided year-round energy savings through micro adjustments in the setpoints and enabled participation in demand response events during the cooling season. Smart thermostats adjust the setpoints of the heating and cooling system to reduce runtimes. We used a difference-in-differences regression model with a control group to estimate the annual energy savings of the thermostats.

The evaluation of these three programs required us to estimate the savings and correctly distribute them among the programs with 90/10 statistical significance. To accomplish this task, we decided to order the programs by the certainty of the estimated energy savings. DI measures have well defined savings and were easy to verify that they are installed. There were enough non-bundled smart thermostats installed through the DR program to perform a statistically significant regression analysis. But the IHEA program had over 90 percent of participants bundled and had to include the bundled DI and DR participants in the regression analysis.

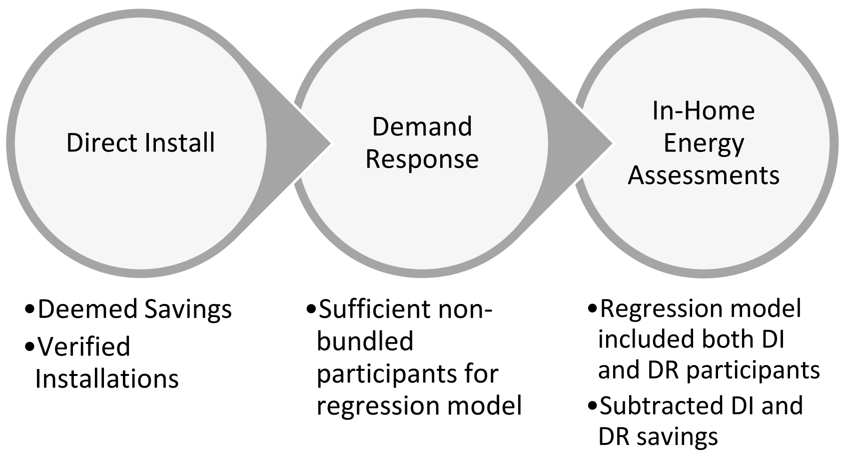

Figure 1. The three bundled programs ordered from left to right based on a hierarchy of certainty for the estimated energy savings.

We separately calculated the savings for each participant in the three programs and then subtracted the DI and DR savings from the IHEA participant savings. This ordered evaluation process accurately distributed the energy savings from the bundled programs for the 2019 program cycle, and we received compliments from the Public Utilities Commission on this approach.

Methods

The following sections describe the verification methodologies used to estimate the ex-post savings for each program and the process for distributing the savings correctly between the three bundled programs.

Direct Install

Our verification work utilized customer lists and program tracking data provided by the utility and field verifications and surveys. Participant surveys were administered to a sample of program participants to determine measure install rates. We conducted a simple random sample of participants for surveying. The participant survey collected verification data regarding measures eligibility for 94 customers. In addition, 17 customers were checked for program eligibility during ride-along verification activities with the program's implementation contractors. We sampled a total of 111 participants (i.e.,94+17), which exceeds the required random sample size of 67 needed for 90/10 statistical significance (EPA 2019).

Based on the results of the online survey, the ride-along visits, and program tracking data, ADM determined the measure-specific verification rates for the measures installed through the program during 2019. The verification rate was applied to the engineering equation based deemed savings for each measure.

Table 1. Direct Install Energy-Efficiency Measures, Related Savings, and Useful Life.

Measure	Annual Energy Savings (kWh)	Effective Useful Life
Air Filter/Furnace Filter	95	0.5
LED (7W A19)	20	6
LED (8W BR30)	50	14
LED (9W A19)	30	10
LED (9W BR30)	49	20
LED (11W BR30)	47	20
LED (11W A19)	37	8
Photocells	3	8
Refrigerator Thermometer	6	3
Air-conditioner Refrigerant Line Insulation	20	10

Residential Demand Response -- Smart Thermostats

We calculated the annual energy savings for the smart thermostats in the Residential DR program through a three-step process: (1) removed cross participants, including the bundled participants, (2) created a matched control group, and (3) conducted difference-in-differences regression modeling.

Removal of Cross Participants

Whenever possible we simply removed cross participants from the regression sample, this allowed the model to estimate the savings for just the smart thermostats. During the 2019 program year, there were enough Residential DR only participants to produce a robust sample for the regression model. Because this is an ongoing program where we evaluate the energy savings for all smart thermostats in the program each year, there are usually many participants to sample from.

Control Group Matching

The utility set aside a group of 50,000 control group homes for the DR program at the beginning of the program nearly a decade ago. We used a two-step process to select a sample of the control group that best matched each of the treatment homes during the full year before the smart thermostat was installed. First, we matched several control homes to each treatment home based on their utility rate codes and zip codes. This procedure was implemented in the R programming language using the "MatchIt" package[1] and matches were selected with replacement. Second, propensity score matching was used to "match" the control group to the treatment group via a propensity score, which is essentially an estimate, derived from observed characteristics of a utility customer's likelihood of participating in the DR program.

The logit model below was used to estimate the propensity scores for all customers.

$$Participation = \alpha + \beta \cdot [kWhN] + \varepsilon$$

[1] https://cran.r-project.org/web/packages/MatchIt/MatchIt.pdf

Where, $Participation$ was a binary variable that is 1 if the customer was a DR program participant and 0 if they are a non-participant; α was the participation intercept; $kWhN$ was a continuous variable that captures the customer's pre-thermostat installation, weather normalized (consumption divided by degree days) average daily consumption; ε was an error term; β was a coefficient showing the changes in propensity to participate in the DR program that occurs for a change in the $kWhN$ variable. After the propensity scores were estimated, for each treatment premise p, a k-nearest neighbors' algorithm was used to find the $k = 1$ closest propensity score from among the control premises.

To ensure the quality of the matching procedure, a Welch's Two Sample t-test was conducted to ensure that the treatment and control group were statistically similar during the pre-installation period. Because the p-value was greater than 0.05 we fail to reject the null hypothesis and conclude that the true difference in means is not statistically significantly different than zero (EVO 2019).

Table 2. Welch's Two Sample t-test for DP23.

Independent Variable	Treatment		Control		Welch Test T-Statistic	Welch Test P-Value
	Pre-Installation Matching Mean	Post-Installation Matching Mean	Pre-Period Matching Mean	Post-Period Matching Mean		
kWhN	1.63	1.63	1.64	1.64	-0.051	0.959

The following figure shows the aggregate monthly usage by the matched treatment and control groups.

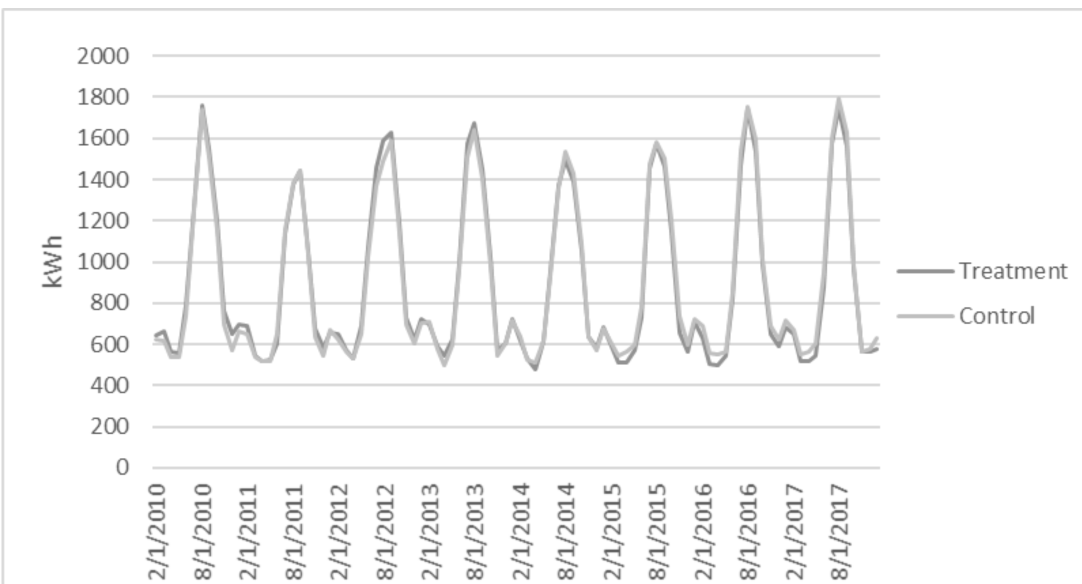

Figure 2. DP23 Pre-Period Average Usage.

Calculation of Energy-Optimization Impacts

With the matched control group, we determined the annual electricity savings resulting from the energy-optimization service by employing difference-in-differences modeling which included the energy savings from the 2019 DR events (EPA 2019). With this method, changes in energy use for customers receiving optimization from the smart thermostat were compared to changes in non-optimized energy use of the control group. Both groups were then compared to a baseline "pre" period occurring prior to the program participants' receipt of their smart devices.

The basic specification for the regression modeling is illustrated as follows. Consider modeling the energy use of a customer who was benefitting from optimization. In simplest terms, average daily electricity use was separated between weather-sensitive and non-weather-sensitive factors. The model to represent this was:

$$AEC_t = \alpha_0 + \alpha_1 CDDperDay_t + \alpha_2 HDDperDay_t + \varepsilon_t$$

Where, AEC_t was average daily use of electricity for period t for a customer; $CDDperDay_t$[2] was cooling degree days during day t; $HDDperDay_t$[3] was heating degree days during day t; ε_t was an error term; α_0 was the intercept term; α_1 and α_2 were regression coefficients showing the changes in use that occurs for a change in either cooling degree days or heating degree days. We ran the pre and post periods of data for the treatment and control groups through the above regression equation, which provided us with four energy consumption estimates for AEC_t.

- $AEC_{TreatmentPre}$ – The average daily consumption of the treatment group prior to the installation of the smart thermostat.
- $AEC_{TreatmentPost}$ - The average daily consumption of the treatment group after the installation of the smart thermostat.
- $AEC_{ControlPre}$ - The average daily consumption of the control group prior to the installation of the smart thermostat.
- $AEC_{ControlPost}$ - The average daily consumption of the control group after the installation of the smart thermostat.

The implicit assumption for the difference-in-differences analysis was that a change in energy use in response to a change in weather conditions (and other time-varying factors like the economy) would be the same for the control group and the treatment group in the absence of the optimization. If this assumption held, then the changes in energy usage of the control group in response to a change in weather conditions can be applied to predict what the (counterfactual) energy use of the treatment group would have been under the changed weather conditions in the absence of the optimization. This allowed the difference between actual post-optimization energy use of the treatment group and the counterfactual predicted energy use to be calculated as the savings attributable to the optimization. The difference-in-differences equation had the following form.

$$AES_{\text{Optimization}} =$$

$$(AEC_{ControlPre} - AEC_{ControlPost}) - (AEC_{TreatmentPre} - AEC_{TreatmentPost})$$

Where, $AES_{\text{Optimization}}$ was the average energy savings attributed to the optimization algorithm.

Standard statistical tests and regression diagnostics were used to evaluate the performance of the models (EVO 2019). Each model was screened for implausible results. The statistical tests and diagnostics included evaluating the t-statistics for estimated coefficients, the adjusted R^2 for equation fit and examining residuals from the fitted models.

In-Home Energy Assessments

ADM used a mixed effects panel regression model to determine daily average electricity savings for treatment group members. With the panel approach, the regression model was applied to monthly billing data for each participant in the sample before and after participation in the program. For program year 2019, a pre/post model was used, which identified the daily savings in the treatment group after controlling for the effects of weather. The model is designated below.

$$AEC_{i,t} = \beta_1 CDD_{i,t} + \beta_2 HDD_{i,t} + \beta_3 Post_{i,t} + \alpha_i \text{Customer}_\text{i} + E_{i,t}$$

[2] The number of cooling degree days per month was calculated using hourly weather data from NOAA.
[3] The number of heating degree days per month was calculated using hourly weather data from NOAA.

Where, subscript i denoted individual customers; t =1, …, T(i) served as a time index, where T(i) was the number of bills available for customer I; $AEC_{i,t}$ was the average daily use of energy (either kWh or therms) for period t for a customer; $Customer_i$ = "1" for customer i and "0" if not; CDD cooling degree days (base temperature 75 degrees Fahrenheit); HDD heating degree days (base temperature 65 degrees Fahrenheit); $Post_i$ = "0" if the monthly period was before the customer received assessment and "1" if not; E_t was an error term; α_1 was a coefficient that represents the grand mean (i.e., mean of the unique customer-specific intercepts); β_1 was a coefficient that adjusts for the customer's cooling season weather-sensitive usage; β_2 was a coefficient that adjusts for the customer's heating season weather-sensitive usage; and β_3 was a coefficient that adjusts for whether customer $\boldsymbol{i}$'s monthly billing data in period $\boldsymbol{t}$ was in the pre or post period.

The model was defined as 'mixed effects' because the model decomposed its parameters into fixed effects for the heating degree days (HDD), cooling degree days (CDD), post variables, and random effects (i.e., the individual customer's base use). A fixed effect was assumed to be constant and independent of the sample, while random effects were assumed to be sources of variation (other than natural measurement error) that were uncorrelated with the fixed effects (EVO 2019). In the model, the first billing period after the beginning of treatment was considered the 'deadband period'. Observations that occurred in the deadband period were not included in the mixed effects panel regression because it was unknown as to how treatment affected consumption within the billing period. The post period began in the first billing period following the deadband period. The post variable was defined as a "0" in the billing periods prior to the beginning of treatment and a "1" for billing periods following the beginning of treatment. We used the R^2 value to evaluate the fit of the regression model.

Distribution of Bundled Savings

The process for distributing the energy savings between the programs was simple once we had chosen the hierarchy of programs (Figure 1) and calculated the annual energy savings for each program separately. We identified the DI program as the highest priority to receive the bundled savings, as the measures in this program have well documented savings and their installation was readily verifiable. The optimization driven energy savings from smart thermostats in the residential DR program was the second priority because there was a statistically significant sample available for the regression model once all cross participants were removed. Lastly, over 90 percent of the 2019 IHEA participants had bundled installations, which meant the regression model had to keep bundled cross participants in the sample to get statistically significant energy savings estimates.

We took the regression results for the IHEA program, which included DI and DR smart thermostat participants, and subtracted the ex-post energy savings for DI and DR smart thermostat cross participants from the total to get the energy savings attributable to the IHEA program.

Results

For the M&V analyses associated with the 2019 programs, the individual program ex-post savings are presented below.

Direct Install

Table 3 provides a summary of the final ex-post verified energy impacts for the average participant in the 2019 DI program.

Table 3. Summary of Average Ex-Post Annual Energy Saving per Program Participant by Measure for the 2019 DI Program.

Measure Type	Average Annual Ex-Post kWh
Air Filter/Furnace Filter	50
Light-Emitting Diode (LED) (7W A19)	4
LED (8W BR30)	19
LED (9W A19)	98
LED (9W BR30)	56
LED (11W BR30)	7
LED (11W A19)	8
Photocell	1
Refrigerator Thermometer	4
Air-conditioner Refrigerant Line Insulation	2
Total	250

Residential Demand Response -- Smart Thermostats

The following tables detail the daily electricity savings associated with the smart thermostat optimization services.

Table 4. Average Daily Electricity Savings (kWh), Optimization, Space Cooling (Jun – Sep).

Treatment Group		Control Group		Estimated Average per Premise Daily Electricity Savings from Optimization
Average Daily Post-Installation Period Consumption (kWh)	Average Daily Pre-Installation Period Consumption (kWh)	Average Daily Post-Installation Period Consumption (kWh)	Average Daily Pre-Installation Period Consumption (kWh)	
46.49	47.96	47.82	47.59	1.71

Table 5. Average Daily Electricity Savings (kWh), Optimization, Space Heating (Oct – May).

Treatment Group		Control Group		Estimated Average per Premise Daily Electricity Savings from Optimization
Average Daily Post-Installation Period Consumption (kWh)	Average Daily Pre-Installation Period Consumption (kWh)	Average Daily Post-Installation Period Consumption (kWh)	Average Daily Pre-Installation Period Consumption (kWh)	
21.97	21.56	23.36	21.63	1.32

The annual per premise energy savings associated with optimization during space cooling was 209 kWh (122 days x 1.71 kWh). The annual per premise energy savings associated with optimization during space heating was 321 kWh (243 days x 1.32 kWh). The total annual per premise energy savings associated with the optimization was 530 kWh (209 kWh + 321 kWh). This savings value included the energy savings that was attributable to the DR events in 2019, we did not subtract out the DR event savings from the annual smart thermostat optimization until after we remove all of it from the IHEA modeled savings.

In-Home Energy Assessments

Table 6 provides the results of the mixed-effects panel regression modeling that was performed on the data. The negative Post variable was the daily energy savings attributed to IHEA participants and contained the annual savings from DI measures and DR[4] program smart thermostat optimization, 832 kWh (365 x 2.28 kWh).

Table 6. IHEA Results of Mixed Effects Panel Regression Modeling.

Programs	Intercept (t-value)	HDD65 (t-value)	CDD75 (t-value)	Post (t-value)	r-squared
IHEA	23.48 (27.41)	0.41 (16.99)	2.17 (120.12)	-2.28 (-10.17)	0.82

The data and steps used to distribute the energy savings attributed to IHEA by removing DI and Residential DR savings are shown in Table 7.

Table 7. Determining IHEA Annual Energy Savings per Participant.

IHEA, Direct Install, and Residential DR	Annual Energy (kWh) Savings per Participant	832
	Count of Participants	6,430
	Annual Energy (kWh) Savings	5,349,760
Cross-Participants	Residential DR Annual Energy (kWh) Savings	530
	Residential DR Participants	1,381
	Residential DR Annual Energy (kWh) Savings	731,718
	Direct Install Annual Energy (kWh) Savings	250
	Direct Install Participants	4,420
	Direct Install Annual Energy (kWh) Savings	1,104,940
IHEA only	Annual Energy (kWh) Savings	3,513,102
	Annual Energy (kWh) Savings per Participant	547

Summary and Conclusions

The IPMVP measurement and verification protocols do not cover bundled programs specifically, so there was room to create our own validation approach in this evaluation (EVO 2016). We were able to separate out the energy savings from the IHEA regression results by subtracting out the verified savings from the DI and DR programs as shown in Table 7. This process could have been more complicated if the smart thermostat group had a higher bundled rate, which would have required benchmarking on the smart thermostat savings as we wouldn't have been able to remove cross participants. Benchmarking could have been pulled from previous years (pre-bundled) evaluated savings or from TRMs in nearby states that have deemed savings for smart thermostats. Even in this situation, the smart thermostats would still have been ranked higher than the IHEA participants because it was a known installed measure.

As utilities are seeking to maximize the energy savings from their EE programs while reducing costs, bundling is becoming a very attractive option. We have shown that it was possible to accurately attribute energy savings between three different residential programs for participants in them. This

[4] Including the energy savings attributable to DR events. DR event savings was later removed from the annual energy savings value.

process of hierarchal ordering based on the certainty of energy savings and verification of measure installations, should be straightforward for other sets of bundled programs in the future.

References

EPA. 2019. *Guidebook for Energy Efficiency Evaluation, Measurement, and Verification: A Resource for State, Local, and Tribal Air & Energy Officials.*

EVO. 2016. *Core Concepts: International Performance Measurement and Verification Protocol.* EVO 10000-1:2016.

EVO. 2019. *Uncertainty Assessment for IPMVP: International Performance Measurement and Verification Protocol.* EVO 10100-1:2019