NAVIGANT ILLUME

Memorandum

To: Massachusetts Program Administrators
Massachusetts Energy Efficiency Advisory Council

From: Molly Podolefsky, Carly Olig, Debbie Brannan, Ken Seiden; Navigant
Anne Dougherty, Eileen Hannigan; Illume

Date: March 31, 2016

Re: Reducing the Size of the Control Group in the Home Energy Report Program

This memorandum addresses the evaluation team's (Navigant and Illume) research and calculations to inform Massachusetts Program Administrators (PAs) and the Energy Efficiency Advisory Council (EEAC) of opportunities to reduce the size of control groups in the Home Energy Report (HER) program. **Power analysis results present the PAs and EEAC with optimal numbers of control group customers which can be removed from each HER cohort and assigned to a new cohort, while taking into consideration the statistical confidence of resulting savings estimation, sensitivity analysis around potential deviations from expected savings values, and other key considerations.**

Key findings include:

- National Grid has the ability to reduce the control group size in six of the eleven HER program cohorts analyzed, resulting in over 100,000 new treatment customers.
- All four of the Eversource Energy (formerly NSTAR) HER program cohorts analyzed were found to be reducible, allowing nearly 50,000 new customers to be transitioned into a new treatment group.

Background

The Massachusetts PAs are interested in exploring the potential for transitioning a portion of the more than 500,000 customers currently assigned to HER control groups into treatment groups. This potential reduction in control groups and associated expansion of treatment groups would increase the number of customers receiving reports, thereby increasing overall program savings. Any reduction in the size of control groups would need to be conducted in a manner consistent with preserving the randomized controlled trial (RCT) structure of cohorts, while ensuring that remaining control groups are sufficient to support statistically significant estimation of program savings. Moreover, the PAs must take into consideration the potential effects of control group reductions on their ability to conduct future research.

1375 Walnut Street
Suite 100 | Boulder, CO 80302
303.728.2500 main
navigant.com

In order to appropriately weigh the risks and benefits of reducing HER program control groups, the PAs must consider all potential implications, including any negative or unforeseen consequences that other jurisdictions may have encountered in undertaking similar efforts.

Navigant consulted directly with Opower, the HER program implementation contractor, regarding other jurisdictions' experiences. To date, Opower has been involved with only two other control group reduction efforts. The manner in which control group reduction was carried out in some cases resulted in the randomization of cohorts not being maintained. This findings suggests that the current control group reduction effort should pay careful attention to ensuring that randomization is maintained in all cases.

Overall, Navigant's research into similar efforts around the country suggests that control group reduction is uncommon. Nevertheless, it has the potential to be implemented without issue provided care is taken to maintain randomization throughout. If the MA PAs decide to proceed with control group reduction, our reduction protocol will ensure randomization is maintained.

Methodology

This section describes (1) the key assumptions and considerations underlying the power analysis, and (2) the power analysis methodology.

Key Assumptions and Considerations

Several key assumptions and considerations underlie the power analysis, including transition method, level of analysis, effect size, statistical precision, and future research. Each of these is described in detail below.

Transition Method. The most basic assumption in conducting the analysis is determining how control group customers will be transitioned into a treatment group. There are three alternatives.

(1) Control group customers are transitioned into a new cohort-specific treatment group.

(2) Add the cohort-specific transitioned controls to the original treatment group, thereby expanding the original treatment group, or

(3) Form one new treatment group composed of all of the transitioned controls, i.e., is not cohort-specific.

Table 1 presents pros and cons of the three different approaches. **The proposed approach is to transition control customers into a new cohort-specific treatment group (option 1) because it minimizes implementation and analytical challenges associated with conducting cohort-specific analysis in the future while providing the most conservative estimate.** Both of the alternative approaches are viable and could allow for greater numbers of control customers to be transitioned. The analysis presented can be extended to either of the alternative methods should the PAs want to consider larger reductions.

Table 1. Pros and Cons of Control Transition Approaches

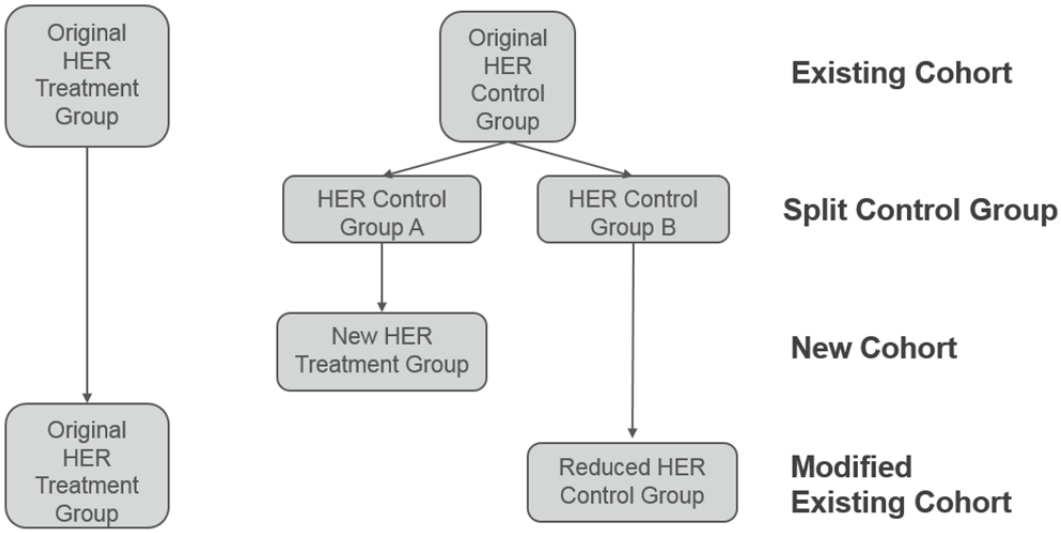
	Description	Pros	Cons
Option 1 (Proposed Approach)	Control group customers are transitioned into a new cohort-specific treatment group	• Provides conservative estimate • Maintains PA ability to conduct future cohort-specific analysis • Introduces few complications for estimating savings for transitioned group	• Potentially allows for smaller reduction of control group
Option 2	Add cohort-specific transitioned controls to the original treatment group, thereby expanding the original treatment group	• Potentially allows for larger reduction of control group	• Results in a treatment group of varying maturity levels introducing implementation and analytical challenges associated with tracking customers if cohort-specific savings estimates are of interest
Option 3	Form one new treatment group composed of all of the transitioned controls		• Requires formation of a new control group with the same distribution of cohorts as the new treatment group resulting in implementation and analytical challenges associated with tracking customers if cohort-specific savings estimates are of interest

Source: Navigant analysis

The modified existing (reduced) control group is maintained as the control group for both the original treatment group and the new treatment group. Using the existing modified control group to control for both the original and new treatment groups economizes the use of controls, avoiding the need to enroll more customers in a new control group.

Error! Reference source not found. illustrates the proposed methodology.

Figure 1. Transition of Control Groups



Source: Evaluation team analysis

Level of Analysis. Power analysis could theoretically be performed at the state level, the PA level, the cohort level, or a sub-cohort level. Recognizing that program impacts are estimated at the level of individual HER program cohorts by each PA, the power analysis is conducted at the cohort level.

The evaluation team excluded specific cohorts from the analysis for several reasons.

- HER cohorts typically experience increasing savings for the first few years, while program savings are ramping up. For this reason, very recent cohorts launched in 2014 and later may not have achieved their full savings potential as of early 2016. Including these cohorts in the power analysis would introduce unnecessary risk and uncertainty around savings, and as such were excluded from the analysis.
- Relatively small cohorts were also excluded as it was unlikely that reductions would be possible.
- Dual fuel cohorts include not just electric and gas customers, but also customers that were originally single fuel but later became dual-fuel customers. These dual fuel cohorts were removed from the analysis to avoid the inherent complications which might not be adequately addressed by the power analysis. However, power analysis could be conducted for these cohorts in the future as deemed appropriate by MA PAs.

Table 2 and

Table *3* detail which existing cohorts were included in the power analysis, which were excluded, and the reasons for their exclusion.

Memorandum

Table 2. Cohort Assignment to Power Analysis—Eligible Cohorts

PA	Cohort Name	Fuel Type	Number of Active Participants	Number of Controls	Eligible for Reduction	Reason
NGRID	2009 Group	Electric	21,187	24,859	Yes	NA
NGRID	2010 Group A	Electric	65,175	32,582	Yes	NA
NGRID	2011 Group A	Electric	83,788	23,387	Yes	NA
NGRID	2011 Group B	Electric	48,870	32,482	Yes	NA
NGRID	2012 Group	Electric	70,408	41,450	Yes	NA
NGRID	2013 Group	Electric	305,547	49,985	Yes	NA
NGRID	2013 Group Email	Electric	44,886	19,998	Yes	NA
NGRID	2009 Group	Gas	21,394	17,139	Yes	NA
NGRID	2011 Group B	Gas	19,466	17,302	Yes	NA
NGRID	2012 Group	Gas	65,299	14,778	Yes	NA
NGRID	2013 Group	Gas	128,416	29,890	Yes	NA
Eversource	2012 Group A	Electric	55,865	20,959	Yes	NA
Eversource	2012 Group B	Electric	17,038	14,999	Yes	NA
Eversource	2013 Group A	Electric	37,815	15,741	Yes	NA
Eversource	2013 Group	Gas	21,394	15,691	Yes	NA

Source: Navigant analysis
Notes: Cohorts denoted as "Eversource" cohorts were NSTAR cohorts prior to 2015.

1375 Walnut Street
Suite 100 | Boulder, CO 80302
303.728.2500 main
navigant.com

Table 3. Cohort Assignment to Power Analysis—Ineligible Cohorts

PA	Cohort Name	Fuel Type	Number of Active Participants	Number of Controls	Eligible for Reduction	Reason
Berkshire	2014 Group	Gas	11,999	9,997	No	New/Small Sample
NGRID	2010 Group B	Electric	20,928	9,958	No	Small Sample
NGRID	2012 Dual Group	Electric	11,421	13,490	No	Dual Fuel
NGRID	2014 Group	Electric	94,874	19,981	No	New
NGRID	2010 Group	Gas	75,911	8,274	No	Small Sample
NGRID	2011 Group A	Gas	77,455	6,529	No	Small Sample
NGRID	2012 Dual Group	Gas	11,398	13,454	No	Dual Fuel
NGRID	2014 Group	Gas	49,741	9,955	No	Small Sample
Eversource	2010 Dual Group	Electric	15,535	18,410	No	Dual Fuel
Eversource	2011 Dual Group	Electric	7,089	6,709	No	Small Sample
Eversource	2013 Group B	Electric	65,826	10,489	No	Small Sample
Eversource	2013 Dual Group	Electric	18,974	15,746	No	Dual Fuel
Eversource	2014 Group	Electric	78,639	10,486	No	Small Sample
Eversource	2010 Dual Group	Gas	19,318	24,270	No	Dual Fuel
Eversource	2011 Dual Group	Gas	19,017	19,770	No	Dual Fuel
Eversource	2013 Dual Group	Gas	18,934	15,716	No	Dual Fuel
WMECo	2014 Group	Electric	113,782	9,997	No	New/Small Sample

Source: Evaluation team analysis
Notes: Cohorts denoted as "Eversource" cohorts were NSTAR cohorts prior to 2015.

Effect Size. A final assumption critical to the power analysis is the anticipated program effect size. Percent savings in a HER program varies by cohort and over time. For this reason, it was determined that a uniform effect size should not be assumed throughout the analysis. Rather, each cohort in the analysis should be assigned its own expected effect size.

Given control group customers are transitioned into a new treatment group, two separate power analyses that must be conducted:

(1) A power analysis of the original treatment group using the modified existing control group. For this group, the most appropriate assumption is to base savings on the most recent impact evaluation.[1]

(2) A power analysis of the new treatment group (the transitioned former control customers) using the modified existing control group. As a new treatment group, it is likely the first year savings will be similar to the original treatment group's first year savings. As such we assume the new treatment group's first year savings rate equals the original treatment group's first year savings rate.[2]

Effect size is a critical assumption in the power analysis. Given the realized savings may be lower than the assumed savings value, the evaluation team conducted a sensitivity analysis whereby four alternate effect sizes were tested: 5% reduction in the assumed effect size, 10% reduction, 0.5 percentage point reduction, and a 1 percentage point reduction.[3] This analysis provides the PAs with increasingly conservative options for reducing the size of the control group while maintaining the ability to estimate savings with statistical precision.

Statistical Significance. Because the power analysis seeks to determine just how far control groups can be reduced while retaining the ability to estimate savings with the desired level of statistical significance, it is important to determine an appropriate level of statistical significance in advance. Accepted significance levels vary widely depending on regulatory requirements and other considerations. While many jurisdictions report HER savings which are statistically significant at the 10% level or lower, others require much more rigorous savings estimates. According to the SEEAction (2012) report on EM&V of residential behavior-based energy efficiency programs, the level of precision considered acceptable is that the savings estimate be statistically significant at the 5% level or lower.[4] In consultation with the evaluation team, the PAs and EEAC have determined that a 5% level of statistical significance should be used throughout the analysis.

---

[1] Massachusetts Cross-Cutting Behavioral Program Evaluation Opower Results, March 9, 2015, prepared by Navigant for the Massachusetts Program Administrators (PAs) and Energy Efficiency Advisory Council (EEAC)

[2] Massachusetts Cross-Cutting Behavioral Program Evaluation, Integreated Report, June 2013, prepared by Opinion Dynamics for the Massachusetts Program Administrators (PAs) and Energy Efficiency Advisory Council (EEAC)

[3] The evaluation team first conducted the sensitivity analysis using absolute reductions in the savings level for both the new treatment group and original treatment group of 0.5 percentage points and 1 percentage point. The results were very sensitive to these reductions as they represented in some cases very large percentage reductions in the effect size. In fact, only three cohorts were reducible under the 0.5 percentage point scenario, and zero were reducible under the 1 percentage point scenario. In order to refine the sensitivity analysis, the evaluation team conducted further sensitivity analysis testing a 10% and 15% reduction on both groups.

[4] *Evaluation, Measurement, and Verification (EM&V) of Residential Behavior-Based Energy Efficiency Programs: Issues and Recommendations*, State and Local Energy Efficiency Action Network (2012)

Future Research. In deciding whether or not to implement control group reductions based on the findings of this study, the PAs will need to weigh these risks against the benefits of increased program participation. For example, reducing the size of control groups may negatively impact the PAs' ability to conduct future studies (e.g., a persistence study). As a result, the PAs may decide to take a conservative approach to control group reduction, or to leave some control groups untouched with the intent of using these intact control groups in future studies.

Power Analysis Methodology

Power analysis is a regression-based simulation exercise that utilizes assumptions on statistical significance, effect size and other considerations to identify the minimum sample size needed to achieve desired results. In the context of reducing the size of HER control groups, power analysis consists of running a series of simulated regressions to evaluate expected savings and confidence intervals for increasingly larger control group reductions (i.e. half, three-quarters, four-fifths).

Customers removed from each cohort's control group will form their own independent, new treatment group. This new treatment group will use the modified existing control group for control, obviating the need to pull additional customers into a new control group. Based on this transition method, for each progressively higher reduction level, two simulated regressions are conducted:

(1) simulate savings and confidence intervals for the original treatment group using the modified existing control group, assuming the percent savings equals the original treatment group's most recent evaluated savings rate

(2) simulate savings and confidence intervals for the new treatment group using the modified existing control group, assuming the percent savings equals the original treatment group's first year savings rate

The power analysis simulations are conducted iteratively until the simulated savings estimate for either group is no longer statistically significant at the 95% significance level. That result identifies the largest feasible reduction in the control group given the desired significance level. In order to provide a conservative recommendation, *the reduction level one lower than the last statistically significant level* is chosen as the optimal reduction.

Power Analysis Results

This section presents the following results (1) main power analysis results, (2) results of the sensitivity analysis, and (3) one detailed example of a cohort-level analysis.

Main Power Analysis

Results of the power analysis suggest that the MA PAs can reduce control groups significantly relative to current sizes and still estimate savings with a high level of statistical rigor. Maintaining the ability to estimate savings for both existing and new treatment groups at the 95% significance level, National Grid and Eversource Energy have the potential to transition approximately 100,000 and 50,000 of their control customers, respectively, into new treatment groups. National Grid would be able to transition approximately 80,000 electric and 21,000 gas control customers into new treatment groups. Eversource Energy would be able to transition approximately 39,000 and 10,000 control customers into new electric and gas treatment groups, respectively.

Table 4 provides a high-level summary of power analysis results by PA and cohort. Blanks represent cohorts that were deemed ineligible for reduction and not included in the power analysis, according to the rationale presented in Table 3. For cohorts included in the power analysis, the table shows details regarding both the original control group and modified existing and new cohorts post-transition.

Table 4. Power Analysis Results Summary

					Modified Existing and New Cohorts		
PA	Cohort Name	Fuel Type	Original Control Group	Control Group Optimal Reduction*	Original Treatment Group	Reduced Control Group	New Treatment Group
NGRID	2009 Group	Electric	24,859	50%	21,187	20,555	4,304
NGRID	2010 Group	Electric	32,581	None	65,175		
NGRID	2011 Group A	Electric	23,387	None	83,788		
NGRID	2011 Group B	Electric	32,482	75%	48,870	13,866	18,616
NGRID	2012 Group	Electric	41,449	83%	70,408	17,453	23,996
NGRID	2013 Group	Electric	49,961	93%	305,547	17,204	32,757
NGRID	2013 Group Email	Electric	19,998	None	44,886		
NGRID	2009 Group	Gas	17,138	75%	21,394	6,902	10,236
NGRID	2011 Group	Gas	17,294	None	19,466		
NGRID	2012 Group	Gas	14,774	None	65,299		
NGRID	2013 Group	Gas	29,880	50%	128,416	18,607	11,273
Total							**101,182**
Eversource	2012 Group A	Electric	20,958	93%	55,865	3,998	16,960
Eversource	2012 Group B	Electric	14,999	86%	17,038	4,327	10,672
Eversource	2013 Group	Electric	15,740	91%	37,815	4,093	11,647
Eversource	2013 Group	Gas	15,690	75%	21,394	6,080	9,610
Total							**48,889**

Source: Evaluation team analysis

**This percentage represents the optimal reduction in the remaining original control group post data cleaning for regression analysis*

Sensitivity Analysis

As described in the Methodology section, the evaluation team conducted a sensitivity analysis to account for the possibility that savings may be lower than the effect sizes assumed in this analysis. Assumed effect sizes were reduced by 10%, 15%, 0.5 percentage points, and 1 percentage point. The results of all four scenarios are summarized in Table 5 below. Cohorts which could not be reduced under the main power analysis results were not tested farther in the sensitivity analysis, and are denoted by blanks in Table 5. Zeros represent cohorts which were included in the sensitivity analysis, but could not be reduced at given levels.

Table 5. Sensitivity Analysis Results Summary

				Modified Existing and New Cohorts				
PA	Cohort Name	Fuel Type	Original Control Group	Original Treatment Group	Reduced Control Group for 10% Sensitivity	New Treatment Group for 10% Sensitivity	Reduced Control Group for 15% Sensitivity	New Treatment Group for 15% Sensitivity
NGRID	2009 Group	Electric	24,859	21,187	20,558	4,301	20,558	4,301
NGRID	2010 Group	Electric	32,581	65,175				
NGRID	2011 Group A	Electric	23,387	83,788				
NGRID	2011 Group B	Electric	32,482	48,870	14,610	17,872	32,482	0
NGRID	2012 Group	Electric	41,449	70,408	14,398	27,051	14,398	27,051
NGRID	2013 Group	Electric	49,961	305,547	32,757	32,591	31,907	28,077
NGRID	2013 Group Email	Electric	19,998	44,886				
NGRID	2009 Group	Gas	17,138	21,394	10,314	6,824	17,138	0
NGRID	2011 Group B	Gas	17,294	19,466				
NGRID	2012 Group	Gas	14,774	65,299				
NGRID	2013 Group	Gas	29,880	128,416	29,880	0	29,880	0
Total						**88,639**		**59,429**
Eversource	2012 Group A	Electric	20,958	55,865	4,215	16,743	4,722	16,236
Eversource	2012 Group B	Electric	14,999	17,038	4,623	10,376	4,623	10,376
Eversource	2013 Group	Electric	15,740	37,815	4,209	11,531	4,351	11,389
Eversource	2013 Group	Gas	15,690	21,394	9,283	6,407	15,690	0
Total						**45,057**		**38,001**

Source: Navigant team analysis

As shown in the sensitivity analysis, successively larger reductions in the assumed savings rate generally result in smaller potential reductions of the control group. For some control groups that were just reducible under the main power analysis scenario, such as National Grid's 2013 gas cohort, any sensitivity analysis resulted in non-reducibility. For other cohorts more robust to changes in savings levels, such as Eversource Energy's (NSTAR) 2013 electric cohort, control groups are reducible even under a 0.5 percentage point drop in savings scenario. Table 6 presents a summary of the sensitivity analysis.

Table 6. Sensitivity Analysis Summary

	None	10%	15%	0.5 point	1.0 point
NGRID	102,673	88,639	63,250	28,077	0
Eversource	48,790	45,057	38,001	20,104	0

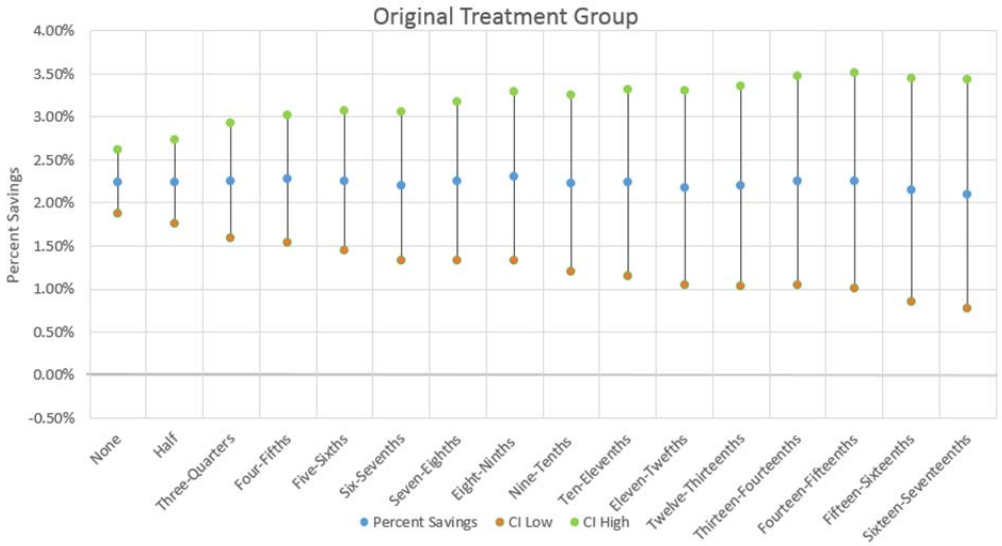
Source: Navigant team analysis

Cohort-Level Detailed Analysis

In order to convey a detailed understanding of the power analysis process, the remainder of this section will examine power analysis and sensitivity analysis results for Eversource Energy's 2012 Electric cohort in detail.

Figure 2 presents the main power analysis results specific to Eversource Energy's 2012 Electric cohort's original treatment group. This power analysis serves to determine the largest reduction in the cohort's control group which will still support savings estimation for the original treatment group significant at the 95% confidence level. As shown in **Error! Reference source not found.**, for all of the reduction levels considered, the 95% confidence interval remains well above zero, suggesting the control group could still be reduced considerably.

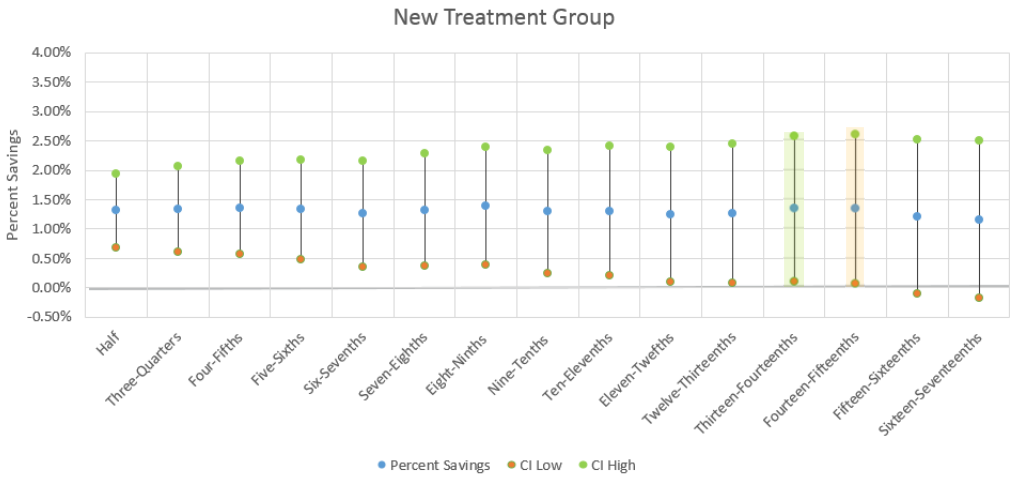
Figure 2. Eversource Energy 2012 Electric – Original Treatment Group



Source: Evaluation team analysis

Figure 3 presents the main power analysis results specific to the new treatment group formed from controls transitioned out of the original control group. In contrast to **Error! Reference source not found.**, this figure illustrates that by the time the control group has been reduced by fifteen-sixteenths, the confidence interval for that simulated savings regression spans zero and is no longer statistically significant at the 95% level. Because fourteen-fifteenths is the last reduction level resulting in statistically significant simulated savings, thirteen-fourteenths (93%) is chosen as the optimal control group reduction level.

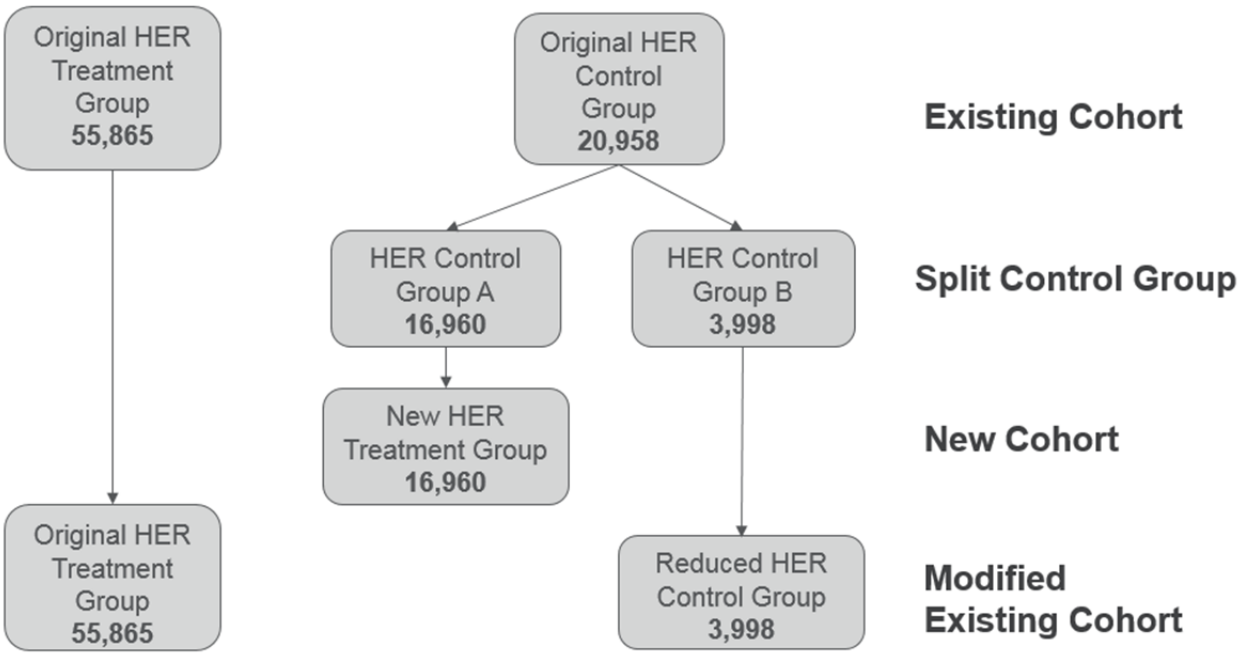
Figure 3. Eversource Energy 2012 Electric – New Treatment Group



Source: Evaluation team analysis

As a result under the most likely savings scenarios, Eversource Energy will be able to reduce their 2012 electric cohort control group by nearly 17,000 customers. Figure 4 illustrates the control group transition process for this cohort in detail under main power analysis results:

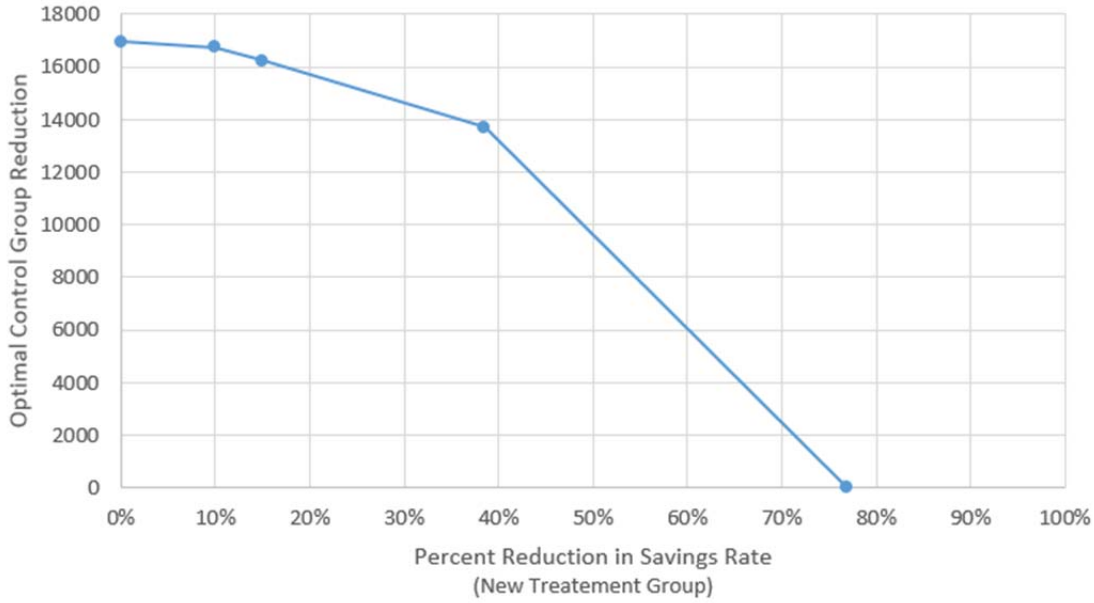
Figure 4. Eversource Energy HER Control Group Reduction



Source: Evaluation team analysis

A final step in analyzing the Eversource Energy 2012 electric cohort's reducibility is to conduct sensitivity analyses assuming savings levels are lower than what is expected based on historical values. Figure 5 examines the results of decreasing assumed savings levels for the new and original treatment groups by successively greater increments. Reducing savings by 10%, 15%, 38% (representing a 0.5 point reduction in assumed savings) and 77% (representing a 1.0 point reduction in assumed savings) allows for successively smaller control group reductions of 16,743, 16,236, 13,698 and zero customers, respectively.

Figure 5. Eversource Energy 2012 Electric Sensitivity Analysis Results

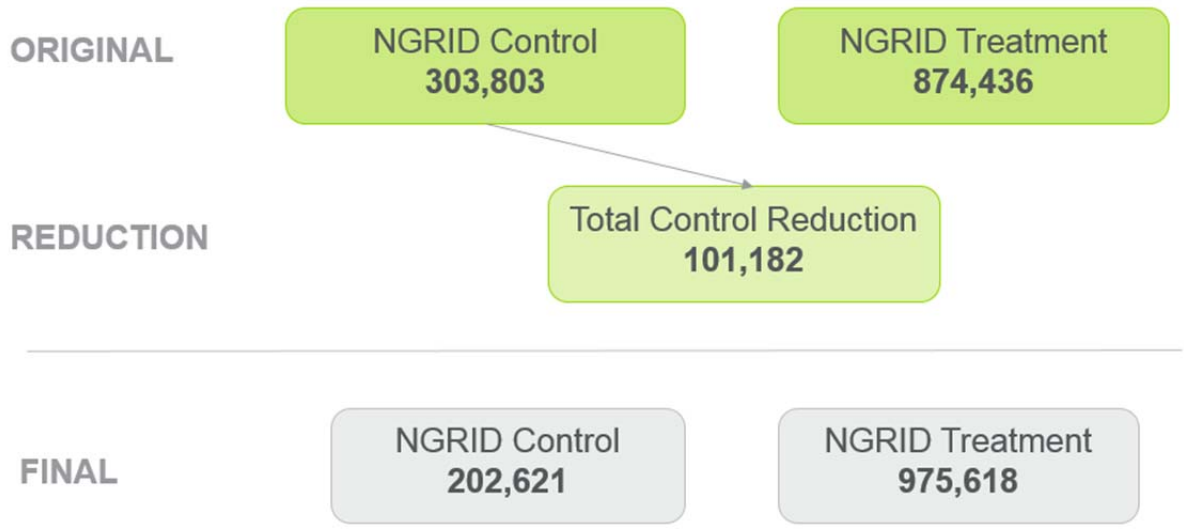


Source: Evaluation team analysis

Summary

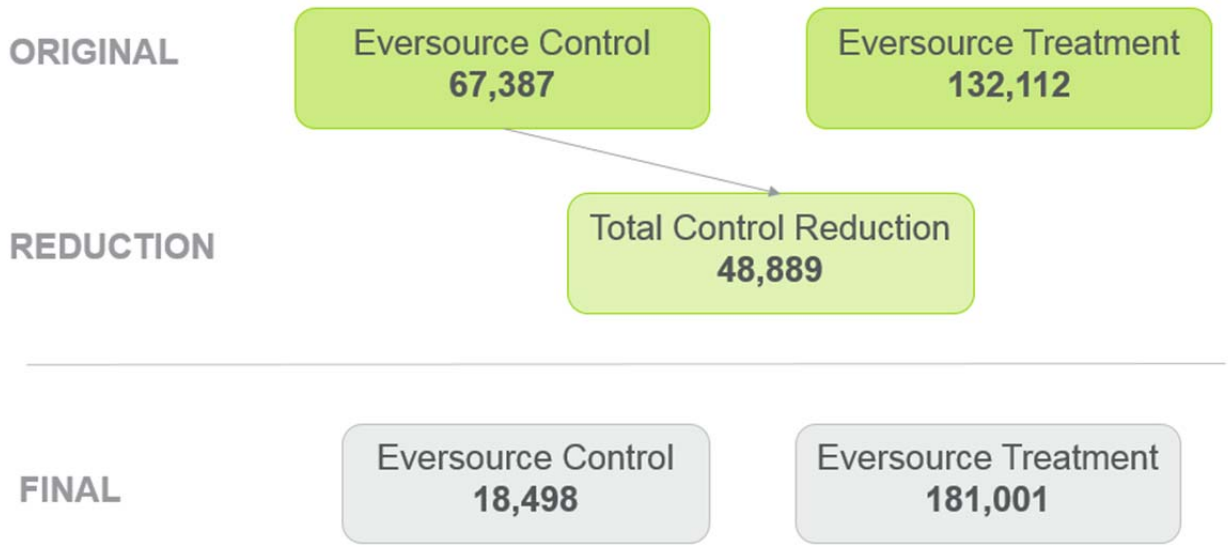
The power analysis suggests that the MA PAs have considerable ability to reduce the size of control groups while maintaining highly statistically significant estimates of savings both for existing and new treatment groups. Figure 6 and Figure 7 illustrate the total reduction in control groups achievable between National Grid and Eversource Energy under the main power analysis results and the resulting increase in program participants.

Figure 6. NGRID Total Control Group Reduction



Source: Evaluation team analysis

Figure 7. Eversource Energy Total Control Group Reduction



Source: Evaluation team analysis