NAVIGANT

ILLUME

1375 Walnut Street
Suite 200
Boulder, CO 80302
303-728-2500

Memorandum

To: Massachusetts Program Administrators (PAs) and Energy Efficiency Advisory Council (EEAC)

From: Navigant Consulting, Inc. and Illume Advising, LLC.

Date: March 9, 2015

Re: Massachusetts Cross-Cutting Behavioral Program Evaluation Opower Results

Executive Summary

This memo was developed by the "evaluation team", including Navigant Consulting, Inc. (Navigant), the prime contractor, and Illume Advising, LLC. (ILLUME) to present results for the impact evaluation of the Home Energy Report (HER) programs for the Massachusetts Cross-Cutting Research in Behavior and Education. In particular, this memo presents results for National Grid and Eversource Energy (formerly NSTAR and Western Massachusetts Electric Company (WMECo)).

National Grid and NSTAR have long-standing HER programs, with both companies dramatically expanding their behavioral programs since the launch of the first program by National Grid in 2009. The last impact evaluation of the HER programs was completed for the 2012 program year and the last process evaluation was completed in 2011.

As of the end of 2014, there were a total of thirty-one evaluable cohorts in the Massachusetts HER programs. This includes fifteen electric cohorts (nine for National Grid, five for NSTAR, and one for WMECo), eight gas cohorts (seven for National Grid and one for NSTAR), and four dual fuel cohorts (one for National Grid and three for NSTAR).

Key Findings

- Total net electric savings from the Massachusetts HER programs are 127,854,643 kWh. Total net gas savings are 643,157 MMBtu.

- The evaluation team recommends that the PAs adopt the following savings estimate ratios[1] in future years when third-party impact evaluations are not completed.
    - National Grid Electric: 95%
    - National Grid Gas: 98%
    - NSTAR Electric: 104%
    - NSTAR Gas: 98%
    - WMECo Electric: 104%

Methodology

This section presents the methodology used for: (1) impact evaluation, (2) channeling analysis and (3) calculation of savings estimate ratios.

Impact Evaluation Methods

The evaluation team conducted multiple analyses to determine cohort-specific savings, including the post-program regression (PPR) and linear fixed effects regression (LFER) models. The PPR model combines both cross-sectional and time series data in a panel dataset. This model uses only the post-program data, with lagged energy use for the same calendar month of the pre-program period acting as a control for any small systematic differences between the participant and control customers. In particular, energy use in calendar month *t* of the post-program period is framed as a function of both the participant variable and energy use in the same calendar month of the pre-program period. The underlying logic is that systematic differences between participants and controls will be reflected in differences in their past energy use, which is highly correlated with their current energy use. The version we estimate includes monthly fixed effects and interacts these monthly fixed effects with the pre-program energy use variable. These interaction terms allow pre-program usage to have a different effect on post-program usage in each calendar month.

Formally, the model is,

Model 1. PPR Model

$$ADC_{kt} = \sum_{J} \beta_{1j} Month_{jt} + \sum_{J} \beta_{2j} Month_{jt} \cdot ADClag_{kt} + \beta_3 Participant_k + \varepsilon_{kt} \ ,$$

where,

ADC_{kt}	= The average daily consumption in kWh or therms for customer *k* during billing cycle *t*. This is the dependent variable in the model;
$Month_{jt}$	= A binary variable taking a value of 1 when *j*=*t* and 0 otherwise;[2]
$ADClag_{kt}$	= Customer *k*'s energy use in the same calendar month of the pre-program year as the calendar month of month *t*;

[1] The saving estimate ratio is the calculated by dividing the modeled savings estimated by the evaluation team by the savings estimated by Opower.

[2] If there are *T* post-program months, there are *T* monthly dummy variables in the model, with the dummy variable $Month_{tt}$ the only one to take a value of 1 at time t. These are, in other words, monthly fixed effects.

$Participant_k$	= A binary variable indicating whether customer k is in the participant group (taking a value of 1) or in the control group (taking a value of 0);
ε_{kt}	= The cluster-robust error term for customer k during billing cycle t. Cluster-robust errors account for heteroscedasticity and autocorrelation at the customer level.[3]

In this model, β_3 is the estimate of average daily energy savings due to the program. Program savings are the product of the average daily savings estimate and the total number of participant-days in the analysis.

To test the robustness of the savings estimate to model specification, the evaluation team also estimated savings using the standard regression approach for estimating electricity savings for a Randomized Controlled Trial (RCT), a linear fixed effects regression (LFER) model. Until the most recent advances in this field using the PPR approach, the LFER approach was viewed as the most appropriate regression approach for RCT programs. According to both LBNL (2012) and a M&V white paper produced by the Brattle Group (2011), this approach was the preferred method for the evaluation of the energy use impacts of behavioral programs.[4] Moreover, it has been used in the academic literature to evaluate other HER programs, and has been used in all evaluations of HER programs done by Navigant as a robustness check on the newer PPR method.[5]

The simplest version of a LFER model, the One-Way LFER model, is one in which average daily consumption of kWh by customer k in bill t, denoted by ADC_{kt}, is a function of two variables: the binary variable $Treatment_k$, taking a value of 1 if household k is assigned to the treatment group, and 0 otherwise; and the binary variable $Post_t$, taking a value of 0 if the observation t is before the *program start date* and 1 if the observation is after the program start date.

---

[3] Ordinary Least Squares (OLS) regression models assume that the data are homoscedastic and not autocorrelated. If either of these assumptions is violated, the resulting standard errors of the parameter estimates are incorrect (usually underestimated). A random variable is heteroscedastic when the variance is not constant. A random variable is autocorrelated when the error term in one period is correlated with the error terms in at least some of the previous periods.

[4] LBNL 2012: *Evaluation, Measurement, and Verification (EM&V) of Residential Behavior-Based Energy Efficiency Programs: Issues and Recommendations*. Prepared by A. Todd, E. Stuart, S. Schiller, and C. Goldman, Lawrence Berkeley National Laboratory. 2012. http://behavioranalytics.lbl.gov.
Brattle Group 2011: *Measurement and Verification Principles for Behavior-Based Efficiency Programs*. Sanem Sergici and Ahmad Fauqui. 2011.

[5] For examples of academic applications of the approach to energy behavioral programs see:
Alcott, Hunt. "Social Norms and Energy Conservation", Working paper, Massachusetts Institute of Technology (MIT), Cambridge, MA, 2009.
Ayres, I., S. Raseman and A. Shih. "Evidence from Two Large Field Experiments that Peer Comparison Feedback Can Reduce Residential Energy Usage", NBER working paper no. 15386, September 2009.
Costa, D.L. and M.E. Kahn. "Energy Conservation "Nudges" and Environmentalist Ideology: Evidence from a Randomized Residential Electricity Field Experiment", NBER working paper no. 15939, April 2010.

Formally, the model is,

Model 2. One-Way LFER Model

$$ADC_{kt} = \alpha_{0k} + \alpha_1 Post_t + \alpha_2 Treatment_k \cdot Post_t + \varepsilon_{kt} .$$

Three observations about this specification deserve comment. First, the coefficient α_{0k} captures *all* customer-specific effects on energy use that do not change over time, including those that are unobservable. Second, α_1 captures the average effect among control customers of being in the post treatment period. In other words, it captures the effects of exogenous factors, such as an economic recession, that affect control customers in the post treatment period but not in the pre-treatment period. Third, $\alpha_1 + \alpha_2$ captures the average effect among treatment customers of being in the post treatment period, and so for these households the effect directly attributable to the program is captured by the coefficient α_2.

Both the PPR and LFER models produce unbiased estimates of the program effect. The evaluation team recommends reporting out results from the PPR model as this model has been shown in past studies to have slightly smaller standard errors on average.

Channeling Analysis Methods

There are two goals to cross-program participation channeling analysis: (1) to document the lift in other program participation due to the behavioral program treatment (participant lift), and (2) to remove the savings co-generated by behavioral and standard programs in order to avoid double counting savings across the portfolio at the program and measure level (savings adjustment).

For this evaluation, it was not possible to do a data-based channeling analysis to document the participant lift or savings adjustment due to the timing of the collection of the program tracking data which is typically not available until several months after year-end. Complete and reconciled program tracking for 2014 will be available later in 2015. Consequently, the evaluation team applied a channeled savings adjustment based on historical data as reported in the last behavioral evaluation report.[6] The last evaluation report cross-referenced utility program databases to calculate the difference in tracked participation in other energy efficiency programs between the treatment and control groups. These differences are the savings that can be attributed to both the HER program and other utility programs.

Using historical values can introduce uncertainty into the estimates. This analysis attempts to minimize the uncertainty by using the most general estimate possible. The team estimated channel savings impacts by taking a weighted average by fuel type of channel impacts for all National Grid and NSTAR cohorts for all program years reported in the 2013 evaluation report. The calculations used each program year's cohort participants as weights.

---

[6] Opinion Dynamics Corporation, Navigant Consulting, and Evergreen Economics (2013). Massachusetts Cross-Cutting Behavioral Program Evaluation Integrated Report. Refer to section 3.1.1 for the channeling analysis methodology.

As shown in Appendix A., the electric weighted averages included 15 cohort-program year groups. Two cohorts spanned three years, two spanned two years, while the remaining five cohorts included only one year of implementation. The gas weighted averages included 15 cohort-program year groups with two cohorts spanning three years, three cohorts spanning two years and the remaining three cohorts comprising one year each.

The channel savings adjustments shown in Table 1 were applied by fuel type to all 2014 NGRID, NSTAR and WMECo cohorts.

Table 1. Historical Weighted Average Channeling Impacts

Fuel- Type	Number of Cohort-Year Groups	Total Number of Participants	Historical Weighted Average % of HH Baseline Usage Saved by Channeling Impacts
Electric	15	711,923	0.075%
Gas	16	613,006	0.053%

Savings Estimate Ratio Calculation Methods

The evaluation team calculated savings estimate ratio for each of the HER program cohorts to examine the differences in savings as measured by the program implementer and the savings verified by our evaluation. For the individual cohorts, the savings estimate ratio is the ratio of estimated savings to the implementer-reported estimates as shown in Equation 1.

Equation 1. Savings Estimate Ratio by Cohort

$$RR_{ucf} = \frac{\text{Evaluation Team Estimated Savings}_{ucf}}{\text{OPOWER Reported Savings}_{ucf}}$$

Where,

RR_{ucf} = The savings estimate ratio for PA u, cohort c, and fuel-type f.

The evaluation team also calculated aggregate savings estimate ratios for each PA and fuel-type. To do this, for each PA and fuel-type we calculated the ratio of total estimated savings from the relevant cohorts to total implementer-reported estimates for the relevant cohorts, as shown in Equation 2. The evaluation team calculated individual cohort and aggregate savings estimate ratios before and after adjusting the estimated savings for channeling.

Equation 2. Aggregate Savings Estimate Ratio by PA and Fuel-Type

$$RR_{uf} = \frac{\sum_{c=1}^{C} \text{Evaluation Team Estimated Savings}_{ucf}}{\sum_{c=1}^{C} \text{OPOWER Reported Savings}_{ucf}}$$

Where,

RR_{uf} = The aggregate realization rate for PA u and fuel-type f.

The evaluation team recommends an approach for applying a savings estimate ratio by PA and fuel type to be used in future years to adjust implementer-reported savings when third-party impact evaluations do not take place.

Results

This section presents (1) impact evaluation results and (2) savings estimate ratio results.

Impact Evaluation Results

Table 2 presents detailed savings estimates for each cohort after the adjustment from the channeling analysis. Baseline usage is presented as both the value in the current TRM[7] and the value from the data for this analysis. A '-' in the TRM Baseline column indicates that the cohort does not have a baseline usage value listed in the TRM.

Two cohorts, NSTAR Group 2010 Dual and NSTAR Group 2011 Dual, have relatively low electric percentage savings estimates, 0.15% and 0.49% respectively. These two groups are relatively small (18,660 and 8,451 treatment customers respectively) and only started receiving dual fuel reports in July 2014. Given that these are estimates of savings after just six months, it is not surprising that these savings estimates are low and it is likely that percentage savings will increase for these cohorts going forward.

[7] Massachusetts Technical Reference Manual. 2013-2015 Program Years.

Table 2. Savings Estimates after Channeling Adjustment, by Cohort

PA	Cohort Name	Fuel-Type	Total Number of Participants	TRM Baseline Usage (kWh/MMBtu)	Modelled Baseline Usage (kWh/MMBtu)	Average Annual Savings per Customer (kWh/MMBtu)*	Percentage Savings*	Total Savings (kWh/MMBtu)*
NGRID	Group 2009	Electric	24,005	11,233	10,669	252.85	2.37%	5,116,541
NGRID	Group 2010	Electric	65,170	12,370	11,815	186.68	1.58%	11,993,567
NGRID	Group 2010 Added	Electric	23,805	15,232	14,682	340.62	2.32%	6,772,880
NGRID	Group 2011	Electric	99,446	9,638	9,415	236.32	2.51%	18,673,469
NGRID	Group 2011 Added	Electric	60,605	6,121	5,986	93.99	1.57%	4,184,350
NGRID	Group 2012	Electric	86,898	6,126	6,003	135.41	2.20%	1,430,655
NGRID	Group 2012 Dual	Electric	12,621	6,239	6,155	93.64	1.56%	5,974,252
NGRID	Group 2013	Electric	324,002	8,036	8,053	105.50	1.31%	28,469,571
NGRID	Group 2013 Email	Electric	46,105	-	7,082	35.41	0.50%	1,443,224
NGRID	Group 2014	Electric	94,874	7,093	7,303	65.72	0.90%	4,033,771
NGRID	Group 2009	Gas	24,790	127.20	139.97	2.09	1.49%	43,727
NGRID	Group 2010	Gas	75,911	31.28	147.01	2.75	1.87%	172,435
NGRID	Group 2011	Gas	100,321	92.90	103.25	1.16	1.12%	86,660
NGRID	Group 2011 Add	Gas	25,673	19.44	86.94	1.03	1.19%	19,356
NGRID	Group 2012	Gas	86,279	81.00	86.73	1.54	1.77%	96,009
NGRID	Group 2012 Dual	Gas	13,416	84.20	95.91	1.05	1.09%	11,553
NGRID	Group 2013	Gas	149,442	76.18	82.67	0.74	0.89%	89,334
NGRID	Group 2014	Gas	49,741	-	112.27	0.92	0.82%	14,558
NSTAR	Group 2010 Dual	Electric	18,660	-	8,127	16.25	0.20%	124,152
NSTAR	Group 2011 Dual	Electric	8,451	-	7,031	39.37	0.56%	132,707
NSTAR	Group 2012a	Electric	55,857	13,027	13,041	281.68	2.16%	15,381,055
NSTAR	Group 2012b	Electric	17,033	11,388	11,085	228.36	2.06%	3,761,491
NSTAR	Group 2013	Electric	37,801	8,423	11,869	153.11	1.29%	5,467,905
NSTAR	Group 2013b	Electric	65,798	-	6,427	71.98	1.12%	4,448,962
NSTAR	Group 2013 Dual	Electric	20,991	-	6,876	107.95	1.57%	915,705
NSTAR	Group 2014	Electric	8,637	-	6,780	53.56	0.79%	2,868,936
NSTAR	Gas Group 2010 Dual	Gas	24,345	102.2	128.92	2.08	1.61%	39,059
NSTAR	Gas Group 2011 Dual	Gas	24,689	89.6	114.51	1.90	1.66%	35,002
NSTAR	Attrition Refill 2013	Gas	38,411	65.5	90.73	0.67	0.74%	21,771
NSTAR	Gas Group 2013 Dual	Gas	20,943	-	73.66	0.78	1.06%	13,693
WMECo	Group 2014	Electric	113,782	-	7,645	67.28	0.88%	6,661,450

Source: Evaluation team analysis
*All savings estimates are after channeling adjustment.

Savings Estimate Ratio Results

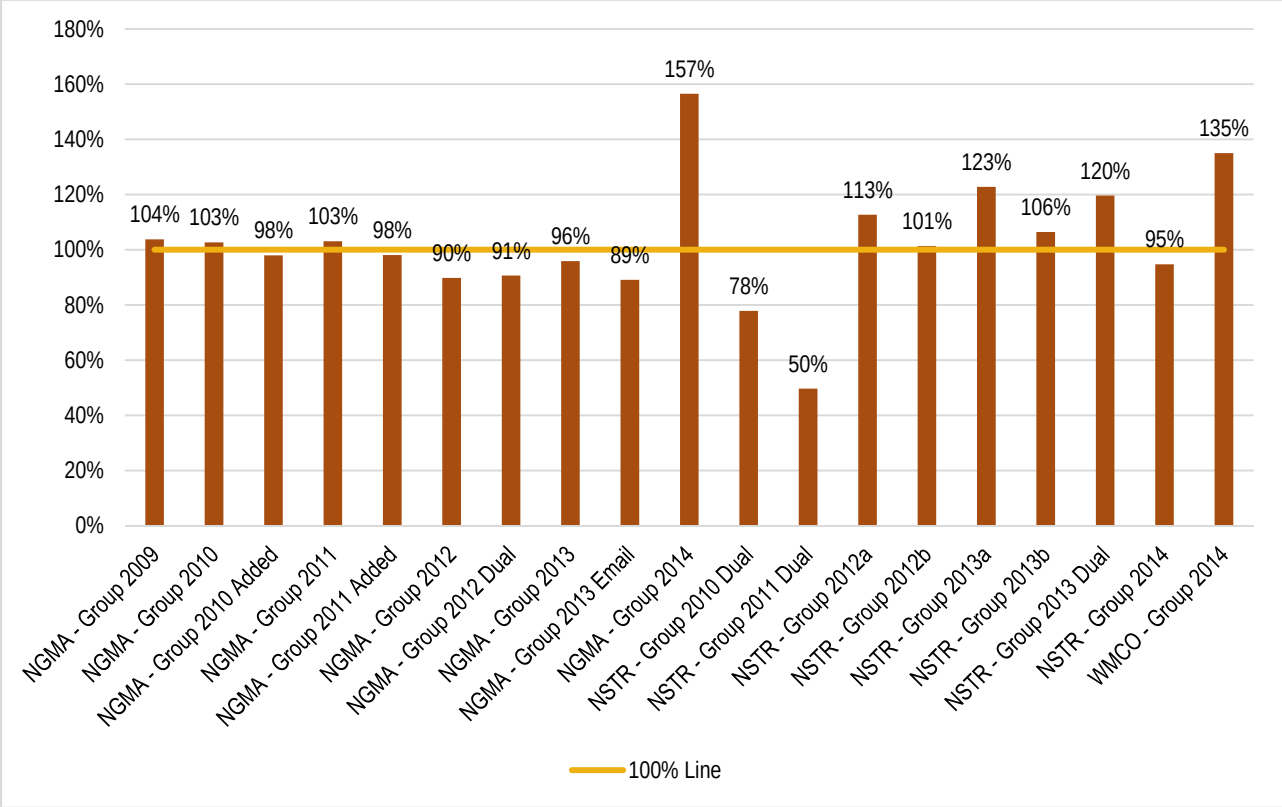
The evaluation team first presents cohort-specific savings estimate ratios followed by aggregate savings estimate ratio by PA and fuel-type.

Cohort-Specific Savings Estimate Ratio Results

Figure 1 and Figure 2 show the savings estimate ratios for each of the electric cohorts. Figure 1 is before the savings adjustment from the channeling analysis, and Figure 2 is after. Two cohorts,

NSTAR Group 2010 Dual and NSTAR Group 2011 Dual, have lower savings estimate ratios than expected, 78% and 50% before channeling respectively. These two groups are relatively small (18,660 and 8,451 treatment customers respectively) and only started receiving dual fuel reports in July 2014. Given the small amount of data for these two cohorts, it is not surprising that small differences in modeling between the evaluation team and Opower result in large differences in the savings estimate ratios. Two cohorts, National Grid Group 2014 and WMECo Group 2014, have relatively high savings estimate ratios, 157% and 135% before channeling respectively. However, these savings estimate ratios are not outside the bounds found for first year electric cohorts in previous evaluations of this program.[8]

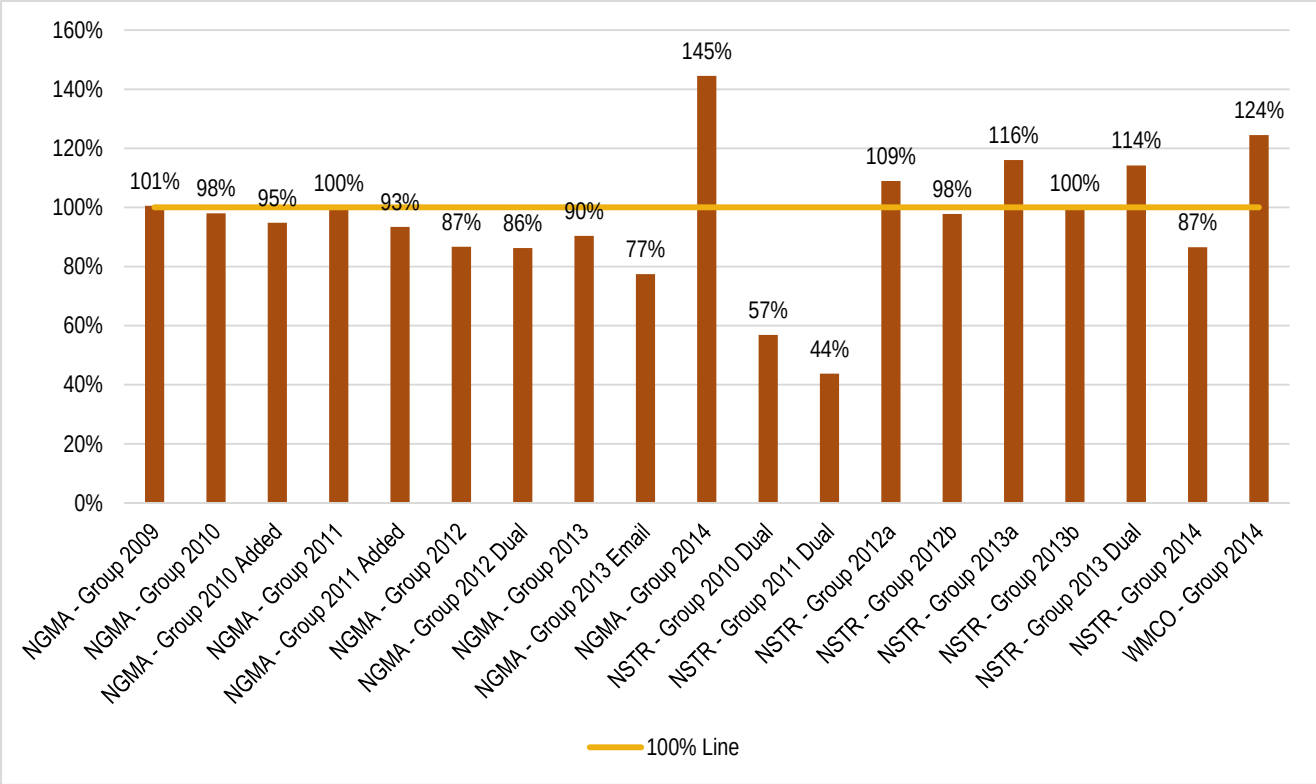
Figure 1. Electric Savings Estimate Ratios before Channeling Adjustment, by Cohort



Source: Evaluation team analysis
NGMA is National Grid Massachusetts, NSTR is NSTAR, and WMCO is WMECo.

[8] Opinion Dynamics Corporation, Navigant Consulting, and Evergreen Economics (2013). Massachusetts Cross-Cutting Behavioral Program Evaluation Integrated Report.

Figure 2. Electric Savings Estimate Ratios after Channeling Adjustment, by Cohort

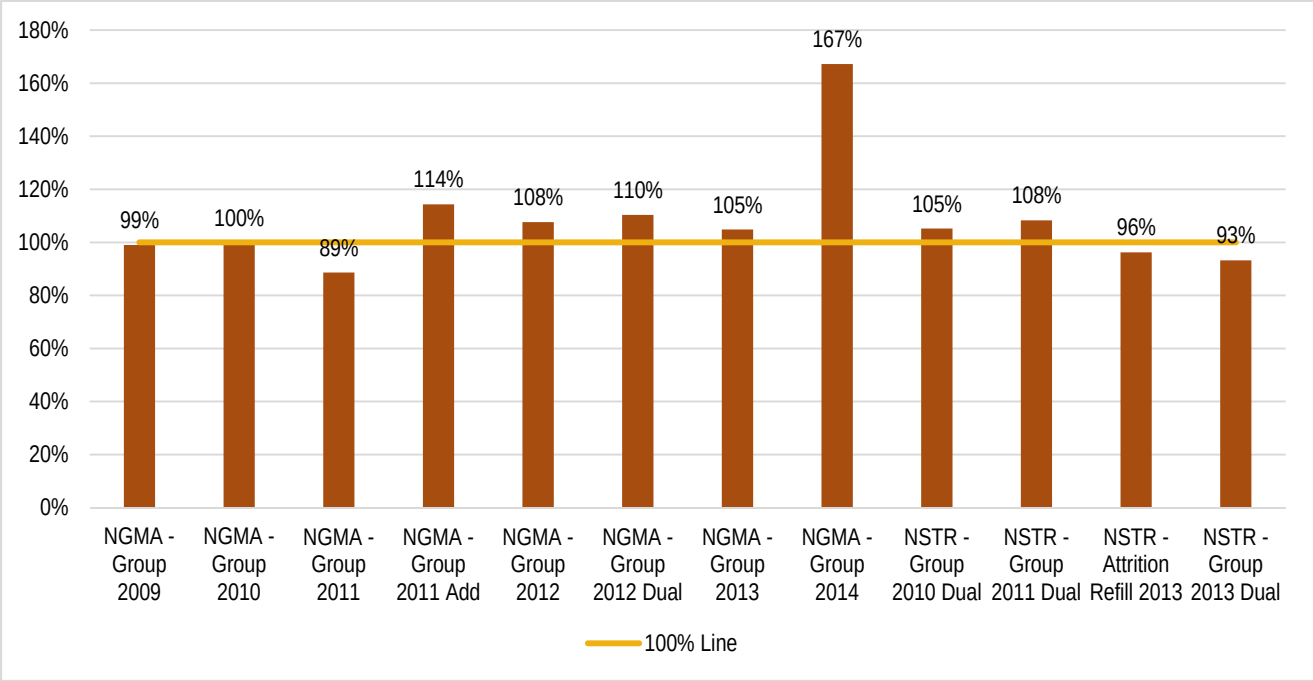


Source: Evaluation team analysis
NGMA is National Grid Massachusetts, NSTR is NSTAR, and WMCO is WMECo.

Figure 3 and Figure 4 show the savings estimate ratios for each of the gas cohorts. Figure 3 is before the savings adjustment from the channeling analysis, and Figure 4 is after. One cohort, National Grid Group 2014, has a relatively high savings estimate ratio of 167% before channeling. However, this savings estimate ratio is not outside the bounds found for first year gas cohorts in previous evaluations of this program.[9]

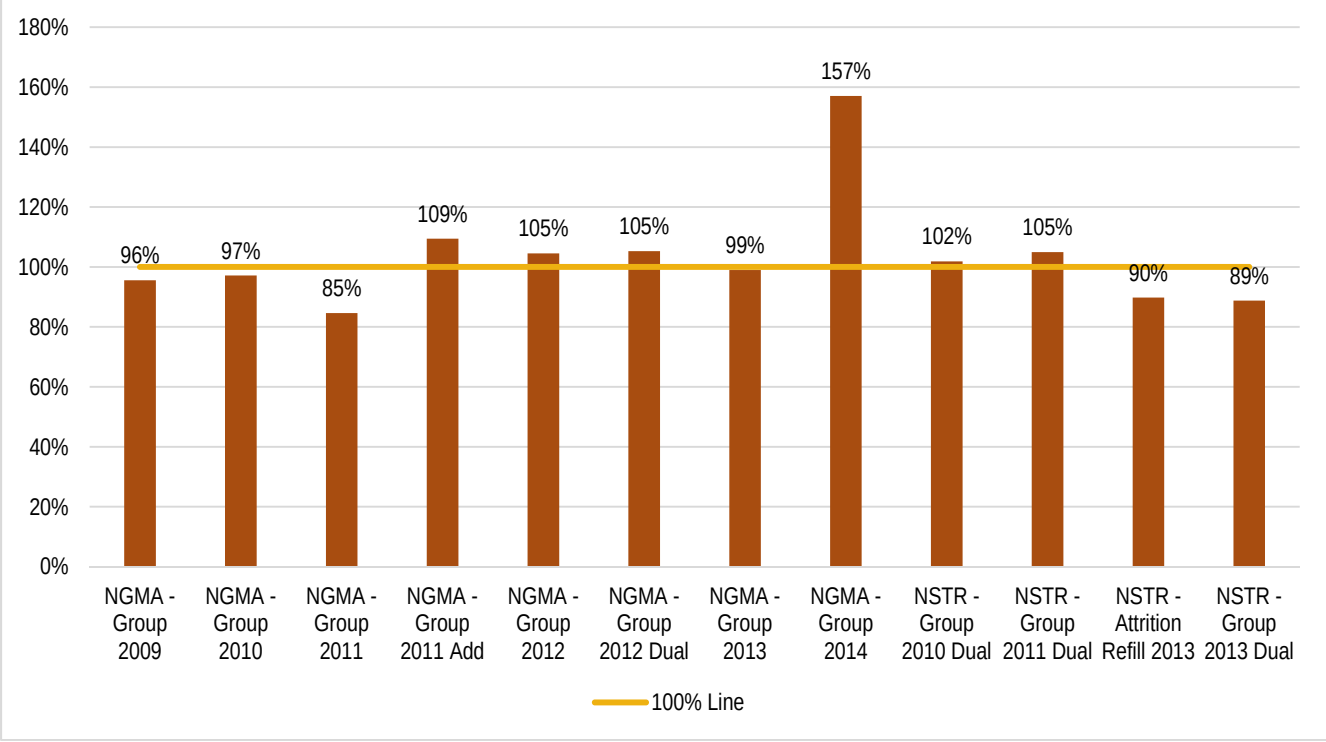
[9] Opinion Dynamics Corporation, Navigant Consulting, and Evergreen Economics (2013). Massachusetts Cross-Cutting Behavioral Program Evaluation Integrated Report.

Figure 3. Gas Savings Estimate Ratios before Channeling Adjustment, by Cohort



Source: Evaluation team analysis
NGMA is National Grid Massachusetts, NSTR is NSTAR, and WMCO is WMECo.

Figure 4. Gas Savings Estimate Ratios after Channeling Adjustment, by Cohort

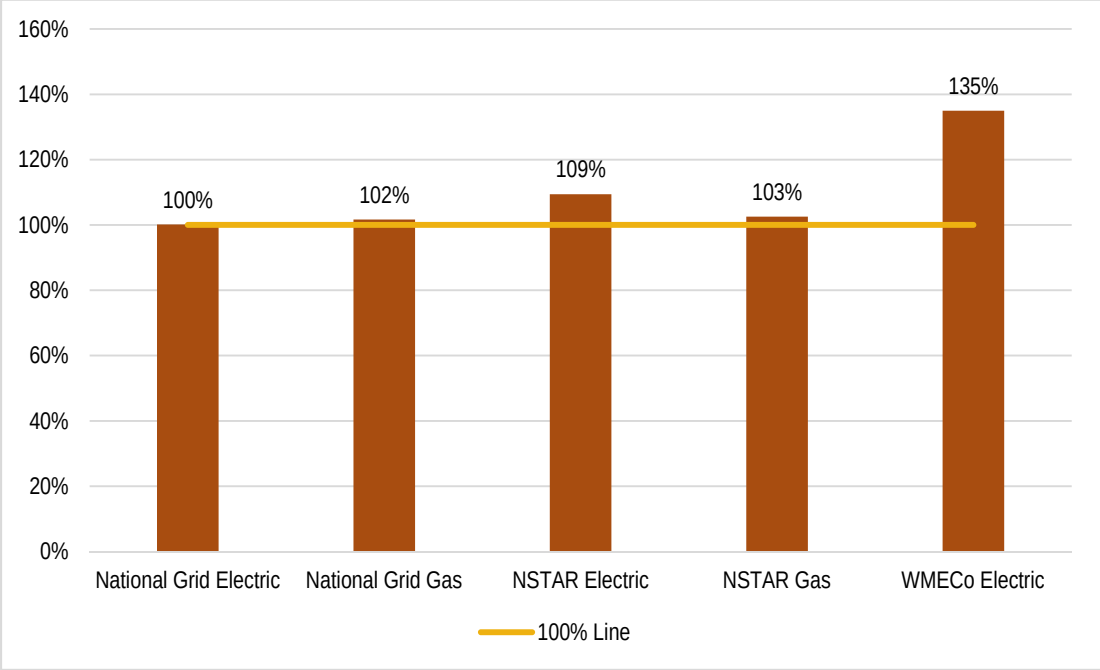


Source: Evaluation team analysis
NGMA is National Grid Massachusetts, NSTR is NSTAR, and WMCO is WMECo.

Aggregate Savings Estimate Ratio Results

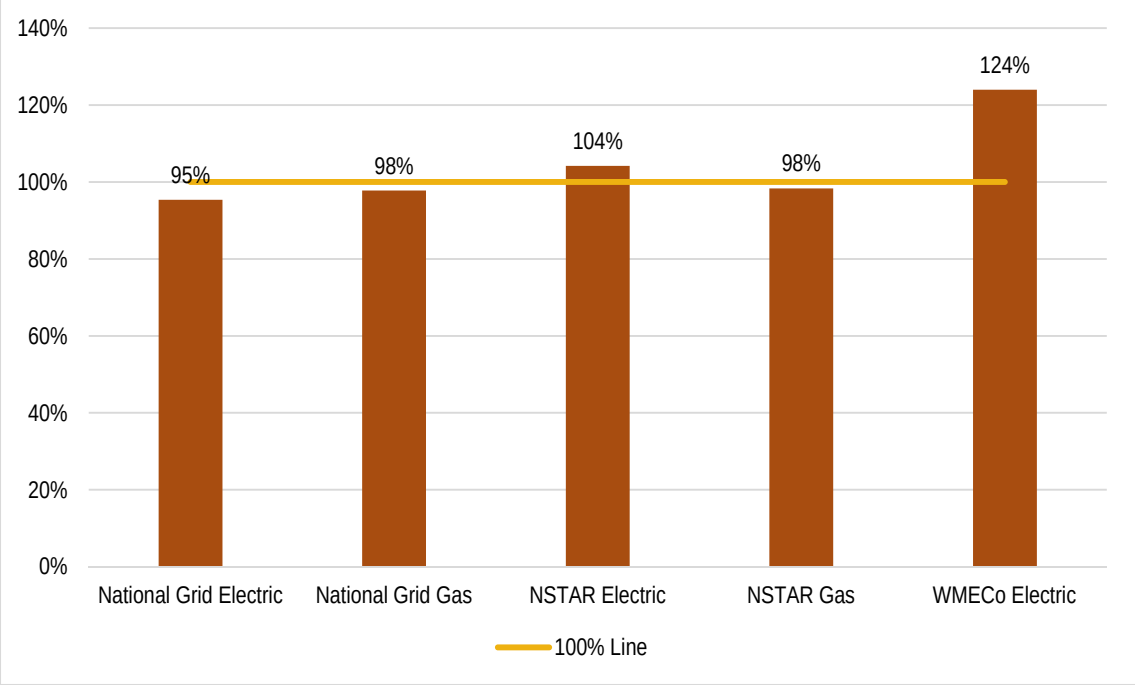
Figure 5 and Figure 6 show the aggregate savings estimate ratio for each PA and fuel-type. Figure 5 is before the savings adjustment from the channeling analysis, and Figure 6 is after. Note, because WMECo only has one cohort, the aggregate savings estimate ratio is equivalent to the cohort-specific savings estimate ratio.

Figure 5. Aggregate Savings Estimate Ratio before Channeling Adjustment, by PA and Fuel-Type



Source: Evaluation team analysis

Figure 6. Aggregate Savings Estimate Ratio after Channeling Adjustment, by PA and Fuel-Type



Source: Evaluation team analysis

Recommended Approach for Applying Savings Estimate Ratio for HER Program

The evaluation team recommends applying the savings estimate ratios presented in Table 3 in future years when third-party impact evaluations do not occur. These ratios should remain in effect until they are supplanted by a future impact evaluation.

Table 3. Savings Estimate Ratio

PA	Fuel Type	Savings Estimate Ratio
National Grid	Electric	95%
National Grid	Gas	98%
NSTAR	Electric	104%
NSTAR	Gas	98%
WMECO	Electric	104%

The evaluation team recommends that National Grid and NSTAR adopt the aggregate savings estimate ratio after channeling as presented in Figure 6 above. For WMECo, the evaluation team recommends applying the NSTAR Electric aggregate savings estimate ratio for use in future years. Given that the calculated savings estimate ratio for WMECo is based on only a partial year of program implementation, it may be inappropriate to apply in future years. In particular, the calculated savings estimate ratio may over-estimate savings due to first year ramping typical of HER programs.

Appendix A. Historic Channeled Savings Estimates

Table A- 1. Electric Channeled Savings Estimate

PA	Cohort Name	Program Year	Total Number of Participants	% per HH incremental savings from other programs
NSTAR	Wave 3	Y1	59,030	0.24%
NSTAR	Wave 4	Y1	17,514	0.23%
NGRID	2009	Y1	24,853	0.00%
NGRID	2009	Y2	23,309	0.00%
NGRID	2009	Y3	21,155	0.17%
NGRID	2010	Y1	68,194	0.06%
NGRID	2010	Y2	67,980	0.10%
NGRID	2010	Y3	62,305	0.01%
NGRID	2010 Add	Y1	23,427	0.05%
NGRID	2010 Add	Y2	21,224	0.22%
NGRID	2011	Y1	94,322	0.00%
NGRID	2011	Y2	82,417	0.06%
NGRID	2011 Add	Y1	55,055	0.08%
NGRID	2012 DF	Y1	12,074	0.06%
NGRID	2012	Y1	79,064	0.05%
			Weighted Average	0.075%

Source: Opinion Dynamics Corporation, Navigant Consulting, and Evergreen Economics (2013). Massachusetts Cross-Cutting Behavioral Program Evaluation Integrated Report.

Table A- 2. Gas Channeled Savings Estimate

PA	Cohort Name	Program Year	Total Number of Participants	% per HH incremental savings from other programs
NSTAR	Wave 1	Pilot Y1	23,247	0.00%
NSTAR	Wave 1	Y1	22,840	0.04%
NSTAR	Wave 1	Y2	21,599	0.01%
NSTAR	Wave 2	Y1	22,108	0.03%
NSTAR	Wave 2	Y2	20,415	0.03%
NGRID	2009	Y1	24,994	0.00%
NGRID	2009	Y2	23,685	0.00%
NGRID	2009	Y3	19,408	0.19%
NGRID	2010	Y1	74,759	0.00%
NGRID	2010	Y2	69,750	0.18%
NGRID	2011	Y1	87,691	0.00%
NGRID	2011	Y2	80,472	0.16%
NGRID	2011 Add	Y1	25,048	0.00%
NGRID	2011 DF	Y1	13,052	0.04%
NGRID	2012	Y1	83,938	0.005%
			Weighted Average	0.053%

Source: Opinion Dynamics Corporation, Navigant Consulting, and Evergreen Economics (2013). Massachusetts Cross-Cutting Behavioral Program Evaluation Integrated Report.