Controlling for Nonprogram Effects in a Statistical Engineering Analysis

Michael P. Gallaher, Stephen A. Johnston, and Andrea S. Thomas
Research Triangle Institute, Research Triangle Park, NC
Alan Bailey; National Grid USA, Westborough, MA

ABSTRACT

This study examined the extent to which engineering estimates of kilowatt-hour (kWh) savings from National Grid USA's Small C&I Program in 2001 are realized in customer billing data. Separate realization rates were estimated by business type to assist in future program design. The analysis focused on lighting measures. Studies of this type are commonly referred to as "realization rate" studies.

Measured changes in energy usage from customer billing data are commonly compared to preliminary savings estimates to develop realization rates for energy conservation programs. An important part of these analyses is controlling for nonprogram-induced changes in energy consumption due to factors such as weather, business fluctuations, or changes in customers' building characteristics. However, limited customer-level data are typically available for program participants' (and nonparticipants') behavioral or building changes that affect changes in kWh usage over the window of analysis. This paper presents a set of innovative statistical techniques used to leverage survey information from a sample of participants and nonparticipants to control for nonprogram usage changes in the analysis population as a whole.

1.0 Introduction

This study examined the extent to which engineering estimates of kilowatt-hour (kWh) savings from National Grid USA's Small C&I Program in 2001 are realized in customer billing data. The Company's Small C&I Program is a direct installation program that serves customers of four of its retail companies: Granite State, Massachusetts, Nantucket, and Narragansett Electric Companies (collectively, the "Company"). Customers pay between 20% and 35% of total installation (material and labor) costs. The program is primarily a lighting program. Because lighting measures represented 94.1% of total program savings, these savings were the focus of the realization rate (RR) analysis.

Separate realization rates were estimated by business type to assist in future program design. Studies of this type are commonly referred to as "realization rate" studies. Measured changes in energy usage from customer billing data are commonly compared to preliminary savings estimates to develop realization rates for energy conservation programs.

An important part of these analyses is controlling for nonprogram-induced changes in energy consumption due to factors such as weather, business fluctuations, or changes in customers' building characteristics. However, limited customer-level data are typically available for program participants' (and nonparticipants') behavioral or building changes that affect changes in kWh usage over the window of analysis. This paper presents a set of innovative statistical techniques used to leverage survey information from a sample of participants and nonparticipants to control for nonprogram usage changes in the analysis population as a whole.

2.0 Methodology

The program savings realization rate was estimated using a statistically adjusted engineering (SAE) model. In this model, the dependent variable (change in measured usage) is regressed against the engineering estimate of savings, which was set equal to zero for nonparticipants. A pre-post, side-by-side sample of participants and nonparticipants was used in the analysis. Prior to estimating the realization rate, the participant and nonparticipant billing data were weather-normalized and screened for inconsistent or inaccurate data. Figure 1 presents an overview of the study methodology. Brief discussions of the surveys and econometric model are presented in Section 2.1 and Section 2.2, respectively. However, the focus of this paper is on the statistical methods used in the screening analysis, which are presented in Section 2.2. Section 3 presents the findings along with a sensitivity analysis for the screening procedures.

2.1 Participant and Nonparticipant Surveys

Telephone surveys of a sample of participants and nonparticipants were used to collect information on changes in business operations and electrical end uses that might affect changes in kWh usage over the time period of analysis. The survey was primarily designed to develop key data elements that are not currently available in the Company customer files. These data were tabulated and used for two purposes in the study: to help screen the initial database for the RR analysis and to help provide future guidance for the Small C&I Program.

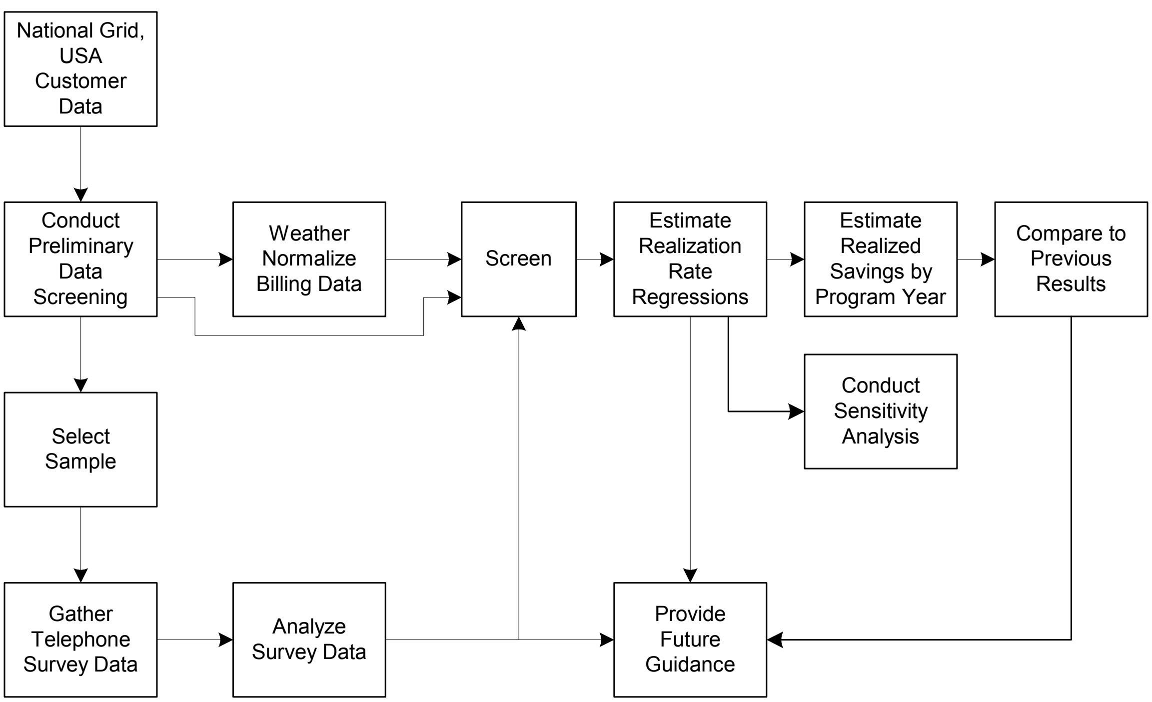

Figure 1. Study Methodology Overview

A sampling strategy was designed to yield a target of at least 200 completed responses from program participants and 200 completed responses from nonparticipants. Each sample was stratified by two kWh consumption levels and four building types. Sample points were allocated to the eight sample strata using an allocation procedure that reflects both the customer population in each stratum and the variability in kWh usage within the customer population in each stratum. Profiles of the participant and nonparticipant populations in terms of these strata are presented in Tables 1 and 2, respectively.

Table 1. Number of Participants in Population Strata

	Small[a]		Large[b]		Total	
Business Type	**Freq.**	**Percent**	**Freq.**	**Percent**	**Freq.**	**Percent**
Office	45	4.1	46	4.2	91	8.2
Retail	120	10.8	201	18.1	321	29.0
Services	211	19.0	145	13.1	356	32.1
Other	178	16.1	162	14.6	340	30.7
Total	554	50.0	554	50.0	1,108	100.0

[a]Small is annual usage less than the median usage of 30,494 kWh/yr.
[b]Large is annual usage more than the median usage of 30,494 kWh/yr.

A primary sample and an alternate sample were selected for each population. The purpose of the alternate sample was to provide a backup sample in case of refusals or other reasons for survey noncompletion that may have occurred with the primary sample. Because participants tend to have higher survey response rates than nonparticipants, the size of the alternate sample for participants was smaller than that for nonparticipants.

Table 2. Number of Nonparticipants in Population Strata

	Small[a]		Large[b]		Total	
Business Type	**Freq.**	**Percent**	**Freq.**	**Percent**	**Freq.**	**Percent**
Office	8,306	15.5	3,675	6.9	11,981	22.4
Retail	5,281	9.9	6,684	12.5	11,965	22.4
Services	9,074	17.0	4,210	7.9	13,284	24.9
Other	10,383	19.4	5,829	10.9	16,212	30.3
Total	33,044	61.8	20,398	38.2	53,442	100.0

[a]Small is annual usage less than the median usage of 30,494 kWh/yr.
[b]Large is Annual Usage more than the median usage of 30,494 kWh/yr.

Telephone survey. The sample survey was conducted by telephone. The survey was brief, requiring approximately 5 minutes to complete. Of those sample members contacted during the survey time period, approximately 65.5% responded with a completed survey.

A total of 239 participants and 258 nonparticipants completed the survey, and most provided usable responses to every question. These completion rates exceeded the target completion rates of 200 in each of the two groups. The additional completions occurred because extra telephone cases had to be worked to achieve or exceed the target number of completions in each of the eight strata. The cases were fielded in waves that included customers in all strata, and when the last stratum achieved its target completion rate, the other strata had exceeded their target completion rates.

A key finding from the survey was the extent to which participants and nonparticipants had undertaken some change action (e.g., employment, hours of operation, electrical equipment) during 2001 that would affect their electrical load and usage (see Table 3). Approximately 38% of all survey respondents reported a major change action, and the percentage of participants who reported a major change action was slightly less than that for nonparticipants.

Table 3. Major Changes Affecting Annual Energy Usage by Program Participation Status

	Participants		**Nonparticipants**		**Total**	
	Frequency	**Percent**	**Frequency**	**Percent**	**Frequency**	**Percent**
No Change	155	64.9	152	58.9	307	61.8
Yes Change	84	35.2	106	41.1	190	38.2
Total	**239**	**100.0**	**258**	**100.0**	**497**	**100.0**

Section 2.2 Realization Rate Model

The program savings RR was estimated with an SAE model. In this model, the dependent variable is regressed against the engineering estimate of savings, which was set equal to zero for nonparticipants. The model was expanded to include building type indicators. Other specifications of the model were also tested, including alternative specifications of the dependent variable.

The dependent variable for the RR analysis is a rate of change variable. It is defined as the change in kWh usage divided by prior kWh usage for each customer. Construction of this dependent variable for program participants and nonparticipants required careful alignment of the data, especially since participants joined the program at different points in 2001. Therefore, some participants could have had as little as one month of post-installation kWh usage data available (customers who joined the program in December 2001), while others had as much as 10 months of post-installation kWh usage data available (customers who joined the program in January 2001).

The final specification of the RR model is as follows:

$$\text{Norm\_}\Delta\text{kWh}_{it} = \alpha + b_1\,(\text{off\_savings}_i) + b_2\,(\text{ret\_savings}_i) + b_3\,(\text{serv\_savings}_i) + b_4\,(\text{oth\_savings}_i) + \delta_1\,(\text{office}_i) + \delta_2\,(\text{retail}_i) + \delta_3\,(\text{service}_i) + e_i$$

where

$\text{Norm\_}\Delta\text{kWh}_i$ = weather-normalized percentage change in aggregate usage for customer I;

off_savings_i = estimated savings as a percentage of weather-normalized usage for office customer i for office participants,
= 0 for office nonparticipants;

$ret\_savings_i$	=	estimated savings as a percentage of weather-normalized usage for retail customer i for retail participants,
	=	0 for retail nonparticipants;
$serv\_savings_i$	=	estimated savings as a percentage of weather-normalized usage for service customer i for service participants,
	=	0 for service nonparticipants;
$oth\_savings_i$	=	estimated savings as a percentage of weather-normalized usage for other customer i for other participants,
	=	0 for other nonparticipants;
$office_i$	=	business type dummy = 1 if customer i is an office building,
	=	0 otherwise;
$retail_i$	=	business type dummy = 1 if customer i is a retail building,
	=	0 otherwise;
$service_i$	=	business type dummy = 1 if customer i is a service building,
	=	0 otherwise; and
e_i	=	model error term.

The final model specification was then used to estimate the RR of program savings: i.e., the percentage of the engineering estimate of kWh savings that is revealed in billing data.

Section 2.3 Screening for Nonprogram Related Changes in Usage

Prior to the SAE analysis, two screening steps were conducted, employing information obtained from participant and nonparticipant surveys and from previous engineering studies. The first screen identified customers with a high likelihood of making nonprogram-related changes. Upper- and lower-bound threshold values for the percentage change in usage were estimated based on frequency distributions from surveyed participants and nonparticipants. Outside the threshold values, participants and nonparticipants had a high probability of having made nonprogram-related changes and were hence excluded from the SAE analysis. The second screen identified program participants whose savings-to-usage ratio was inconsistent with their likelihood of having electric heat or significant electric cooling.

The data problems are reflected in the following descriptive statistics:

1. Several participants had "unrealistically" high ratios of savings to annual usage. For participants with savings/usage greater than 0.50, we reviewed account numbers to identify obvious savings to usage mismatches. Of the 132 customers reviewed, 45% had account mismatch problems. In these "account mismatch" situations, measures were installed in building areas not served by the meter represented in the account billing data, thus leading to an underrepresentation of annual usage and overestimation of the saving/usage ratio.
2. The average percentage change in usage from 2000 to 2001 for participants was approximately –5.4%. This is far less than the average estimated savings as a percentage

of usage, 18.6%. It is likely that a subset of participants made other changes (in addition to the program measures installed) that influenced their usage. For example, the telephone surveys found that 12% of participants increased their number of employees from 2000 to 2001 (see Section 5, Table 5-3).

The two issues presented above stem from missing information. The main types of missing information are the following:

- information used to link participants' savings to the associated account numbers/location IDs (locations associated with the meter readings), and
- information on additional changes participants and nonparticipants made to increase (or decrease) usage over the time period of the analysis.

We addressed these issues by using the two screening procedures described below.

Screen on the ratio of savings to annual usage. To address the issues of possible additional "account mismatch" problems, we developed reasonable limits on the ratios of savings to annual usage based on engineering judgment. It was suspected that several customers could have missing accounts in their usage estimates. For example, if lighting measures were installed in several different buildings but recorded on a single rebate form as all being associated with the master account, then not all the baseline usage associated with the total area where the measures were installed would be included. This would lead to an underestimation of annual usage and hence an underestimate of an associated change in usage. As a result, the ratio of estimated savings to change in usage would appear to be "unreasonably" high. The actual problem may have been that not all the space in which the measures were installed was included in the analysis.

To determine the likelihood of this problem, we reviewed 132 participants with ratios of savings to usage greater than 0.50 (i.e., the ratio indicated that this group of participants would reduce their current usage by 50%). Of the participants reviewed, 59 were found to have account mismatch problems. After identifying the customer's additional accounts, their annual usage was recalculated.

Time and resource constraints prevented reviewing all 1,393 eligible program participants for account mismatch problems. As an alternative, we developed savings-to-usage ratio screens using the participant survey data and engineering judgment.

Based on average energy intensity for specified uses in the Northeast, we identified the following reasonable limits for lighting savings as a share of total energy usage. These ratios are presented in Table 4. The ratios vary dependent on the electrical end uses an individual customer may have. For example, if a participant's main electrical load is lighting, implying that he does not have electric cooling or electric heating, then it may be reasonable for this customer to realize a usage reduction of up to 30% by installing high-efficiency lighting. However, if the customer also has electric cooling, but no electric heating, then it is unlikely that installing high-efficiency lighting can reduce the total energy usage by more than 14.8%. Conversely, if a customer has electric heating but no electric cooling, it is unlikely that installing high-efficiency lighting can reduce the total energy usage by more than 8.6%. Similarly, if a customer has electric cooling and electric heating, then it is unlikely that installing high-efficiency lighting can reduce the total energy usage by more than 6.6%.

Unfortunately, this procedure could not be duplicated for the nonsurveyed population because it was not possible to predict the electrical end uses that customers may have. Instead, we estimated a single cut based on the average energy intensity per square foot of space in the Northeast for heating, cooling, lighting, and water heating and the percentage of the surveyed population with heating and/or cooling. Based on the survey of 239 participants, 21.8% of participants in the Company's service territory have only electric lighting, 62.6% have only electric lighting and cooling, 2.4% have only electric lighting and heating, and 13.1% have a combination of electric lighting, heating, and cooling. These results were consistent between participants and nonparticipants. Based on the survey, 18.1% of nonparticipants in the Company's service territory have only electric lighting, 61.2% have electric lighting and cooling, 3.5% have electric lighting and heating, and 17.2% have a combination of electric lighting, heating, and cooling.

Table 4. Savings to Annual Usage Ratio Screen

Customers' Electrical End Uses	Upper Bounds	Percentage of Participants	Percentage of Nonparticipants
Lighting (i.e., No Cooling or Electric Heating)	Savings/Usage = 30.0%	21.8	18.1
Lighting and Cooling (i.e., No Electric Heating)	Savings/Usage = 14.8%	62.6	61.2
Lighting and Heating (i.e., No Electric Cooling)	Savings/Usage = 8.6%	2.4	3.5
Lighting Cooling and Electric Heating	Savings/Usage = 6.6%	13.1	17.2

We applied the energy intensity data assuming that lighting cannot reduce usage by more than 30% to the information gathered from the survey concerning the percentage of the surveyed participants with heating and/or cooling. Thus we determined that savings to usage should not exceed 16.9% for nonsurveyed participants. Appendix B details the calculations used for both the nonsurveyed and surveyed populations. As a result of applying this screen an additional 646 participants were removed from the final RR sample.[1]

Screen on change in annual usage. To address the "additional changes" issue, we developed a set of upper- and lower-bound screens for percentage changes in annual usage to remove participants and nonparticipants that seemed to have made major changes in business operations, occupancy, or employment, or that seemed to have undertaken measures in addition to those installed under the program. For example, if a customer had a very large increase (decrease) in usage between the first and second year of our analysis period, this is a good indication that the customer made major changes in addition to installing the program measures. The difficulty in addressing this problem is that we have no information on nonsurveyed customers' behavior or changes in business operations. The upper- and lower bound screens on percentage changes in annual usage were designed to compensate for this missing information.

To develop the upper- and lower-bound screens for changes in annual usage, responses from the participant and nonparticipant surveys were used to identify when customers were most likely to have made nonprogram changes as a function of their actual percentage change in usage. These nonprogram changes included

- increases or decreases in building size,
- changes in major electricity-using equipment,
- changes in building occupancy,
- changes in operating hours, and

[1]No nonparticipants were removed as a result of applying this screen because this screen does not apply to them (since it is based in part on energy savings estimates).

- changes in number of employees.

Thirty-five percent of participants surveyed and 41% of nonparticipants surveyed answered yes to one or more of these questions. Section 5 provides more details on the frequency of specific changes.

Based on the survey responses, frequency distributions were developed for participants and nonparticipants who indicated that changes were made that increased and/or decreased usage. The distribution of participants and nonparticipants who made "changes" was then compared to similar distributions of participants and nonparticipants who indicated they made "no changes." By comparing the "change" and "no change" distributions, upper- and lower-bound screens on percentage change in usage were developed that identified the thresholds where participants and nonparticipants become more likely to have made changes. Estimating separate thresholds for participants and nonparticipants also helped control for differences in average lighting loads between the two groups.

The upper- and lower-bound screens for participants and nonparticipants are shown in Table 5. The screens are interpreted as follows:

- For the upper-bound participant screen, if participants had a change in usage greater than 0.0%, there was a higher probability that participants made a nonprogram change to increase usage. If the change in usage was less than 0.0%, there was a higher probability that participants did not make changes to increase usage.
- For the participants' lower-bound screen of –30%, if a participant's usage decreased by more than 30%, they were likely to have made a nonprogram change that affected their usage over the period of analysis.
- Nonparticipants' upper- and lower-bound screens were interpreted similarly.

Table 5. Upper- and Lower-Bound Screens for Percentage Change in Usage

Screen	Percentage Change
Participants Upper Bounds	0%
Participants Lower Bounds	–30%
Nonparticipants Upper Bounds	20%
Nonparticipants Lower Bounds	–10%

Final population used in the RR analysis. Table 6 indicates the number of participants and nonparticipants that were removed as a result of applying the savings-to-annual-usage ratio screen and the upper- and lower-bound percentage-change-in-usage screen. The resulting final population used in the RR regression was 415 participants and 1,028 nonparticipants.

Table 6. Participants and Nonparticipants Removed by the Two Final Screens

Screen	Participants	Nonparticipants
Preliminary Eligibility Population	1,393	1,393
Removed by Savings to Annual Usage Ratio Screen	646	0
Removed by Upper-and Lower-Bound Percentage Change in Usage Screen	332	365
Final Sample for RR Regression Analysis	415	1,028

3.0 Findings

A cross-sectional regression model was estimated for the combined final sample of 415 participants and 1,028 nonparticipants. The dependent variable in the regression is the weather-normalized aggregate percentage change in usage. The independent variables in the regression are the estimated energy savings, by business type, as a percentage of weather-normalized usage and business type indicators ("dummy variables").

3.1 RR Regression Results

Table 7 presents the RR obtained with the final model specification. The F value indicates that the model is statistically significant at the 99% confidence level. The estimated coefficients for energy savings business types (the business type RRs) range from –0.785 to –1.108. The standard errors for the business type RR estimates range from 0.121 to 0.062. All of these parameter estimates are significantly different from 0 and only the "retail" RR is significantly different from –1 at the 95% confidence level. The adjusted R2 for the model is 0.314 and the F value is 94.1.

The intercept term is also significantly different from zero at the 95% confidence level. The estimated intercept indicates that "other" businesses had a 1.6% growth trend over the period of analysis. The other three parameter estimates for the dummy variables show how the growth trends in these business sectors differ from the omitted dummy category. For example, "retail" businesses show a general trend of 0.3% (1.6 – 1.3) change in usage. The "retail" parameter estimate is significant at the 95% confident level, and the "service" parameter estimate is significant at the 90% confident level. The "office" parameter estimate is not significant, indicating the usage growth trends for the "other" and "office" categories are not significantly different.

To estimate a single RR that can be applied to total program savings, the business-type RRs were weighted by the share of total program savings attributable to each business type. The savings weights and the single weighted RR estimate of 0.97 are shown in Table 8. This weighted RR estimate indicates that 97% of the engineering estimates of program savings are realized in the customer's billing usage, and is statistically significant at the 95% confidence level.

This weighted average RR of 97% for program year 2001 is up sharply from the RR of 71% estimated for program year 1999. A portion of this RR increase can be attributed to fewer data quality problems. With better data, we were able to estimate a more detailed regression model than before, and we found the model results to be statistically significant. The improved RR results may also be the result of more participants who are in the manufacturing sector. Lighting savings estimates tend to be more accurate for manufacturers than for nonmanufacturers, because their lighting hours of operation are more directly related to business hours of operation.

Applying the RR result to the engineering savings estimate for all 2001 Small C&I Program lighting participants of 18.0 million kWh yields a realized lighting savings of 17.5 million kWh. If a similar RR holds for all program participants (including participants who installed nonlighting measures), the realized program savings are 18.5 million kWh.

Table 7. Energy Usage Regression

Dependent Variable: Normalized Aggregate Percentage Change in Usage
Average Dependent Variable: –0.0205
Number of Observations: 1443
Adjusted R^2: 0.3138
F Value: 94.1
Prob > F: 0.0001

Variable	Parameter Estimate	Standard Error	t-stat for Ho: Parameter = 0	t-stat for Ho: Parameter = –1
Intercept	0.016[a]	0.004	4.12	
RR on Office Savings	–0.889[a]	0.121	–7.35	–0.92
RR on Retail Savings	–0.785[a]	0.075	–10.50	–2.87[a]
RR on Service Savings	–0.888[a]	0.062	–14.38	–1.81[b]
RR on Other Savings	–1.108[a]	0.066	–16.75	–1.64[b]
Baseline: Office	–0.008	0.008	–0.99	
Baseline: Retail	–0.013[a]	0.006	–2.28	
Baseline: Service	–0.011[b]	0.006	–1.91	

[a]Significant at 95% level.
[b]Significant at 90% level.

Table 8. Weighted Realization Rate

Business Type	Parameter Estimate	Savings (kWh)	Savings Weight	Weighted RR
Office	–0.889	333,440	15%	
Retail	–0.785	442,012	19%	
Service	–0.888	508,965	22%	
Other	–1.108	1,010,918	44%	
Total		**2,295,335**	**100%**	**0.97**

3.2 Sensitivity Analysis For Screening

Sensitivity tests were conducted for the two types of screens described in Section 2.3. The sensitivity of the weighted RR results to a change in the percentage change in usage screen is summarized in Tables 9 and 10. The sensitivity of the weighted RR results to the savings/usage screen is summarized in Table 11.

Table 9. Weighted RR Sensitivity Analysis to Upper-Bound Percentage Change in Usage Screen

Participants	**Nonparticipants**		
	10.0%	**20.0**%	30.0%
–10.0%		–1.45	
0.0%	–0.81	**–0.97**	–1.03
10.0%		–0.63	

Table 10. Weighted RR Sensitivity Analysis to Lower-Bound Percentage Change in Usage Screen

Participants	**Nonparticipants**		
	–20.0%	**–10.0**%	0.0%
–40.0%		–1.05	
–30.0%	–0.85	**–0.97**	–1.24
–20.0%		–0.84	

Table 11. Savings to Annual Usage Screen Sensitivity Analysis

Savings to Usage Upper Bound	**Weighted RR**
13.4%	–1.06
16.9%	**–0.97**
20.4%	–0.85

In Tables 9 through 11, the final weighted RR estimates (presented in Section 3.1) are shown in bold in the center of the tables, and the original screens are shown in bold in the table labels. The off-center entries in the table indicate how the weighted RR estimate changes as the screen parameters are varied, holding all other screens constant at their original values.

The sensitivity analyses for the percentage change screens presented in Table 9 show that the RR estimate increases as the nonparticipant upper bound increases, and decreases as the participant upper bound increases. This is because increasing the nonparticipant upper-bound screen increases the overall population trend in the percentage change in usage. And because "real" savings are measured relative to the overall population trend, a greater trend increases the RR estimate.

Table 9 also shows that the weighted RR estimate decreases as the participant upper bound increases. This is because increasing the percentage change upper bounds for participants included more participants in the regression with increased usage (i.e., no apparent savings).

Table 10 shows the participants' and nonparticipants' percentage change lower-bound sensitivity analysis. The weighted RR estimate is less sensitive to the lower-bound screens compared to the upper-bound screens. For example, as the nonparticipant lower-bound screen decreases to 0.0%, the weighted RR estimate increases because the overall population percentage change in usage trend decreases.

The sensitivity analyses presented in Table 11 show that the weighted RR estimate changes as the savings-to-usage screens used for nonsurveyed participants change. The weighted RR estimate increases as the ratio of savings-to-usage screens increases or decreases (i.e., as screens become more stringent). This may indicate that the ratio of savings to usage screens in the final regression analysis is accomplishing the intended purpose of controlling for account mismatches.

4.0 Conclusions

National Grid USA uses the estimated realization rates to adjust the savings estimates they report to their regulatory commission. The business-specific realization rate estimates can also be used to refine program design and marketing efforts.

Using the primary data from the surveyed population, it was estimated that approximately 8% of the participants and 15% of nonparticipants made significant nonprogram-related changes in their behavior or building characteristics that led to changes in kWh usage. Excluding these participants and nonparticipants from the SAE analysis increased the realization rate from 55% to 97%. Realization rates were also estimated by building type (office, retail, service, other). Building type estimates ranged from 0.79 for retail to 1.10 for other building types. All of the estimates were statistically significant at the 95% confidence level.

Realized savings for similar lighting measures estimated using monitoring techniques (lighting loggers) typically range from 80% to 100%. After applying our screening procedures, realization rates from the SAE analysis for this program were within the typical range. In addition to nonprogram-related changes in customer usage, issues associated with matching billing data with specific meter accounts also lead to data inconsistencies. The screening approaches discussed in this paper were also designed to address this issue.