Rhode Island Behavioral Program and Pilots Impact and Process Evaluation

Contact Name:

Anne Dougherty
Founding Advisor
ILLUME Advising, LLC
608 561 2019
anne@illumeadvising.com

Table of Contents

1. EXECUTIVE SUMMARY

Illume Advising, LLC (ILLUME) with subcontractor Navigant Consulting (Navigant) (henceforth the ILLUME Team), is pleased to present National Grid Rhode Island with our impact results for the Rhode Island Home Energy Reports Program (the program) and the associate rewards and thermostat pilots.

1.1 Program and Pilot Descriptions

The first of its kind, the Rhode Island Home Energy Reports (HER) program is administered across National Grid's entire Rhode Island customer base treating electric only, gas only and dual fuel customers. There are multiple program components as well as two pilot efforts, including the following: (1) home energy reports (HERs) offered to multiple population segments, (2) an initiative to offer HERs to new home owners, (3) an online web portal, (4) a rewards pilot offered to HER participants, (5) a programmable communicating thermostat (PCT) pilot offered to HER participants, and (6) mass media promotional and public relations activities. This evaluation focuses on the first five listed program components. The evaluation effort covers the first year of the program and pilot efforts implemented from April 2013-May 2014.

1.2 Impact Findings

The HER program efforts consisted of six total gas and electric treatment cohorts (or groups of customers): high usage electric only customers, gas only customers, dual fuel customers (those who have both electric and gas meters served by National Grid), and three new movers groups for electric only, gas only, and dual fuel customers. We summarize the impact findings below.

1.2.1 HER Savings

Below we present the savings for the core HER program cohorts, including electric and gas savings results obtained from the electric only, gas only, and dual fuel cohorts. Note these values do not include the savings associated with new movers, but do include any savings that were generated through the rewards and thermostat pilots (See the next section, 1.2.2 for more detail).[1]

The HER program saved 20,066,543 kWh during the first 14 months of the program, amounting to .98% savings per household across the high usage electric only and dual fuel groups. Notably, the dual fuel electric metered households

[1] "New movers" are defined as those customers who have recently activated or reactivated an account with National Grid. This group was treated with a separate HER initiative described in more detail in Section 2.1, Introduction to the Program.

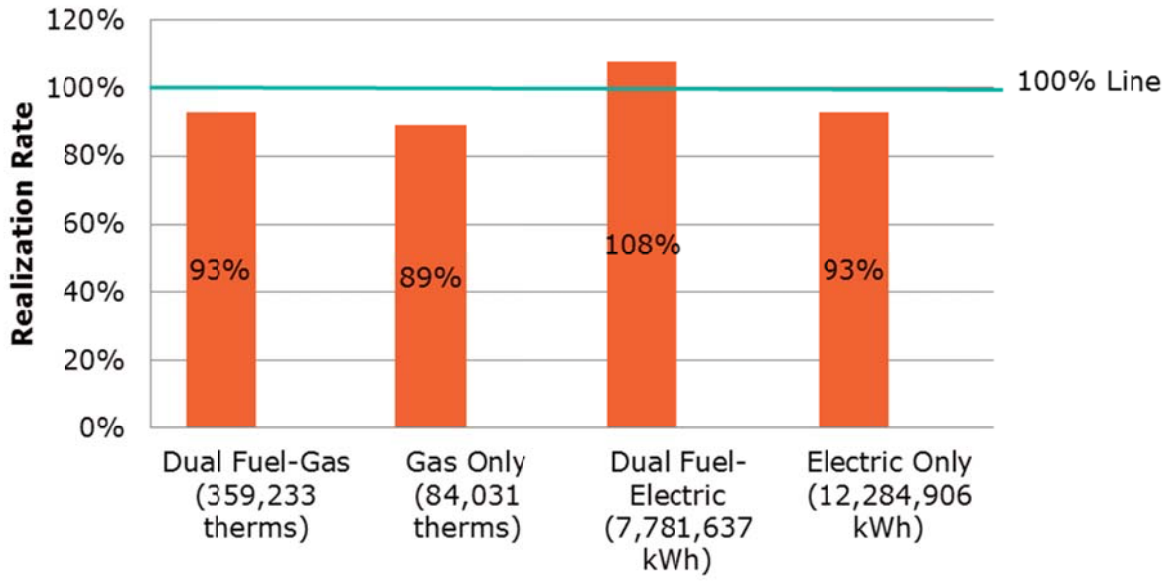
performed better than the electric only households on the realization rate, which measures the difference between the vendor estimate of savings and the evaluation estimate (Figure 1). At present, it is unclear what is driving those differences in savings performance.

The HER program saved 443,264 therms during the first 14 months of the program, amounting to .37% savings per household across the gas only and dual fuel groups. A number of factors contributed to the lower-than-expected savings for gas metered customers, including: (1) mistakes in the initial gas savings forecasts made by the program implementer, (2) fewer than expected dual fuel customers who contribute to the overall gas savings, and (3) a general tendency for gas HER programs to underperform relative to goal in the first year of the program due to a savings "ramp" effect.

Across fuels, there were very few channeled savings achieved through the HER program (savings due to participation in other program) with 695,735 kWh (3.35% of total HER savings) and 3,005 therms (.67% of total HER savings), generated through other programs. Notably, the majority of the cross-program savings were generated through the EnergyWise program.

The figure below demonstrates the Gas and Electric HER performance against the measured savings of the program implementer (in this case Opower). The goal of this ratio is to determine how greatly the savings measured by the third party evaluation vary from the savings measured by the implementation team.

Figure 1. Gas and Electric Savings Estimated Realization Ratios by Cohort*



*Note: Does not include results from the new movers initiative.

1.2.2 New Movers Initiative and Rewards and Thermostat Pilots

In addition to the core program, National Grid experimented with a new mover's HER initiatives as well as a rewards and thermostat pilot. Here, we present the results of these efforts.

The new movers initiative had small samples sizes at the time of the evaluation and thus statistically insignificant results. However, our best estimates of the initiative's effects is that the new movers initiative achieved .51% gas savings and electric usage increase of 0.83% per household total (52,193 therms and -716,522 kWh). The new movers initiative should be re-evaluated after a longer treatment period with the program implemented as originally designed.

The rewards pilot achieved.98% electric savings per household *in addition to the HERs* per household to total 520,741 kWh. Our best estimate of the gas rewards pilot is that it achieved .43% reduction incremental to the HERs and a savings of 8,345 therms, though the gas value is not statistically significant.[2] This total incremental savings is already accounted for in the total HER program savings values. However there is clear evidence that the rewards portion of the program is effective in generating savings above the HER treatment.

The best estimate of the *incremental* savings for the thermostat pilot is 2.31% in gas savings and .88% in electric savings per household, amounting to a total of 3,902 therms and 11,592 kWh overall. These values improve in the heating and cooling season. Notably, the electric and gas savings estimates improve in the cooling and heating season respectively[3], with incremental electric savings at 2.42% in the cooling season and incremental gas savings at 2.35% in the heating season. To date, these values are statistically insignificant, but do provide an indicator of the incremental effect and value of the thermostat effort.

1.3 Process Findings and Recommendations

The RI HER program and pilots are, by design, aimed at targeting all National Grid customers in the state. In most HER programs nationwide, the program design focuses on high-energy users and does not treat new mover populations. As a result, the program and pilot's first year was largely exploratory; the program aimed to identify ways to successfully reach all of National Grid's customers as cost-effectively as possible. In the process, there have been a number of key process and design findings.

---

[2] The rewards pilot electricity impact was statistically significant while the gas impact was not. That said, these were the best, unbiased estimators of impact even though the 90% confidence interval around the gas estimate does include 0.

[3] The cooling season is defined as June-August and the heating season as September-April.

A core mission for this program and its pilots is to enhance customer engagement and satisfaction across the state. While the program aims to generate savings associated with its efforts, National Grid also sought to increase customer engagement and satisfaction by providing enhanced service and support through the HERs, rewards, and thermostats. Overall, the program team has reported it has been successful in this respect.

The gas savings for the program underperformed due to a number of planning-related challenges. First, savings were overestimated due to errors in Opower's forecasting models and difficulties in successfully identifying dual fuel customers. Both of these issues have since been resolved in program plans. Further, the savings goals did not fully account for a traditional "ramping" year for gas programs. Often, gas programs do not achieve their expected savings in the first year. Since gas programs are heavily based on winter savings, they often need a year or two to ramp up to full savings. Finally, based on recommendations from the Massachusetts evaluation, National Grid modified the computation of gas savings to include months that have negative savings in the annual savings calculation.

- **Recommendation:** The program team should consider having implementer-derived savings forecasts reviewed by a third party in the future to avoid similar planning errors.

- **Recommendation:** The gas savings first year "ramp" should be factored into program decisions on whether or not to continue the program.

New movers definitions were too broad to inform a targeted outreach strategy. Due to customer data tracking limitations, new movers were identified broadly, including those who were new customers to National Grid as well as those who had delinquent and then reactivated accounts –two very different populations. National Grid now has an indicator in their customer database to distinguish true new movers from reactivations.

- **Recommendation:** Since this is a distinctive population not typically targeted by programs, we recommend examining this program again after it has been implemented as designed. We also recommend considering a strategy of outreach for delinquent and re-activated customers, who may benefit from the educational elements of the program.

Opt-in HER component did not generate enough interest to comprise an evaluable cohort. The opt-in component targeted lower electricity users, a group not typically included in opt-out programs. However, marketing and outreach efforts did not spur enough sign-ups to evaluate the program. National Grid concluded there was not enough interest to justify the cost and has discontinued the initiative within the HER program.

Program design and implementation details were not carefully documented. Fully interpreting and contextualizing impact and process analysis findings, particularly for a program with this complexity, requires understanding program design details.

- **Recommendation:** The program vendor should develop a single decision-making document and database to clearly delineate the program design and avoid loss of information over time due to staffing changes.

Randomized encouragement design (RED) design for opt-in efforts did not have sufficient participation levels, and statistical power, to be evaluable. The impacts of the rewards pilot were then calculated using a matching methodology, yielding similar results that were statistically significant for electric savings. The RED design may not be the best design for evaluating programs with small impacts and low participation levels.

- **Recommendation**: We recommend discontinuing the use of the RED design for the pilot rewards initiatives and using a matched comparison group for evaluation instead. Our results indicate the method is appropriate and accurate relative to the RED. Further, the matching method can support a territory-wide roll out of the rewards initiative if desired.

2. OVERVIEW OF THE EVALUATION

Illume Advising, LLC (ILLUME) and Navigant Consulting, Inc. (Navigant) are pleased to present National Grid with our evaluation of the Rhode Island Home Energy Reports Program Evaluation and the rewards and thermostat pilot efforts. The first of its kind, this program aims to treat all of Rhode Island's customers with an Opower-implemented behavioral program.

2.1 Introduction to the Program

The Rhode Island Home Energy Reports program has multiple components, including the following: (1) home energy reports (HERs) offered to multiple population segments, (2) an initiative to offer HERs to new home owners, (3) an online web portal, (4) a rewards pilot offered to HER participants, (5) a programmable communicating thermostat (PCT) pilot offered to HER participants, and (6) mass media promotional and public relations activities. Our evaluation focused on the first five listed program components. This evaluation effort covers the first year of the program's efforts implemented April 2013-May 2014.

National Grid Rhode Island identified five key goals of the program that was launched with contracting implementer Opower:

1) To contribute to Rhode Island's aggressive energy savings goals while staying within budget and maintaining cost-effectiveness test;

2) To implement broad outreach to everyone in the state, not just high energy users;

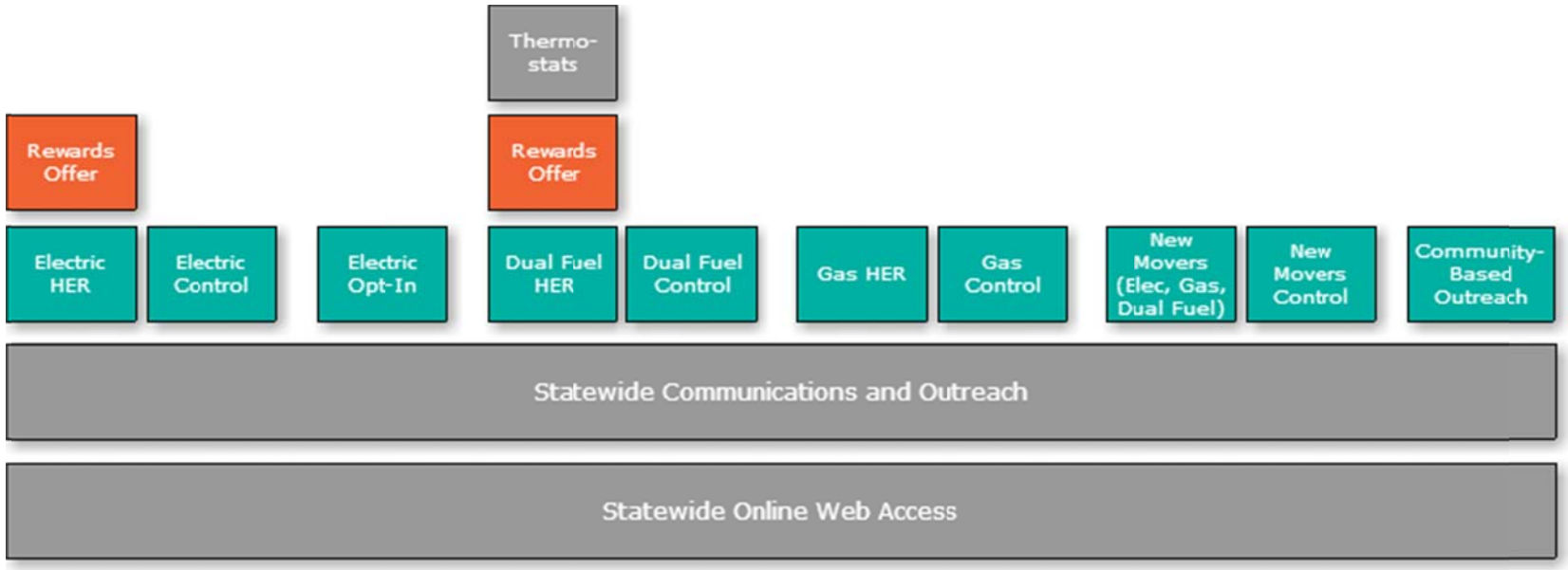
3) To tie in with other National Grid programs and to drive participation in energy efficiency programs;

4) To provide a positive customer experience and improve customer satisfaction with National Grid; and

5) To showcase National Grid's commitment to innovative by trying new approaches and techniques in behavioral programs.

This program was offered or administered across % National Grid's 425,000 Rhode Island residential customers, including established electric only customers, gas only customers, dual fuel customers, and new movers to the territory across all three fuel sources. The diagram below captures the overall program design. On page 10, we provide a model representation of the program design. Each findings section further details the program design, outlines the number of evaluated participants, and indicates treatment periods specific to each group.

Figure 2. Rhode Island Statewide Program Design Model



2.2 Evaluation Goals

The primary goal of the evaluation is to conduct an impact evaluation of the HER program with a secondary focus on process insights focused on improving the evaluability of the program through design modifications. The HER program is designed as a set of randomized control trials, where different target customers in Rhode Island were randomly assigned to treatment and control groups. Treatment customers receive a HER, while control customers do not receive a HER or any other program components. The core HER program consists of electric only, gas only, and dual fuel HER groups.

In select cases, HER recipients were further subdivided into two pilot efforts: (1) a rewards effort and a thermostat effort. The rewards treatment group and a no-rewards control group, where the rewards treatment customers were encouraged to opt-in to the online rewards portal and receive rewards for conserving energy (a randomized encouragement design). The thermostat customers were recruited from the dual fuel group and did not have a retained control group.

In addition to these efforts, the program has a new movers initiatives, which treated electric only, gas only, and dual fuel customers who recently activated service. These customers were subdivided into a three fuel-specific groups and not combined with the core HER program due to a lack of pre-period usage data.

Drawing on this core design, we developed the evaluation to measure the following:

- **Overall Net Savings:** This is the savings value measured for the behavioral program using experimental or quasi-experimental designs. This includes:
    - ***Total program savings*** gained for each treatment group
    - ***Incremental rewards savings*** gained through the rewards component using an interaction term in the overall impact models and utilizing a matching method.
    - ***Incremental savings associated with thermostat efforts*** where the program introduced programmable thermostats.
- **Cross-Program Effects or "Channeling" Effects:** This analysis examines the impact of the behavioral program on driving program participation. We estimate the impact of the behavior program on participation in other energy efficiency programs.
- **Overall Adjusted Net Savings:** This value measures the final net savings associated with the behavioral program. This is the net savings value minus any savings gained by the treatment group through other programs. The adjustment removes the energy savings gained from participation in other energy efficiency programs from the estimate of the behavior program savings.

- **Realization Rate:** This realization rate value is the expected savings for the program based on the difference between the third-party evaluation and Opower's own evaluation of the program.

We describe our evaluation approach in greater detail in the following methods section.

3. EVALUATION METHODOLOGY

3.1 Overview of our Approach

In this section we describe the evaluation activities conducted for the Rhode Island Behavioral Program and Pilots Impact and Process evaluation.

The program utilizes a randomized control trial (RCT) design for each of the HER cohorts. For the rewards groups, the program utilized a random encouragement design (RED). For the HER savings, we utilized the RCT design. In estimating the rewards savings our team initially exploited the random encouragement design of the program, but due to the low encouragement effect and low savings, confidence in the estimates were low. Our team then used a matching method to estimate program savings associated with rewards to obtain an alternative estimate. For the thermostat analysis, we utilize the same matched comparison group approach as the rewards component to estimate the impacts for this element of the program, incremental to the HERs.

3.2 Interviews with the Program Team

The first step in our evaluation was to conduct in-depth interviews with the National Grid and Opower staffs. In total, the ILLUME team conducted four formal interviews and numerous on-going conversations. The focus of the interviews were as follows:

1. Determine program goals and objectives, both formal and informal;
2. Document details on the program design to inform data requests and the evaluation methodology;
3. Determine and document the interim reporting approach and associated savings estimates to develop realization rates; and
4. Identify areas for improvement in the program implementation.

3.3 Impact Approach

Error! Reference source not found. summarizes the approach used to calculate net savings for each program component. Detailed descriptions of the methods follow the table.

Table 1. Summary of Net Savings Analysis Methods by Program Component

Treatment Type	Program Group	Evaluation Design	Net Savings Analysis Approach	Description
Opt-out	Dual Fuel, HER	RCT	Post Program Regression	We conducted fuel-specific impact analyses for each

Treatment Type	Program Group	Evaluation Design	Net Savings Analysis Approach	Description
Opt-out	High Usage Electric Only, HER	RCT		of the treatment groups in the program. Each of these target groups had adequate bill histories to conduct a Post Program Regression (PPR) analysis.
Opt-out	Gas Only, HER	RCT		
Opt-out	New Movers	RCT		
Opt-in	Low Usage Electric Only	NA	NA	We did not evaluate this program. NGRID dropped the program in Jan 2014.
Rewards Opt-in with Encouragement	High Usage Electric Only, Dual Fuel	RED	Post Program Regression	We used this approach to estimate incremental savings associated with rewards.
Rewards Opt-in with Encouragement	High Usage Electric Only, Dual Fuel	Matched Comparison	Matching with Post Program Regression	To add to the analysis done using the RED, the Illume team also used a matched comparison group approach as outlined below.
Thermostat Opt-in	Programmable Thermostats	Matched Comparison	Matching with Post Program Regression	To measure the impacts of the T-stats, we used a matched comparison group approach as outlined below.
Channeling Analysis	All HER Groups	Difference in Differences	--	We used a difference in differences comparison between treatment and control group participation and savings through other programs.

3.3.1 Opt-out Home Energy Report (HER) Impact Assessment

A key feature of the RCT design of the HER program is that the analysis estimates net savings, not gross savings. While some customers receiving reports may have taken energy conserving actions or purchased high efficiency equipment in the absence of the program, the random selection of program participants (as opposed to voluntary participation) assures that on average their behavior in this regard would have been no different in the absence of the program than the actual average behavior of the control group. Thus, there is no free ridership, and no "net-to-gross" adjustment is necessary.

The post-program regression (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 usage 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;[4]
$ADClag_{kt}$	= Customer *k*'s energy use in the same calendar month of the pre-program year as the calendar month of month *t*;
$Participant_k$	= 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);
	= The cluster-robust error term for customer *k* during billing cycle *t*. Cluster-robust errors account for heteroscedasticity and autocorrelation at the customer level.[5]

In this model, 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.

3.3.2 HER New Movers Impact Analysis Method

For new movers the pre-period typically does not have the same calendar months as the post period, and so we estimate a form of the PPR model in which the pre-program usage variable, $ADClag_k$, does not vary with the post-program month t, instead pertaining to the same fixed pre-period for all customers in the model. Consequently the model is cohort-based, run separately on each cohort, with cohorts defined by the month of the first received HER. Electric-only, dual fuel electric, and dual fuel gas customers each had 6 new mover cohorts, one for each month of enrollment, November 2013 through April 2014.

[4] In other words, 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.

[5] 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.

These cohorts had, on average, only one month of pre-program energy use data. Gas-only customers had two new mover cohorts, November 2013 and February 2014. These gas-only cohorts included 3 months of pre-program energy use data for the November 2013 cohort and 6 months of pre-program energy use data for the February 2014 cohort.[6]

Average energy savings for new movers are calculated as a cohort-weighted average, where weights are based on total observations (participant-days per cohort). The standard error on the average is calculated as a similarly weighted average based on each cohort's estimated value for the variance of β_3.

An obvious alternative is to combine all new movers for a particular customer type (such as dual fuel electric) in a single equation and form $ADClag_k$ as the average pre-program energy use for each customer. Keeping in mind that the pre-program months in the data set differ across customer cohorts, this approach essentially assumes that the average effect of pre-program energy use on post-program energy use is insensitive to its composition. In other words, it assumes that the effect of pre-program summer months on energy consumption in month *t* of the post program period is the same as the effect of pre-program winter months. This is unlikely, and in any case it is an unnecessary assumption to impose, because even if true it remains appropriate to use the cohort-based approach described above.

3.3.3 Rewards Program Impact Analysis Method

The rewards program employed a random encouragement design (RED) in which a subset of treatment customers were randomly assigned to receive encouragement to access the rewards program web portal. The rewards program component was administered to electric only customers and dual fuel customers.

The theoretical advantage of an RED design for an opt-in program is that its structure provides the opportunity to easily address selection bias. The encouragement is effectively an instrumental variable, correlated with the treatment variable (a matter of empirical observation), but by design not correlated with energy use.

Estimation of incremental energy savings using an RED design proceeds in two steps. First, the effect of receiving encouragement is estimated using regression analysis. We use the PPR model presented previously (Model 1), where in this case "participant" is defined as those customers receiving the encouragement. The estimated encouragement effect (in Model 1) is then used to develop a Wald statistic, whereby the estimated encouragement effect is divided by the difference between the proportion of encouraged customers who enter the program and the proportion of non-encouraged customers who enter the program. Under two reasonable statistical assumptions, the Wald statistic generates the "Local Average Treatment Effect" (LATE); that is, the average energy savings by those in the encouraged group who enroll in the program but would not have enrolled in the absence of

[6] The additional pre-program energy use data for the gas-only cohorts reflects the fact that this group had significantly more pre usage data than the other new mover cohorts.

the encouragement (often called “compliers”). In the case at hand, where the probability of enrollment in the program in the absence of encouragement is virtually zero, the statistic is more generally interpreted as the average treatment effect on the treated (ATT); that is, the average treatment effect on any customer who accesses the web portal. Keeping in mind that the RED is applied only to a subset of customers already receiving home energy reports, it is important to emphasize that in this case the ATT is properly understood to be the average *incremental* energy savings due to web access by the subset, because it pertains to energy savings above and beyond those generated by the home energy reports. It also should be kept in mind that the estimated average saving apply to the entire period after encouragement, including the period before web access.

Although the RED design and associated analysis has the advantage of generating an unbiased estimate of energy savings, in practice it can generate poor results if (a) the treatment effect is relatively small relative to the variation in the dependent variable, and (b) the proportion of encouraged customers who actually enroll in the program is very low. Both are typically the case with energy behavioral programs. Suppose, for instance, that the true average savings from accessing the web portal is 4% - a value that is higher than usually found for behavioral programs. If only 5% of customers enroll in the program, then working backwards using the Wald statistic, and assuming that all customers accessing the web portal do so as soon as they are encouraged to do so (an obviously too generous assumption that we make to keep things simple) it is clear that the treatment effect to be estimated in regression analysis is only (5%) x (4%)=0.2% of average daily energy use - a very small value that, to properly identify with any reasonable confidence, would typically require an extremely large sample.

Given these issues with the estimation of the effect of rewards encouragement via the RED, we also estimate the incremental effect using the same matching methods described in section 3.3.4.3. In this case, a match is selected for each customer who enrolled in the rewards program. Matching was done for the 2466 dual fuel participants and the 4880 electric only participants who enrolled in the rewards program and had at least 12 observations in the 16-month period covering both the matching period and the test period. The pool of non-participant households available for matching consisted of 116,188 dual fuel customers and 100,837 electric only customers who received HERs and also had at least 12 months of account history. Significantly, the measured effect of the rewards program is the incremental effect above and beyond the effect of the HER.

In total, the matches consisted of 4602 electric only participants with 4447 unique comparison customers, 2322 dual fuel electric participants with 2288 unique comparison customers, and 2321 dual fuel gas participants with 2263 unique comparison customers.

Error! Reference source not found.5 and **Error! Reference source not found.**6 present the average energy use of dual fuel rewards participants and their matches over the period *t*-16 to *t*-1, for electric and gas, respectively. Figure 7 presents the same for electric only participants and matches. The figures illustrate that on average the energy use by matches is very similar to that of program participants for all three groups. Any differences in pre-program usage that remain after matching are accounted for by the regression analysis used to estimate program savings.

Figure 3. Electricity usage for dual fuel participants and their matches in the 16 months before program enrollment, months -16 to -5 are the matching period while months -4 to -1 are the IW test period.

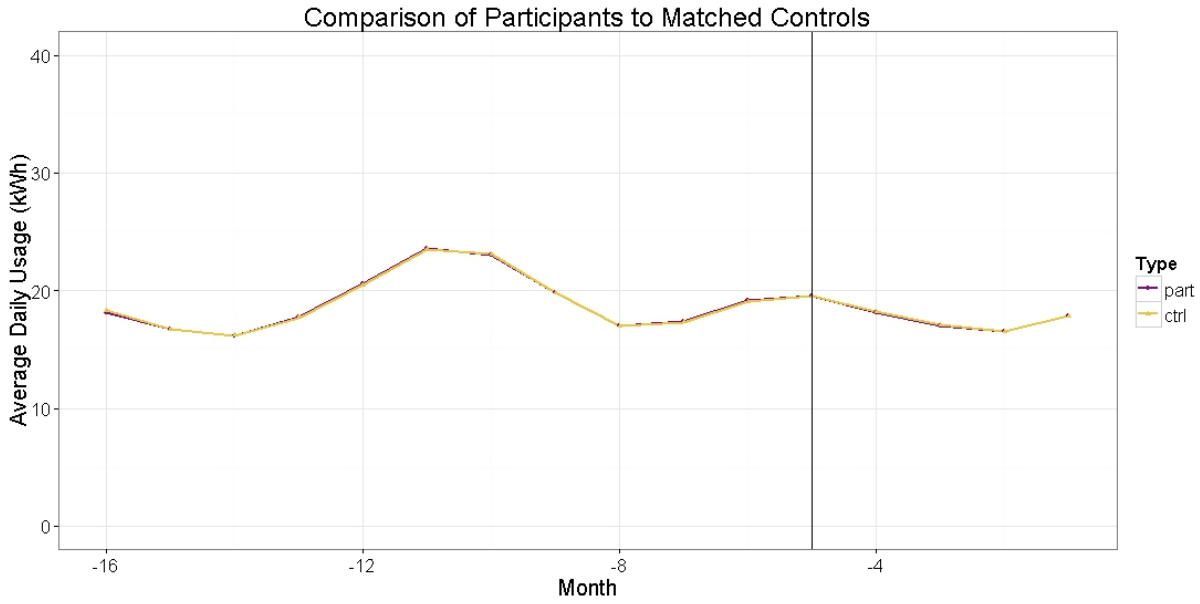


Figure 4. Gas usage for dual fuel participants and their matches in the 16 months before program enrollment, months -16 to -5 are the matching period while months -4 to -1 are the IW test period.

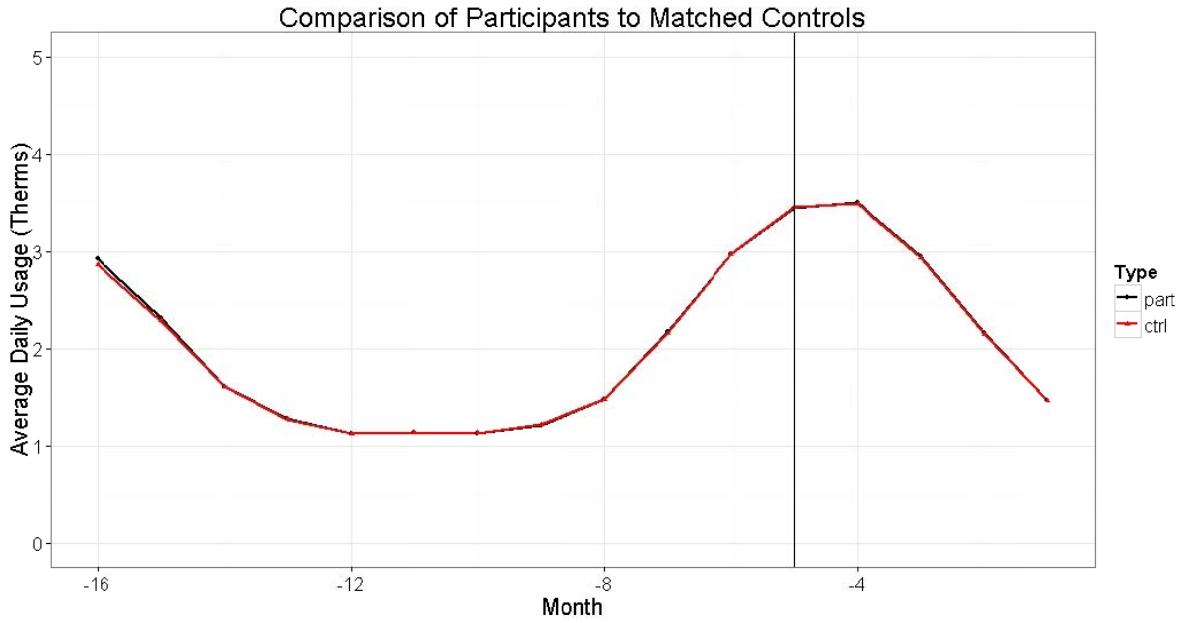
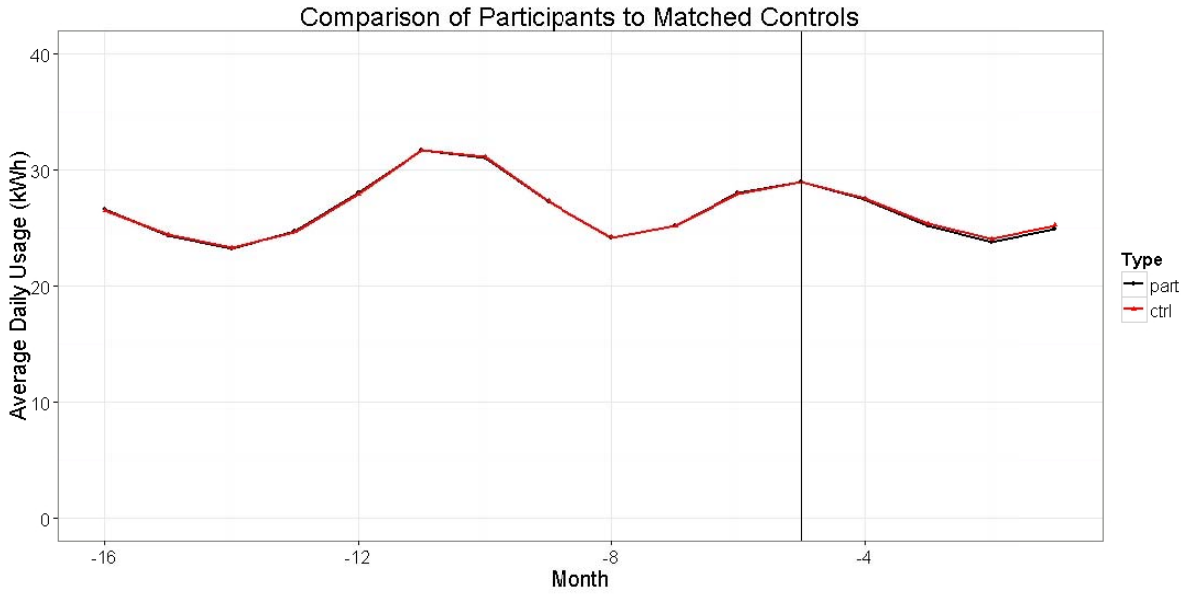


Figure 5. Electricity usage for electric only participants and their matches in the 16 months before program enrollment, months -16 to -5 are the matching period while months -4 to -1 are the IW test period.



3.3.4 Programmable Thermostat Pilot Impact Evaluation Methods

The Illume team used two versions of a matching method to estimate *incremental* savings from the thermostat program. Both versions use the same set of program enrollees and their 1:1 non-program matches, but are distinguished by the statistical analysis used to estimate program impacts. The first of these is regression with pre-program matching

(RPPM) described in Ho, Imai, King, and Stuart (2007).[7] The other is matching with bias correction (MBC) discussed by Abadie and Imbens (2011).[8] In reporting total savings we use results from the RPPM approach.

Matching was done for the 112 participants with at least 12 observations in the 16-month period covering both the matching period and the test period (which is described below). The pool of non-participant households available for matching consisted of 49,799 customers receiving HERs who also had at least 12 months of account history and were not selected to receive rewards encouragement. Since the matches and the participants both received HERs, this analysis estimates the incremental savings from the thermostat program given the receipt of an HER. Thermostats were installed in April and May of 2013, which was the same month or one month later than the customers received their first HER.

For each program participant with monthly billing data extending back at least 16 months before program enrollment, energy consumption in each month in the period spanning 5 to 16 months before program enrollment (a twelve-month period) was compared to that of all customers in the pool of potential matches with billing data over the same 12 months who received their first HER in the same month as the participant. For the sake of expositional clarity below, we denote by t_k=0 the month t in which customer k enrolled in the program, with t_k-1 denoting the month before enrollment, t_k+1 denoting the month after enrollment, and so on. In this notation, the matching period is [t_k-16, t_k-5].

The basis of the matched comparison is the difference in monthly energy use between a participant and a potential match, D_{PM} (Difference between Participant and potential Match). The quality of a match is denoted by the Euclidean distance to the participant over the 12 values of monthly D_{PM} used for matching; that is, denoting by SSD the sum of squared D_{PM} over the matching period, it is denoted by $SSD^{1/2}$. The non-participant customer with the shortest Euclidean distance to a participant was chosen as the matched comparison for the participant.[9] Program participants were dual fuel customers, and so we allowed each customer to have different matches for electric and gas. Matching was done with replacement, and so, after excluding observations based on screening criteria explained in the next section, there were 112 electric participants with 110 unique comparison customers and 110 gas participants with 110 unique comparison customers.

---

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

[8] Abadie, Alberto, and Guido Imbens. 2011. Bias-corrected matching estimators for average treatment effects. *Journal of Business and Economic Statistics* 29(1): 1-11.

[9] This is a form of nearest-neighbor matching. Typically, the distance is normalized using sample standard deviations or the full covariance matrix of the variables used for matching (Mahalanobis distance), because matching variables are measured in different units. In this analysis all matching variables are measured in the same units (kWh or therms, depending on the analysis), and standardized distances are nearly the same across months, and so we use the "raw energy distance", i.e., the actual sum of squared differences in energy use over the 12-month matching period.

It is not possible to statistically test for selection bias, but Imbens and Wooldridge (2009) present a quasi-test that is suggestive (hereafter called the "IW test").[10] In the current context, the logic of the test is that in the absence of selection bias there should be no difference between participants and matches in average energy use outside of the matching period and outside of the program period. A simple implementation of the test is to determine whether, given matching based on months t_k-5 to t_k-16, average D_{PM} in the four months before program enrollment, energy use in months t_k-1 to t_k-4, is practically or statistically different than zero. If not, the analyst gains some confidence that selection bias is not a critical issue in the analysis. **Error! Reference source not found.**3 and **Error! Reference source not found.**4 present average energy use of participants and their matches over the period t-16 to t-1, for electric and gas, respectively. The figures illustrate that on average the energy use by matches is very similar to that of program participants. Mean differences in energy use between participants and their matches are not statistically different in either the 12-month matching period or in the 4-month test window. Moreover, during the test window there is no clear pattern in the mean differences.

Figure 6. Electricity usage for participants and their matches in the 16 months before program enrollment, months -16 to -5 are the matching period while months -4 to -1 are the IW test period.

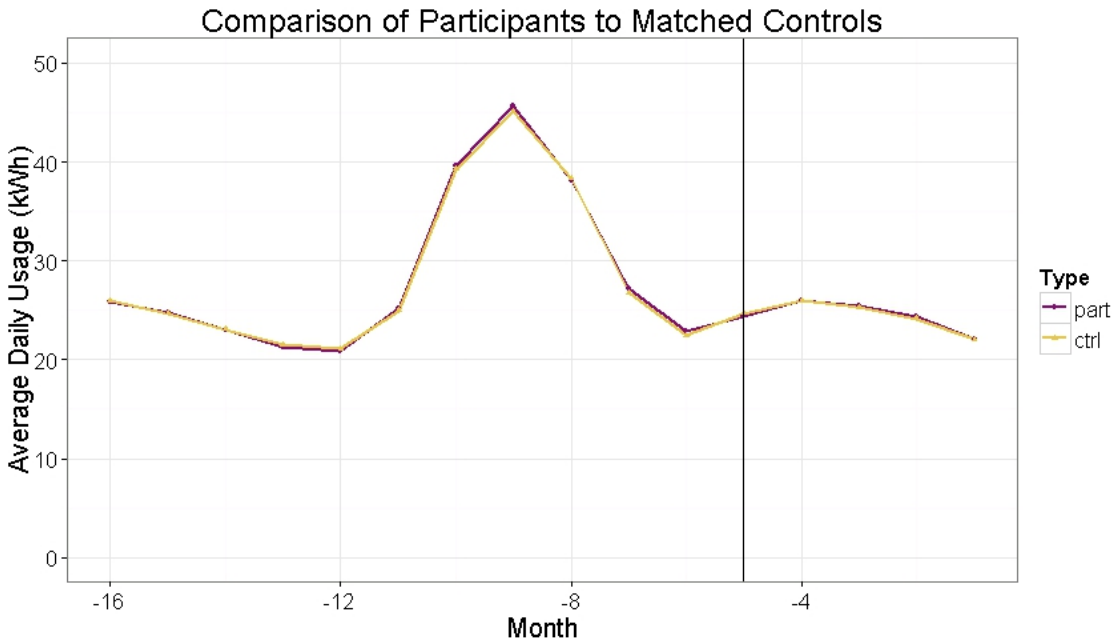
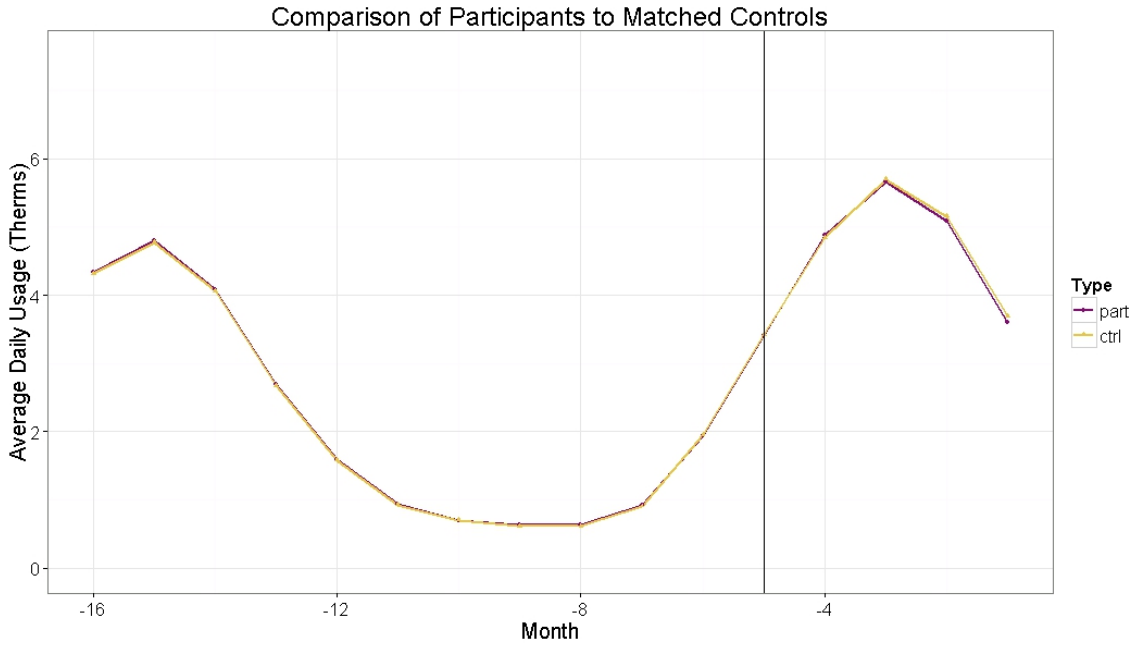


Figure 7. Gas usage for participants and their matches in the 16 months before program enrollment, months -16 to -5 are the matching period while months -4 to -1 are the IW test period.

[10] Imbens, Guido W., and Jeffrey M. Wooldridge. 2009. "Recent Developments in the Econometrics of Program Evaluation." *Journal of Economic Literature*, 47(1): 5-86.



Details of the statistical analysis performed on participants and matches in the RPPM and MBC approaches are presented in Appendix A. The two approaches draw on the same set of matches for the comparison group, but differ in their use of a structural model to estimate program savings. The MBC approach is less parametric, using regression analysis to correct for the differences in post-program energy use between participants and their matches that can be attributed to differences in their pre-program energy use. The RPPM method, by contrast, treats matching as a "pre-processing" stage of regression modeling in which the sample distribution of past energy use for participants is matched to that of a group non-participants.

3.4 Participation Lift and Savings Adjustment Findings

In order to understand the full energy efficiency impact of behavioral programs, it is important to examine the programs' effects on other National Grid programs across the treatment and control groups. The goal of this analysis is to examine the extent to which the program is driving participation, and savings, through other program initiatives through the statewide behavioral program.

Notably, behavioral programs have been demonstrated to prompt a wide range of actions, including measure installations, both within and outside of other National Grid programs, such as Massachusetts. This task specifically measures increased participation rates, and subsequent savings impacts, among the treatment group compared to the control group in existing Rhode Island programs.

In this program database review, our team gathered Rhode Island program databases to examine program participation among the treatment and control groups over the pre- and post- treatment period. Using the program databases and deemed savings values associated with program measures, we examined two key questions:

1. **Participation Lift:** Does the statewide program treatment have an incremental effect on participation in other energy efficiency programs (treatment above control)?

2. **Savings Lift and Adjustment:** What portion of savings from the Home Energy Reports program has been obtained through Rhode Island's other energy efficiency efforts?

To answer these questions, we measured the incremental difference between the treatment and control groups in their participation rates and subsequent savings using program databases.

Participant Lift: Using participation flags, we calculated a participation rate for each program year, based on the number of accounts that initiated participation in other tracked energy efficiency programs after the first report date. The analysis includes efficiency programs that track participation by individual or household and does not include upstream programs, such as lighting, that do not capture information on participants. [1112] The program participation rate captures how many customers engaged in a utility program after exposure to the behavioral program. The difference in treatment and control participation across the pre- and post-treatment period is participation lift.

Savings Lift and Adjustment: Drawing on our core database of participation information, we then estimated the savings associated with participation in other Rhode Island energy efficiency programs. To do this, we used the same pre- and post-treatment analysis between the treatment and control groups, focused instead on estimating savings rather than participation rates. We did this through a number of steps, as follows:

1. Collected deemed net savings from all measures installed in the pre- and post-period;

2. Adjusted annual deemed net savings for each measure installation by the number of days per year in which a measure was installed for both the treatment and control group in the pre- and post-period;

3. Determined the average annual net savings from other programs as the average of the sum of savings for both the treatment and control group; and

11 The energy efficiency programs included in the analysis were: Income eligible single family, Energy Wise Single Family, ENERGY STAR Products ENERGY STAR HVAC, and ENERGY STAR Lighting.

12 Note that ENERGY WISE Multifamily was not included as most measures are installed in common areas. In addition, 2014 data for ENERGY WISE gas measures were not available at the time of the analysis. As mentioned, upstream programs also could not be counted. Consequently, the channeling analysis is missing a portion of the energy efficiency program activity that occurred during the pre- and post-period.

4. Used a difference-of-differences pre-post treatment-control estimate to identify the incremental savings gained by the treatment group above the control group in the treatment period above the pre-period.

Once this estimate was determined, we then reduced the overall savings estimated in the billing analysis by the final estimated incremental savings of the treatment group to avoid double counting. We present these results in each of the HER-specific impact savings sections.

4. ELECTRIC HER IMPACT FINDINGS

4.1 Electric Home Energy Report Overview

The Home Energy Reports (HERs) included six total cohorts (or program groups), including a high-usage electric only cohort, a gas only cohort, a dual fuel cohort, and three new movers cohorts. The first three treatment groups follow a standard HER program model as implemented in other states. National Grid also implemented several innovations in the HER program including reaching out to populations not typically included in behavior programs, such homeowners who have recently moved, hence the "new movers" group.

We briefly describe each group below:

1. **Electric Only:** The electric only cohort is comprised of high electricity users who received reports on their electricity consumption only. This group started receiving reports in April 2013.
2. **Duel Fuel:** The dual fuel cohort received information on both their gas and electric use and began receiving reports in April 2013. Reports received in September through April focused more on gas use while reports received in summer months targeted electricity use.
3. **New Movers Initiative: Electric Only and New Movers-Dual Fuel:** The new movers initiative is composed of customers with recently activated accounts. New movers are subdivided into dual fuel, gas only, and electric only based on the categorization of the premise into which they moved. Notably this group includes reactivated customers as well as those who are "true new movers." National Grid has since established a flag in their database that distinguishes between new movers and reactivations. Going forward the program will target true new movers with messaging that emphasizes this major event. Customers trickled into the program on a rolling basis, however, reports were mailed in batches after enough customers were accumulated. The first batch of reports was mailed in November 2013.

In this section, we present our electric impact results associated with the electric only, dual fuel, new movers – electric only, and new movers – dual fuel cohorts.

4.2 Electric Home Energy Report Impacts

Table 2 displays the electric impact results for the electric only and dual fuel cohorts. As a proportion of energy use, the electric only cohort saves more energy than the dual fuel electric cohort, 1.10% compared to .92%. Notably, the electric only group consistently produced the most consistent and stable findings of all treated HER groups when testing our results using multiple savings impact models. These findings are expected, as high-usage electric cohorts have demonstrated the most robust savings in similar climates, such as those measured in the state of Massachusetts, as compared to lower usage electric customers and gas customers in general.

Accounting for savings due to other program results in small adjustments to overall savings (see section 4.5 for more detail on the participation lift associated with these savings).

Table 2. Home Energy Report Electric Savings (kWh)

	Dual Fuel-Electric	Electric Only
Treatment Period	April 2013 - May 2014	April 2013 - May 2014
First Report Date	2-Apr-13	2-Apr-13
Total Evaluated Participants	114,228	105,139
Baseline Usage (average daily kWh)	19.23	27.87
Net Savings (kWh per HH per day)	0.18	0.31
Net Program Savings from PPR (% per HH)	0.92%	1.10%
90% Confidence Interval Lower Bound	0.54%	0.75%
90% Confidence Interval Upper Bound	1.30%	1.45%
Incremental Savings from Other Programs (% per HH)	0.002%	0.06%
Final Adjusted Net Savings (% per HH)	0.92%	1.04%
Final Adjusted Net Savings (kWh per HH per day)	0.18	0.29
Total Adjusted Net Savings (kWH)	7,781,637	12,284,906
Implementer Reported Savings	7,183,012	13,149,758
Realization Rate	108%	93%

4.3 Electric Home Energy Report Impacts - New Movers Initiative

Impact findings for the electric only- and dual fuel-new movers are significantly more variable than for established customers and statistically insignificant (Table 3). As a

percentage of usage, the dual fuel cohort saved an unadjusted average of 1.64% per household while the electric only group gained an average 2.24% per household. However, confidence in these estimates is low. The dual fuel analysis has a 90% confidence bound of [-.95%, 4.11%] and the electric only analysis has a confidence bound of [-4.25%, .35%].

As mentioned above, the new movers program struggled to clearly distinguish between new customers and customers that were reactivating after a service disconnection. Consequently, plants to target the messaging to those moving into a new home were not implemented and the data sets do not allow for analysis of each distinct group (new movers vs. service re-connects). In addition, new movers began to receive treatment at varying points in time after activation as customers did not start receiving reports until there were sufficient numbers to assign to treatment and control groups. Both factors likely contributed to the large standard error on, and thus low confidence in, estimated savings.

Since the program could not be implemented as intended, the program implementer did not report savings for the program through May 2014 and this report represents the first evaluated savings for this treatment group.

Table 3. Home Energy Report Electric Savings – New Movers (kWh)

	Dual Fuel - Electric	**Electric Only**
Treatment Period	Nov 2013- May 2013	Nov 2013- May 2013
First Report Date	19-Nov-13	19-Nov-13
Total Evaluated Participants	16,916	25,545
Baseline Usage (average daily kWh)	14.25	19.07
Net Savings (kWh per HH per day)	0.238	-0.40
Net Program Savings from PPR (% per HH)	1.64%	-2.24%
90% Confidence Interval Lower Bound	-0.95%	-4.25%
90% Confidence Interval Upper Bound	4.11%	0.35%
Incremental Savings from Other Programs (% per HH)	0.08%	-0.13%
Final Adjusted Net Savings (% per HH)	1.56%	-2.37%
Final Adjusted Net Savings (kWh per HH per day)	0.24	(0.39)
Total Adjusted Net Savings (therms or kWh)	493,573	(1,210,095)

It is important to note here that the new movers' savings estimates may improve as the program matures (and customers are treated by the program for greater lengths of time) and as the data intelligence improves with National Grid's new methods for identifying new movers vs. service reactivations.

4.4 Opt-in Home Energy Reports

National Grid offered an opt-in web-based Home Energy Reports program to customers with low electricity use. These customers are typically not included in opt-out behavior programs, but National Grid wanted to offer low-usage customers a way to participate. Despite outreach and marketing efforts aimed at this group, the program did not generate enough interest to provide an evaluable cohort. National Grid ceased marketing outreach to this group in January 2014.

4.5 Overall Electric Participation Lift

As noted earlier, it is important to estimate the full energy efficiency impact of behavioral programs by examining the programs' effects on other National Grid programs across the treatment and control groups. Through this work, we can identify participation impacts associated with the program's activities. While these effects are directly attributable to the program, the savings associated with this lift in participation is removed from the overall program impacts to avoid double counting with other programs. That said, participation lift is an important metric when examining the impact of behavioral programs on the residential savings portfolio.

Table 4 below details the participation lift between electric-metered treatment and control customers. Note this includes electric only, dual fuel, and new mover cohorts. Overall, the program produced an incremental increase in program participation of 0.08% among electric-metered customers, resulting in a total increase of 1,242 additional participants in Rhode Island's residential programs due to the Home Energy Report program's activities. Figure 8 further demonstrates the participation lift that occurred after the Home Energy Report program's launch.

Table 4. Electric-Metered Customer Participation Rates (Treatment vs. Control) and Overall Program Lift

Cohort	**Electric Customers**	
	Control	**Treatment**
Treatment group size (n)	27,651	270,015
Before Behavioral Program		
Participants in other EE programs	1,274	12,673
Participation rate	4.61%	4.69%
Difference in Participation Rate	0.08%	
P-value of difference	0.5190	
Incremental Participants	n/a	
After Behavioral Program		
Participants in other EE programs	1,216	13,131

Cohort	Electric Customers	
	Control	Treatment
Participation rate	4.40%	4.86%
Difference in Participation Rate	0.46%	
P-value of difference	0.001	
Incremental Participants	1,242*	

*Incremental Participants is equal to Difference in Participation Rate time Treatment group size.

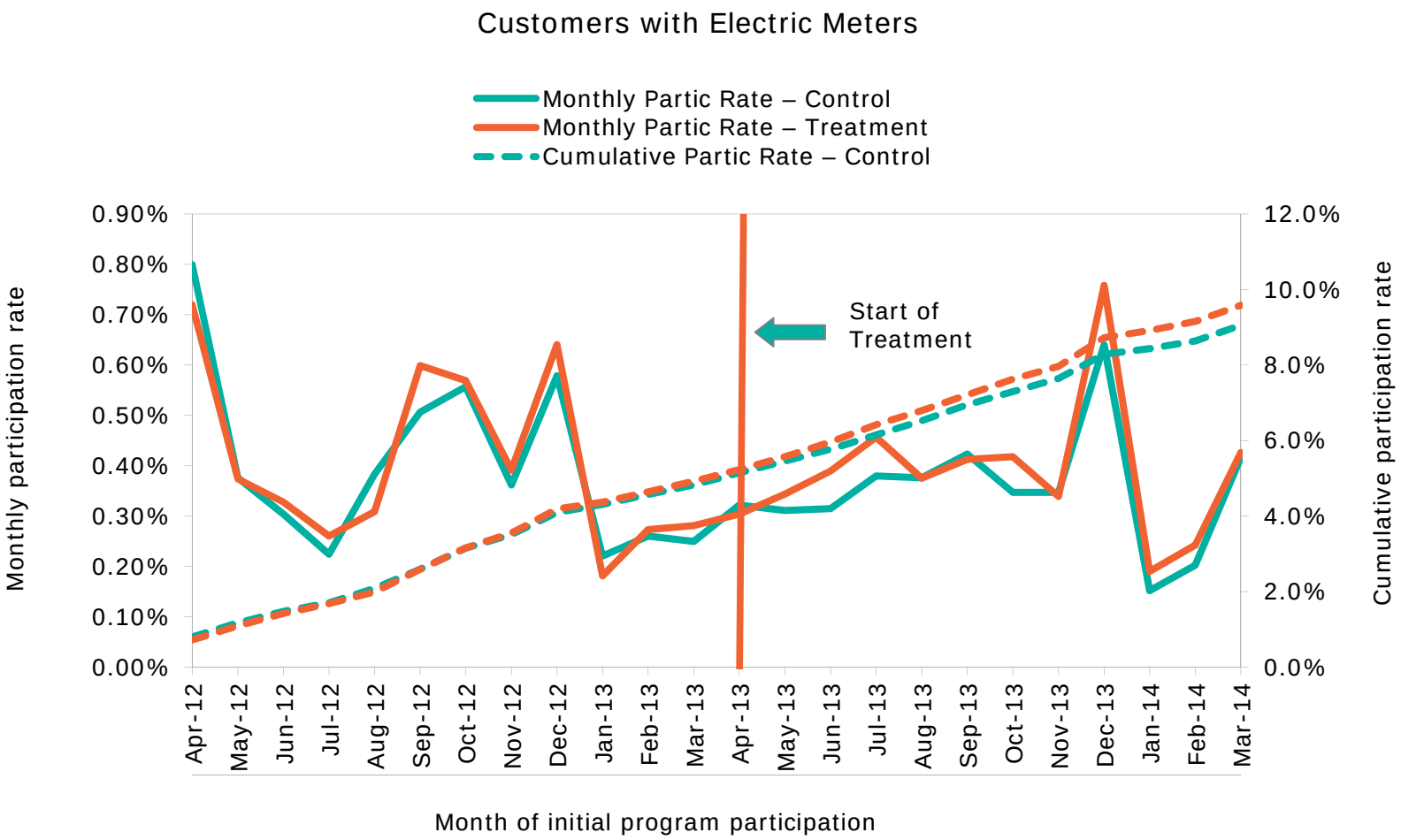
Table 5. Participation Lift for Program with Electric-Saving Measures

	Post-Behavioral Program Participation		
Energy Efficiency Program	Control	Treatment	Participation Lift* (%)
Income Eligible Single Family	189	1,361	-0.07%
Energy Wise Single Family	659	6,516	0.31%
ENERGY STAR Products	400	4,175	0.09%
ENERGY STAR HVAC	163	1,741	-0.01%
ENERGY STAR Lighting**	17	381	0.09%

*Overall participation lift for program including participation in gas- or electricity saving measures for programs that include both.

**Includes the mail-in rebate program only and does not include any bulbs purchased through the retail buy-down program. Bulbs purchased through the buy-down program are not tracked by participant.

Figure 8. Electric-Metered Customer Monthly and Cumulative Participation (Treatment vs. Control)



5. GAS HER IMPACT FINDINGS

5.1 Gas Home Energy Reports Overview

The statewide program implemented a gas only, duel fuel, and new movers HER initiative for gas-metered customers.

We briefly describe each below:

1. **Gas Only:** The gas only cohort received reports pertaining only to their gas use. Gas only customers started receiving reports in September 2013, just prior to the start of the heating season. To create the gas only cohort, customers with gas and electric use were first assigned to the dual fuel cohort to maximize the number of dual fuel customers. Then, a few areas of the state with high gas usage and a high density of gas customers were targeted for selecting customers to comprise the gas only group. To avoid sending gas reports to electric control customers, the gas areas were not included in the electric only cohort.
2. **Dual Fuel:** As noted earlier, customers in the dual fuel cohort received information on both their gas and electric use and began receiving reports in April 2013, where reports received in September through April focused more on gas use to target the heating season.
3. **New Movers Gas Only and New Movers Dual Fuel:** The program used a similar criteria for assignment for gas-metered customers who fell into the new movers designation as described for electric-metered new movers.

We detail our impact findings for the gas-metered customers below.

5.2 Gas Home Energy Reports Impacts

Savings by gas customers were relatively lower than savings by electric customers, with unadjusted savings per household of 0.34% and 0.67%, respectively. It is important to note that gas programs do not typically perform at the level of electric programs in HER programs and these savings differences align with impacts measured in other jurisdictions.

Table 6. Home Energy Report Gas Savings (therms)

	Dual Fuel - Gas	**Gas Only**
Treatment Period	Apr 2013 - May 2014	Sep 2013- May 2014
First Report Date	02-Apr-13	09-Sep-13
Total Evaluated Participants	114,228	16,191

	Dual Fuel - Gas	Gas Only
Baseline Usage (average daily therms)	2.4115	3.3387
Net Savings (therms per HH per day)	0.0083	0.0224
Net Program Savings from PPR (% per HH)	0.34%	0.67%
90% Confidence Interval Lower Bound	0.12%	0.31%
90% Confidence Interval Upper Bound	0.56%	1.02%
Incremental Savings from Other Programs (% per HH)	0.001%	0.02%
Final Adjusted Net Savings (% per HH)	0.34%	0.66%
Final Adjusted Net Savings (therms per HH per day)	0.01	0.02
Total Adjusted Net Savings (therms)	359,233	84,031
Implementer Reported Savings	386,995	94,330
Realization Rate	93%	89%

5.3 Gas Home Energy Reports Impacts – New Movers

Similar to the electric-metered customers receiving new mover treatment, the results for the gas-metered new movers have a large confidence interval and are statistically insignificant. We recommend that the program reevaluate these savings after new mover cohorts have had a longer treatment period.

Table 7. Home Energy Report Gas Savings – New Movers (therms)

	Dual Fuel-Gas	Gas Only
Treatment Period	Nov 2013 - May 2013	Nov 2013 - May 2014
First Report Date	19-Nov-13	19-Nov-13
Total Evaluated Participants	14,189	4,293

	Dual Fuel-Gas	Gas Only
Baseline Usage (average daily therms)	3.69	3.45
Net Savings (therms per HH per day)	0.018	0.023
Net Program Savings from PPR (% per HH)	0.49%	0.66%
90% Confidence Interval Lower Bound	-0.92%	-1.13%
90% Confidence Interval Upper Bound	1.86%	2.39%
Incremental Savings from Other Programs (% per HH)	0.05%	-0.05%
Final Adjusted Net Savings (% per HH)	0.44%	0.71%
Final Adjusted Net Savings (therms per HH per day)	0.02	0.02
Total Adjusted Net Savings (therms)	37,742	14,451

5.4 Overall Gas Participation Lift and Savings Findings

Gas customers have a smaller increase in program participation (when measured as a group) than electric customers. This may be due to a number of factors, including: (1) a pre-existing difference between treatment and control customers in program participation, where treatment customers participated at a higher rate in the pre-period than their control group customers, reducing their opportunities to participate in the post-period and (2) fewer total gas-focused actions to take through existing Rhode Island programs.

Table 8. Gas-Metered Program Participation Rates and Overall Program Lift (Treatment vs. Control)

Cohort	Gas Customers	
	Control	Treatment
Treatment group size (n)	28,425	167,123
Before Behavioral Program		
Participants in other EE programs	824	6,193
Participation rate	2.90%	3.71%
Difference in Participation Rate	0.81%	
P-value of difference	<.0001	
Incremental Participants	n/a	
After Behavioral Program		
Participants in other EE programs	627	4,333
Participation rate	2.21%	2.59%
Difference in Participation Rate	0.38%	

Cohort	Gas Customers	
	Control	Treatment
P-value of difference	0.0001	
Incremental Participants*	635	

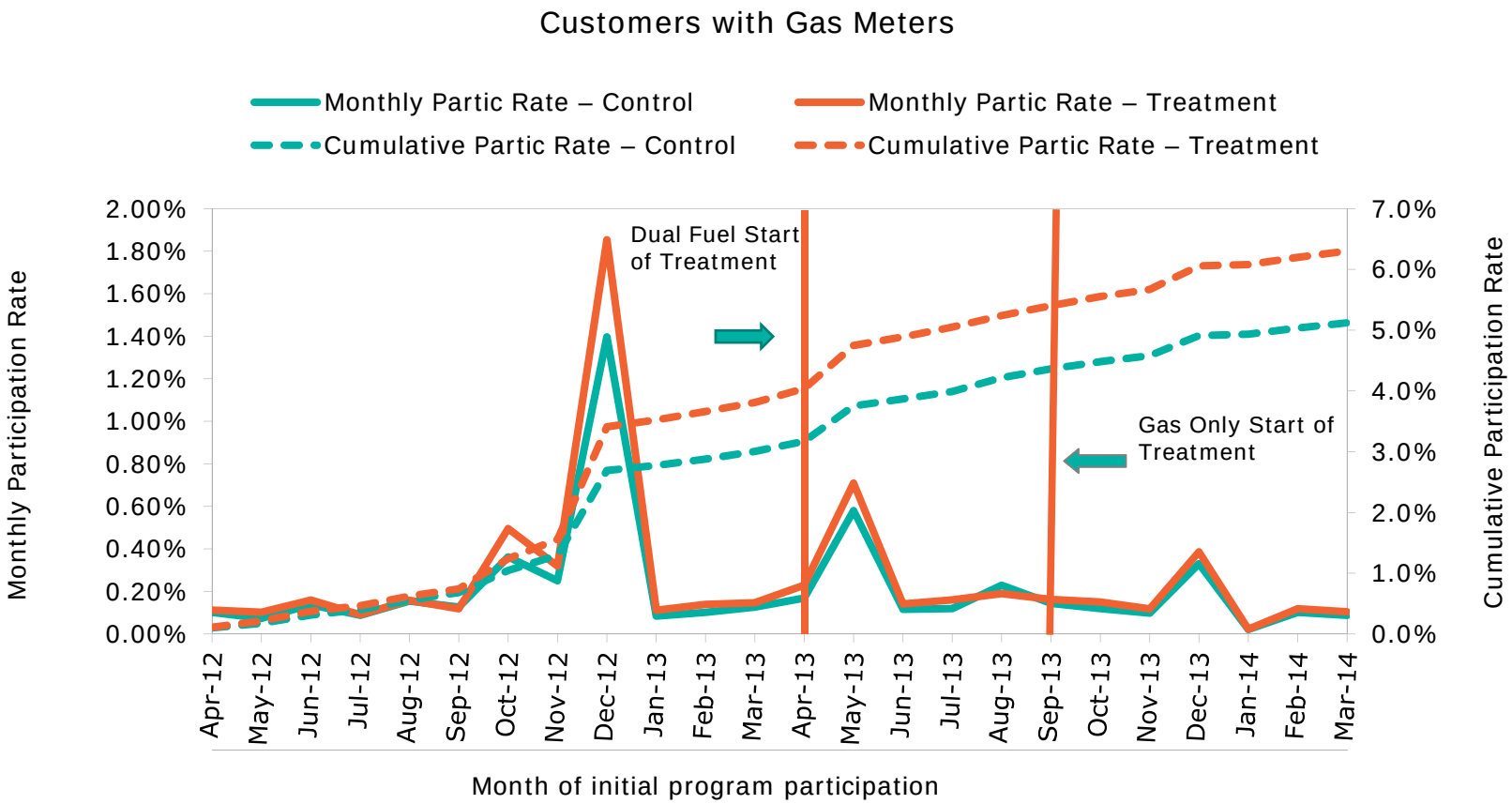
*Incremental Participants is equal to Difference in Participation Rate time Treatment group size.

Table 9. Participation Lift for Programs with Gas-Saving Measures

Energy Efficiency Program	Post-Behavioral Program Participation		Participation Lift* (%)
	Control	Treatment	
Income Eligible Single Family	189	1,361	-0.07%
Energy Wise Single Family	659	6,516	0.31%
ENERGY STAR HVAC	366	2,610	-0.28%

*Overall participation lift for program including participation in gas- or electricity saving measures for programs that include both.

Figure 9. Gas-Metered Customer Monthly and Cumulative Participation (Treatment vs. Control)



6. REWARDS PILOT IMPACT FINDINGS

6.1 Overview of the Rewards Pilot

A subset of the dual fuel and electric only groups were selected to participate in an online rewards pilot. The rewards pilot allowed customers to earn points for reductions in their energy use and other actions. Customers redeemed the points for gift cards and charitable contributions.

Dual fuel and electric only customers who received HERs were subdivided into treatment and control groups for rewards encouragement. Ninety percent of electric-only customers receiving HERs were assigned to the treatment group and ten percent were assigned to the control group. For dual fuel customers, rewards assignment was trickier.[13]

While the rewards portal was live as of May 2013, rewards treatment group customers did not receive notification of the pilot until June 2013. Initial pilot notifications included instructions for logging onto the rewards portal that required a separate sign-on using the customer's billing account number. Simply signing on the first time earned customers enough points to redeem a $1 donation to charity.

Throughout the pilot, customers received regular emails about points they earned. They were also informed of special offers, which included the following: limited-time double points for saving energy; matching donations to charities; points for signing up for electronic bills or for completing online audits.

The pilot experienced relative positive response and participation rates for opt-in programs administered to a large population. About five percent of eligible customers overall signed up, while over 10% of customers contacted via email signed up. Over 4,000 rewards have been redeemed and rewards pilot staff report that email communication open rates are over 50 percent.

Opower and National Grid have identified improvements to ease access to the rewards pilot. The primary improvement, scheduled to launch soon after the delivery of this report, is to implement a single-sign-on for the rewards portal so that customers will not have to log on to the online portal and the rewards portal separately. The initial sign on for customers in 2013 required knowledge of their billing account number.

---

13 Dual fuel customers with at least 12 months of data were randomly assigned to one of three groups: 1) Rewards treatment group; 2) Thermostat pilot treatment group; or 3) Control group. The Control group acted as the control for both the Rewards treatment group and the Thermostat treatment group.

Figure 10. Rewards Enrollment Sign-on

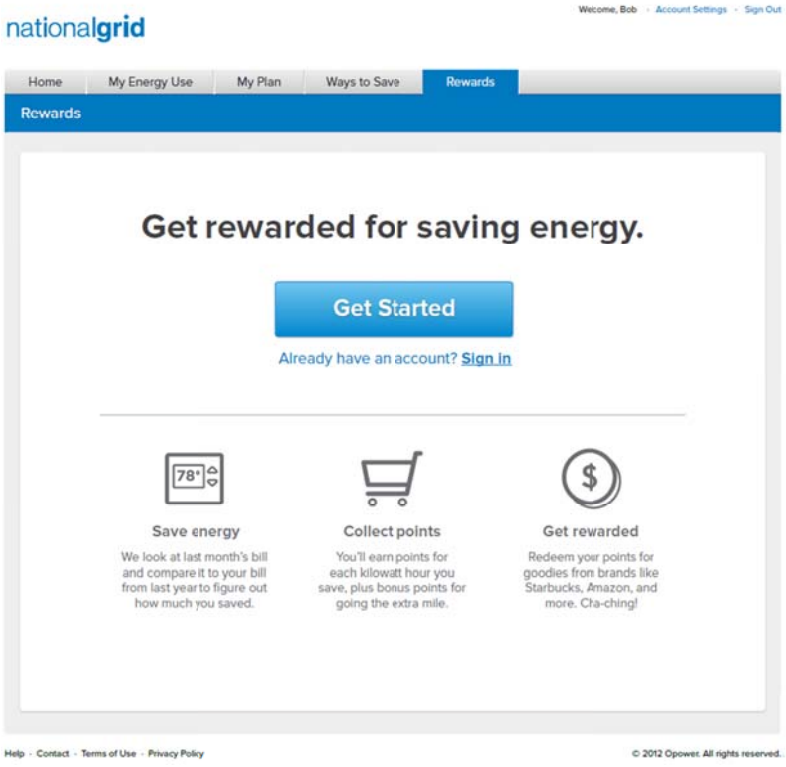
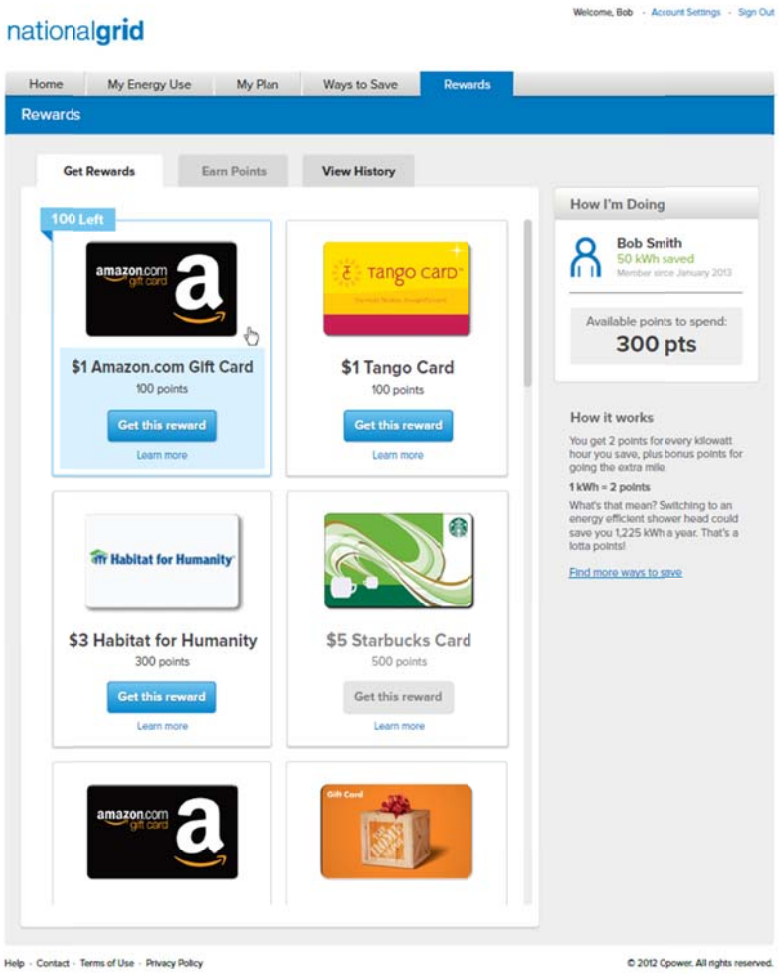


Figure 11. Example Rewards Offerings



The ILLUME team conducted an impact evaluation using two methods for the rewards efforts: (1) an evaluation using experimental design as originally developed by the pilot through randomized encouragement, and (2) a quasi-experimental design using a matched comparison group method (see Section 3.3.3 for more detail on our methodologies). The second approach was necessary because although the RED approach corrects for selection bias, it involves the use of regression to discern the overall average savings by all encouraged customers –whether they actually participated in the rewards pilot or not— which means, in the case where average savings conditional on rewards participation is modest, and the rate of response to the encouragement is low, attempting to discern a very small change in energy use. The matching method is not guaranteed to avoid selection bias, but directly compares energy savings by rewards participants and nonparticipants, yielding a more precise estimate of savings.

6.2 Electric Rewards Impacts

As measured by the matching approach, electric-metered customers across both the electric only and the dual fuel cohorts showed a significant increase in savings through the rewards pilot at .85% and 1.35% respectively. These savings values are very close to the estimates obtained with the parameter estimates produced using the RED approach, indicating that they are not contaminated by selection bias and significantly increasing confidence in the estimates.

Table 10. Rewards Electric Savings Comparison (kWh)

	Dual Fuel-Electric - RED	**Dual Fuel - Electric - Matching**	**Electric Only - RED**	**Electric Only - Matching**
Treatment Period	June 2013 - May 2014	June 2013 - May 2014	June 2013 - May 2014	June 2013 - May 2014
First Rewards Enrollment Date	15-Apr-13	15-Apr-13	04-May-13	04-May-13
Total Evaluated Participants	45,745	2,466	95,323	4,880
Baseline Usage (average daily kWh)	20.02	18.50	29.09	27.06
Net Savings (kWh per HH per day)	0.34	0.25	0.26	0.23
Net Pilot Savings (% per HH)	1.68%	1.35%	0.88%	0.85%
90% Confidence Interval Lower Bound	-5.01%	0.56%	-6.21%	0.26%
90% Confidence Interval Upper Bound	8.92%	2.13%	8.26%	1.44%

Total Net Savings	249,466	184,641	374,326	336,100

6.3 Gas Rewards Impacts

Similar to our findings for HERs, rewards savings are relatively lower for gas than for electric, with the matching method generating estimated savings of 0.012% per household. This value, while statistically significant, does not align as closely with the estimated savings obtained with the RED method, which is negative and has a very large confidence bound that encompasses 0.012%. We recommend using the savings estimate from the matching method for claimed pilot savings for both electric and gas into the future.

Table 11. Rewards Gas Savings Comparison (therms)

	Dual Fuel - Gas - RED	**Dual Fuel - Gas - Matching**
Treatment Period	June 2013 - May 2014	June 2013 - May 2014
First Rewards Enrollment Date	15-Apr-13	15-Apr-13
Total Evaluated Participants	45,745	2,466
Baseline Usage (average daily therms)	2.4269	2.6435
Incremental Net Savings (therms per HH per day)	-0.0529	0.0115
Incremental Net Pilot Savings (% per HH)	-2.23%	0.43%
90% Confidence Interval Lower Bound	-6.66%	-0.12%
90% Confidence Interval Upper Bound	1.85%	0.98%
Total Net Savings	-43,377	8,345

7. SMART THERMOSTAT IMPACT FINDINGS

7.1 Overview of the Smart Thermostat Pilot

National Grid launched a Smart Thermostat Pilot program within the Home Energy Report program to provide customers mobile access to their heating and cooling systems. The pilot was offered to dual fuel HER participants. The pilot used a Honeywell Wi-Fi Thermostat paired with Opower's Thermostat Management software. The thermostat offered mobile and web applications to control the thermostat, and provided real-time energy-saving tips to customers to reduce heating and cooling usage. For the pilot, 125 participants were recruited and given a professionally installed Honeywell VisionPRO thermostat at no charge, and asked to download the free Opower App that accesses the customer's energy usage data and controls their new thermostat. We provide images of the mobile application below.

Figure 12. Example Thermostat Opower Application

7.2 Thermostat Pilot Impact Results

The Illume team conducted an impact evaluation aimed at estimating the incremental savings the—savings above and beyond those generated by the HERs—associated with the thermostat pilot by developing a comparison group for thermostat participants among dual fuel HER customers. Of the 125 pilot participants, 123 had sufficient data to be included in the analysis.

Results indicate that dual fuel customers reduce electricity use by 2.42% in the cooling season, and reduce gas use by 2.35% in the heating season. The latter estimate is statistically significant at the 90% confidence level.

If the program team is interested in determining estimates of thermostat savings alone, it should consider creating a thermostat-only treatment group in addition to a joint thermostat-HER group, and expanding the overall treatment group size in the next iteration of pilot tests.

Table 12. National Grid Thermostat Pilot Savings - Electric

	Dual Fuel Electric Cooling Season	**Dual Fuel Electric Non-Cooling Season**
Treatment Period	June 2013 - August 2013	Apr 2013 - May 2013 & Sep 2013 - May 2014
Total Evaluated Participants	123	123
Baseline Usage (average daily kWh)	38.0214	24.1331
Incremental Net Savings (kWh per HH per day)	0.9411	0.0259
Incremental Net Program Savings from PPR (% per HH)	2.42%	0.11%
90% Confidence Interval Lower Bound	-0.60%	-2.89%
90% Confidence Interval Upper Bound	5.25%	2.93%
Total Savings (kWh)	10650	943

Table 13. National Grid Thermostat Pilot Savings - Gas

	Dual Fuel Gas Heating Season	Dual Fuel Gas Non-Heating Season
Treatment Period	Apr 2013 & Sep 2013 - Apr 2014	May 2013 - Aug 2013 & May 2014
Total Evaluated Participants	123	123
Baseline Usage (average daily therms)	3.7485	1.1005
Net Savings (therms)	0.0902	0.0237
Net Program Savings from PPR (% per HH)	2.35%	2.11%
90% Confidence Interval Lower Bound	0.62%	-1.04%
90% Confidence Interval Upper Bound	4.02%	5.06%
Total Savings (therms)	2662	430

8. OVERARCHING PROCESS & DESIGN FINDINGS

Below we present several overarching process and design findings synthesized from our interviews of program staff and impact analysis work.

Gas savings were over-estimated and thus planning goals were unattainable. Over-estimations resulted from errors in Opower's forecasting models as well as National Grid's difficulties in successfully identifying dual-fuel customers. Due to legacy IT systems, National Grid has separate electronic databases and billing systems for gas and electric customers. To identify which customers received both fuels from National Grid, Opower worked with the databases to match records on premise address and other key fields, but this resulted in fewer dual fuel customers than originally predicted. Both of these issues have since been resolved in program plans.

New movers definitions were too broad to inform a targeted outreach strategy. Due to customer data tracking limitations, new movers were identified broadly, including those who were new customers to National Grid as well as those that had delinquent and then reactivated accounts –two very different populations. As a result, the program could not be implemented in 2013 as intended and outreach and messaging could not focus on moving as a major event in customers' lives and impetus for saving energy. National Grid now has an indicator in their customer database to distinguish true new movers from reactivations. Since this is a distinctive population not typically targeted by programs, we recommend examining this program again after it has been implemented as designed.

Opt-in HER component did not generate enough interest to comprise an evaluable cohort. The opt-in component targeted lower electricity users, a group not typically included in opt-out programs. However, marketing and outreach efforts did not spur enough sign-ups to evaluate the program. National Grid concluded there was not enough interest to justify the cost and has discontinued the program.

Program design and implementation details were not carefully documented. To fully interpret and contextualize impact and process analysis findings, particularly for a program with this complexity, requires understanding program design details. However, details surrounding aspects of the program and pilots including program design, cohort requirements for inclusion in the program, and the launch dates for rewards initiatives were not carefully documented resulting in confusion internally at Opower and adding additional hours to the evaluation effort. We recommend that Opower develop documentation including relevant program dates and cohort inclusion requirements for each program/pilot and cohort.

RED design for opt-in efforts did not have sufficient participation levels, and statistical power, to be evaluable. The impacts of the rewards pilot were then calculated using a matching methodology, yielding similar, but statistically significant results. The RED design may not always be the best design for evaluating programs with small impacts and

low participation levels. We recommend including a matching methodology in future evaluations of pilot programs.

APPENDIX A: MATCHING METHODOLOGY

In program evaluation, the basic logic of matching is to balance the participant and non-participant samples by matching on the exogenous covariates known to have a high correlation with the outcome variable. Doing so increases the efficiency of the estimate and reduces the potential for model specification bias. Formally, the argument is that if the outcome variable Y is independently distributed conditional on X and D (conditional independence assumption), where X is a set of exogenous variables and D is the program variable, then the analyst can gain some power in the estimate of savings and reduce potential model specification bias by assuring that the distribution of X is the same for treatment and control observations.

In this evaluation, the outcome variable is monthly post-program energy use, and the available exogenous covariate with by far the greatest correlation with this outcome variable is energy use in the same month of the pre-program period, $ADClag_{kt}$, where k indexes the customer and *t* indexes the month. Both the RPPM and MBC approaches can be interpreted as using regression analysis to further control for any remaining imbalance in the matching on this variable. If, for instance, after matching the participants use slightly more energy on average in the pre-program period than their matches – they are higher baseline energy users, in other words — then for both the RPPM and the MBC approaches, including $ADClag_{kt}$ as an explanatory variable in a regression model predicting monthly energy use during the post-program period prevents this remaining slight difference in baseline energy use from being attributed to the program.

The RPPM approach

In the RPPM approach the development of a matched comparison group is viewed as a useful "pre-processing" step in a regression analysis to assure that the distributions of the covariates (i.e., the explanatory variables on which the output variable depends) for the treatment group are the same as those for the comparison group that provides the baseline measure of the output variable (see footnote 7). This minimizes the possibility of model specification bias. The regression model is applied only to the post-treatment period, and the matching focuses on those variables expected to have the greatest impact on the output variable.

The regression model used is the same PPR model used for the HER analysis (see Model 1).

The MBC approach

The second matching method follows the approach summarized in Imbens and Wooldridge (see footnote 9) and applied in Abadie and Imbens (see footnote 6). In this model, the effect of the program in month *t* is the difference between the energy use of participant *k* and their estimated counterfactual (baseline) consumption. The estimated counterfactual consumption is the average consumption of household *k*'s match amended to reflect differences between participants and their matches in the covariates ***X*** affecting energy use. Formally we have,

Model 3: Abadie-Imbens Model

$$Savings_{kt} = ADC_{kt} - ADC_{kt}^{C}$$
$$ADC_{kt}^{C} = ADC_{kt}^{M} + \hat{\alpha}\left(\mathbf{X}_{kt} - \mathbf{X}_{kt}^{M}\right)$$

where:

ADC_{kt} = the average daily energy consumption by household *k* during month *t;*

ADC_{kt}^{C} = the estimated counterfactual energy consumption by household *k* during month *t*;

ADC_{kt}^{M} = the energy consumption by household *k*'s match during month *t*;

$\mathbf{X}_{kt}$ = the values for household *k* in month *t* of the independent variables **X** affecting energy use;

X_{kt}^{M} = the values of **X** in month *t* for household *k's* match;

$\hat{\alpha}$ = the factors used to adjust household *k*'s energy use to reflect differences between household *k* and its match in the value of **X.**

The values of the adjustment factors $\hat{\alpha}$ used in Model 3 are derived from a regression model applied to the post-program period, estimated using *only* the matched comparison households. In the current analysis the regression model used for adjustment purposes is identical to the PPR model (Model 1) except that the variable $Participant_k$ is dropped, as the model is applied only to the matched comparison households. Formally,

$$ADC_{kt} = \sum_j \beta_{1j} Month_{jt} + \beta_2 ADClag_{kt} + \varepsilon_{kt}$$

To apply this regression equation to Model 4, we define $\mathbf{X}_{kt} = ADClag_{kt}$, and $\hat{\alpha} = \hat{\beta}_2$, so that Model 3 becomes in our specific context,

Model 3a: Abadie-Imbens Model (current context)

$$Savings_{kt} = ADC_{kt} - ADC_{kt}^{C}$$
$$ADC_{kt}^{C} = ADC_{kt}^{M} + \hat{\beta}_2\left(ADClag_{kt} - ADClag_{kt}^{M}\right)$$