Con Edison EEPS Programs - Impact Evaluation of Large Commercial & Industrial Program Group

December 11, 2014

energy & resource solutions

Corporate Headquarters:
120 Water St., Suite 350
North Andover, Massachusetts 01845
(978) 521-2550

Contents

1 Executive Summary

The Consolidated Edison Company of New York (CECONY) Large Commercial & Industrial (C&I) Equipment Rebate suite of programs was launched in 2009. It promotes the installation of high-efficiency equipment by C&I customers in existing facilities by providing customers with financial incentives to offset the higher purchase cost of energy efficient equipment and information on the features and benefits of energy efficient equipment. Qualifying equipment eligible for prescriptive incentives includes packaged air conditioners, variable speed drives, motors, lighting fixtures, boilers, and controls. High efficiency equipment not specified in CECONY's prescriptive rebate program may participate in the custom efficiency programs. Energy Efficiency Portfolio Standard (EEPS) funding is separate for each of the prescriptive electric, prescriptive gas, custom electric, and custom gas measures, but CECONY markets the programs to customers and administers them together. CECONY has contracted with Lockheed Martin to administer the programs and supports program operation with dedicated internal staff.

1.1 Impact Evaluation Objectives

The goals of this group impact evaluation are to:

- Evaluate the program's performance by developing gross savings realization rates (RRs) and a net-to-gross ratio (NTGR) for installed projects committed in 2010 or 2011 and acquired by May 23, 2012 for the four CECONY programs.
- Provide actionable recommendations for improving the program's implementation as a result of these assessments.

The report includes estimates of gross and net impacts (annual kWh, annual therms, and summer coincident peak kW) from all the program participants and measures.

1.2 Research Approach

The realization rates were developed from on-site measurement and verification (M&V) of a statistically representative sample of sites using the stratified ratio estimation (SRE) method. Post-stratification analysis was used to examine results by measure, particularly for variable frequency drive (VFD), lighting, and energy management system (EMS) measures.

This report also includes a study of free ridership (FR) and spillover (SO). In a separate effort, using an independently drawn sample, evaluators measured program attribution and thus net impact, using a self-reported approach with a telephone survey of an independent random sample of participants. Evaluators conducted trade ally surveys to verify the FR responses and explore potential nonparticipant spillover (NPSO). NPSO's potential impact was examined to determine if future investigation is warranted but was not used as a component of the net-to-gross (NTG) results.[1]

This evaluation plan complies with the requirements of the August 7, 2008 (updated February 2012) Evaluation Guidelines issued by the DPS[2] and is intended to provide robust, timely, and transparent results. The impact methods are aligned with the guidelines of the National Action Plan for Energy Efficiency's Model Energy Efficiency Program Impact Evaluation Guide (NAPEE Guide).

1.3 Results

The evaluation team calculated gross and net energy and demand savings separately for the four CECONY programs.

1.3.1 Gross Program Impacts

Each site in the M&V sample was visited by an engineer who installed metering equipment appropriate for the measure. Customers were interviewed to determine year-round operation and the operation of baseline equipment. At the conclusion of the metering period, loggers were pulled from each site and the data was incorporated into algorithms to calculate first-year savings for the measures. The engineers produced a site report documenting their findings and conclusions for each site included in the sample.

Program level savings were calculated by applying the site sample weights to both the tracking and evaluated savings. The ratio of the sum of the weighted evaluated savings divided by the sum of the weighted tracking savings is the program RR.

Tables 1-1 and 1-2, respectively, present the gross energy and peak demand results of the electric program evaluation. Table 1-3 presents the gross energy impacts of the gas program evaluation. The results did not quite achieve the sample design's targeted 10% relative precision (RP) at 90% confidence for each program's energy savings RR.

All programs and savings types had gross RRs of 70% or greater, with the exception of the gas rebate program, which had a gross RR of 48%.

---

[1] NPSO is expected to be evaluated in a separate statewide study.

[2] Including subsequent documents when finalized: "Spillover DPS Guidance 7-11-12," "Appendix A: Sampling DPS Guidance 7-11-12," and "Appendix B: Reporting and Accessing DPS Guidance 7-11-12."

Table 1-1. Electric Program Gross Energy Impacts

Program	Population	Sample	Tracking Savings (kWh/yr)	Gross Energy RR	RP
Electric Rebate	632	56	67,813,154	0.70	12%
Electric Custom	232	47	27,132,715	0.72	9%
Total	**864**	**103**	**N/A**	**N/A**	**N/A**

Table 1-2. Electric Program Gross Peak Demand Impacts

Program	Population	Sample	Tracking Savings (kW)	Gross Peak Demand RR	RP
Electric Rebate	632	56	8,368	0.76	35%
Electric Custom	232	47	3,512	0.92	14%
Total	**864**	**103**	**N/A**	**N/A**	**N/A**

Table 1-3. Gas Program Gross Energy Impacts

Program	Population	Sample	Tracking Savings (Therms/yr)	Gross Energy RR	RP
Gas Rebate	66	24	422,630	0.48	18%
Gas Custom	4	1	326,219	1.01	N.D[1]
Total	**70**	**25**	**N/A**	**N/A**	**N/A**

[1] The RP for the Gas Custom program was not determined (N.D.) because the evaluators were able to recruit only one site for their on-site analysis.

1.4 Attribution Results

The evaluation team estimated the FR and SO rates for CECONY's program using results from surveys conducted with the program participants and participating trade allies of the four C&I CECONY programs. These surveys followed the AAPOR guidelines for conducting the surveys. As part of the survey effort, the evaluation team contacted all program participants (census attempt) associated with projects completed between early 2010 and May 2012 and asked them a series of structured and open-ended questions about the program's influence on their decision to make energy efficient improvements. Interviews with participating trade allies further validated the FR results.

As a result of the research efforts, the evaluation team was able to develop independent FR rates for CECONY's Electric Rebate and Electric Custom program components. For the Gas Rebate program, a small number of interviews (n=3) were completed, which yielded a FR value of 0.24. Even though the evaluation team had results from the three surveys for the Gas Rebate program, the evaluation team deemed them to be insufficient to accurately represent the program.

Due to the small number of projects (n=5) for the Gas Custom component, the evaluation team was unable to complete any interviews and as a result, the evaluation was not able to develop a FR estimate specific to the Gas Custom program.

The evaluation team also estimated SO at the overall program level. The evaluation team found that additional energy savings occurring as a result of the program were very limited. The final FR, SO, and the resulting NTGR results are presented in Table 1-4. As can be seen in the table, FR ranges from 0.26 for the Electric Rebate program to 0.38 for the Electric Custom program.

Since the attribution survey was a census attempt, the concept of sampling error does not apply. The evaluation team developed relative precision estimates to provide insight on what they would have been if the effort was a sampling effort and not a census attempt, but did not propagate the relative precision to the relative precision around net impacts.

Table 1-4. Free Ridership, Spillover, and Net-to-Gross Estimates

Program Component	Count	FR	SO	NTG	NTG RP[1]
Electric Rebate	84	0.26	0.0	0.74	0.244
Electric Custom	26	0.38	0.0	0.62	0.188
Gas Rebate	3	I.D	N.D	0.77	0.074
Gas Custom	0	N.D	N.D	N.D	N.D

[1] Since the survey was a census attempt, the concept of relative precision, which represents sampling error, does not apply. The relative precision values presented above cannot be propogated to the rest of the population due to a large number of non responses in the survey.

I.D. – Insufficient data. This requires further study to accurately estimate this value.

N.D. – No data was available to develop estimates for these variables.

The FR rates found as part of this evaluation are not unusual and generally mimic those for other similar programs in the region and across the country. Table 1-5 provides an overview of the FR rates found in other parts of the country for similar programs.

Table 1-5. Free Ridership Rates in Other Jurisdictions

Program	FR	Study Year
Northeast Utilities Electric C&I program	0.28	2011
Massachusetts gas programs	0.31	2012
Massachusetts electric programs	0.16	2011
Efficiency Maine Electric Business Incentive Program	0.34	2011
NYSERDA Existing Facilities Program	0.31	2010–2011

Quantifying NPSO and vendor off-site SO was not in the scope of this evaluation but was explored during the trade ally interviews. Indications are that trade allies are implementing high efficiency projects outside of the program due to the program's time-consuming paper work and the installation of equipment/measures that is not currently listed in the program's approved category.

1.5 Program Net Savings

The net program results are calculated by multiplying the gross program results by the NTGR. Tables 1-6 and 1-7 present the net energy and peak demand values for the CECONY electric programs.

Table 1-6. Net Electric Programs Energy Summary

Program	Tracking Savings (kWh/yr)	Gross RR	Evaluated Gross Savings (kWh/yr)	NTGR	Evaluated Net Savings (kWh/yr)	Net RP
Electric Rebate	67,813,154	0.70	47,469,208	0.74	35,127,214	27%
Electric Custom	27,132,715	0.72	19,535,555	0.62	12,112,044	21%

Table 1-7. Net Electric Programs Peak Demand Summary

Program	Tracking Savings (kW)	Gross RR	Evaluated Gross Savings (kW)	NTGR	Evaluated Net Savings (kW)	Net RP
Electric Rebate	8,368	0.76	6,360	0.74	4,706	41%
Electric Custom	3,512	0.92	3,231	0.62	2,003	26%

The evaluation team did not present the net program savings results for the two CECONY gas programs because the attribution survey, which was a census attempt, did not yield sufficient data (3 surveys for Gas Rebate and no surveys for Gas Custom) to accurately estimate the attribution factors for the two CECONY gas programs.

1.6 Findings and Conclusions

The findings and conclusions are presented through two perspectives:

1. Sources of savings discrepancies
2. Equipment performance, particularly relative to New York Technical Reference Manual (NYTM) values

1.6.1 Sources of Savings Discrepancies

The evaluation team identified reasons why the evaluated savings were different from the applicant estimates of savings for each site. In general the applicant's load profile and operations profile estimates were good. The NYTM assumptions likewise proved to be unbiased estimators. The program's post-implementation quality control also appeared to be working well because little deviation was observed in the installed equipment from what had been proposed in the application.

The biggest contributors to the RRs being less than 1.0 are:

- **Boiler controls savings calculation methods** – A majority of the boiler system controls estimates tended to be based on rules of thumb that resulted in high savings relative to evaluated findings. For boiler control measures in particular, the savings estimate was repeatedly based on a vendor claim that the installed controls equipment would save 35% of the baseline annual gas use (site L006).
- **VFD replacement** – Evaluators found instances where applicants received incentives to replace failed VFDs in systems that could not sustain operation without them. In some of the failed VFD projects, we found the simple payback to be below the threshold defined by the program guidelines, which represented an additional level of some these measures being ineligible from the financial perspective. Because there was no long-term alternative to a VFD-driven system, the baseline was a VFD system; hence the savings were evaluated as zero.
- **Administrative matters** – This category accounts for such issues as typographical errors, tracking savings reflecting the results from an early – not final – application revision, and extraction of the incorrect savings value from an analysis spreadsheet. While the errors had no pattern, their net effect was significant overestimation compared to the evaluated savings.
- **Baseline** – The evaluation team adjusted savings downward in instances where there was judged to be an end-of-life replacement instead of a retrofit. In such cases the baseline was new standard efficiency equipment rather than the preexisting equipment. The former tends to be more efficient than the latter, resulting in lower savings. An example of this situation was an EMS controls project that involved simple space temperature setbacks associated with a gut rehab of a space. By the definition of a gut rehab project, the evaluators considered the applicable energy codes to be the baseline for this project and not the current conditions or the conditions before the gut rehab.

The body of this report includes detailed discussions of the reasons for the discrepancies with examples and quantification of the effect of each factor, positive and negative. Some of the deviations are not uncommon for recently started programs, and the evaluation team would expect the RR to increase over time as these issues are addressed.

1.6.2 Equipment Performance Findings

Although the impact evaluation was not designed to provide specific measure results, some measures were sufficiently represented in the program to provide an indication of measure-level performance. For rebate measures, key NYTM parameters, such as hours of operation, were compared to the evaluated results.

Tables 1-8 and 1-9 summarize the results by measure for electric and gas measures, respectively.

Table 1-8. Equipment Performance – Electric Measures

Program	Measure Category	Count	Gross RR	Error Ratio
Electric Custom	EMS and controls	4	71%	0.06
	Lighting	29	88%	0.29
	Motors and VFDs	5	34%	0.72
	Other	10	49%	0.50
Electric Rebate	EMS and controls	13	50%	0.61
	HVAC	4	99%	0.12
	Lighting	26	77%	0.49
	Motors and VFDs	18	71%	0.48
	Other	2	64%	0.26

Table 1-9. Equipment Performance – Gas Measures

Program	Measure Category	Count	Gross RR	Error Ratio
Gas Custom	EMS and controls	1	101%	N/A
Gas Rebate	EMS and controls	12	39%	0.45
	HVAC	1	76%	N/A
	Other	2	150%	0.66
	Tune-up	9	49%	0.27

N/A – Not applicable. Error ratio is not relevant for a sample quantity of 1.

Of the measures with a significant number of observations (>5) the EMS and controls equipment group is the only measure group with a low RR (<60%). As noted above in the discrepancy section, the evaluation team perceives the underperformance of this measure subgroup to be related more to the savings estimation methodology than to equipment underperformance. The shell measures performed significantly worse than other gas measures and below average for electricity savings. The sample size is too small to treat the observation as other than anecdotal, but it fits a pattern that the evaluation team has observed in other NY State and New England utility program evaluations.

The custom estimated lighting hours tended to be unbiased and more accurate than the applicable NYTM hours, suggesting that the NYTM hours should only be used as a back-up for the lighting projects. Therefore, for this program, the evaluators agree with the NYTM guidance that self-report hours should be used when available. Interestingly, it should be noted that this finding was exactly opposite to the evaluation teams finding associated with the SBDI program where the applicant hours were found to be greater than the NYTM hours while the evaluated hours were found to be closer to the NYTM hours.

VFDs had high variability in RRs but overall, if not for the failure replacement issue described in the discrepancy section, would have had a strong 93% RR.

The peak kW estimates had a high degree of variability as evidenced in the estimate presented in the table above. This data highlights the need for the program staff to pay attention to peak demand savings calculations and to accurately capture these estimates in the tracking system.

1.7 Program Recommendations

The recommendations are divided into those relating to equipment, those relating to program administration, NYTM recommendations, and recommendations relating to the evaluation itself.

1.7.1 Specific Equipment Recommendations

The evaluation team recommends that the program administrators consider the following equipment-related changes.

Lighting

Revise the application documentation to require inclusion of the detailed space-by-space inventory of the fixtures affected in a given space. During the course of this evaluation, the investigating engineers had to rely on a supporting scope of work or site-provided documentation to understand the overall scope of the lighting projects. The supplied XACT project files did not convey details regarding the fixtures affected by space for a given project. If such details were incorporated into the project files, it would also enable the program staff to accurately conduct pre- and post-installation site inspections, which would ultimately help improve the overall accuracy of the tracking savings estimates that get entered into the LM Captures database. In addition, space-by-space inventories are not usually unduly burdensome for contractors to provide, since they are usually required for a cost estimate.

Electric – EMS and Controls

- On larger EMS projects (>250,000 kWh/yr or based on affected square footage or based on cost effectiveness to the program), collect additional data related to baseline operating conditions and conduct post-installation inspections to verify that the planned EMS strategies are implemented correctly. The engineering analysis should also be specific to the site and should account for the weather effects. The engineering analysis document should clearly document the baseline and proposed EMS strategies and the equipment that will be affected by the EMS. The baseline verification process should involve documenting the actual status of the current EMS strategies. The post installation inspection process should verify and document that the control strategies indicated in the initial scope of work are functioning. Preferably trend plots indicating conformance should be collected for future reference.

- For smaller sized EMS projects (<250,000 kWh/yr or based on affected square footage or based on cost effectiveness to the program) implemented on standard building types (offices, schools, etc.), we recommend developing a simple analysis tool for the program staff to estimate EMS savings based on the equipment affected, the planned EMS control strategies, and historic energy use. Such a tool could be a spreadsheet based model relying typical operating profiles that could be adjusted for the particular facilities energy use. It should have a menu based system to enable easy picking of the various EMS strategies that could then be used to estimate the savings for that particular project. There are no publicly available tools of this nature, but the evaluators have seen one in use developed by the Northeast Utilities for their use. An alternative to the above recommendation could involve using building energy simulation models to come up with savings fractions for typical buildings (offices, schools, hospitals etc.) and the most commonly applied control actions in these buildings.

Gas – EMS and Controls/Tune-Up

On boiler tune-up projects, apply actual measured pre- and post-tune-up combustion efficiency values along with appropriate annual heating hours in the savings calculations. In the absence of site-specific measured efficiency values, the evaluation team recommends using a 2% rather than a 5% savings factor. Two percent is the average efficiency improvement measured for those sampled tune-up projects where both pre- and post-tune-up data were available. This was not a focused study on boiler tune-ups; hence we did not have sufficient data points to suggest a change to the NYTM. In general, the 2% savings estimate is a better default savings estimate to use going forward, unless specific pre- and post-measured data is available. We also recommend crosschecking the savings with the billed usage to verify that the savings do not exceed the billed usage.

1.7.2 Program Administration Recommendations

This is an impact not a process evaluation, but some of the evaluation team's observations made while conducting M&V suggest opportunities for administrative improvement. Based on observations from the site reports, the evaluation team recommends the following to improve program administration:

- **Administrative errors** – Data entry and update errors in the tracking system had a substantial negative impact on the RR. We recommend that a combined Con Edison/Lockheed Martin team convene to examine the data management process. An example review process could involve the following:
  - Periodic crosschecking of Lockheed Martin and Con Edison tracking savings
  - Automated crosschecking of the modification date of the most recent XACT file with a last update field in tracking to ensure the most recent version of savings
  - Automated range checks, such as savings per fixture or savings per kWh per measure type, which might help identify misplaced decimals
- **Application review policy, procedures, and training** – The nature and extent of the sources of discrepancies suggest that the application reviewers would benefit from additional support and training that provides the policy background, procedures, and reference materials to enhance application review. Reviewers also need adequate time to review the application package and ask follow-up questions if necessary. The evaluation team particularly recommends additional reviewer training:
  - Baseline: Look out for new construction or major renovation projects for which the baseline is dictated by prevalent code, not the preexisting conditions.
  - Peak demand: Focus more on peak electric demand savings, as it is expected to gain more importance in the future.
  - Trending: We recommend that Con Edison have a process for identifying vendor trends, reviewing their savings calculation methods on a routine basis, and advising them on acceptable calculation methods. Some of these recommendations are further specified later in this report.
- **Establish measure review benchmarks** – It is clear that certain measures are underperforming. With a new program, there are fewer benchmarks by which to judge the performance of a measure; however, this evaluation has provided feedback on specific measures that should be incorporated into the application review process. Specific benchmarks (e.g., boiler controls, boiler tune-ups, and EMS savings fractions) are further specified later in this report. As an example, we found that typical boiler controls and

tune-up projects save approximately 2% of the baseline system operations instead of the 5% to 15% range claimed by the applicant in the current evaluation sample of projects. Providing such benchmarks to the reviewers and training them on watching out for these factors would help improve the overall accuracy of the savings estimates entered in the program database.

- **Reduce the effect of FR** – The FR estimates investigated through this effort were in the range of 0.23 to 0.38 and are within the typical range observed for similar LC&I programs offered throughout the country. However, further action could be taken to avoid future erosion of savings due to FR. These steps could involve the following:
  - Determining how customers are coming in to the program and not depending on "walk-ins"; seeking out participants rather than having them seek out the program.
  - Providing key technical assistance early in the project. Informing clients about the savings and rebate amounts *after* the project has begun does not seem to be having an effect on what is installed.

1.7.3 Recommendations for New York Technical Manual

For boiler tune-up projects, our metered sample suggests that a 2% savings estimate is a better default savings factor to use in the absence of metered data than the 5% used by Con Edison. Our study was not focused on boiler tune-ups so we do not have sufficient data to recommend a change to the NYTM. If the policy makers choose to create a new measure for commercial boiler tune-up, then the evaluators suggest using a 2% savings factor as a default and using our data along with data from other evaluations to develop a number for large commercial and industrial customers. The current NYTM reference refers to furnace tune-ups for residential customers. Due to the limited sample size, the evaluators do not recommend making a change in the current NYTM savings factor associated with residential tune-ups.

1.7.4 Program Evaluation Recommendations

Upon completing this evaluation, the evaluation team has a few recommendations for ways to improve future Large C&I evaluations.

- **Emphasize evaluation commitment** – During the initial on-site recruiting phase of this evaluation, some of the largest savings projects declined to participate in the M&V effort. The evaluation team requested assistance from the program staff and were eventually informed that these customers were high value and hence could not comply with the follow-up evaluation M&V due to a variety of reasons. Even though the overall evaluation RP did not suffer drastically, the loss of these sites was regrettable and contributed to the reduction in the overall RP values of this evaluation. Therefore, the evaluation team would like to stress that the program staff may want to add language to agreements that encourages participation in evaluations.

- "**Continuous evaluation approach**" – Having evaluation activities (and more specifically, FR research) occur shortly after the implementation of the project can greatly increase the accuracy of the results and reduce nonresponse bias. Asking complex and inherently hard-to-answer counterfactual NTG questions several years after the completion of a project can greatly reduce respondent ability to make the needed estimates and consequently diminish the accuracy of the estimates. Decision-makers might not clearly remember what the influence of the program was on their decision to implement high efficiency improvements installed a while ago. Furthermore, over time, decision-makers might no longer be available for interviewing (e.g., they could have left the company, assumed a different position within the organization, etc.), which might result in nonresponse error. It is our understanding that CECONY is currently in the process of making the transition to the "continuous evaluation approach."

The CECONY programs have come a long way in a short amount of time. The kinds of issues highlighted in this section are indicative of a new program going through its growing pains. Addressing the administrative and procedural types of recommendations mentioned in this section should help improve the overall performance of the CECONY programs significantly and could be implemented in a fairly short order.

2 INTRODUCTION

Consolidated Edison Company of New York (CECONY) and Orange & Rockland Utilities (O&R), collectively "the Companies," have completed the delivery of the first cycle (2009 – 2011) of a portfolio of Energy Efficiency Portfolio Standard (EEPS) utility administered programs, as ordered by the New York Public Service Commission (PSC). This document presents a detailed impact evaluation of the four CECONY Large Commercial and Industrial (Large C&I) Electric and Gas Rebate and Custom programs. The results of the O&R Existing Buildings program are presented in a separate report.

2.1 Program Background and Objectives

CECONY presents the four DPS-ordered programs to the market together as the CECONY Large C&I Equipment Rebate program. The group of programs promotes the purchase and installation of specific high-efficiency equipment by C&I customers in existing facilities. End-of-life equipment replacements are also incentivized. These programs provide customers with information on the features and benefits of energy efficient equipment along with financial incentives to offset the higher purchase cost of that equipment (prescriptive rebates). Qualifying equipment includes electric cooling, ventilation, motors, and lighting. CECONY customers interested in installing high efficiency equipment not specified in its Equipment Rebate program may participate in the utility's Custom Efficiency program and receive a corresponding custom rebate.

CECONY uses a combination of dedicated internal staff and third-party implementation contractors to market, manage, and administer the program and uses market partners for measure installation. CECONY has contracted Lockheed Martin to administer the program in its service territory.

Rebate measure savings are calculated using a spreadsheet tool called XACT, which has embedded in it the assumptions and calculation methods of the New York Technical Reference Manual (NYTM). It also has a tab for capturing custom measure information for screening purposes. The tool is used for both electric and gas measure savings assessments.

Rebate incentives are designed to pay up to 70% of either the measure cost or the incremental measure cost (difference in cost between the efficient measure chosen and the baseline). Tiered incentives are offered under the Gas and Electric Custom programs for those customers exceeding certain efficiency thresholds.

The four CECONY programs processed applications for over 950 unique accounts, and many accounts participated in multiple programs offered by CECONY. Lighting, motors and VFDs, and EMS/controls measures represented over 87% of the total savings attributed to the electric programs offered by CECONY while EMS/controls and tune-up measures represent 93% of the total savings attributed to CECONY's gas programs.

2.2 Evaluation Objectives

The intent of this group impact evaluation of the four CECONY Large C&I Gas and Electric programs is to provide the following:

- A general assessment of each program's performance in total during the 2009–2011 period
- A focused and more robust assessment of select sites based on primary data collection, including telephone surveys, customer bills, and on-site measurement and verification (M&V).
- Actionable recommendations for improving the program's implementation as a result of these assessments

The report includes estimates of gross and net impacts (kWh and kW) of all the program participants and measures. The overall evaluation scope and objectives are identified in the Table 2-1.

Table 2-1. Evaluation Scope and Objectives

Objective	Definition
Evaluation scope	Assessment of the four CECONY Large C&I Electric and Gas programs and the O&R Existing Buildings program.Primary data collection activities will be focused on a statistically representative sample of sites, where all program measures that were installed at the site will be assessed.
Gross energy impacts	Report annualized first-year gross energy savings based on electric (kWh) and natural gas savings (therms) at the customer meter (gross savings) using on-site logging and custom engineering assessments. Applicable results will be weather normalized to a typical year using typical meteorological year (TMY3) weather data.
Gross demand impacts	Report the electrical demand impact at the customer meter; electrical demand impact is defined as the energy reduction during the hottest non-holiday summer (June through August) weekday during the hour from 4 p.m. to 5 p.m. Report the gas demand impact at the customer meter; gas demand impact defined as the therms reduction during a 24-hour period starting at 10 a.m. in which the average temperature for the day is -9ºF for both upstate and downstate.The definitions are consistent with the NYTM requirements.
Program attribution	Estimate free ridership (FR) and participant spillover (SO) using self-reported responses from telephone surveys. Trade ally surveys will be conducted to verify the FR responses.
Precision	The sample designs will target 10% precision at the 90% confidence level for a program energy savings realization rate (RR) and for the net-to-gross (NTG) factor as directed by the DPS Evaluation Guidelines for each program administrator. Subsector precisions will be less precise.
NYTM factors	Expectations are noted in Tables 4-1 and 4-2 in Section 4.

3 Evaluation Methodology

This evaluation plan complies with the requirements of the August 7, 2008 (updated February 2012) Evaluation Guidelines issued by the DPS[3] and is intended to provide robust, timely, and transparent results. The impact methods are aligned with the guidelines of the National Action Plan for Energy Efficiency Model's Energy Efficiency Program Impact Evaluation Guide (NAPEE Guide).

Realization rates (RR) were developed from on-site measurement and verification (M&V) of a statistically representative sample of sites. Post-stratification analysis was used to examine results by measure, particularly for variable frequency drive (VFD), lighting, and/or energy management system (EMS) measures.

This report also includes a study of program attribution, including both free ridership (FR) and spillover (SO). Participant FR and SO were measured through surveys with program participants. The evaluation team also conducted interviews with trade allies to further validate FR as well as explore the occurrence of nonparticipant spillover (NPSO). The potential impact of NPSO was examined to determine if future investigation is warranted, but NPSO **was not used** as a component of the net-to-gross (NTG) results. The product of the attribution assessment includes estimated and statistically valid NTG factors.

While subject to a common plan and methodology, the evaluation team conducted a separate analysis of each utility company incorporating individual program results. The evaluation methodology is presented in the following section.

3.1 Gross Savings Evaluations Methods

The impact evaluations of the four CECONY Large C&I Electric and Gas Rebate and Custom programs are organized into a single group effort to increase efficiency and to capture synergies between the evaluation efforts. The groups consist of four programs serving the same sector, with similar measures, evaluation approaches, and timing.

Gross savings were estimated for each selected sample site through inspection and verification of conditions and equipment at the site and measurement of one or more key parameters characterizing the savings. Analysis was then performed to extrapolate the findings to a year resulting in a typical first-year savings estimate.

[3] Including subsequent documents when finalized: "Spillover DPS Guidance 7-11-12," "Appendix A: Sampling DPS Guidance 7-11-12," and "Appendix B: Reporting and Accessing DPS Guidance 7-11-12."

3.1.1 Baseline Characterization

Evaluators considered the appropriate baseline for each custom and prescriptive measure. The baseline was defined according to one of three conditions:

1. New construction baseline (market opportunity, normal replacement, incremental, replace on failure)
2. Retrofit baseline
3. Dual baseline

New Construction Baseline

The new construction baseline was applied to any new construction, gut renovation, or process expansion and to any measure for which the removed equipment exceeded its effective useful life (EUL).[4] For such measures, the applicable code, in force at the time of the application approval, was used to define baseline efficiency.

New York City has unique considerations regarding the triggering of code as opposed to the "grandfathering" of preexisting conditions. In many jurisdictions including New York City, prior to July 1, 2010, renovation projects were only required to comply with the code in effect at the time of permit application if the project substantially affected more than 50% of the building space. The current New York City Energy Conservation Code (NYCECC) now requires buildings "to meet the most current energy code for any renovation or alteration project."[5] Historic buildings and landmarks are allowed exceptions. If no code applies, the baseline will be defined by the least efficient equipment commonly used in the market.[6]

---

[4] This interpretation complies with the NYTM "Appendix M: Guidelines for Early Replacement Conditions," page 5.
http://www3.dps.ny.gov/W/PSCWeb.nsf/96f0fec0b45a3c6485257688006a701a/06f2fee55575bd8a852576e4006f9af7/$FILE/Appendix%20M%20final%205-05-2011.pdf
It is the default option. The appendix puts the burden of proof on the program administrator to substantiate that the age of the replaced equipment is less than its EUL. If not documented the measure is considered new construction.

[5] Local Law 85 is part of a series of four laws collectively referred to as the NYCECC, and it includes this policy. Quoted text from the NYC Green Buildings & Energy Efficiency website, http://www.nyc.gov/html/gbee/html/plan/ll85.shtml, 7/3/2012.

[6] "Standard practice" and "least efficient commonly used" typically are synonymous terms. On rare occasions a material minority of the market may employ a system less efficient than that considered "standard." In such instances the lower efficiency is the baseline.

Retrofit Baseline

This scenario was most traditionally considered for non-new construction efficiency programs. For such measures the first-year savings was calculated by comparing the previously installed equipment energy use with that which was installed, and the savings were projected for the entire measure EUL.

Dual Baseline

The dual baseline concept was applied to measures that replaced older inefficient equipment relatively early in their expected life.[7] Evaluators considered the dual baseline applicability for LC&I measures and did not find any instance where it applied. The evaluation team followed the general and special circumstance protocols described in Appendix M to analyze savings for dual baseline measures. Further details regarding the dual baseline process are described in detail in the evaluation work plan.

3.1.2 Approach to Sampling and Data Collection

The evaluation team used a combination of on-site metering and attribution phone surveys on sampled sites to estimate the net program impacts. The on-site metering effort was used to determine the gross realized energy savings and peak demand savings for a representative sample of program participants. The attribution surveys were conducted to develop the FR and SO estimates, which were then used to determine the net-to-gross ratio (NTGR).

Sample Design Approach – Gross Impact Assessment

The sampling approach required both telephone and on-site M&V samples for each of the four programs as well as a within-site sampling protocol for selecting equipment for logger placement. The on-site M&V samples were selected from customers' projects that committed savings to the program during 2010 or 2011 and were acquired by May 23, 2012.

The sample designs reflect the regulatory requirement for results with ±10% precision at 90% confidence by program. As such, the sampling unit is each unique combination of site and program. While results for key measure types within programs are also of interest and reported on, they are not captured explicitly in the proposed designs, since doing so would increase the sample sizes significantly.

The on-site M&V selections were based on a model-based statistical sampling (MBSS) method to support the stratified ratio estimation of results. This method ensures that sample sites are

[7] Defined in CASE 07-M-0548 – Proceeding on Motion of the Commission Regarding an Energy Efficiency Portfolio Standard, *Order Approving Consolidation and Revision of Technical Manuals*, issued and effective October 18, 2010, page 12, and referenced in the subsequent Appendix M of the NYTM,

allocated efficiently across the population of program participants in order to target precise estimates of M&V parameters such as RRs and NTG factors. It involved the assignment of participants to segments of interest and then to size strata based upon gross savings (kWh and therms). For M&V, sample sizes required to meet precision and accuracy goals are calculated based on the expected variance of the evaluated savings to the tracking savings (the error ratio). A precision of ±10% at 90% confidence or better was targeted for overall program energy savings. A brief summary of the MBSS methodology is provided in Appendix A.

While the MBSS technique is applied to design and select both the on-site and telephone survey samples, the precision targets, segmentation categories, required sample sizes, and selection criteria are independent. A more detailed description of the sampling approach along with details on the targeted and achieved samples can be found in the Sample Design and Final Sample Disposition section of Appendix A. The sampled on-site strata are shown in Table 3-1. Table 3-1 does not include the counts of the very small saver sites that were excluded from the on-site sample population as their contribution to the overall program savings was minimal. Their impacts are, however, included in the final program reported savings in the executive summary.

Table 3-1. Sampled On-Site Strata

Program	Stratum	Program Population	Planned On-Site Sample	Achieved On-Site Sample
Electric Custom	1	101	9	9
Electric Custom	2	38	9	9
Electric Custom	3	21	9	9
Electric Custom	4	14	9	8
Electric Custom	5	10	8	9
Electric Custom	6	4	4	3
Electric Rebate	1	258	10	10
Electric Rebate	2	101	10	8
Electric Rebate	3	57	10	8
Electric Rebate	4	28	10	12
Electric Rebate	5	15	9	10
Electric Rebate	6	9	9	4
Gas Custom	1	1	1	1
Gas Custom	2	2	2	0
Gas Rebate	1	21	4	6
Gas Rebate	2	8	4	8
Gas Rebate	3	5	4	3
Gas Rebate	4	5	4	3
Gas Rebate	5	6	6	2
Total		**704**	**131**	**122**

Sample Design Approach – Net Impact Assessment

Net impacts of CECONY's Large C&I program were estimated using a survey with program participants and interviews with participating trade allies. The goal of the telephone survey with program participants was to collect data that would allow the team to estimate program attribution. The goal of the sample design was to achieve a target of 90% confidence with 10% precision on program impacts per the evaluation guidelines.[8] The analysis of the program tracking data identified 591 unique contacts across 985 unique projects for CECONY. Small participant population required a census attempt across all programs and all program participants. The evaluation team completed 74 interviews. In addition, there were three respondents who terminated mid-interview whose responses the evaluation team could nevertheless use in the analysis. This brought the total number of completed interviews to 77.

Following data collection efforts, the evaluation team post-stratified the frame and completed sample with the same stratification method used for the on-site sample to develop sample weights for expansion of results.

Table 3-2 provides a breakdown of the unique contacts and projects in the population, the total number of projects for which the evaluation team completed interviews, and the total number of contacts with whom the evaluation team completed interviews.

Table 3-2. Overview of Projects in Net Impact Assessment Population (Projects Completed as of the End of 2011)

Program Type	Total Number of Unique Projects	Total Number of Unique Contacts[2]	Total Number of Unique Projects for Which Interviews Were Completed	Total Number of Completed Interviews[3]
Electric Rebate	759	433	87	51
Gas Rebate	87	69	3	3
Electric Custom	265	211	23	23
Gas Custom	5	5	0	0
Total[1]	**985**	**591**	**113**	**77**

[1] Note that since a single project can include measures rebated through different programs, the total number of unique projects (in the population as well as for which interviews were completed) is lower than the sum of unique projects by program. The same is true for unique contacts and the total number of unique completed interviews.

[2] Note that the number of unique contacts does not match what appears in the disposition reports due to different/additional contact information that is uncovered during the interviewing process.

[3] Includes three mid-interview terminate completes.

[8] August 7, 2008 (updated November 2012) Evaluation Guidelines issued by the DPS through the NYS Evaluation Advisory Group (EAG), http://www3.dps.ny.gov/W/PSCWeb.nsf/96f0fec0b45a3c6485257688006a701a/766a83dce56eca35852576da006d79a7/$FILE/EVALGUIDE.11.12.pdf

The Large C&I programs use various tactics to influence the way trade allies (equipment vendors, installation contractors, designers, engineers, etc.) specify projects to customers, and C&I customers can be heavily influenced by their trade allies in the decision to install high efficiency equipment. For these reasons the evaluation team completed interviews with trade allies to capture and integrate the influence of the program on trade allies into the final NTGR.

The trade ally interviews were tied to participant responses to the telephone survey – the evaluation team only completed interviews with trade allies involved in projects where the participant rated the influence of the trade ally on the recommendation of high efficiency equipment as high. From the participant survey, the evaluation team arrived at a sample frame of twenty-three trade allies. The evaluation team made multiple attempts to contact every trade ally on the list and were able to complete nine interviews.

Appendix A of this report provides a detailed sample design description and disposition reports for the participant survey and trade ally interviews.

Data Collection Methods Overview

For the general data collection effort the evaluation team conducted the following:

1. Rigorous on-site metering at 122 sampled CECONY sites to develop the ex post energy and peak demand savings estimates
2. A phone survey to determine the FR and SO associated with the program participants

The data collection process is highlighted in the following sections.

On-Site Measurement and Verification Approach

The 122 sites included in the CECONY on-site metering sample were metered from August 2012 through April 2013. The engineering team deployed over 1,000 loggers during the M&V process, leaving most in place for at least a month to collect data over varying load conditions. At each site, savings were estimated for each selected sample site through inspection and verification of conditions and equipment at the site and measurement of one or more key parameters characterizing the savings.

For each site, the evaluation team developed an M&V report tailored to the specifics of the project. The evaluation team applied the International Performance Measurement & Verification Protocols (IPMVP) options A, B, C, or D depending on the specific situation at the site. Lighting and a select list of prescriptive measures typically fell under the IPMVP option A (partially measured retrofit isolation) classification. For custom projects (compressed air, heat recovery etc.), the IPMVP option B (retrofit isolation) approach was applied. On a few projects, IPMVP option D (whole building simulation) was also used. Weather normalized billing analysis (IPMVP option C) was applied on a number of tune-up and boiler/furnace replacement projects.

Details about the meters used on-site can be found in the Metering Equipment Details section of Appendix B. Appendix C presents the site-specific M&V reports.

3.1.3 Approach to Gross Data Analysis

The analysis for this evaluation included conducting on-site metering at the sampled sites of one or more key parameters affecting the measure savings. Analysis was then performed to extrapolate the findings to a year. The typical steps involved in this process were-

1. Review project files.
2. Review the baseline.
3. Recruit the site.
4. Develop site-specific M&V plan.
5. Conduct site visit – Deploy loggers and verify equipment counts, nameplate efficiencies, and operating hours. For sites with numerous buildings or pieces of equipment, develop sample plan to monitor specific buildings or equipment.
6. Retrieve loggers after pre-determined time period (minimum 4-week deployment).
7. On specific sites request pre- and post-billing data from the program staff.
8. Analyze logger/billing data – Develop correlations with outside air temperature, time, occupancy, or other relevant variables that will be used to extrapolate the results to the rest of the year.
9. For measures affected by weather, normalize for weather by applying the nearest TMY3 weather data and extrapolate the base-case and post-case energy calculations for an entire year.
10. Develop NYTM-specified peak demand savings.
11. Convert the site-specific M&V plan into a report.

3.2 Potential Gross Impact Error and Bias

The following sources of error discussed below were considered over the course of this impact study, and precautions were taken to mitigate or address them.

3.2.1 Sample Bias

The evaluation team worked closely with CECONY staff to identify a data set of record for sampling and computing purposes to minimize frame errors. The program savings from this detailed record were reconciled with the scorecard information to ensure a correct, complete, and consistent sample frame. The sample was drawn randomly, which reduced the potential for sample bias.

3.2.2 Nonresponse Bias

Nonresponse bias is always an issue when conducting surveys of voluntary participants. The evaluation team employed industry standard techniques for mitigating the impact of nonresponse bias. These included stratifying the sample, sending advance letters, making site recruitment calls at varying times of day and evening, calling sampled participants at least seven times before removing them from consideration, and offering alternative time slots for on-site data collection. CECONY program staff assistance was also used for recruitment of critical sites.

3.2.3 Measurement Error

The first step taken to address measurement error was to develop rigorous field protocols and internal oversight. Rigorous field training was conducted in order to ensure that field staff were collecting data and recording it in an accurate and consistent manner. In the final data analysis, rigorous quality control was performed on all spot measurements and metered data. Data that was out of range or that had some other fault was removed from the final savings analysis.

Care was taken in the selection and use of spot measurement and metering equipment. Only high accuracy spot measurement equipment was used, and devices were checked against a reference at the beginning of the study and at least weekly after that point. In addition, only metering equipment that met rigorous sensor and time stamp accuracy was used.

3.2.4 Modeler and Baseline Error

For each sampled site the metered data was used to calculate the annual energy savings. Potential errors might arise from the way the metered data was used to model the annual savings.

The period of metering of the key parameter of interest typically fell between 4 and 8 weeks. Regression or other techniques were utilized to correlate the key parameter to an independent variable such as outdoor weather conditions, time of day or day of week, or level of production. Using this information the short-term data was extrapolated to the entire year to estimate the annual energy savings and to create peak-day hourly profiles.

The error associated with this process is minimized by collecting sufficient data to span the expected variation in the independent variable. For example, for heating measures, it was important to measure performance in the winter over a span of outdoor temperatures from swing season through much colder weather.

While the evaluator can directly observe the installed measure performance, the baseline cannot be directly observed. Errors associated with baselines were mitigated through selection of the correct baseline standard, measurement of existing conditions, and customer interviews.

The correct baseline standard can be applied either from an existing condition or measure or from a code that is currently in place. When the pre-installation usage was the baseline, the evaluator ascertained that baseline through customer interviews about the prior operation and measurement of installed conditions. Typically, this usage was calculated as a function of the reported prior equipment efficiency (for example, lighting fixture wattage or boiler efficiency) and the measured operating profile of the installed equipment.

3.3 Attribution

Program attribution accounts for the portion of the gross energy savings associated with a program-supported measure or behavior change that would not have been realized in the absence of the program. The program-induced savings, indicated as an NTGR, is made up of FR and SO and is calculated as $(1 - FR + SO)$. FR is the portion of the program-achieved verified gross savings that would have been realized absent the program and its interventions. SO is generally classified into participant and nonparticipant spillover. Participant spillover (PSO) occurs when participants are influenced by the program interventions to take additional energy-saving actions but did not receive program support. Nonparticipant spillover (NPSO) is the reduction in energy consumption and/or demand by nonparticipants due to the influence of the program.

As part of this evaluation, the evaluation team focused on the estimation of FR and participant SO. Through interviews with trade allies, the evaluation team also explored the presence of NPSO and whether additional research was justified to accurately quantify NPSO. Quantifying savings from NPSO activities is a challenging task that warrants a separate study and is outside of the scope of this evaluation effort.

Below is a general overview of the method for developing FR and PSO results. Appendix D of this report contains further detail on the NTG estimation method.

3.3.1 Free Ridership

Free riders are program participants who would have implemented the incented energy efficient measure(s) even without the program. In other words, FR represents the percentage of savings that would have been achieved in the absence of the program. The FR component of the NTGR was derived from self-reported information from telephone interviews with program participants and further adjusted through the interviews with participating trade allies.

Initial Free Ridership Estimation through Participant Survey

Using the survey instrument developed for this evaluation, program participants were interviewed and asked a series of structured and open-ended questions about the influence of the program and its various components on the decision to have high efficiency equipment installed at their business. More specifically, program participants were asked about any preexisting plans to implement the program measure(s), influence of the various program components, such as marketing, incentives, technical assistance, on the decision to install high efficiency equipment and the likelihood of them taking the same action absent the program. Program participants were also asked about the program influence on the timing or the scope of the high efficiency project. Appendix E presents the details associated with the participant survey instrument.

The goal of most incentive-based energy efficiency programs is to influence customer decision-making regarding energy efficient improvements. Programs can do this by changing *what* customers install, *when* they install it, and *how much* they install. In other words, programs influence the *efficiency, timing,* and *quantity* of customers' energy-using equipment installations.

The bulk of program savings is typically achieved by encouraging customers to install higher-efficiency equipment than they would have installed on their own. Programs may also encourage early replacement of still-functioning equipment that is less efficient, thus impacting the timing of the installation so that savings can be realized earlier. The incentive may also make it more affordable for customers to install a greater number of high efficiency measures.

The FR algorithm outlined here combines estimates of each of these concepts:

- Program influence on the efficiency level of the installed equipment (FRE)
- Program influence on the timing of the installation of high efficiency equipment (FRT)
- Program influence on the quantity of the high-efficiency equipment installed (FRQ)

To arrive at the FRE estimate, the evaluation team asked respondents about the influence of the following program components on the decision to install high efficiency equipment:

- Program incentives

- ❑ Information and recommendations provided as the result of the energy study
- ❑ Interactions with program staff, including implementation partner staff
- ❑ Utility account executive endorsement
- ❑ Program marketing and outreach
- ❑ Previous experience with the program

We also asked participants to estimate the likelihood of completing a high efficiency project if the program had not been available.

To arrive at the FRT estimate, the evaluation team asked participants whether the timing of the project would have been the same or if the project would have been completed later, and if later, the evaluation team asked them to estimate how much later.

Finally, to arrive at the FRQ estimate, the evaluation team asked participants whether the program influenced the scope and, if so, what percent of the scope would have happened even in the absence of the program.

Each concept took a value between 0 and 1. The values were expressed in FR terms, with 0 meaning no FR and 1 meaning full FR. To calculate an overall estimate of program influence, the algorithm first multiplied the estimates of efficiency (FRE) and quantity (FRQ); then averaged the resulting estimate with the estimate of timing (FRT), but *only* in cases where the FRT value was lower than the product of FRE and FRQ ($FRE \times FRQ$). In cases where the FRT value was higher, the timing component was not included as part of the algorithm. This was done in order not to penalize the program for influencing the timing of the high efficiency project to a lesser degree than influencing the efficiency and quantity of the project. Furthermore, since the concepts of timing and quantity are conditional on efficiency (or the probability of the high efficiency installation taking place), the FRT and FRQ elements were only incorporated in the FR estimate when the FRE component was 0.5 or higher (50% or higher probability of a high efficiency installation taking place).

Thus, the formula to calculate FR can be expressed as:

$$If\ FRE < 0.5, FR = FRE$$

$$If\ FRE \geq 0.5\ and\ FRT < (FRE \times FRQ), FR = Average((FRE \times FRQ);\ FRT)$$

$$All\ other\ cases$$

$$FR = FRE \times FRQ$$

Because respondents can sometimes give inconsistent answers, the survey instrument included consistency checks to clarify these responses. As part of the data analysis, the evaluation team carefully studied those responses and adjusted either FR scores or individual component scores accordingly.

Respondents can sometimes provide "don't know" responses to one or more questions that are critical to FR estimation. Sometimes, respondents can simply refuse to answer some questions. This leads to item nonresponse. To overcome any biases associated with item nonresponse, the evaluation team inputted data based on the responses that other similar participants gave.

Free Ridership Adjustment through Interviews with Trade Allies

It is our understanding that the Large C&I programs used various tactics to influence the way trade allies (equipment vendors, installation contractors, designers, engineers, etc.) specify projects to customers (including marketing and outreach to trade allies, trade ally training, etc.). As such, there was a concern that assessment of FR solely through participant research would not credit the program for influences not visible to customers, primarily program outreach to trade allies. To address this concern, the final FR estimates incorporated the results of trade ally interviews, in addition to the research with program participants.

The interviews were completed with a subset of participating trade allies who, based on the results of the participant survey, had influenced end-user decisions to install program rebated high efficiency equipment.

The interviews with trade allies were project-specific; that is, they explored trade ally involvement with each individual project in question as well as program influence on trade ally recommendations pertaining to each individual project. The interviews were in-depth, which allowed for deeper exploration and understanding of the various factors that influenced trade ally recommendations.

Using the results of the interviews, the evaluation team adjusted, where appropriate, participant-derived individual project FR rates downward to account for indirect program influence on customer decisions through trade allies.

3.3.2 Spillover

PSO represents additional savings (expressed as a percentage of total program savings) that were achieved without program rebates but would not have happened in the absence of the program. PSO was assessed through interviews with participating customers by asking about non-program efficiency actions that participants took as a result of participating in the program. The actions could have taken place at the same facility that received the program-funded upgrades or at another site. SO questions covered both program-like measures that were installed without program rebates as well as energy efficiency measures not offered as part of the program design. The survey instrument contained checks to ensure consistency of response.

While PSO can result from a variety of measures, survey length did not allow for estimation of PSO across all possible measures or scenarios. Given the types of businesses that participated in the program, the evaluation team included measures that experiences with the program could reasonably influence. As such, PSO was measured for the following categories:

- Lighting equipment
- Cooling equipment
- Refrigeration equipment
- Kitchen equipment
- Motors
- Heating and water-heating equipment
- Other

Participants were asked if they made any of the above-listed improvements but did not receive incentives for them through the program. Those who did were asked if the program was of any influence and, if so, what the degree of influence was. Respondents were also asked to explain in their own words exactly how the program influenced their decision to make specific additional improvements.

Respondents were also asked a few equipment-specific questions that allowed the calculation of savings associated with the installed equipment. The equipment details were limited by the survey length as well as by what the evaluation team believed respondents could reliably answer.

As part of the SO calculation, the evaluation team applied savings values to the measures installed outside of the program. We estimated savings for each measure using the most recent NYTM values supplemented by engineering assumptions. We determined the program-level SO factor by dividing the estimated savings of the measures installed by survey respondents outside of the program (but influenced by the program) by the savings the survey respondents realized through the program.

$$\text{PSO} = \frac{\textit{Respondent energy savings from measures installed outside the program}}{\textit{Respondent energy savings from measures installed through the program}}$$

Neither NPSO nor vendor off-site SO was in the scope of this evaluation. NPSO is expected to be evaluated in a separate statewide evaluation in 2014.

3.3.3 Trade Ally Interviews

Because the Large C&I programs use various strategies to influence the way trade allies specify projects to customers, and because C&I customers can be heavily influenced by their trade allies in the decision to install high efficiency equipment, it was important to capture and integrate the influence of the program on trade allies into the final NTGR.

To do that, the evaluation team completed interviews with trade allies. The interviews were triggered by participant responses to the telephone survey – the evaluation team only completed interviews with trade allies that were involved in projects where participant-rated influence of the trade ally on the recommendation of high efficiency equipment was high. At the completion of the participant survey, the evaluation team arrived at the sample frame of twenty-three trade allies. We tried to contact every trade ally from the list multiple times and were able to complete nine interviews. Appendix F presents details associated with the trade ally interview process.

3.3.4 Potential Attribution Error and Bias

The sources of error discussed below were considered over the course of this impact study, and precautions were taken to mitigate or address them.

Sample Bias

Sample bias was not an issue for this study, because the evaluation team attempted a census of all program participants.

Nonresponse Bias

Nonresponse bias is always an issue when conducting surveys of voluntary participants. The evaluation team employed industry standard techniques for mitigating the impact of nonresponse. These included making phone survey calls at varying times of day and evening, calling sampled participants at least twelve times before removing them from consideration, offering cash incentives for participation in the survey, and offering alternative time slots for on-site data collection, such as evenings and weekends.

Measurement Error

To mitigate measurement error, the evaluation team employed thorough testing of the survey instrument and pretesting of the survey with a small sample of participants to ensure that all questions were understood as intended, as well as thorough interviewer training. It should be noted that none of these measurement uncertainties are believed to have inherent bias that will affect the RR estimate in a particular direction.

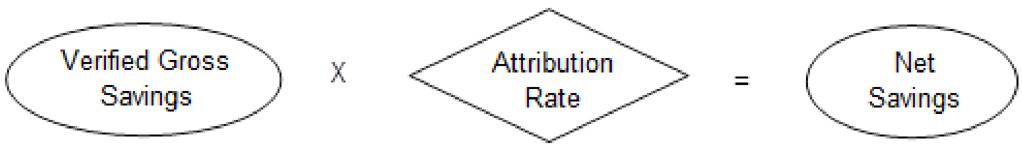
3.4 Program Net Impact Analysis

This task involved combining the site-specific results with the attribution survey results to develop the final program level gross and net impacts.

This analysis involved the following general steps -

- The NTGR factors were calculated using the individual customer attribution survey results.
- The individual site-specific gross RR results were used to determine the verified gross savings.
- The evaluation team used the overall attribution rate and applied it to the verified gross savings for the program to calculate the program's net savings, as shown in Figure 3-1.

Figure 3-1 Net Savings Algorithm



4 RESULTS

The gross and net results of the program evaluation are shown in the sections below.

4.1 Gross Savings Results

The following section presents the program-level savings results for the four CECONY programs.

4.1.1 Electric Custom Program-Level Savings Results

Table 4-1 presents the gross evaluation results for the Electric Custom program. The results for the energy portion exceeded the evaluation goal of 90/10 confidence and precision requirements of the stipulated state guidelines. The relative precision (RP) estimate for the peak demand was slightly above the required guideline but is still fairly close to the desired target. The higher RP for peak kW is indicative of the high variability found in the accuracy of the program tracking peak kW estimates.

Table 4-1. Electric Custom Program – Gross Energy and Peak Demand Impacts

	Total No. of Sites	Sampled No. of Sites	Tracking Gross Savings	Realization Rate	Evaluated Gross Energy Savings	RP	Error Ratio
Energy (kWh/yr)	232	47	27,132,715	72%	19,535,555	9%	0.38
Peak demand (kW)			3,512	92%	3,231	14%	0.55

Table 4-2 presents a summary of the interactive energy impacts associated with the electric measures for the Electric Custom program. The interactive impacts are largely attributed to the lighting projects which typically involve a heating penalty. This table provides the interactive energy associated with the sampled projects. This information could be used by the program planners to estimate the non-electric impacts associated with their Electric Custom program.

Table 4-2. Electric Custom Program - Interactive Impacts

Gross Tracking Savings (kWh/yr)	Interactive Evaluated Natural Gas Savings (Therms/yr)	Interactive Savings Factor (Therms/kWh)
27,132,715	-124,195	-0.0046

Figure 4-1 presents a plot of the gross evaluated energy savings versus the tracking energy savings associated with the Electric Custom program. The diagonal line in the plot depicts the 100% realization rate (RR) line. Therefore, the points that fall below the line represent projects that underperformed while the points that appear above the line over-performed compared to the tracking savings estimates, and the points appearing on the line represent the projects that achieved a 100% RR.

Figure 4-1. Electric Custom Program – Gross Evaluated Energy vs. Tracking Energy

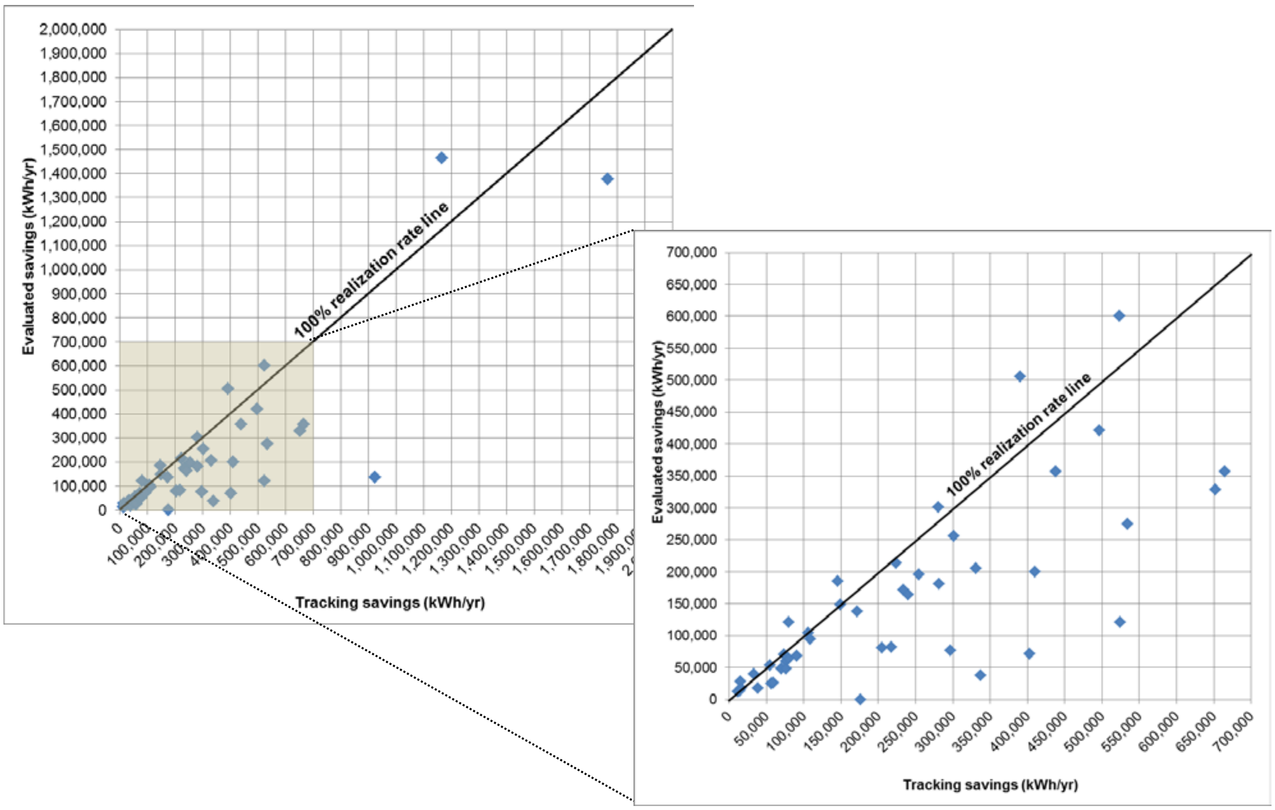


Figure 4-2 presents a plot of the gross evaluated peak demand savings versus tracking peak demand savings.

Figure 4-2. Electric Custom Program – Gross Evaluated Peak Demand vs. Tracking Peak Demand

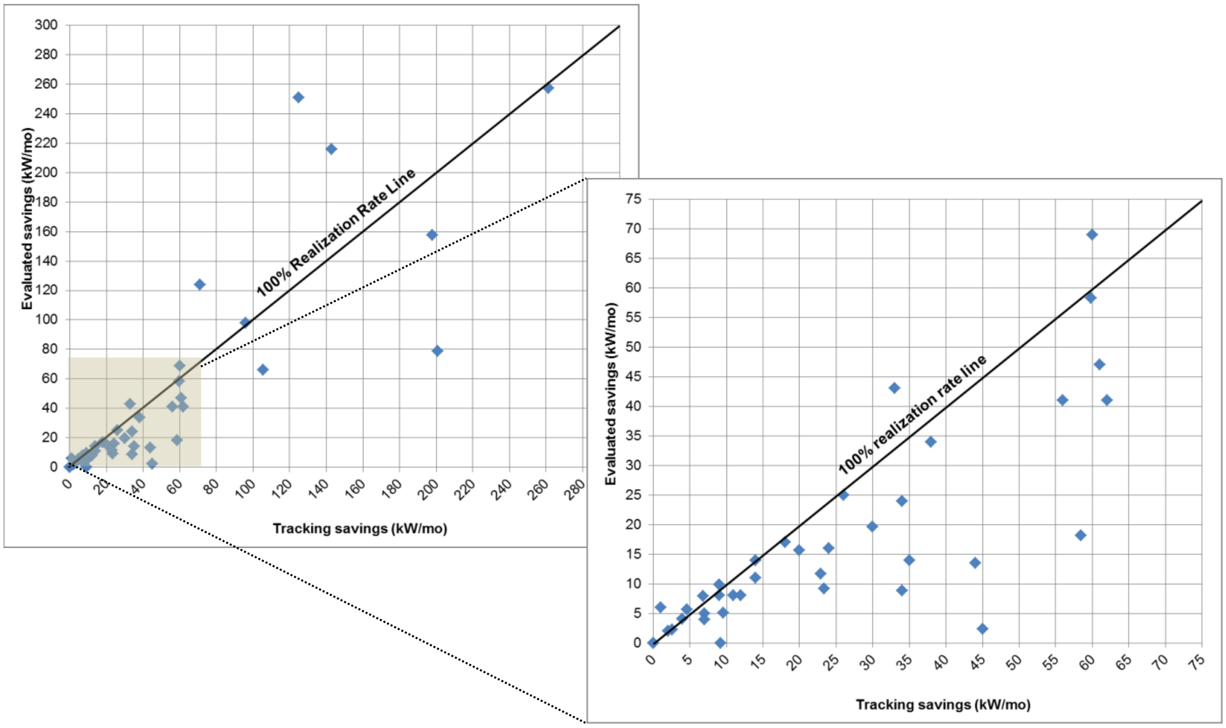
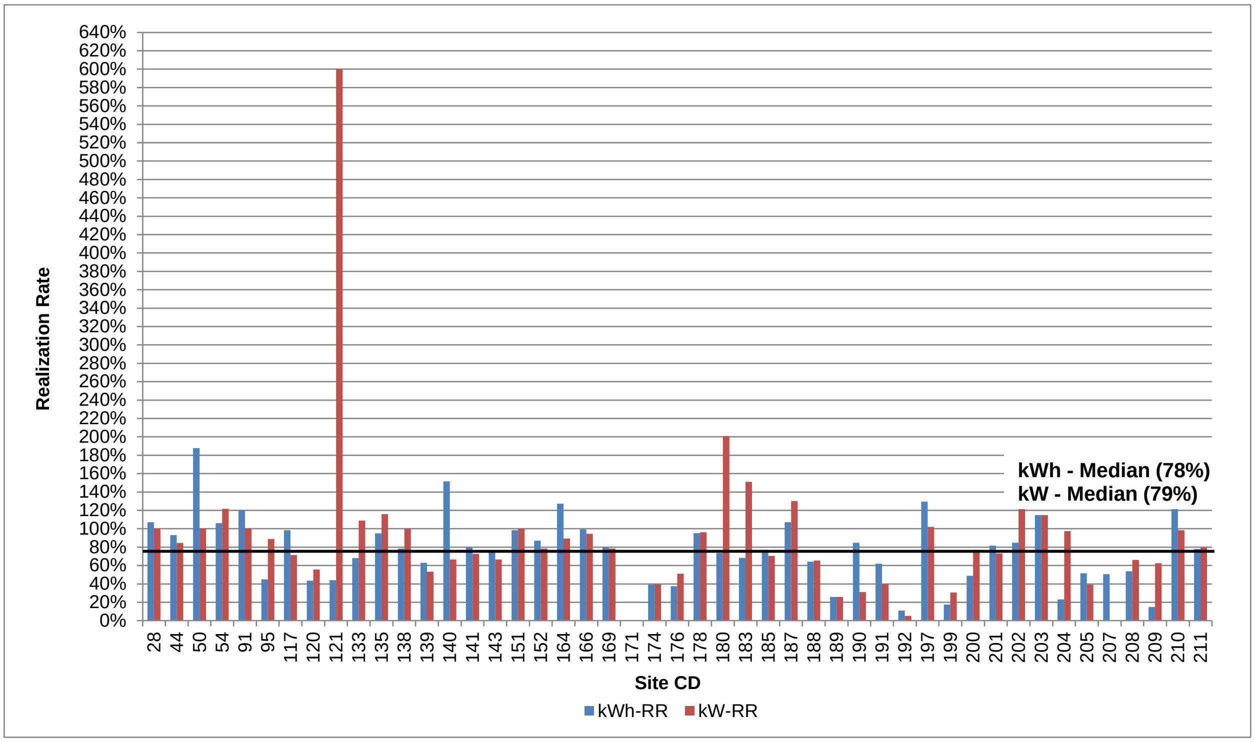


Figure 4-3 presents a plot of the energy and peak demand RR for the individual sampled site under the Electric Custom program. The x-axis is represented by the unique sampling site ID that was assigned to each site as a part of the on-site work. This number is different from "L" coded site ID that was used for the on-site planning purposes. The relationship of the unique sampling site ID presented in the plot and on-site work site ID (L-XYZ) is presented in Appendix C. The blue bars represent the energy realization rate information while the red bars represent the peak demand realization rate for the same site. Site 171 is not missing data but it represents a project that was evaluated to have no savings, hence it has no data in the plot. Site 121 was a lighting project where the applicant had significantly understated the peak demand savings resulting in the 600% peak demand realization rate for that project.

Figure 4-3. Electric Custom Program – Individual Site kWh and kW Realization Rates



4.1.2 Electric Rebate Program-Level Savings Results

Table 4-3 presents the gross evaluation results for the Electric Rebate program. The precision results for the energy savings were slightly above the evaluation goal of 10%. The error ratio for the energy value was slightly greater than the sample estimate of 0.5 while the peak demand error ratio was 1.52 indicating an exceptionally large degree of variability found between the evaluation and tracking savings estimates. The tracking peak demand results represent all types of technologies and the high ER value indicates a need to closely monitor the development of these estimates, which would fall under the improvement of quality control practices.

Table 4-3. Electric Rebate Program – Gross Energy and Peak Demand Impacts

	Total No. of Sites	Sampled No. of Sites	Tracking Gross Savings	RR	Evaluated Gross Energy Savings	RP	Error Ratio
Energy (kWh/yr)	632	56	67,813,154	70%	47,469,208	12%	0.53
Peak demand (kW)			8,368	76%	6,360	35%	1.52

Table 4-4 presents a summary of the interactive energy impacts determined by the evaluation team for the measures implemented under the Electric Rebate program. The interactive impacts are largely represented by the lighting projects, which typically involve a heating penalty. This table provides the interactive energy associated with the sampled projects. This information could be used by the program planners to estimate the non-electric impacts associated with their Electric Rebate program

Table 4-4. Electric Rebate Program – Interactive Impacts

Tracking Gross Savings (kWh/yr)	Evaluated Interactive Natural Gas Savings (Therms/yr)	Interactive Savings Factor (Therms/kWh)
67,813,154	-110,844	-0.0016

Figure 4-4 presents a plot of the gross evaluated energy savings vs. tracking energy savings.

Figure 4-4. Electric Rebate Program - Gross Evaluated Energy vs. Tracking Energy

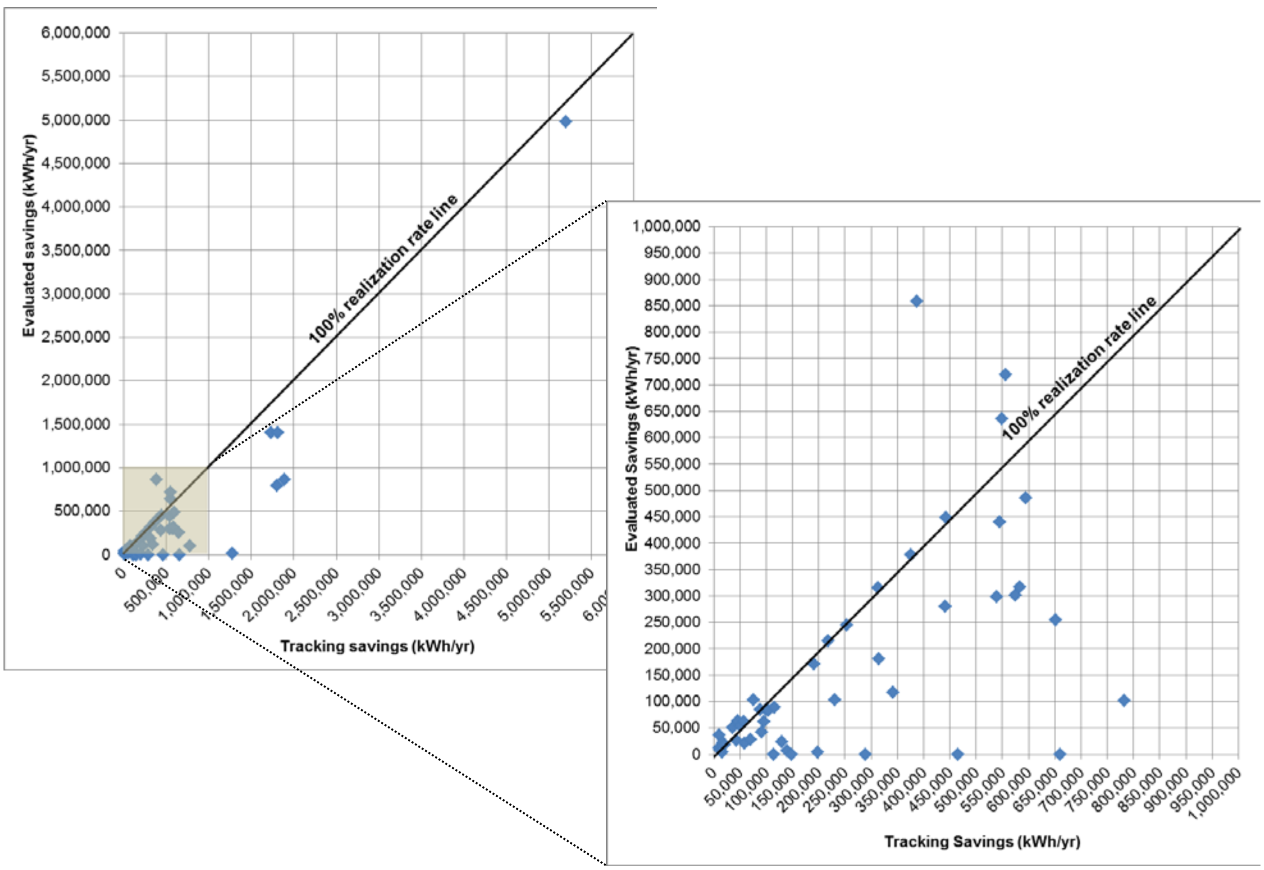


Figure 4-5 presents a plot of the gross evaluated peak demand savings versus tracking peak demand savings.

Figure 4-5. Electric Rebate Program - Gross Evaluated Peak Demand vs. Tracking Peak Demand

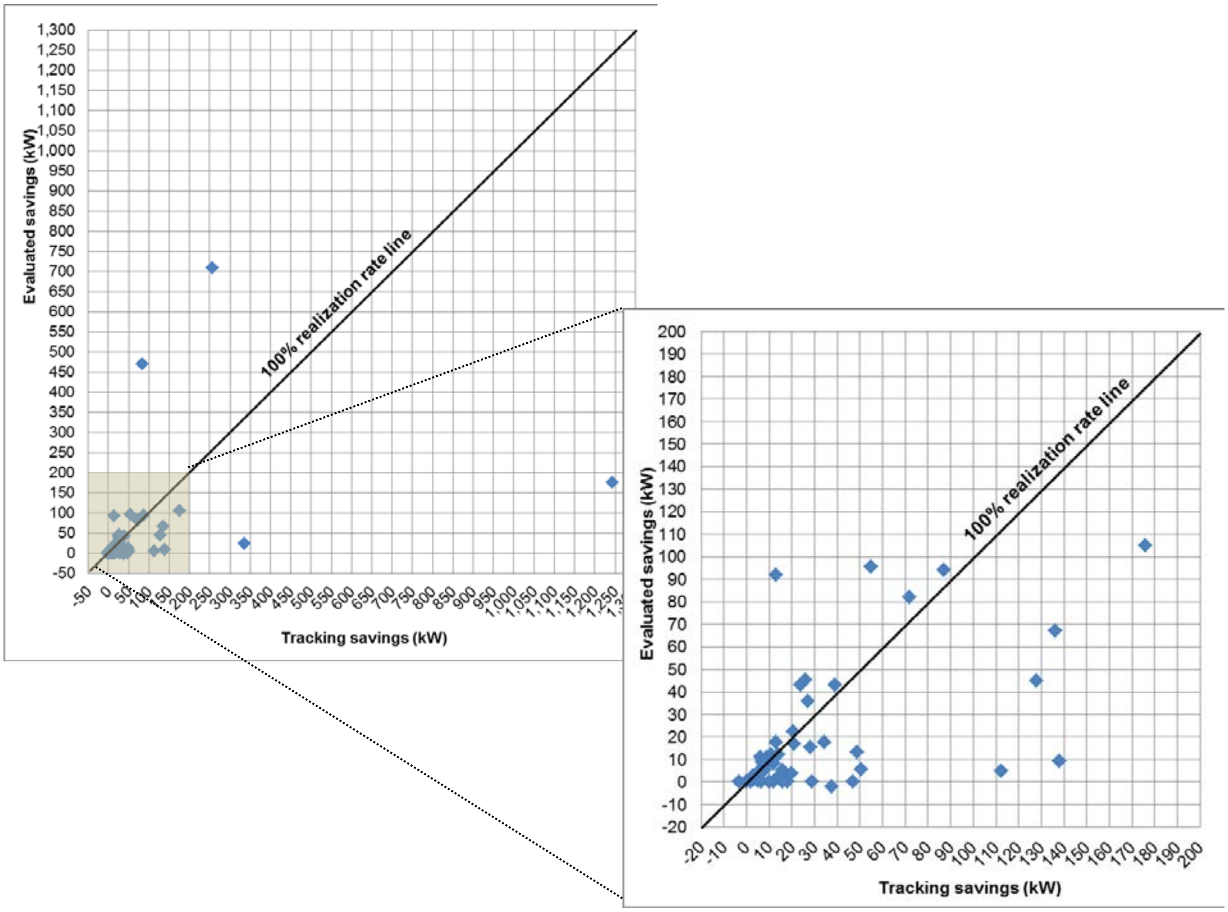
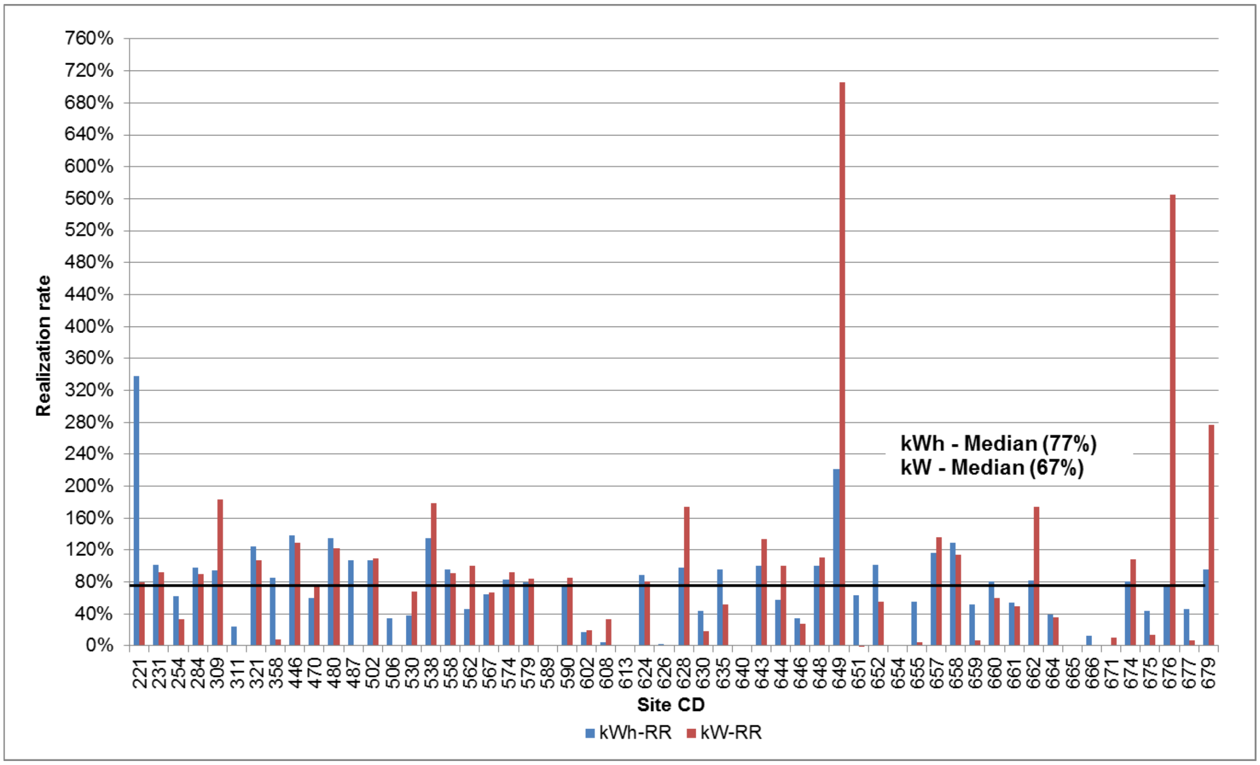


Figure 4-6 presents a plot of the energy and peak demand realization rate for each sampled site under the Electric Rebate program. Sites 649 and 679 were predominantly VFD projects where the applicant had used the required 2010 NYTM algorithm, which resulted in the savings being significantly understated for these projects. Site 676 was a comprehensive chiller controls upgrade project where the tracking peak demand savings were significantly understated. Site 221 was a lighting project at a retail store where the evaluated hours were found to be significantly greater than the hours used by the applicant. Site 613 was a lighting project where the savings were double counted for the same measure by the two Con Edison programs. The measure in this project aligned better with the custom program, hence the evaluation team assigned the savings to the Electric Custom program, which resulted in zero savings for the Electric Rebate program. Site 626 was a VFD project in which the applicant used an incorrect NYTM savings factor, which resulted in overstating the savings. Sites 640, 654, 665, and 671 were VFD replacement sites where the evaluation team assigned zero savings because the base case was determined to be VFD controlled systems. Appendix G presents evaluators comments on this issue of projects involving the replacement of VFDs.

Figure 4-6. Electric Rebate Program – Individual Site kWh and kW Realization Rates



4.1.3 Gas Custom Program Level Savings Results

Table 4-5 presents the gross evaluation results for the Gas Custom program. The results for this program are based on evaluating the impacts for one site. The evaluation team was not able to recruit the remaining sites in this program group; hence the supporting RP, standard error, and error ratio values were not computed for this program.

Table 4-5. Gas Custom Program – Gross Energy Impacts

	Total No. of Sites	Sampled No. of Sites	Tracking Gross Savings	RR	Evaluated Gross Energy Savings	RP	Error Ratio
Energy (therms/yr)	4	1	326,219	101%	329,481	N.D[1]	N.D[1]

[1] The RP and error ratio for the Gas Custom program were not determined (N.D.) because the evaluators were able to recruit only one site for their on-site analysis.

This program was represented by a lone site that achieved an RR of 101%. It was for installing EMS controls at a school building that set back space temperature during unoccupied times.

4.1.4 Gas Rebate Program Level Savings Results

Table 4-6 presents the gross evaluation results for the Gas Rebate program. The RP results for this program were fairly close to the evaluation goal of 10% RP at the 90% confidence interval.

Surprisingly, even with the low RR, the evaluated error ratio of 0.51 closely matched the evaluation plan estimate of 0.5.

Table 4-6. Gas Rebate Program – Gross Energy Impacts

	Total No. of Sites	Sampled No. of Sites	Tracking Gross Savings	RR	Evaluated Gross Energy Savings	RP	Error Ratio
Energy (therms/yr)	66	24	422,630	48%	202,862	18%	0.51

Table 4-7 presents a summary of the interactive energy impacts determined by the evaluation team for the measures implemented under the Gas Rebate program. The impact presented here is from one gas roof-top unit (RTU) tune-up project, which was assessed for natural gas savings (site L112 or site cd-687); however, the evaluation team found that the RTUs at this site did not utilize natural gas heating. Therefore, the evaluation team credited this gas tune-up project with electric savings instead because the cooling coils were cleaned.

Table 4-7. Gas Rebate Program – Interactive Impacts

Tracking Gross Savings (Therms/yr)	Interactive Evaluated Electric Savings (kWh/yr)	Interactive Savings Factor (kWh/therms)
422,630	49,722	0.1176

Figure 4-7 presents a plot of the gross evaluated energy savings versus tracking energy savings. With the exception of three projects, all Gas Rebate program-funded projects underperformed.

Figure 4-7. Gas Rebate Program – Gross Evaluated Energy vs. Tracking Energy

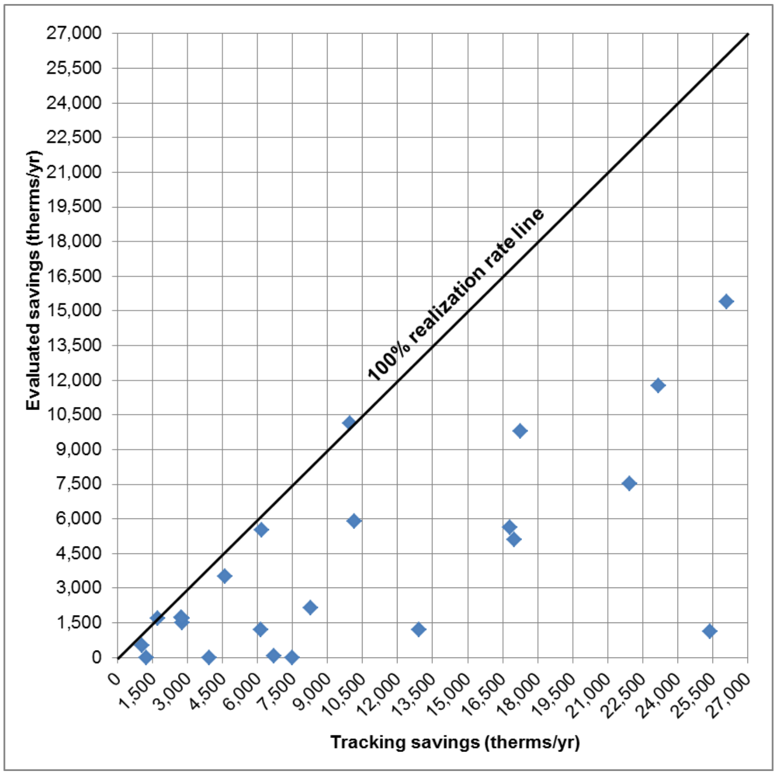
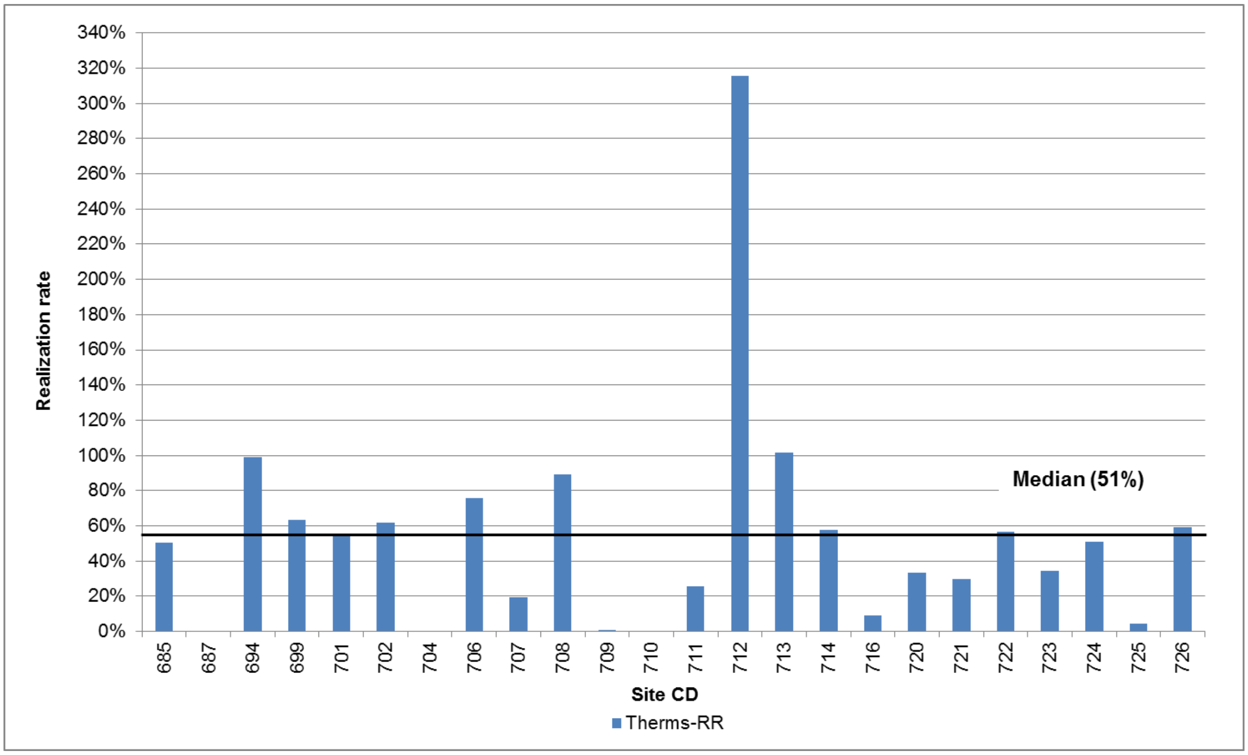


Figure 4-8 presents a plot of the energy based RR for each sampled site under the Gas Rebate program. The site 687 project involved an RTU tune-up but the evaluation team found that the RTUs at this site used electric heat instead of gas for heating; hence this project resulted in zero gas savings. The site 710 project involved installing controls on the boiler plant to optimize the boiler operations for which 35% savings were claimed compared to the baseline gas consumption. A review of the pre- and post-billing data for this site revealed that no savings were being realized from this project. The project at site 709 involved tune-up of multiple RTUs. Evaluators found that one RTU did not use gas for heating and that the savings claimed for this project exceeded the annual natural gas consumption for this site. The project at site 712 involved multiple weatherization and water conservation measures at a college campus. This project resulted in a significantly high RR (>270%) due to a shift in the fuel mix used by the facility. When the project application was submitted, the campus fuel use breakdown was approximately 24% gas and 76% oil. However, at the time of the evaluation site visit, the campus had shifted its use completely to natural gas and had indicated plans to do the same in the foreseeable future, increasing the gas savings significantly for this site.

Figure 4-8. Gas Rebate Program – Site-by-Site Realization Rates



4.2 Attribution

The sections below provide a detailed overview of the attribution results, including FR, SO, and the resulting NTGR.

4.2.1 Free Ridership

The evaluation team was able to develop independent FR rates for CECONY's Electric Rebate and Electric Custom program components. Despite substantial efforts to reach every program participants through a variety of channels, for the Gas Rebate program, a small number of interviews (n=3) were completed, which yielded a FR value of 0.24. Even though the evaluation team had results from the three surveys for the Gas Rebate program, the evaluation team deemed them to be insufficient to accurately represent the program.

The interviewers were unable to complete any interviews with the project managers associated with the five Gas Custom projects. As a result, the evaluation team was not able to develop any attribution estimates specific to the Gas Custom program.

Appendix A of this report discusses in greater detail the efforts that the evaluation team undertook in an attempt to complete interviews with all program participants. The efforts included placing multiple calls at varying times of the day and week over an extended period of time, searching for alternative contact information, leaving multiple voicemails, following up via email, and offering incentives for survey completion. As observed through the disposition reports for the gas program participants, core reasons for our inability to complete interviews included respondent refusal to participate and non-working or wrong phone numbers, which accounted for over half of available customer contacts. Other dispositions included continuously reaching answering machine, no answer, and gatekeeper call back. Underlying reasons for these dispositions include the evaluation taking place months, and in some cases years after project completion. During this time, project decision makers might no longer be employed at the same company. Those who are still employed might not be as engaged due to the amount of time elapsed since the implementation of the project. The reasons for why such response patterns are observed among gas program participants and not electric program participants, however, are unknown.

The final FR rates are presented in Table 4-8. As can be seen in in this table, FR ranges from 0.26 for the Electric Rebate program component to 0.38 for the Electric Custom program components.

Since the survey was a census attempt, the concept of sampling error does not apply. However, there might be other sources of error and bias, such as a nonresponse error, that could be influencing the results.

Table 4-8. Free Ridership, Spillover, and NTG Estimates

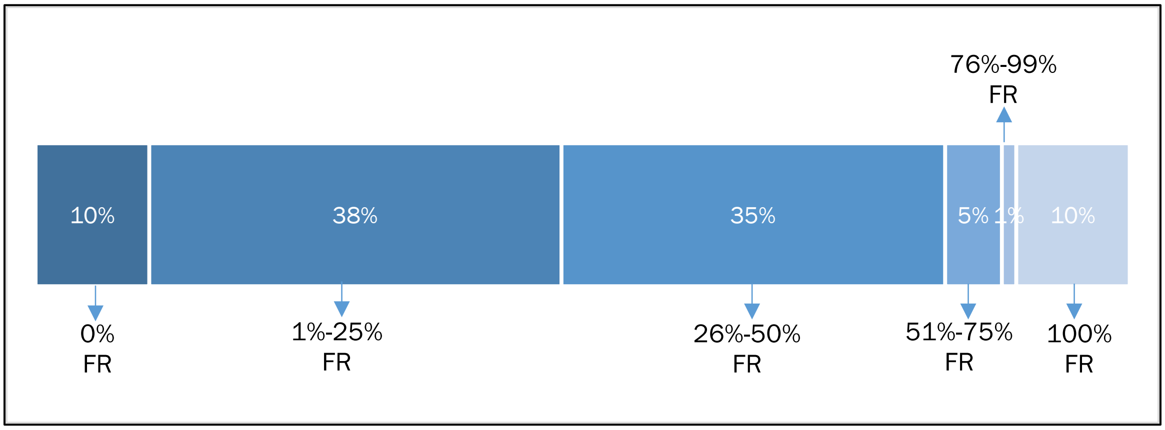
Program	Count	FR
Electric Rebate	84	0.26
Electric Custom	26	0.38
Gas Rebate	3	I.D
Gas Custom	-	N.D

I.D. – Insufficient data. This requires further study to accurately estimate this value

N.D. – No data was available to develop estimates for these variables

Figure 4-9 shows the breakdown of the survey respondent FR values. Overall, only 10% of the survey respondents were complete free riders and only 16% had an FR of more than 50%.

Figure 4-9. Breakdown of the Survey Respondent Free Ridership Values



Few projects were either complete free riders or completely driven by the program. Eighty percent of respondents were partial free riders.

Interestingly, in addition to influencing the decision to install high efficiency equipment, the program is having a considerable influence on advancing the timing and increasing the scope of the project. Based on the survey responses, over a quarter of respondents (29%) reported that the program accelerated the completion of the project by a year or more, while 19% reported that the program increased the scope of the project.

While a considerable percentage of respondents placed great importance on program incentives and support in driving the high efficiency project(s), a considerable number cited other non-program factors of influence, likely contributing to the current level of program FR. More specifically, 60% of respondents mentioned that they either wanted to reduce their energy bills or save energy by installing high efficiency measures. Furthermore, over half of respondents (58%) reported being influenced by the standard practice in their business or industry, and close to half (49%) reported being influenced by their company's corporate policy.[9]

The participant self-assessments about the program influence were consistent with the trade ally assessments and as a result there was minimal impact of the trade ally assessments on the overall attribution results.

The FR rates found for CECONY's LC&I program are not unusual and generally mimic those for other similar programs in the region and across the country. Table 4-9 provides an overview of the FR rates found in other parts of the country.

[9] A rating of 5, 6, or 7 on a scale of 1 to 7, where 1 is not at all influential, and 7 is very influential.

Table 4-9. Free Ridership Rates in Other Jurisdictions

Program	FR	Study Year
Northeast Utilities Electric C&I program	0.28	2011
Massachusetts gas programs	0.31	2012
Massachusetts electric programs	0.16	2011
Efficiency Maine Electric Business Incentive Program	0.34	2011
NYSERDA Existing Facilities Program	0.31	2010–2011

4.2.2 Spillover

The following sections describe the evaluation of the components of spillover.

Participant Spillover

The evaluation team asked Large C&I program participants about the SO associated with installing a variety of equipment options without receiving incentives through the program. Measures included lighting, motors, HVAC, refrigeration, heating and water heating, and kitchen equipment.

Three out of seventy-seven survey respondents indicated that they made additional installations after participating in the program and highly rated the program's influence on their decision to proceed with those installations. All of the SO measures included additional lighting installations (LED light fixtures, motion sensors, and T8s). Total energy savings from those measures amounted to 14,293 kWh, 3.94 kW, and -0.05 therms. The SO savings were calculated using NYTM savings assumptions. The participant SO values were very small (0.001 to 0.0007) and hence are not reported.

Nonparticipant Spillover

Quantifying NPSO and vendor off-site SO was not in the scope of this evaluation but was explored during the interviews with the trade allies that participated in the Con Edison Large C&I programs. These interviews revealed the potential presence of NPSO. Of the nine trade allies interviewed, six were asked about NPSO, and five confirmed that they had completed high efficiency projects outside of the program. Three of the five said that the program requirements, particularly the paperwork and site visits, were too time consuming. One trade ally reported that if customers weren't looking for rebates, he purposefully didn't mention the program to them so as to avoid more paperwork. The other two of the five trade allies that mentioned completing high efficiency projects outside of the program said that they installed high efficiency equipment that wasn't on the approved measures list.

4.2.3 Net-to-Gross Ratio

Estimates of FR and participant SO were used to derive the final NTGRs. Table 4-10 presents the electric and natural gas energy related NTG values associated with the four CECONY programs. Please note that due to insufficient attribution survey data, the evaluation team did not present attribution results for the two gas programs.

Because the attribution survey effort was a census attempt, the concept of relative precision (which is associated with sampling) does not apply. Nevertheless, the evaluation team developed relative precision estimates to provide insight on what they would have been if the effort was a sampling effort and not a census attempt, but did not propagate the relative precision to the relative precision around net impacts.

Table 4-10. Program Energy (kWh or Therms) NTG

Program Component	Count	FR	SO	NTG	NTG RP[1]
Electric Rebate	84	0.26	0.0018	0.74	0.244
Electric Custom	26	0.38	0.0007	0.62	0.188
Gas Rebate	3	I.D	N.D	N.D	0.074
Gas Custom	0	N.D	N.D	N.D	N.D

[1]Since the survey was a census attempt, the concept of relative precision, which represents sampling error, does not apply. The relative precision values presented above cannot be propogated to the rest of the population due to a large number of non responses in the survey.

I.D. – Insufficient data. This requires further study to accurately estimate this value.

N.D. – No data was available to develop estimates for these variables.

4.3 Net Program Level Results

Net program level results, calculated using the verified gross impacts and the NTG, are presented in the following section.

4.3.1 Electric Custom Program Net Impacts

The net annual Electric Custom program impacts are 12,112,044 kWh and 2,003 kW. These results are shown in Table 4-11.

Table 4-11. Electric Custom Program – Net Energy Impacts

	Total No. of Sites	Sampled No. of Sites	Tracking Gross Savings	Gross RR	Evaluated Gross Energy Savings	Net RR	Net Energy Savings
Energy (kWh/yr)	232	47	27,132,715	72%	19,535,555	45%	12,112,044
Peak demand (kW)			3,512	92%	3,231	57%	2,003

4.3.2 Electric Rebate Program Net Impacts

The net annual Electric Rebate program impacts are 35,127,214 kWh and 4,706 kW. These results are shown in Table 4-12.

Table 4-12. Electric Rebate Program – Net Energy Impacts

Program	Total No. of Sites	Sampled No. of Sites	Tracking Gross Savings	Gross RR	Evaluated Gross Energy Savings	Net RR	Net Energy Savings
Energy (kWh/yr)	632	56	67,813,154	70%	47,469,208	52%	35,127,214
Peak demand (kW)			8,368	76%	6,360	56%	4,706

4.3.3 Gas Custom Program Net Impacts

The net annual Gas Custom program impacts are not presented because the evaluation team was not able to develop attribution estimates specific to this program which is a critical factor in estimating the net program impacts. As indicated earlier, the attribution survey was a census attempt, therefore, the evaluation team reached out to all available contacts included in this evaluation.

4.3.4 Gas Rebate Program Net Impacts

The net annual Gas Rebate program impacts are not presented because the evaluation team was not able to develop credible attribution estimates specific to this program which is a critical factor in estimating the net program impacts. As indicated earlier, the attribution survey was a census attempt, therefore, the evaluation team reached out to all available contacts included in this evaluation.

4.4 Measure Level Results

The following sections present information on the measure categories that contributed the most to the electric and gas programs in the on-site survey population. Based on the tracking energy savings contribution, three measures –- motors and VFDs, lighting, and EMS and controls – contributed over 87% of the savings in the sampled population of sites. Table 4-13 presents a summary of the measures represented in the electric programs based on energy savings.

Table 4-13. Electric Programs – Measure Breakdown Based on kWh

Program	Measure Category	Count	Gross RR	RP	Error Ratio
Electric Custom	EMS and controls	4	71%	7%	0.06
	Lighting	29	88%	9%	0.29
	Motors and VFDs	5	34%	68%	0.72
	Other	10	49%	29%	0.50
Electric Rebate	EMS and controls	13	50%	30%	0.61
	HVAC	4	99%	15%	0.12
	Lighting	26	77%	17%	0.49
	Motors and VFDs	18	71%	20%	0.48
	Other	2	64%	115%	0.26

With the exception of the motors and VFDs and "other" measure categories in the Electric Custom program and the EMS and controls and "other" measure categories in the Electric Rebate program, the remaining measure level results indicate good RP values (<20%). Therefore, the results for the measures with less than 20% RP could be reliably applied to the same measure types in the rest of the program population.

Table 4-14 presents the measure specific peak demand results associated with the electric programs. Please note that peak demand counts do not line up with the energy values because certain projects did not report a peak demand (kW) value but reported an energy (kWh) value. In addition, because a site could have had multiple measures, the measure counts presented in Tables 4-13 and 4-14 do not match with the site counts.

Table 4-14. Electric Programs – Measure Breakdown Based on kW

Program	Measure Category	Count	Gross RR	RP	Error Ratio
Electric Custom	EMS and controls	4	134%	31%	0.26
	Lighting	28	86%	8%	0.25
	Motors and VFDs	5	38%	71%	0.75
	Other	9	69%	18%	0.29
Electric Rebate	EMS and controls	9	49%	88%	1.43
	HVAC	4	124%	31%	0.26
	Lighting	26	58%	44%	1.33
	Motors and VFDs	18	120%	52%	1.27
	Other	2	47%	114%	0.26

The RP values for the peak demand savings estimates indicate a higher degree of variability between the evaluated peak demand savings values and the tracking values. With the exception of the lighting measure category in the Electric Custom program, the peak demand savings impacts associated with the rest of the measure categories displayed a significantly higher amount of variability as reflected by the RP and error ratio values. This information highlights the need to develop tools and protocols to accurately record the peak demand savings information in the program database.

Table 4-15 presents the measure level results associated with the gas programs.

Table 4-15. Gas Programs – Measure Breakdown

Program	Measure Category	Count	Gross RR	RP	Error Ratio
Gas Custom	EMS and controls	1	101%	N.D	N.D
Gas Rebate	EMS and controls	12	39%	23%	0.45
	HVAC	1	76%	N.D	N.D
	Other	2	150%	294%	0.66
	Tune-up	9	49%	17%	0.27

N.D. – These parameters were not determined due to a lack of sufficient statistical data points.

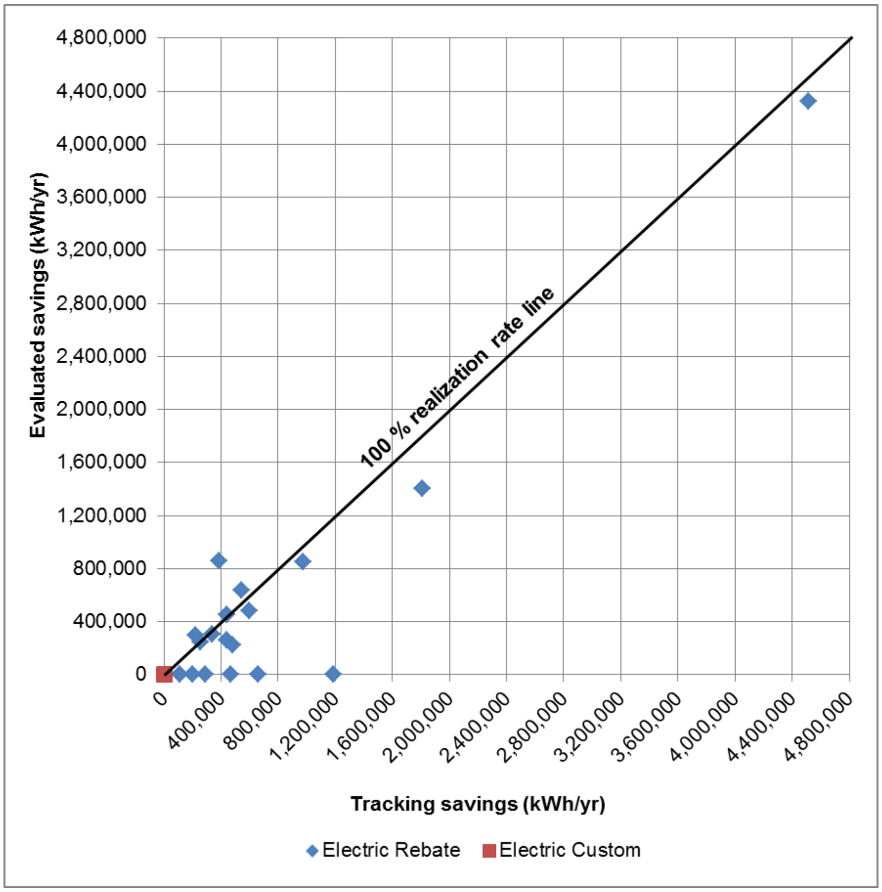
For the Gas Rebate program, the measure level data indicates decent RP values (<25%) for the two measures – "EMS and controls" and "tune-up" – that contribute the most to the program savings. However, these two measures also demonstrate poor RRs (39% and 49%) as seen in the information presented in the table.

In the gas programs, the EMS and controls measure and the tune-up measure contributed over 93% of the savings analyzed in the on-site sample.

4.4.1 Electric – Motors and VFDs

The motors and VFDs measure principally represents projects involving the installation of VFDs on fan and pumping applications across a variety of end uses. There were a total of twenty-two VFD projects at eighteen sites funded by the Electric Rebate program and one site funded by the Electric Custom program. Figure 4-10 below shows the overall performance of the electric energy savings associated with the motors and VFDs measure funded under the Electric Rebate and Custom programs.

Figure 4-10. Electric Rebate and Custom Motors and VFDs Measures kWh Performance



The majority of the VFD projects did not achieve the savings that were projected. The motors and VFD group had an average RR of 34% in the Electric Custom program and 71% in the Electric Rebate program.

There were eight sites at which the evaluation engineers found that the funded VFDs replaced preexisting VFDs. Evaluators assessed whether or not the preexisting VFD should be the baseline technology for the projects or whether regression to a non-VFD baseline was plausible.[10] For measures at six of the eight projects the evaluation team concluded that the VFD was baseline, resulting in zero evaluated savings and a 0% RR. At the two remaining sites, the evaluation team found a small subset of applications to have a plausible non-VFD baseline; hence the savings associated with those installations were calculated.

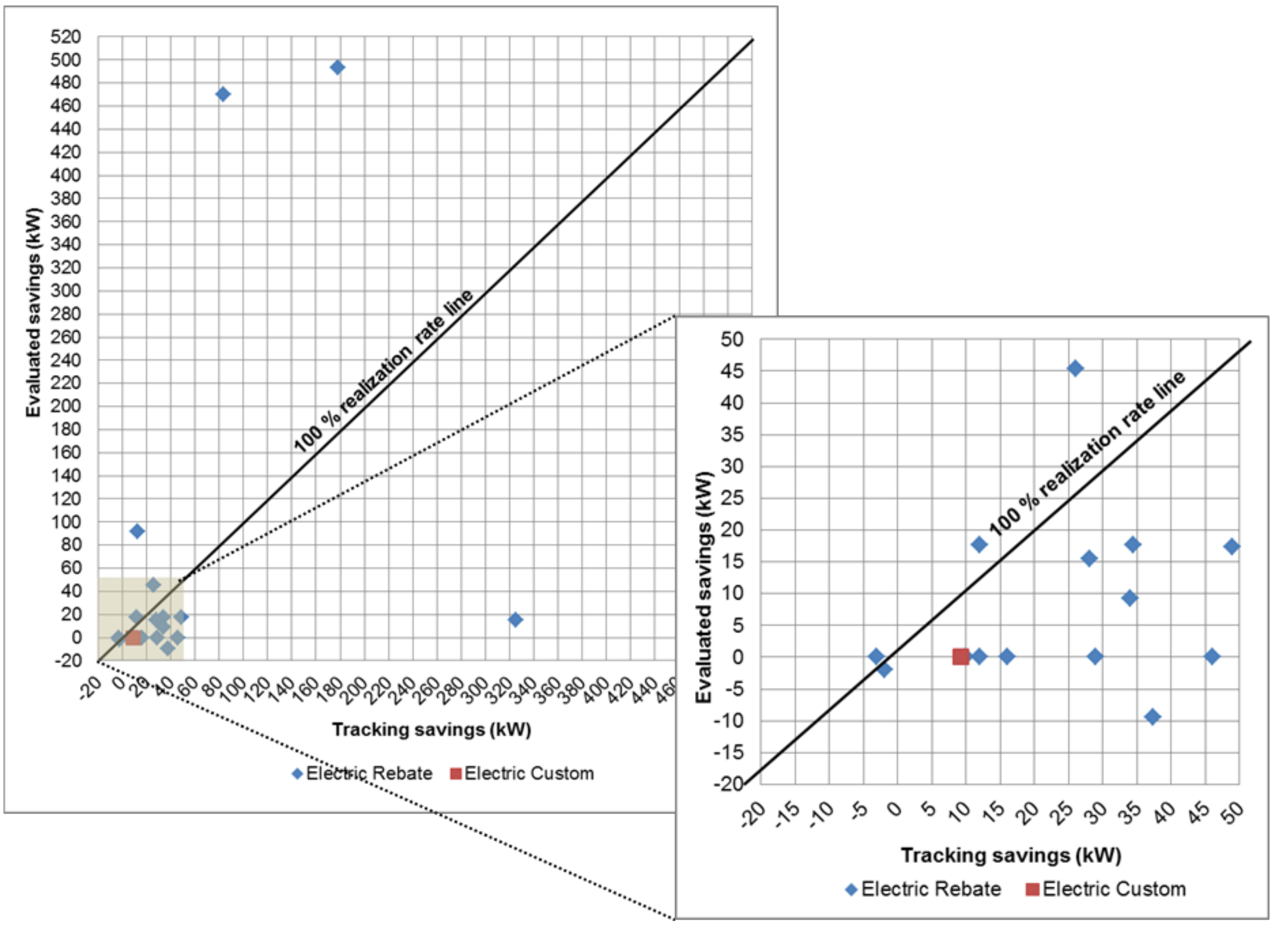
Excluding the zero-saver VFD sites, the RR for this measure group increases from 34% to 41% for the Electric Custom program and from 71% to 88% for the Electric Rebate program. This means that the baseline issue explains over half of the Rebate Program's deviation in the

[10] See Appendix G for the methodology and battery of questions used in this exercise.

realization from 100%. Appendix G also presents our recommendations related to the zero-saver VFD finding associated with this evaluation.

Figure 4-11 below shows the overall performance of the peak demand savings associated with the motors and VFDs measures funded under the Electric Rebate and Custom programs. The inserted plot shows the overall performance of the peak demand savings associated with the motors and VFDs measures with savings of up to 50 kW. The plot indicates that for a majority of the projects the evaluated peak demand values were less than the tracking values.

Figure 4-11. Electric Rebate and Custom Motors and VFDs Measures kW Performance



Based on a survey of the above VFD projects under the rebate and custom programs, Table 4-16 highlights the evaluation findings associated with the parameters listed in the NYTM for the different facility types and their associated VFD applications. This table does not include the data associated with the zero-saver VFD sites.

Table 4-16. NYTM Relevant Parameters for VFD Measures

Site	Application	Count	Avg. Evaluated kWh/HP	Avg. NYTM kWh/HP	Avg. Evaluated kW/HP	Avg. NYTM kW/HP
Hospital	RF	1	1,215	1,801	0.065	0.111
Hospital	SAF	1	947	2,137	0.026	0.070
Hotel	HWP	2	137	6,603	0.000	0.498
Multifamily	CHWP	1	386	N/A	0.044	N/A
Multifamily	CT fan	1	605	N/A	-0.024	N/A
Office	CHWP	4	2,303	1,176	0.256	0.061
Office	CT fan	3	665	407	0.022	-0.025
Office	CWP	7	1,487	1,183	0.099	0.061
Office	HWP	1	1,302	845	0.000	0.498
Office	SAF	2	888	1,605	0.074	0.078

RF= return fan, SAF= supply air fan, HWP = hot water pump, CHWP= chilled water pump
CT fan = cooling tower fan, CWP = condenser water pump

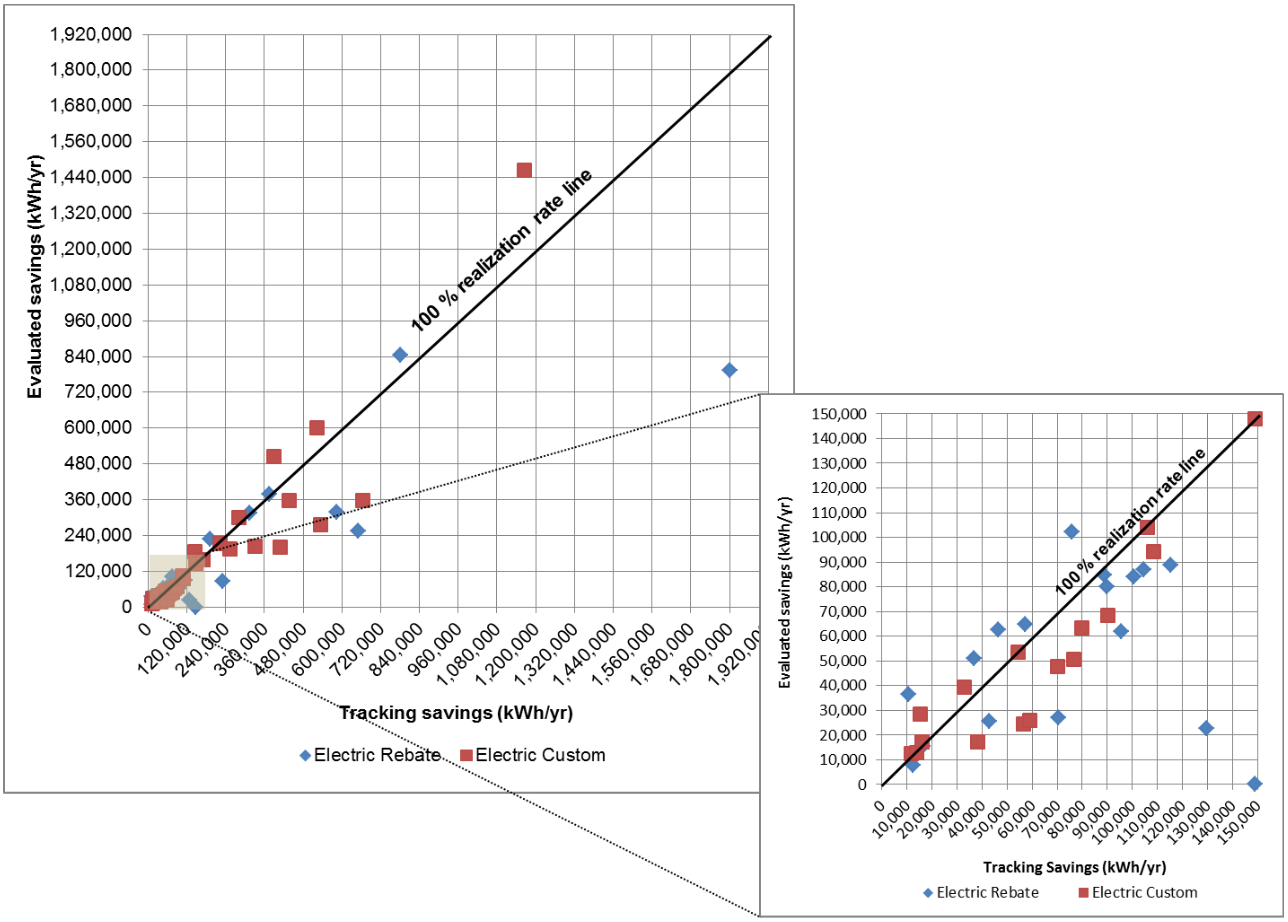
The following conclusions can be drawn from the limited number of data points for which this data was collected.

- The evaluation team found that in Con Edison's territory the VFDs controlling hot water pumps in hotel applications are saving markedly less than the NYTM predicts. While the sample is small (two) and there is an odd anomaly in evaluators' finding (only 137 kWh saved per hp in a hotel application), Con Edison may want to work with the DPS staff to adjust the values, perhaps by pooling these results with those of a recent NEEP study or discounting the savings from the NYTM estimates within the program's tracking system to protect against lower RRs in the future.
- For office buildings, the evaluated kWh/hp and kW/hp factors associated with the chilled water pumps, condenser water pumps, and cooling tower fans were observed to be greater than the NYTM specified values. In all of these instances, these were installations in the New York City area that had equipment operating for a significantly greater number of hours. Central cooling plant pumps appear to be rich targets for additional savings.
- For the supply fan VFDs in the office buildings, the evaluated energy savings factor (888 kWh/hp) was less than 60% of the savings factor listed in the NYTM (1,605 kWh/hp). The evaluated demand savings factor for supply fan VFDs was in conformance with the NYTM demand savings factor.

4.4.2 Lighting

There were thirty-three verified lighting projects funded by the Electric Custom program and twenty-six verified projects funded by the Electric Rebate program. Figure 4-12 below shows the performance of the electric energy savings associated with the lighting measures under the two programs. The subset plot shows results for projects with savings of up to 150,000 kWh per year. Over 63% of the lighting sites in the on-site evaluation sample had tracking savings of less than 150,000 kWh per year. The average energy RR for this group was 93% while the peak demand RR was 100%.

Figure 4-12. Rebate and Custom Lighting Measures kWh Performance (All sites)



The average energy RR for the entire portfolio of Electric Rebate and Custom lighting projects was 89% and the peak demand RR was 93% for this group.

Figure 4-13 presents the information on the peak demand savings associated with the lighting projects. The subset plot shows results for projects under 100 kW.

Figure 4-13. Rebate and Custom Lighting Measures kW Performance (All Projects)

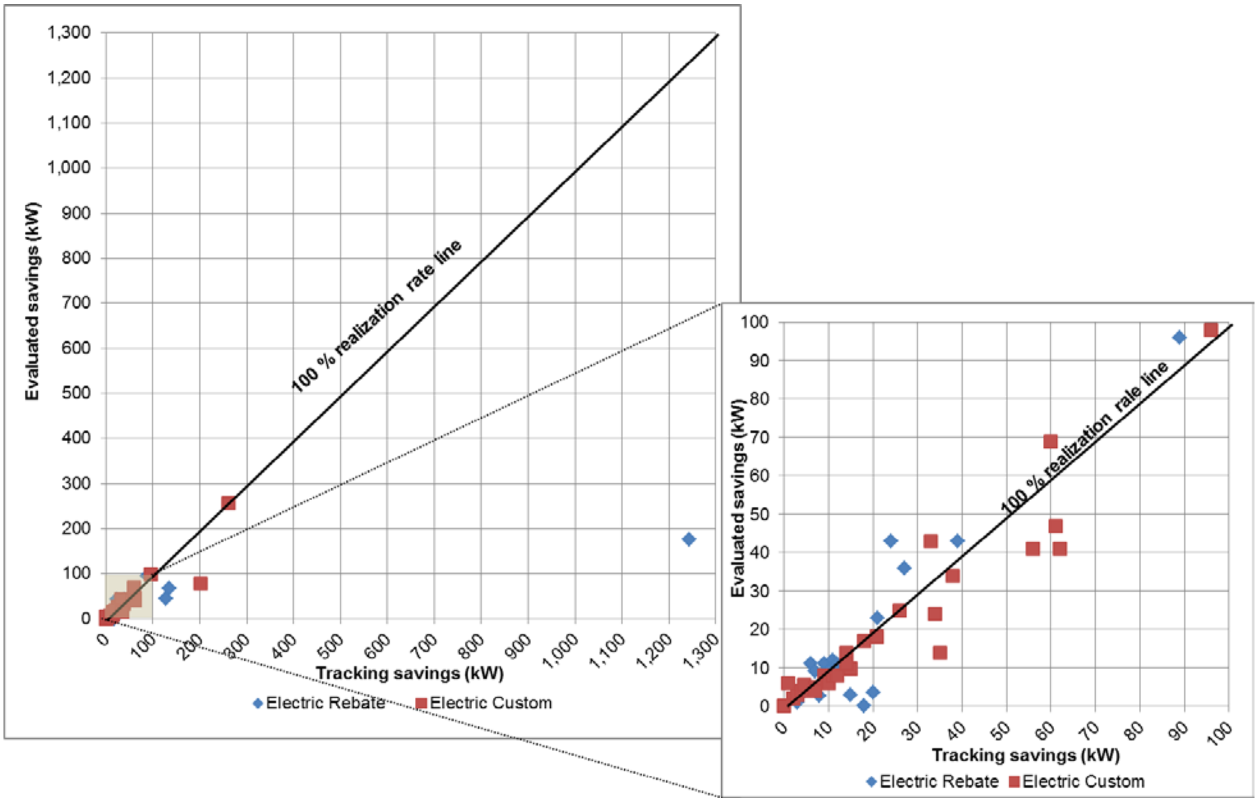


Table 4-17 tabulates the evaluation's findings associated with the parameters listed in the NYTM for the different facility types. This table only includes information for the facility types that were represented more than one time in the on-site sample.

Table 4-17. NYTM Relevant Parameters for Lighting Measures

Rebate Lighting	Count	Avg Applicant Hours	Avg Evaluated Hours	Avg NYTM Hours	Avg. Evaluated kWh Cooling Factor	Avg. NYTM kWh Cooling Factor	Avg. Evaluated Heating Factor	Avg. NYTM Heating Factor	Evaluated CF	NYTM CF
Small retail	8	7,156	7,904	4,057	1.103	1.130	-0.012	-0.022	1.00	1.00
Parking garage	2	8,760	8,760	4,368	0.000	0.000	0.000	0.000	1.00	1.00
Office	1	8,760	3,744	3,100	1.065	1.033	-0.011	-0.022	0.79	1.00
Multifamily	2	8,760	8,563	7,655	1.000	1.101	-0.007	-0.021	0.99	1.00
Hotel	2	5,550	5,700	3,064	0.526	0.517	-0.007	-0.011	0.75	1.00
Wtd average		**7,477**	**7,535**	**4,382**	**0.862**	**0.887**	**-0.009**	**-0.017**	**0.95**	**1.00**
Custom Lighting	**Count**	**Avg Applicant Hours**	**Avg Evaluated Hours**	**Avg NYTM Hours**	**Avg. Evaluated kWh Cooling Factor**	**Avg. NYTM kWh Cooling Factor**	**Avg. Evaluated Heating Factor**	**Avg. NYTM Heating Factor**	**Evaluated CF**	**NYTM CF**
Small retail	2	3,854	4,901	4,057	1.060	1.077	-0.012	-0.023	1.00	1.00
Large retail	2	5,400	7,165	4,057	1.055	1.023	-0.014	-0.024	0.91	1.00
Parking garage	3	8,760	8,694	4,368	0.333	0.000	0.000	0.000	0.99	1.00
Office	12	7,213	6,416	3,100	0.939	0.931	-0.003	-0.011	0.86	0.92
Multifamily	3	8,040	7,389	7,655	0.335	1.067	-0.005	-0.009	0.71	1.00
Wtd average		**7,066**	**6,790**	**4,068**	**0.795**	**0.844**	**-0.005**	**-0.012**	**0.87**	**0.95**
CF = coincidence factor, Heating factor = therms/kWh										

The following conclusions can be drawn for the rebate lighting projects:

- The lighting hours data indicates that the applicants consistently used a greater number of hours than those specified in the NYTM. Since the program allows applicants to use site-specific hours, the custom hours used by the applicants were applicable (except one office site) to the projects as the evaluated hours also closely matched the custom hours used by the applicants. This limited data indicates that the applicants should continue to use custom site-specific hours, and the NYTM hours should be used only when site-specific hours are not available.[11]
- The evaluated interactive cooling energy factor was similar to the factors specified in the NYTM for the sites listed in the table.
- The evaluated interactive heating penalty factor was found to be consistently lower than the value specified in the NYTM for all facility types.
- The evaluated coincidence factor was markedly less than one for offices and hotels, indicating a need to develop a coincidence factor table specific to at least those different building types, similar to the operating hours table presented in the NYTM.

[11] This advice contradicts that offered for Con Edison's SBDI program. Before making any change the staff for the two programs should discuss implementation to determine if the apparently contradictory recommendations are valid for each.

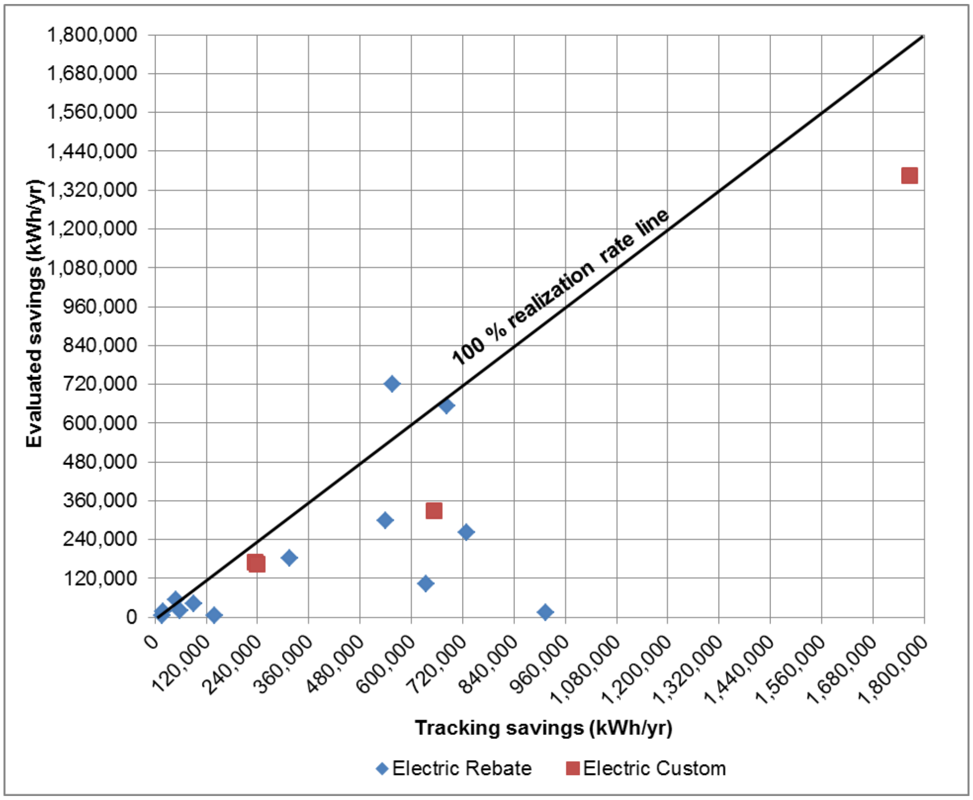
The following conclusions can be drawn for the custom lighting projects:

- The lighting hours data indicates that the number of lighting hours used by the applicants was higher than that specified in the NYTM, but again the evaluation team found the custom hours to be mostly applicable to these projects as the evaluated hours were observed to be more closely matched to the applicant hours than to the NYTM hours.
- On average, the evaluated interactive cooling energy factors were also consistent with the factors specified in the NYTM for the different facility types.
- The evaluated interactive heating energy factors were found to be consistently lower than the values specified in the NYTM for all facility types. This suggests that the buildings and spaces affected by lighting projects in this evaluation in the New York City area tended to a lower balance point temperature; hence offering more opportunities for economizer cooling.
- With the exception of the small retail sites, the evaluated coincidence factors were also always less than one, indicating a need to develop a coincidence factor table specific to the different building types, similar to the operating hours table presented in the NYTM.

Electric – EMS and Controls

The electric EMS and controls projects in the on-site sample included projects involving upgrading old EMS systems with new direct digital controls (DDC)-based EMS systems and installing programmable thermostats, timer controls, or occupancy-based controls on HVAC and lighting systems. The electric programs funded a total of seventeen EMS and controls projects: thirteen were funded by the Electric Rebate program and four were funded by the Electric Custom program. Figure 4-14 below shows the overall performance of the EMS and controls measures under the Electric Rebate and Custom programs.

Figure 4-14. Electric Rebate and Custom Programs – EMS and Controls Measures kWh Performance



The vast majority of the electric EMS and controls projects underperformed. The average RRs for the Electric Rebate and Electric Custom program-funded EMS and controls measures were 53% and 67%, respectively. The RR for the peak demand was 49% for the Electric Rebate program and 108% for the Electric Custom program.

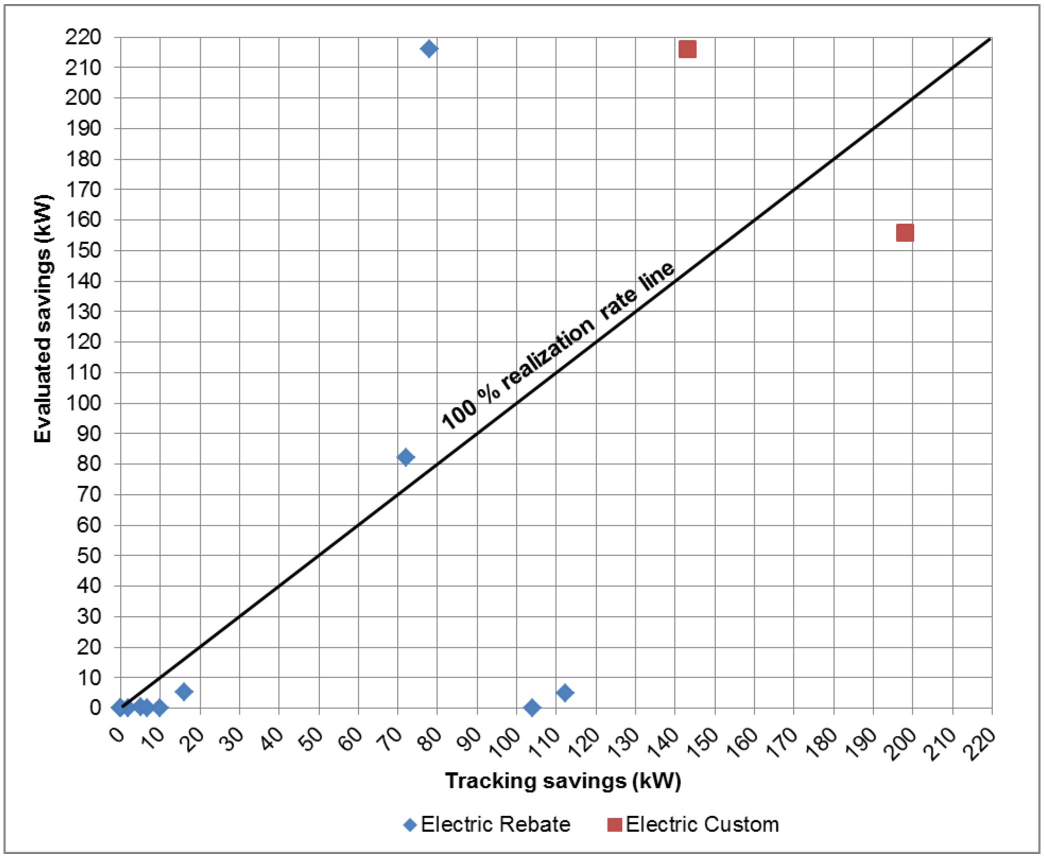
Figure 4-15 shows the overall performance of the peak demand associated with the EMS and controls measures under the electric and custom programs.

For the most part, the EMS and controls projects underperformed due to the following issues:

- The projects lacked sufficient detail associated with the baseline operating conditions.
- Savings on many projects were determined using unsubstantiated savings factors.
- The systems were not commissioned in accordance with the information presented in the project documents.

The solutions to these issues are listed in the Recommendations section.

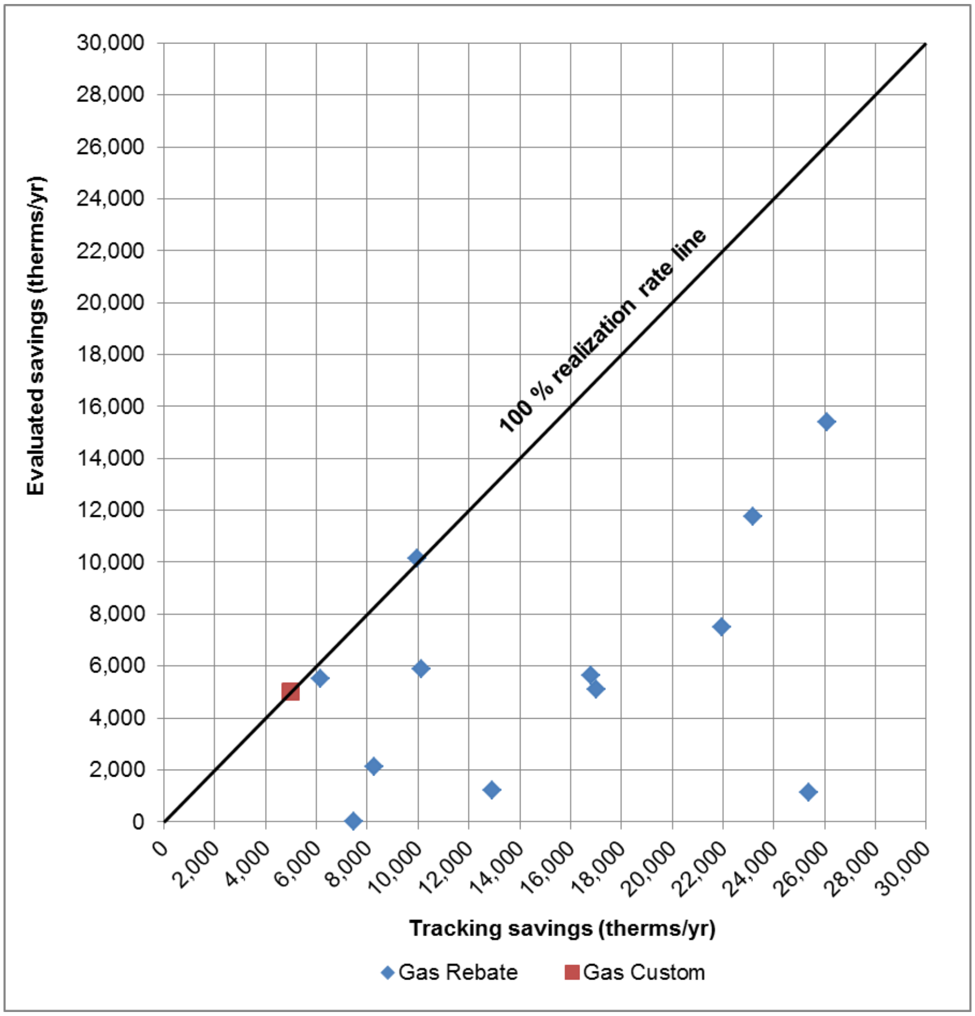
Figure 4-15. Electric Rebate and Custom Programs – EMS and Controls Measures kW Performance



4.4.3 Gas – EMS and Controls

The gas EMS and controls projects in the on-site sample included projects involving upgrading old EMS systems with new DDC-based EMS systems and installing programmable thermostats, timer controls, or occupancy-based controls on HVAC systems. The gas programs funded a total of thirteen EMS and controls projects – twelve rebate and one custom. Figure 4-16 shows the overall performance of the gas EMS and controls measures.

Figure 4-16. Gas Rebate and Custom Programs – EMS and Controls Measure Performance

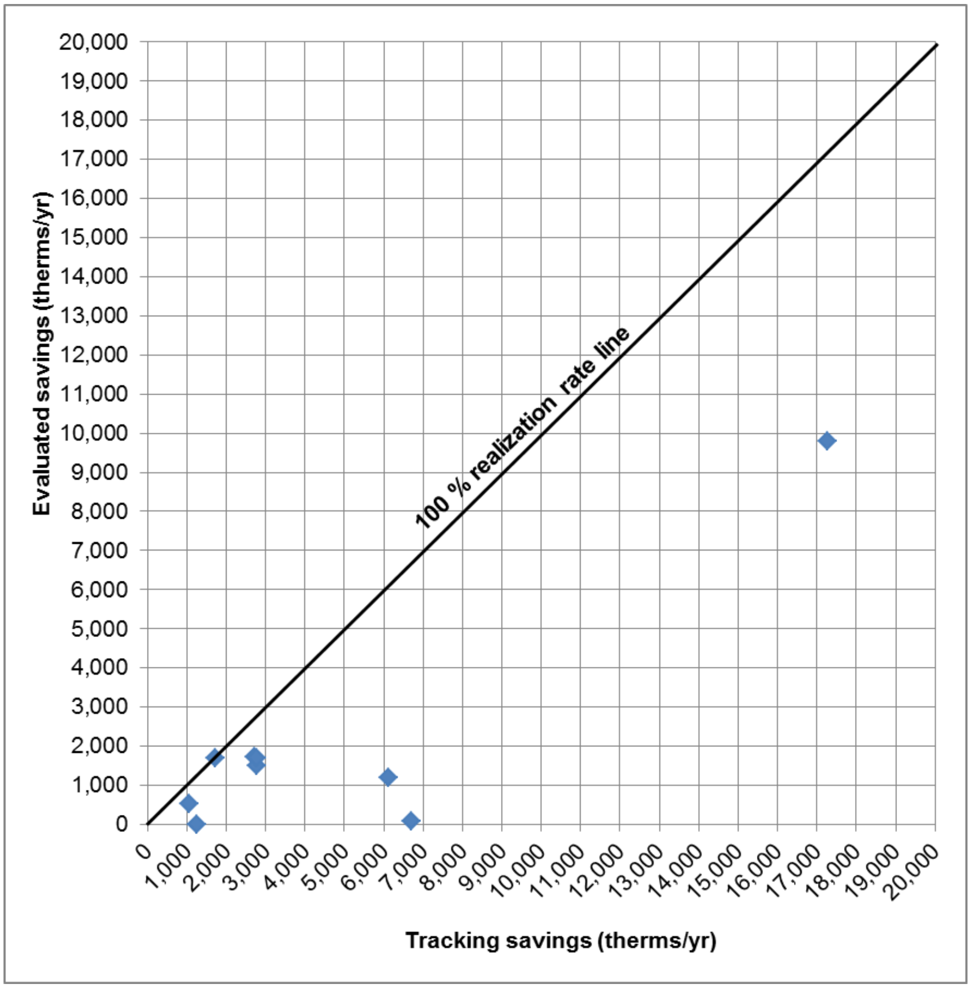


The gas EMS projects involved installing controls on boiler systems. Savings for a vast majority of these projects were derived from unsubstantiated savings factors. This situation can be immediately remediated by increasing the focus on reviewing the project applications and their associated savings calculations.

4.4.4 Gas – Tune-Up

The gas tune-up projects in the on-site sample included tuning-up the combustion controls on boilers and furnaces in a variety of facility types. The gas programs funded a total of nine tune-up projects, all of which were funded by the Gas Rebate program. Figure 4-17 shows the overall performance of the gas tune-up measures. The plot shows that with the exception of one project, the rest of the tune-up measures significantly underperformed.

Figure 4-17. Gas Rebate Program – Tune-Up Measures Performance



The evaluation team looked closely at this suite of measures and conducted a survey of the boiler and furnace tune-up projects. Table 4-18 highlights the evaluation findings associated with the parameters listed in the NYTM for the different facility types.

Table 4-18. NYTM Relevant Parameters for VFD Measures

Site Type	Count	Applicant Savings Factor	Evaluated Savings Factor	Avg. NYTM EFLH	Avg. Evaluated EFLH
Dormitory	3	4.6%	1.7%	681	856
Gymnasium	1	3.6%	2.7%	799	604
Multifamily	2	6.1%	2.2%	589	800
Retail	2	4.0%	1.8%	625	835
Total/avg.	**8**	**4.6%**	**2.1%**	**673**	**774**

Performance differences between building types was not significant. The pattern that did emerge was the consistently lower evaluated savings across all types. For most of the projects applicants used the default NYTM specified 5% energy savings factor along with custom site heating hours to estimate the savings.[12] Through a review of numerous tune-up projects in the current evaluation sample and other literature sources, the evaluation team believes that the 5% savings factor significantly overestimates the savings associated with the tune-up projects. The evaluation team recommends that the program staff either base their savings on custom pre- or post-tune-up measurements or use 2% as a savings factor, which is found to be more reasonable and representative of savings that can be achieved from these tune-ups. If the policy makers choose to create a new measure for commercial boiler tune-up, then the evaluators suggest using 2% savings factor as a default. Due to the limited sample size, the evaluators do not recommend making a change in the current NYTM savings factor associated with residential furnace tune-up.

In addition to the above recommendation, the evaluation team also suggests that the program staff pay extra attention to projects involving the tune-up of facilities with multiple boilers or furnaces because it was found that the same equivalent full load heating hours (EFLH) were being applied to all boilers even though some of the boilers were purely serving as backups.

The EFLH were also found to be slightly greater than the values stipulated in the NYTM, but overall there was too much variation in the EFLH values found at each site to draw broad conclusions.

4.5 Discrepancy Results

The evaluation team identified reasons for the evaluated savings being different from the applicant estimates of savings for each measure and then aggregated the results together to determine the largest source of positive (evaluated savings are higher than applicants') and negative (evaluated savings are lower than applicants') discrepancies. The goal was to help identify corrective actions the program can take to improve the accuracy of savings estimates going forward.

---

[12] The NYTM presents the 5% value in the context of a residential furnace tune-up measure, but the study on which it was based was actually of commercial boilers and furnaces. It is reasonable for Con Edison administrators to have accepted the 5% basis. The 5% value is presented with the observation that "Energy savings on the order of 2% - 5% were realized from a boiler tune-up program in the Pacific Northwest." See Dethman and Kunkle, *Building Tune-Up and Operations Program Evaluation*. Energy Trust of Oregon, 2007. In that study, however, the evaluation team stated that "The PMC and Energy Trust staff were all surprised and disappointed at the boiler tune-up savings, which were significantly below program projections. Based on research showing typical boiler tune-up savings of 2% to 5%, the program conservatively assumed 2% savings. The actual savings were about half that amount. "A high degree of measurement uncertainty" was acknowledged (p. 27). This evaluation also has admittedly high uncertainty regarding the savings basis, as an ex post-only evaluation such as this must rely on non-observed performance data or secondary sources to characterize pre-measure conditions.

The underlying reasons for these discrepancies vary widely, from incorrect values entered into the tracking database, or correct values entered into the incorrect data column, to subtle differences in the operating profile of a chiller. The results are presented separately below for the electric and gas programs.

4.5.1 Electric Measures Discrepancy Categories

Figure 4-18 defines the electric measures discrepancy categories with a description of the discrepancy category and its overall impact on the gross evaluation results.

4.5.2 Electric Measures Discrepancies

Broadly speaking, the discrepancies appear to show the administrative pangs of rolling out a new program, with a fairly high rate of administrative errors and an application review process that failed to screen out ineligible measures and to consistently assign correct baselines. The process for evaluating applicant claims of savings may also be immature, where a more skeptical eye may have questioned the high savings claims made for EMS (as indicated by the EMS and controls category).

However, the applicant estimates of equipment performance, as indicated by the load profile and operations profiles was quite reasonable, showing some overestimation and some underestimation, which is to be expected given the difficulty of forecasting as-built operation. The analysis also shows that using NYTM assumptions did not adversely affect the results, revealing some overestimating and some underestimating, with a small net underestimate of savings.

The program's post-implementation quality control appears to be working well because little deviation was observed in the installed equipment from what had been proposed in the application.

Figure 4-18 summary of the program-wide impact on RRs for each of the categories. The figure shows the impact of both discrepancies that decrease the RR (shown in red) and those that increase the RR (shown in blue) as well as the number of measures where the discrepancy was found. Some of the categories are further grouped by where in the implementation cycle the discrepancies are likely to arise. For example, in an application review, the reviewer has the opportunity to consider whether the applicant baseline is correct, the measure is eligible, and the savings are reasonable.

Figure 4-18. Electric Measures Sources of Discrepancies

Category	Description	Negative Impact on RR	Positive Impact on RR
Administrative: 46 measures	This category accounts for typographic errors, failure to update tracking with application revisions, incorrect extraction of savings from spreadsheets and the like. (L126)	-8%	2%
Baseline: 33 measures	This category accounts for a baseline adjustment from early replacement to normal replacement. (L050)	-6%	1%
VFD screening method: 10 measures	This category accounts for those VFD measures that did not meet program eligibility criteria or clearly violate program tenets. (L077)	-8%	0%
Unsubstantiated savings claims: 7 measures	This category accounts for savings estimates based on a claimed savings fraction which was not supported with site based analysis, measurement, or evidence from an applicable study. (L097)	-2%	0%
Inoperable Equipment: 1 measures	This category accounts for the equipment discovered to not be operational or controlled as described in the project documents. (LR21)	-2%	0%
Quantity or size: 41 measures	This category accounts for the differences in the quantity or size of an installed measure when compared with the project documents. (L062)	-1%	0%
Technology: 27 measures	This category accounts for the differences in the actual baseline and installed technologies observed by the evaluators versus the project documented technologies. (L134)	-1%	0%
Applicant used deemed value: 16 measures	This category accounts for differences in the NYTM specified parameters (usually hours of operation) and the evaluated value. (L134)	-2%	3%
Interactivity: 38 measures	This category accounts for the interactive effects of measures. A good example of measure interactivity is lighting which typically has interactive cooling and heating effects. (LR02)	0%	1%
Operation/Load Profile: 81 measures	This category accounts for the deviations in the projected equipment load profile (part load factor, part load profile, or temperature profile) or the operational schedule. (L106)	-8%	3%
Other/Weather: 2 measures	This category accounts for the weather normalization applied by the evaluators. (L084)	-1%	0%

4.5.3 Gas Measures Discrepancy Categories

Figure 4-19 defines the gas measures discrepancy categories with a description of the discrepancy categories and its overall impact on the gross evaluation results.

4.5.4 Gas Measures Discrepancies

The gas programs show trends that are similar to those observed in the electric programs. The gas programs show a fairly high rate of administrative errors and an application review process that failed to screen out ineligible measures and to question savings claims that were unsupported by site-specific information or other evidence. This was often observed for boiler control measures where the savings estimate was, for example, based on a vendor claim that the equipment would save "a high percentage (15%-30%) of the bills."

Similar to the electric programs, the applicant estimates of equipment performance, as indicated by the load profile, and operations profiles, and the NYTM assumptions did not adversely affect the results.

The program's post-implementation quality control appears to be working reasonably well because little deviation was observed in the installed equipment from what had been proposed in the application.

Figure 4-19 presents a summary of the program-wide impact on RRs for each of the categories. The figure shows the impact of both discrepancies that decrease the RR (shown in red) and those that increase the RR (shown in blue) as well as the number of measures where the discrepancy was found. Some of the categories are further grouped by where in the implementation cycle the discrepancies are likely to arise.

Figure 4-19. Gas Measures Sources of Discrepancies

Category	Description	Negative Impact on RR	Positive Impact on RR
Administrative: 16 measures	This category accounts for typographic errors, failure to update tracking with application revisions, incorrect extraction of savings from spreadsheets and the like. (L062)	-9%	5%
Baseline: 3 measures	This category accounts for a baseline adjustment from early replacement to normal replacement. (L029)	-1%	1%
Ineligible measure: 2 measures	This category accounts for measures that the project files reported affecting equipment using gas (RTUs) while the evaluators found that the affected equipment does not use gas. (L112)	-2%	0%
Unsubstantiated savings claims: 6 measures	This category accounts for savings estimates based on a claimed savings fraction which was not supported with site based analysis, measurement, or evidence from an applicable study. (L006)	-11%	0.3%
Quantity or size: 3 measures	This category accounts for the differences in the quantity or size of an installed measure when compared with the project documents. (LR50)	-1%	0%
Technology: 6 measures	This category accounts for the differences in the actual baseline and installed technologies observed by the evaluators versus the project documented technologies. (L029)	-3%	0.3%
Applicant used deemed value: 10 measures	This category accounts for differences in the NYTM specified parameters (usually hours of operation) and the evaluated value. (L107)	-3%	2%
Operation/Load Profile: 14 measures	This category accounts for the deviations in the projected equipment load profile (part load factor, part load profile, or temperature profile) or the operational schedule. (LR42)	-2%	2%
Other/Weather: 4 measures	This category accounts for the weather normalization applied by the evaluators. (L099)	-2%	0%

5 CONCLUSIONS AND RECOMMENDATIONS

The evaluation team presents several conclusions and recommendations for the four CECONY programs. These are given in the following sections.

5.1 Conclusions

The findings and conclusions are presented through two perspectives:

1. Sources of savings discrepancies
2. Equipment performance and New York Technical Reference Manual (NYTM) findings

The equipment performance and NYTM findings are combined, since the NYTM values are closely linked to the performance of a particular technology.

5.1.1 Sources of Savings Discrepancies

The evaluation team identified reasons why the evaluated savings were different from the applicant estimates of savings for each site. In general the applicant load profile and operations profile estimates were good. The NYTM assumptions likewise proved to be unbiased estimators. The program's post implementation quality control also appeared to be working well because little deviation was observed in the installed equipment from what had been proposed in the application.

The biggest contributors to the realizations rates (RRs) being less than 1.0 are:

- **Boiler controls savings calculation methods** – A majority of the boiler system controls estimates tended to be based on rules of thumb that resulted in high savings relative to evaluated findings. For boiler control measures in particular the savings estimate was repeatedly based on a vendor claim that the installed controls equipment would save 35% of the baseline annual gas use (site L006).
- **VFD replacement** – Evaluators found instances where applicants received incentives to replace failed VFDs in systems that could not sustain operation without them. In some of the failed VFD projects, we found the simple payback to be below the threshold defined by the program guidelines, which represented an additional level of some these measures being ineligible from the financial perspective. Because there was no long-term alternative to a VFD-driven system, the baseline was a VFD system; hence the savings were evaluated as zero.
- **Administrative matters** – This category accounts for such issues as typographical errors, tracking savings reflecting the results from an early – not final – application revision, and extraction of the incorrect savings value from an analysis spreadsheet. While the errors

had no pattern, their net effect was significant overestimation compared to evaluated savings.

- **Baseline** – The evaluation team adjusted savings downward in instances where there was judged to be end-of-life replacement instead of a retrofit. In such cases the baseline was new standard efficiency equipment rather than the preexisting equipment. The former tends to be more efficient than the latter, resulting in lower savings.

The previous sections of this report included detailed discussions of the reasons for the discrepancies with examples and quantification of the effect of each factor, positive and negative. Some of the deviations are not uncommon for recently started programs and the evaluation team would expect the RR to increase over time as these issues are addressed.

5.1.2 Equipment Performance Findings

Although the impact evaluation was not designed to provide specific measure results, some measures were sufficiently represented in the program to provide indication of measure-level performance. For rebate measures, key NYTM parameters, such as hours of operation, were compared to the evaluated results.

Tables 5-1 and 5-2 summarize the results by measure for electric and gas measures, respectively.

Table 5-1. Equipment Performance – Electric Measures

Program	Measure Category	Count	Gross RR	Error Ratio
Electric Custom	EMS and controls	4	71%	0.06
	Lighting	29	88%	0.29
	Motors and VFDs	5	34%	0.72
	Other	10	49%	0.50
Electric Rebate	EMS and controls	13	50%	0.61
	HVAC	4	99%	0.12
	Lighting	26	77%	0.49
	Motors and VFDs	18	71%	0.48
	Other	2	64%	0.26

Table 5-2. Equipment Performance – Natural Gas Measures

Program	Measure Category	Count	Gross RR	Error Ratio
Gas-Custom	EMS and Controls	1	101%	N/A
Gas Rebate	EMS and Controls	12	39%	0.45
	HVAC	1	76%	N/A
	Other	2	150%	0.66
	Tune-up	9	49%	0.27

N/A – Not applicable. Relative precision is not relevant for a sample quantity of 1.

Of the measures with a significant number of observations (greater than 5), the EMS and controls equipment group is the only measure group with a low RR (below 60% in general). As noted above in the discrepancy section, the evaluation engineers perceive the underperformance of this measure subgroup to be related more to the savings estimation methodology than equipment underperformance. The shell measures performed significantly worse than other gas measures and below average for electricity savings. The sample size is too small to treat the observation as other than anecdotal, but it fits a pattern that the evaluation team has observed in other NY State and New England utility program evaluations.

The custom-estimated lighting hours tended to be unbiased and more accurate than the applicable NYTM hours, suggesting that the NYTM hours should be used on a back-up basis for the LC&I programs.

VFDs had high variability in RRs but overall, if not for the failure replacement issue described in the discrepancy section, would have had a strong 93% RR.

The peak kW estimates had a high degree of variability as evidenced in the estimate presented in the table above. This data highlights the need for the program staff to pay attention to peak demand savings calculations and to accurately capture these estimates in the tracking system.

5.2 Program Recommendations

The recommendations are divided into those for the program itself relating to equipment and program administration, NYTM recommendations, and those concerning methods for improving future evaluations.

5.2.1 Specific Equipment Recommendations

The evaluation team recommends that the program administrators consider the following equipment-related changes:

Lighting

Revise the application documentation to require inclusion of the detailed space-by-space inventory of the fixtures affected in a given space. During the course of this evaluation, the investigating engineers had to rely on a supporting scope of work or site-provided documentation to understand the overall scope of the lighting projects. The supplied XACT project files did not convey details regarding the fixtures affected by space for a given project. If such details were incorporated into the project files, it would also enable the program staff to accurately conduct pre- and post-installation site inspections, which would ultimately help improve the overall accuracy of the tracking savings estimates that get entered into the LM Captures database. In addition, space-by-space inventories are not usually unduly burdensome for contractors to provide, since they are usually required for a cost estimate.

Electric – EMS and Controls

- On larger EMS projects (>250,000 kWh/yr or based on affected square footage), collect additional data related to baseline operating conditions and conduct post-installation inspections to verify that the planned EMS strategies are implemented correctly. The engineering analysis should also be specific to the site and should account for the weather effects. The baseline verification process should involve documenting the actual status of the current EMS strategies. The post installation inspection process should verify and document that the control strategies indicated in the initial scope of work are functioning. Preferably trend plots indicating conformance should be collected for future reference.
- For smaller sized EMS projects (<250,000 kWh/yr or based on affected square footage or based on cost effectiveness to the program) implemented on standard building types (offices, schools, etc.), we recommend developing a simple analysis tool for the program staff to estimate EMS savings based on the equipment affected, the planned EMS control strategies, and historic energy use. Such a tool could be a spreadsheet based model relying typical operating profiles that could be adjusted for the particular facilities energy use. It should have a menu based system to enable easy picking of the various EMS strategies that could then be used to estimate the savings for that particular project. There are no publicly available tools of this nature, but the evaluators have seen one in use developed by the Northeast Utilities for their use.

Gas – EMS and Controls/Tune-Up

- On boiler tune-up projects, apply actual measured pre- and post-tune-up combustion efficiency values along with appropriate annual heating hours in the savings calculations. In the absence of site-specific measured efficiency values, the evaluation team recommends using a 2% rather than 5% savings factor. Two percent is the average efficiency improvement measured for those sampled tune-up projects where both pre- and post-tune-up data were available. We also recommend crosschecking the savings with the billed usage to verify that the savings do not exceed the billed usage.

5.2.2 Program Administration Recommendations

This is an impact not process evaluation, but some of the team's observations made while conducting measurement and verification (M&V) suggest opportunities for administrative improvement. Based on observations from the site reports, the evaluation team recommends the following steps to improve program administration:

- **Administrative errors –** Data entry and update errors in the tracking system had a substantially negative impact on the RR. We recommend that a combined Con Edison/Lockheed Martin team convene to examine the data management process. An example review process could involve the following:
  - Periodic crosschecking of Lockheed Martin and Con Edison tracking savings
  - Automated crosschecking of the modification date of the most recent XACT file with a last update field in tracking to ensure the most recent version of savings
  - Automated range checks, such as savings per fixture or savings per kWh per measure type, which might help identify misplaced decimals
- **Application review policy, procedures, and training** – The nature and extent of the sources of discrepancies suggest that the application reviewers would benefit from additional support and training that provides the policy background, procedures, and reference materials to enhance application review. Reviewers also need adequate time to review the application package and ask follow-up questions if necessary. The evaluation team particularly recommends additional reviewer training:
  - Baseline: Look out for new construction or major renovation projects for which the baseline is dictated by prevalent code, not the preexisting conditions.
  - Peak demand: Focus more on peak electric demand savings, as it is expected to gain more importance in the future.
  - Trending: We recommend that Con Edison have a process for identifying vendor trends, reviewing their savings calculation methods on a routine basis, and advising them on acceptable calculation methods.
- **Establish measure review benchmarks** – It is clear that certain measures are underperforming. With a new program, there are fewer benchmarks by which to judge the performance of a measure; however, this evaluation has provided feedback on specific

measures that should be incorporated into the application review process. Specific benchmarks (e.g., boiler controls, boiler tune-ups, and EMS savings fractions) are further specified later in this report. As an example, we found that typical boiler controls and tune-up projects save approximately 2% of the baseline system operations instead of the 5% to 15% range claimed by the applicant in the current evaluation sample of projects. Providing such benchmarks to the reviewers and training them on watching out for these factors would help improve the overall accuracy of the savings estimates entered in the program database.

- **Reduce the effect of free ridership (FR)** – The FR estimates investigated through this effort were in the range of 0.26 to 0.38 and are within the typical range observed for similar LC&I programs offered throughout the country. However, further action could be taken to avoid future erosion of savings due to FR. These steps could involve the following-
  - Determining how customers are coming in to the program and not depending on "walk-ins"; seeking out participants rather than having them seek out the program.
  - Providing key technical assistance early in the project. Informing clients about the savings and rebate amounts *after* the project has begun does not seem to be having an effect on what they install.

5.2.3 Recommendations for New York Technical Manual

For boiler tune-up projects, our metered sample suggests that a 2% savings estimate is a better default savings factor to use in the absence of metered data than the 5% used by Con Edison. Our study was not focused on boiler tune-ups so we do not have sufficient data to recommend a change to the NYTM. If the policy makers choose to create a new measure for commercial boiler tune-up, then the evaluators suggest using a 2% savings factor as a default and using our data along with data from other evaluations to develop a number for large commercial and industrial customers. The current NYTM reference refers to furnace tune-ups for residential customers. Due to the limited sample size, the evaluators do not recommend making a change in the current NYTM savings factor associated with residential tune-ups.

5.2.4 Program Evaluation Recommendations

Upon completing this evaluation, the evaluation team has identified a few recommendations for ways to improve future Large C&I evaluations.

- **Emphasize evaluation commitment** – During the initial on-site recruiting phase of this evaluation, some of the largest savings projects declined to participate in the M&V effort. The evaluation team requested assistance from the program staff and were eventually informed that these customers were high value and hence could not comply with the follow-up evaluation M&V due to a variety of reasons. Even though the overall evaluation RP did not suffer drastically, the loss of these sites was regrettable and contributed to the reduction in the overall RP values of this evaluation. Therefore, the evaluation team would like to stress that the program staff may want to add language to agreements that encourages participation in evaluations.

- **"Continuous evaluation approach"** – Having evaluation activities (and more specifically, FR research) occur shortly after the implementation of the project can greatly increase the accuracy of the results and reduce nonresponse bias. Asking complex and inherently hard-to-answer counterfactual NTG questions several years after the completion of a project can greatly reduce respondent ability to make the needed estimates and consequently diminish the accuracy of the estimates. Decision-makers might not clearly remember what the influence of the program was on their decision to implement high efficiency improvements installed a while ago. Furthermore, over time, decision-makers might no longer be available for interviewing (e.g., they could have left the company, assumed a different position within the organization, etc.), which might result in nonresponse error. It is our understanding that CECONY is currently in the process of making the transition toward a "continuous evaluation approach."

The CECONY programs have come a long way in a very short amount of time. The kinds of issues highlighted in this section are symbolic of a new program going through its growing pains. Addressing the administrative and procedural types of recommendations mentioned in this section should help improve the overall performance of the CECONY programs significantly and could be implemented in a fairly short order.

APPENDIX A – SAMPLING METHODOLOGY

Participant Survey Sample Design

The goal of the telephone survey with program participants was to collect data that will allow an estimate of program attribution. The goal of the sample design is to achieve a target of 90% confidence with 10% precision on program impacts per the evaluation guidelines[13]. When developing a sampling approach, the evaluation team took into account the following factors, attempting to strike a balance between all of them:

1. Total size of the sample frame in terms of completed projects and unique participants (defined and companies and phone numbers)
2. Desired level of rigor
3. Difference in the decision-making process as related to various technologies

The analysis of the program tracking data identified 591 unique contacts across 985 unique projects for CECONY. Small participant population required a census attempt across all programs and all program participants.

Table A-1. Overview of Projects Completed as of the End of 2011

Program Type	Total Number of Unique Projects	Total Number of Unique Contacts
CECONY C&I Electric - Rebate	759	433
CECONY C&I Gas - Rebate	87	69
CECONY C&I Electric - Custom	265	211
CECONY C&I Gas - Custom	5	5
Total CECONY*	**985**	**591**

*Note that since a single project can include measures rebated through different programs, the total number of unique projects is lower than the sum of unique projects by program. The same is true for unique contacts.

The sampling unit for the survey was the project contact. Within the framework of the Large C&I program, a single contact can be associated with multiple projects, which in turn could be comprised of multiple end uses. The table below provides an overview of the quantity of end use types per project.

[13] August 7, 2008 (updated November 2012) Evaluation Guidelines issued by the DPS through the NYS Evaluation Advisory Group (EAG), http://www3.dps.ny.gov/W/PSCWeb.nsf/96f0fec0b45a3c6485257688006a701a/766a83dce56eca35852576da006d79a7/$FILE/EVALGUIDE.11.12.pdf

Table A-2. Quantity of End Use Types per Project

Program Type	CECONY
Number of projects with one end use	935
Number of projects with two end uses	26
Number of projects with three or more end uses	24
Average number of end uses per project	**1.08**

The table below provides an overview of quantity of projects per unique contact.

Table A-3. Quantity of Projects per Unique Contact

Program Type	CECONY
Number of unique contacts with one project	467
Number of unique contacts with 2 projects	72
Number of unique contacts with 3 or more projects	51
Average number of projects per unique contact	**1.7**

Because decision making varies by the type of project, the facility it is performed at, as well as the end use, to obtain the most accurate measurement of program influence, for each participant the evaluation team focused the interview on a specific end use completed as part of a specific project. We classified all measures into lighting, motors, HVAC, EMS, and other categories. Such breakdown is not too granular and will provide greater detail into the decision-making process.

For participants that were responsible for completing multiple projects across a variety of end uses, the evaluation team attempted, to the degree possible and feasible to capture the information across as many projects as possible. We asked those participants questions about whether the decision-making process was the same for all end uses or all projects in question. If so, the evaluation team were able to capture the savings associated with the end use categories/projects as part of a single interview. In complex cases, where a single contact is associated with multiple projects comprised of a variety of end uses, the evaluation team completed in-depth interviews in order to obtain as much detail on the decision-making process across as many projects and end uses as possible without placing too much burden on the respondent. The in-depth interviews generally followed the structure of the survey instrument provided with this document.

In cases when survey length and the resulting respondent fatigue was an issue, the evaluation team prioritized projects that were a part of the least populous program. If the contact had more than one project that fell under the same program type, the evaluation team selected the project at random. If a participant installed equipment that belonged to more than one end use category, the evaluation team randomly assigned the end use that the interview focused on.

Following data collection efforts, post-stratification weights were applied to align the sample with the frame and the participant sample.

The table below provides a breakdown of the unique contacts and unique projects in the population, as well as the total number of projects for which the evaluation team completed interviews, as well as the total number of contacts with whom the evaluation team completed interviews.

Table A-4. Overview of Projects Completed as of the End of 2011

Program Type	Total Number of Unique Projects	Total Number of Unique Contacts**	Total Number of Unique Projects for Which Interviews were Completed	Total Number of Completed Interviews***
CECONY C&I Electric - Rebate	759	433	87	51
CECONY C&I Gas - Rebate	87	69	3	3
CECONY C&I Electric - Custom	265	211	23	23
CECONY C&I Gas - Custom	5	5	0	0
Total CECONY*	**985**	**591**	**113**	**77**

*Note that since a single project can include measures rebated through different programs, the total number of unique projects (in the population as well as for which interviews were completed) is lower than the sum of unique projects by program. The same is true for unique contacts and the total number of unique completed interviews.

**Note that the number of unique contacts does not match what appears in the disposition reports due to different/additional contact information that is uncovered during the interviewing process.

***Includes three mid-interview terminate completes.

Trade Ally Sample Design

Because the Large C&I program used various tactics to influence the way trade allies (equipment vendors, installation contractors, designers, engineers, etc.) specify projects to customers, and because commercial and industrial customers can be heavily influenced by their trade allies in the decision to install high efficiency equipment, it was important to capture and integrate the influence of the program on trade allies into the final NTG ratio.

To do that, the evaluation team completed interviews with trade allies. The interviews were triggered by participant responses to the telephone survey – the evaluation team only completed interviews with trade allies that were involved in projects where participant-rated influence of the trade ally on the recommendation of high efficiency equipment was high. At the completion of the participant survey, the evaluation team arrived at the sample frame of 23 trade allies. We tried to contact every trade ally from the list multiple times and were able to complete nine interviews.

Program Participant Survey Sample Disposition

The program participants were surveyed from July 22, 2013 through August 28, 2013. The evaluation team completed a total of 77 interviews. The telephone interviews were conducted using a Computer-Assisted Telephone Interviewing (CATI) system. Table A-5 shows the final survey dispositions for the participant survey.

To minimize the measurement error, the survey was tested internally for comprehension. Additionally, it was pre-tested with several participants to ensure that survey questions are interpreted correctly and answered in a consistent manner.

Table A-5. LC&I Program Participant Survey Dispositions

Disposition	N
Completed interviews	74
Partial completes	20*
Eligible non-interviews	304
Refusals	165
Break off	0
Telephone answering device	84
Respondent never available	52
Language problem	3
Not eligible	82
Fax/data line	3
Non-working	25
Wrong number	22
Business/government	1
No eligible respondent	28
Duplicate number	3
Unknown eligibility non-interview	99
Not dialed/worked	8
No answer	88
Busy	3
Call blocking	0
Total participants in sample	579

*Note that the evaluation team included three of these interviews in the analysis, as they were complete enough to allow for analysis.

Table A-6 provides the response and cooperation rates. The survey response rate is the number of completed interviews divided by the total number of potentially eligible respondents in the sample. The evaluation team calculated the response rate (RR3) using the standards and formulas set forth by the American Association for Public Opinion Research (AAPOR).[14]

The evaluation team also calculated a cooperation rate, which is the number of completed interviews divided by the total number of eligible sample units actually contacted. The cooperation rate gives the percentage of participants who completed an interview out of all of the participants with whom the evaluation team actually spoke. This evaluation used AAPOR Cooperation Rate 1 (COOP1).

Table A-6. LC&I Program Participant Survey Response and Cooperation Rates

AAPOR Rate	%
Response rate (RR3)	19%
Cooperation rate (COOP1)	29%

There are multiple sources of non-sampling error that can impact survey results, including non-response error and resulting coverage bias. The evaluation team tried to mitigate the non-response bias through the fielding process by taking the following steps:

- We called every program participant at least six times at varying times of the day and week over an extended period of time. Most participants received over 10 calls during the fielding process. We left multiple voicemails in cases when we got an answering machine.
- For customers with non-working phone numbers, we requested alternative numbers from Con Edison. If unavailable, we searched the Internet to find alternative contact information for program participants.
- For participants with email addresses, we followed up via email to invite them to participate.
- We offered participants $100 incentives for completing the survey with us.
- We fielded the survey for over a month.

In addition, in the NTG calculations we applied post-stratification weights to account for non-response bias.

Trade Ally Interview Dispositions

The evaluation team conducted interviews with nine trade allies who were rated by surveyed Large C&I program participants as being very influential in the participant decision to install

[14] *Standard Definitions: Final Dispositions of Case Codes and Outcome Rates for Surveys*, AAPOR, 2011. http://www.aapor.org/AM/Template.cfm?Section=Standard_Definitions2&Template=/CM/ContentDisplay.cfm&ContentID=315

high efficiency equipment. Interviews took place from September 26, 2013 through October 23, 2013. The sample frame consisted of 23 trade allies. Trade allies were offered $100 incentives for completing an interview.

On-Site Sample Stratification and Design

The on-site sample supports the evaluation of measure savings installed by CECONY as part of their Large C&I programs. Evaluator engineers verified measure installation and installed data loggers to measure key parameters appropriate for verifying savings. The installation verification supported the general assessment of each program's performance for the 2009–2011 program cycle.

The evaluation team selected a stratified random sample of sites according to the sample design outlined below. To comply with NY DPS guidelines the sample must be sufficient to estimate realization rates and other savings parameters for each C&I program by Company - with at least ±10% precision at a 90% confidence level (90%/10% precision).

Sample Size and Allocation for On-Site M&V

A key determinant of sample sizes and anticipated precisions is the amount of variability, often described as an error ratio, which is likely to exist from site to site in the parameter being estimated. For the on-site sample, the parameter being estimated is the realization rate for kWh (or therms) savings. Since this is the first time that the Companies are evaluating the Large C&I programs, there is no history upon which to base an estimate of the error ratio. Therefore, the evaluation team examined prior studies in other jurisdictions to get a sense of what the error ratio might be. Previous evaluations have indicated that lighting measures tend to have fairly low error ratios (around 0.4) while other custom measures can be more variable. Several custom gas and electric program evaluations completed by the evaluation team yielded error ratios between 0.5 and 0.6. The master evaluation plan assumed an error ratio of 0.7. After examining the detailed project files evaluators concluded that this assumption is likely to be excessively conservative, as 45% of the Con Edison and 72% of the O&R projects' acquired savings to date are lighting. An even higher percentage of the projects are lighting. An error ratio of 0.6 was assumed for this design. The true error ratio will not be known until the analysis is conducted.

Given an assumed error ratio of 0.6, several alternative designs were considered. The evaluation team uses the "site" as the sampling unit. The "site" is defined as all measures funded by a unique program at a unique facility during the sample frame period.[15] The team recommended

[15] For example if Joe's Garage at 123 Main Street received one prescriptive lighting rebate in 2010, a prescriptive motor rebate 2011 and received a gas rebate for a new boiler, this constitutes three separate rebate applications but two "sites" for sample design purposes - one for the electric prescriptive rebate

a sample of 140 sites to produce acceptable results (at least ±10 precision at 90% confidence) for the five programs. The MBSS algorithms were applied to allocate the 140 sites to the five programs (with the gas rebate and custom programs combined) in a way that ensures precision targets will be met. Within each program, sites are first sorted by size (kWh or therms), from smallest to largest savers. Second, the sites in each program are grouped into one of five or six strata for each program, with the grouping based on site's annual kWh savings (where gas savings are converted to therms for presentation purposes). The second column of Table A-7 shows the strata. The stratification boundaries are set so that the total kWh savings in each stratum are somewhat similar. Thus each site is grouped with other similarly sized projects in each stratum. The "Maximum kWh Savings" per Project column shows the size boundaries, and the "Total kWh Savings" column illustrates the similar total savings per stratum within program. Third, the sample is allocated. As with savings, the sample points are allocated so that the number of sample points and the total kWh savings are spread somewhat evenly across the size strata. The sites with the largest savings in each program are selected with certainty. The "Planned Sample" column illustrates this allocation. In an ideal exercise, the consequence of making the total energy associated with each stratum the same and also sampling an equal number of sites from each stratum would be that each evaluated sample site would have about the same influence on the final results as the other sampled sites. The last column, Inclusion Probability, is simply the Planned Sample (n for the stratum) divided by the Sites (N for the stratum). It reflects the likelihood that any given project in the stratum will be selected in the random sampling process.

program and one for the gas rebate program. If Joe's Garage on Broadway, owned by the same firm but at a different location, also received a prescriptive lighting rebate in 2011, it would constitute a third site. This approach maximizes the efficiency of engineering resources by enabling them to evaluate multiple measures at one premise while still maintaining separate analyses for each program.

Table A-7. Sample Stratification and Allocation

Program	Stratum	Maximum kWh* Savings	Sites	Total kWh* Savings	Planned Sample	Inclusion Probability
O&R Existing Buildings Electric	1	76,303	10	364,882	3	0.3000
	2	150,523	5	544,260	2	0.4000
	3	228,414	3	586,819	2	0.6667
	4	362,563	2	651,864	2	1.0000
	5	1,410,400	4	3,688,396	4	1.0000
CECONY Electric Rebate	1	43,031	258	5,481,584	10	0.0388
	2	98,959	101	6,829,368	10	0.0990
	3	218,370	57	7,759,339	10	0.1754
	4	546,746	28	9,407,557	10	0.3571
	5	1,281,227	15	11,522,531	9	0.6000
	6	6,626,924	9	26,037,877	9	1.0000
CECONY Electric Custom	1	61,989	101	2,890,706	9	0.0891
	2	140,095	38	3,540,438	9	0.2368
	3	245,452	21	4,170,173	9	0.4286
	4	392,524	14	4,612,893	9	0.6429
	5	665,139	10	5,258,729	8	0.8000
	6	2,608,854	4	6,465,499	4	1.0000
CECONY Gas Rebate	1	115,237	21	1,203,175	4	0.1905
	2	274,600	8	1,549,413	4	0.5000
	3	381,838	5	1,687,094	4	0.8000
	4	506,099	5	2,327,826	4	0.8000
	5	1,785,278	6	5,458,854	6	1.0000
CECONY Gas Custom	1	145,416	1	145,416	1	1.0000
	2	5,310,537	2	9,409,959	2	1.0000
Total			**728**	**121,594,652**	**144**	

*Gas therms savings were converted to equivalent kWh at 29.3 kWh per therm.

Table A-8 provides the anticipated precisions for this design. Anticipated precisions are based on the current data available at this time. The final results may achieve better or worse precision levels, depending on the actual variability of the sites selected for evaluation.

Table A-8. Anticipated Precisions

Program	Sites	Total kWh* Savings	Error Ratio	Confidence Level	Planned Sample Size	Anticipated Relative Precision	Error Bound
Electric Rebate	468	67,038,254	0.6	90%	58	±9.22%	6,178,214
Electric Custom	188	26,938,438	0.6	90%	48	±9.52%	2,563,604
Gas Rebate	45	12,226,363	0.6	90%	22	±9.99%	1,221,543
Gas-Custom	3	9,555,375	0.6	90%	3	±0.00%	0
Total	**728**	**121,594,652**	**0.6**	**90%**	**144**	**±5.61%**	**6,823,787**

* Gas therms savings were converted to equivalent kWh at 29.3 kWh per therm

For planning and budgeting purposes a preliminary sample was selected. At that time, the evaluation team also looked at the distribution of measures included at these sites and compared them to the population. The measure counts and savings are included as Table A-9 and Table A-10 respectively. The overall population and sample counts in these tables are higher than previously shown because one unique site and program row may have multiple measure types installed. Tables A-9 and A-10 are a combination of CECONY sites by measure type used to help identify general trends and illustrate that post-stratification is going to be reasonably representative of the measures in the programs; these tables are not used for the actual sample design process.

Table A-9. Distribution of Sites by Measure Type

Measure Type	Population Count	% of Population	Sample Count	% of Sample	Sample % of Population Count
Chillers and unitary	4	0.5%	4	2.3%	100.0%
EMS and controls	107	14.4%	38	21.5%	35.5%
Furnaces and boilers	19	2.6%	11	6.2%	57.9%
Lighting	500	67.5%	80	45.2%	16.0%
Motors and VFDs	69	9.3%	31	17.5%	44.9%
Other	15	2.0%	3	1.7%	20.0%
Process	7	0.9%	0	0.0%	0.0%
Shell	6	0.8%	5	2.8%	83.3%
Tune-up	13	1.8%	4	2.3%	30.8%
Unknown	1	0.1%	1	0.6%	100.0%
Total	**741**	**100.0%**	**177**	**100.0%**	**23.9%**

Table A-10. Distribution of Savings by Measure Type

Measure Type	Population kWh* Savings	% of Population Savings	Sample kWh* Savings	% of Sample Savings	Sample % of Measure Savings
Chillers and unitary	238,033	0.2%	92,027	0.1%	38.7%
EMS and controls	31,022,873	26.8%	26,314,872	33.6%	84.8%
Furnaces and boilers	3,470,636	3.0%	2,833,010	3.6%	81.6%
Lighting	37,505,749	32.4%	15,596,032	19.9%	41.6%
Motors and VFDs	36,238,163	31.3%	29,615,806	37.8%	81.7%
Other	2,463,340	2.1%	715,255	0.9%	29.0%
Process	864,649	0.7%	-	0.0%	0.0%
Shell	2,386,366	2.1%	2,271,129	2.9%	95.2%
Tune-up	1,316,976	1.1%	623,445	0.8%	47.3%
Unknown	251,644	0.2%	251,644	0.3%	100.0%
Total	**115,758,430**	**100.0%**	**78,313,220**	**100.0%**	**67.7%**

*Gas therms savings were converted to equivalent kWh at 29.3 kWh per therm.

These tables illustrate several points. The vast majority of projects and savings are in three major measure categories: EMS and controls, lighting, and motors/VFDs. Lighting projects dominate the counts, but represent a lower percent of savings. The samples selected represent these three measure types well. The sample allocation, which gives large projects a higher probability of selection, has resulted in a sample that represents 70% of the population savings. In the three major measure types, the percent of savings represented ranges between 47% and 85%. For these subgroups, the results produced by post-stratification should be reliable and informative, although the precision levels may be above ±10%.

This section provides a brief summary of the model-based statistical sampling (MBSS) methodology. This methodology has been applied to the evaluation of energy efficiency program impacts by DNV KEMA for over 20 years and has proven to be effective and efficient. It is also widely accepted by the industry as the foundation for load research sampling. In energy efficiency impact evaluations, the goal is typically to estimate savings realization rates as the ratio between evaluated savings and preliminary estimates of savings recorded in tracking system data.

Statistical Methodology

This section provides a brief summary of the model-based statistical sampling (MBSS) methodology. This methodology has been applied to the evaluation of energy efficiency program impacts by DNV KEMA for over 20 years and has proven to be effective and efficient. It is also widely accepted by the industry as the foundation for load research sampling. In energy efficiency impact evaluations, the goal is typically to estimate savings realization rates as the ratio between evaluated savings and preliminary estimates of savings recorded in tracking system data. Conventional methods for sample design and estimation are documented in standard texts such as Cochran's *Sampling Techniques*[16]. MBSS is grounded in theory of model-assisted survey sampling developed by C.E. Sarndal[17] and others[18]. MBSS methodology has been applied in load research and impact evaluation for more than 30 years. This fusion of theory and practice has led to important advances in both model-based theory and practice, including the use of the error ratio for preliminary sample design, the model-based methodology for efficient stratified ratio estimation, and effective methods for domains estimation.

MBSS and conventional methodologies are currently taught in the AEIC Advanced Methods in Load Research seminar. MBSS methodology is also documented in *The California Evaluation Framework*[19]. It has been used in countless load research and program evaluation studies and has been examined in public utility hearings and in numerous EPRI studies.

The Role of the Statistical Model

As the evaluation team have stated, MBSS uses a statistical model to guide the study planning and the sample design. The parameters of the model, especially the error ratio, are used to represent prior information about the population to be sampled. The model describes the nature of the variation in the relationship between any target y variable of the study and one or more x variables that can be developed from known tracking data and other supporting information. The y variable can be any of the measurements taken at the evaluated site or any function thereof. In impact evaluations, the x variable is usually the tracking system savings estimate. The model is used to help choose the sample size n, to assess the expected statistical precision of any sample design and to help formulate a sample design that is efficiently stratified for ratio estimation.

The model is used as a guide to the sample design, but the results of the study itself are not strongly dependent on the accuracy of the model[20]. Once the sample design is selected, the subsequent analysis of the data is usually based only on the sample design and not on the model used to develop the sample design. In particular, conventional stratified-sampling techniques can be used to analyze the sample data collected from an MBSS sample design. The

---

[16] W. G. Cochran, *Sampling Techniques,* 3rd ed. (Wiley, 1977).

[17] Carl Erik Sarndal, Bengt Swensson and Jan Wretman, *Model Assisted Survey Sampling (*Springer-Verlag, 1992).

[18] Wright, R. L., "Finite Population Sampling with Multivariate Auxiliary Information," *Journal of the American Statistical Association,* 78, (1983): 879-884.

[19] The report can be downloaded from the site http://www.calmac.org/calmac-filings.asp

[20] Other methods, called "model-dependent sampling" approaches, are much more dependent on the accuracy of the model. Such methods are not commonly used in load research or evaluation applications since they are more difficult to defend than MBSS and conventional methods.

resulting estimates will be essentially unbiased in repeated sampling and the confidence intervals will also be valid, provided that the sample design has been followed to select the sample customers. The results will be consistent with traditional sampling theory as found in texts such as Cochran's Sampling Techniques and consistent with standard load and market research practice.

Stratified Ratio Estimation

We assume that an impact evaluation study is to be conducted of a given population of N projects in a given program. In the study, the sample sites will be monitored and savings determined based on actual loads and observed operations. We let y denote any characteristic to be determined from the on-site evaluation, and the evaluation team let x denote any suitable characteristic of the site.

We define the population ratio B by the equation:

$$B = \frac{\sum_{i=1}^{N} y_i}{\sum_{i=1}^{N} x_i}$$

Here the summations are over the entire N units (e.g., accounts, measures or projects) in the target population. We note that the population mean or total of y is equal to B times the population mean or total of x. The latter is assumed to be known from the tracking data.

We assume that a sample of n sites is selected following a stratified sample design. For each sample customer, the evaluation team defines the case weight w to be equal to the number of sites in the target population within the stratum containing the given site divided by the number of sites in the sample within the given stratum. The case weight is used to avoid any bias that might otherwise arise from the different sampling fractions used from one stratum to another.

Using the case weight, the evaluation team defined the combined ratio estimator of B by the equation:[21]

[21] This equation gives the same result as the conventional stratum-weighted equation: $b = \frac{\sum_{h=1}^{L} N_h \bar{y}_h}{\sum_{h=1}^{L} N_h \bar{x}_h}$

$$b = \frac{\sum_{i=1}^{n} w_i \, y_i}{\sum_{i=1}^{n} w_i \, x_i}$$

Then, if desired, the population mean or total of *y* can be estimated as *b* times the population mean or total of *x*, known from the tracking data.

Using the case weights, the evaluation team calculated the relative precision at the 90% level of confidence in three steps:

1. Calculate the sample residual for each unit in the sample.
2. Calculate[22] $se(b) = \frac{\sqrt{\sum_{i=1}^{n} w_i \left(w_i - 1\right) e_i^2}}{\sum_{i=1}^{n} w_i \, x_i}$
3. Calculate $rp = \frac{1.645 \, se(b)}{b}$

A 90% confidence interval for *B* is calculated using the equation. A confidence interval for the mean or total can be calculated in a similar way.

We can also use the sample data to estimate a measure of population variability called the error ratio, denoted *er*. The error ratio, defined in the next section, is the key determinant of the expected relative precision, along with the sample size *n*. We estimate the error ratio from the sample using the following equation:

$$\hat{er} = \frac{\sqrt{\left(\sum_{i=1}^{n} w_i \, e_i^2 / x_i^{\gamma}\right)\left(\sum_{i=1}^{n} w_i \, x_i^{\gamma}\right)}}{\sum_{i=1}^{n} w_i \, y_i}$$

[22] The conventional equation is $se(b) = \frac{1}{\sum_{h=1}^{L} N_h \, \bar{x}_h} \sqrt{\sum_{h=1}^{L} N_h^2 \left(1 - \frac{n_h}{N_h}\right) \frac{s_h^2(e)}{n_h}}$ where $s_h^2(e) = \frac{1}{n_h - 1} \sum_{i=1}^{n_h} \left(e_i - \bar{e}\right)^2$. Our equation assumes that $\frac{1}{n_h - 1} \sum_{i=1}^{n_h} \left(e_i - \bar{e}\right)^2$ is approximately equal to $\frac{1}{n_h} \sum_{i=1}^{n_h} \left(e_i\right)^2$ in each stratum.

The parameter γ (gamma) is also defined in the next section. In load research and evaluation applications it is usually taken to be 0.8. We will not attempt to interpret the preceding equation here, but the evaluation team will define both the error ratio and gamma in the following section.

A key advantage of the MBSS methodology is the ease of domains estimation. A domain is any identifiable subset of the population, e.g., the sites in a particular region or having a particular appliance or end use. Domain estimation is the process of obtaining the results of interest for one or more domains. With the MBSS methodology, domains estimation is very straightforward. We usually calculate the case weights for each sample site to reflect the sample design and current population and then regard them as fixed for any domains analysis. Then the evaluation team simply evaluate the preceding equations for the sample sites that are included in each domain.[23]

The Ratio Model

The ratio model is used to choose the appropriate sample size *n*, to assess the expected statistical precision of any stratified sample design, and to develop an efficiently stratified sample design. The ratio model describes the relationship between *y* and *x* for the set of all units in the population. The model consists of two equations called the primary and secondary equations respectively.[24]

$$y_i = \beta\, x_i + \varepsilon_i$$

$$\sigma_i = sd(\varepsilon_i) = \sigma_0\, x_i^{\gamma}$$

Here *i* denotes any customer, account, or premise in the target population. $x_i > 0$ is usually known throughout the population. The primary equation describes the relationship between the *y variable* of interest and the *x* variable used in the ratio estimate, i.e., annual use. Since the evaluation team assume that $E(\varepsilon_i) = 0$, the primary equation can also be written as $\mu_i = E(y_i) = \beta\, x_i$. Here μ_i denotes the expected value of *y* for unit *i*. The primary equation says that under the model, the expected value of y_i is equal to a fixed constant β times the known x_i

[23] In the software, a domain is any class or sector.

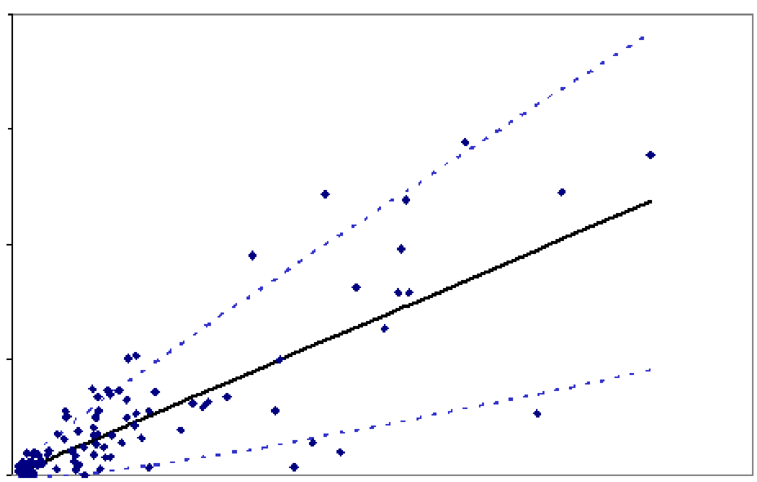
[24] The x variable in the primary equation is sometimes different than the x variable in the secondary equation. In the SAS modules, we refer to the latter as the stratification variable. For simplicity, we will not make this distinction in the theoretical discussion given here.

The quantity, $\varepsilon_i = y_i - \mu_i$, is called the residual. The N residuals are considered to be N independent random variables. The standard deviation of ε_i is denoted as σ_i. We refer to σ_i as the residual standard deviation of each customer *i*. The secondary equation is used to estimate the residual standard deviation and to guide the development of an efficient sample design.

To summarize, under the ratio model, the target variable y_i is a random variable with expected value μ_i and standard deviation σ_i. The expected value μ_i is determined by the primary equation of the model. The standard deviation σ_i is determined by the secondary equation of the model. There are three parameters in the model: β (beta), σ_0 (sigma-naught), and γ (gamma).

Figure A-1 shows an example. The points of the scatterplot represent the values of (*x*, *y*) for each site in the population. The solid line represents the equation $y = \beta\ x$, i.e, the expected value of *y* given *x*. This is a line through the origin with slope given by the parameter β. The two dashed lines represent the equation $y = \beta\ x \pm \sigma$, i.e, the one-standard deviation interval around the expected value. Here $\sigma = \sigma_0\ x^{\gamma}$ so the dashed lines are determined by the two parameters σ_0 and γ.[25]

Figure A-1. The Ratio Model



[25] The role of gamma can be seen by rewriting this equation as $\log(\sigma) = \alpha + \gamma \log(x)$ where $\alpha = \log(\sigma_0)$. This shows that for each site in the population the log of sigma is a constant plus gamma times the log of the value of x for the site. Gamma is the slope in the relationship between the log of x and the log of sigma.

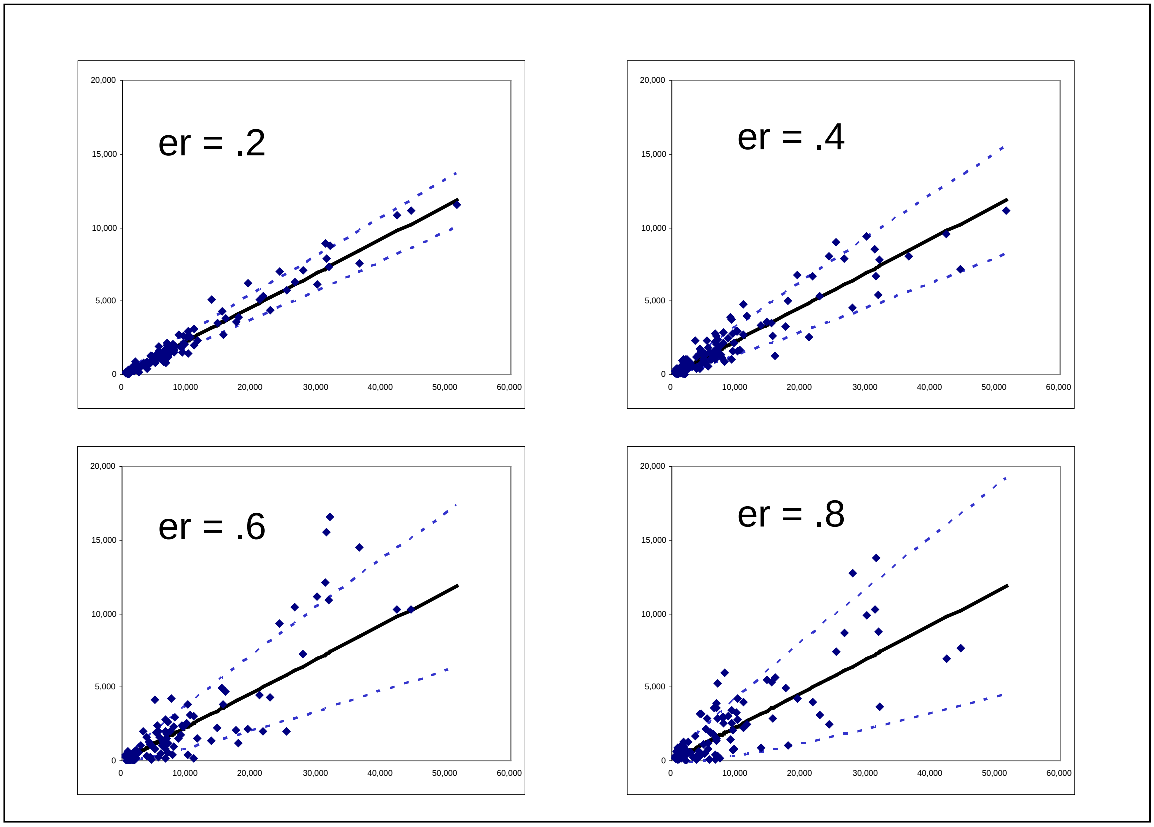
The error ratio is defined by the equation: $er = \frac{\sum_{i=1}^{N} \sigma_i}{\sum_{i=1}^{N} \mu_i}$

The error ratio can be regarded as an alternative parameter to σ_0 since under the preceding ratio model σ_0 can be calculated from the error ratio using the equation:

$$\sigma_0 = er \frac{\sum_{i=1}^{N} \mu_i}{\sum_{i=1}^{N} x_i^{\gamma}}$$

The error ratio is the key measure of variability when stratified ratio estimation is to be used to analyze the data. Figure A-2 shows some examples. If the error ratio is close to zero, there is a strong relationship between x and y. If the error ratio is larger, the relationship is weaker.

Figure A-2. Examples of Error Ratios



Choosing the Sample Size

We assume that the ratio model provides a reasonably accurate description of the relationship between *y* and *x* in the target population. We also assume that the sample design will be efficiently stratified as discussed in the following section and that the analysis will use stratified ratio estimation.

Under these assumptions and the added assumption that the population size *N* is large, then the expected relative precision is given by the equation:

$$rp = 1.645 \frac{er}{\sqrt{n}}$$

If the population is relatively small, the finite population correction factor can be added, giving

$$rp = 1.645 \sqrt{1 - \frac{n}{N}} \frac{er}{\sqrt{n}}.$$

If the desired relative precision *D* is specified, then the preceding equations can be solved to determine the required *n*. If the population size *N* is large, we have

$$n = \left(\frac{1.645\ er}{D} \right)^2$$

If the population is small, the sample size can be calculated in two steps. First, calculate $n_0 = \left(\frac{1.645\ er}{D} \right)^2$. Then calculate $n = \frac{n_0}{1 + n_0 / N}$

These equations are generally sufficient to develop a preliminary plan. However, additional issues must typically be addressed. First, there are usually many *y variable*s of interest, e.g., annual kWh savings, connected load savings, or savings in a peak hour. Second, it is often necessary to consider the expected statistical precision in various segments of the target populations. Third, there are usually limits on the sample size or other resource constraints.[26]

[26] When domains estimation is involved we may use two added results: (1) The standard error of the total of y across two or more mutually exclusive domains is the square root of the sum of the squared standard error of the total of y within each of the individual domains, and (2) for estimating the total of y across two or more mutually exclusive domains, the optimal allocation of the sample to each domain is proportional to the sum of the σ_i within each domain. The latter is equal to the error ratio within the domain times the expected value of the total of y within the domain.

Model-Based Stratification

The preceding results assume that the sample is efficiently stratified. Under the ratio model, an efficiently stratified sample design for ratio estimation can be developed in the following steps:[27]

1. Use the sampling frame and the assumed model to calculate σ_i for each customer in the population.
2. Choose the desired number of strata[28]
3. Sort the sampling frame by increasing σ_i.
4. Choose stratum cut points to divide the sum of the σ_i approximately equally between the strata.
5. Allocate an equal number of sample customers to each stratum.
6. Make added adjustments if the sample size exceeds the population size in any stratum.

Under the ratio model, σ_i is determined by the x variable together with the value of γ. Methods are available for estimating γ from a sample. Indeed, the evaluation team has estimated γ in numerous studies. The evaluation team has found that the estimated values are clustered around 0.8. The evaluation team has also found that the key results are not very sensitive to γ. Therefore, in load research and evaluation applications, the evaluation team generally recommends the use of $\gamma = 0.8$ both in constructing strata as discussed in this section and in estimating the value of the error ratio from a given sample.

Evaluating the Precision of Any Design

For any sample design, the evaluation team define the inclusion probability of each site in the population, denoted π_i, to be the probability that the site is included in the sample. For a

---

[27] This methodology is the model-based version of the Dalenius-Hodges method of constructing strata combined with optimal allocation of the sample using the within-strata population standard deviation of the e_i. However, Dalenius-Hodges stratification is approximately optimal for stratified mean per unit estimation whereas model-based stratification is approximately optimal for stratified ratio estimation. Moreover, with conventional methods it is common to calculate the required sample size from the within-stratum population standard deviation of x_i. This practice can yield very misleading results and cannot be recommended.

[28] With MBSS methodology we can systematically assess the gain from increased stratification. These studies indicate that five annual-use strata are usually sufficient in most load research applications. Some applications may call for added stratification by seasonal use, customer load factor, etc.

stratified sample design, the inclusion probability is the sampling fraction in each stratum, i.e., n_h / N_h .

Under the ratio model discussed previously and any sample design, the expected relative precision of the stratified ratio estimator is

$$rp = z\sqrt{\sum_{i=1}^{N}\left(\pi_i^{-1}-1\right)\sigma_i^2} \Big/ \sum_{i=1}^{N} \mu_i$$

Here $z = 1.645$ for the 90% level of confidence.

This key result has the following mathematical implications:

1. For any given sample size n, a sample design is said to be efficient if the sample design minimizes the expected relative precision. For any efficient sample design, $\pi_i = \frac{n}{\sum_{i=1}^{N}\sigma_i}\sigma_i$ provided that the right-hand side is less than 1.
2. If the right-hand side is greater than 1, the site should be included with certainty.
3. If the sample design is efficient and the population is large, then the expected relative precision is $rp = \frac{z\ er}{\sqrt{n}}$.
4. The model-based sample design, described in Section 3, is practically efficient as long as the number of strata is large enough.

The preceding equation can be also used to calculate the expected statistical precision of any sample design in any domain of interest. For example, this result can be used to calculate the expected relative precision as the number of strata is increase, e.g, from 5 to 6 to 7, etc. This type of analysis has led us to conclude that 5 strata are usually enough in most cases.

Summary

Extensive experience indicates that stratified ratio estimation is very effective in almost all load research and evaluation applications. MBSS methods are generally grounded on the same principles as conventional sampling methods such as Dalenius-Hodges stratification and Neyman allocation, but MBSS methods are specifically tailored to ratio estimation. Some methods for calculating sample sizes that have been used in the past can provide badly misleading results for ratio estimation. The MBSS approach addresses these problems and provides a coherent, consistent approach to both sample design and analysis. The MBSS methodology follows the life cycle of load research and evaluation studies very nicely.

A bonus of MBSS methodology is its strength for multiple y variables and domains estimation. At the sample design stage, MBSS provides straightforward methods for assessing the statistical precision expected for various y variables and domains of interest from the associated error ratios. At the analysis stage, MBSS again provides straightforward methods for developing estimates and their statistical precision for various y variables and domains, and for estimating the associated error ratios. In the past it has been thought to be risky to report results for domains that were not factored into the sample design. MBSS methodology has shown that meaningful results can generally be developed for questions that arise later in the study, much after the planning stage.

APPENDIX B – METERING EQUIPMENT DETAILS

Equipment Type	Manufacturer	Model #	Accuracy
Current clamps (AC/DC)	Amprobe spot power meter	ACD-31P	Current AC: ±1.5% rdg +5 digits; Current DC: ±1.5% rdg +4 digits; AC voltage: ±1% rdg +4 digits; DC Voltage: ±0.5% +3 digits; resistance: ±0.8% rdg +6 digits
Digital light meter	Amprobe spot power meter	LM631A	±3% rdg + 10 digits
Split core current transformer	Dent CTs	C100A	±1% at 10% to 130% of rated current
Current transformer	Dent CTs	C150a	2 to 80A: 1% of reading; 80 to 150A: 1.5% of reading
Large split core current transformer	Dent CTs	CT-SC-L-1000	±1% at 10% to 130% of rated current
Clamp on current transformer	Dent CTs	500A	10-600A: ±2.5% ±0.6A; 48-440 Hz 10-600A: ±3.5% ±0.6A 440-1000Hz
MilliVolt output flex current transformers	Dent CTs	CT-RMV-16-1000	<1% typical at 2% to 500% of rated current
MilliVolt output flex current transformers	Dent CTs	CT-RMV-24-2000	<1% typical at 2% to 500% of rated current
Small split core current transformer	Dent CTs	CT-SCS-0050	±1% at 10% to 130% of rated current
Small split core current transformer	Dent CTs	CT-SC-S-0100	±1% at 10% to 130% of rated current
Recording poly phase power meter	Dent CTs	ELITEpro	<1% of reading , exclusive of sensor (0.2% typical)
Infrared camera	FLIR IR camera	FLIR T200	±2°C (±3.6°F) or ±2% of reading
Portable ultrasonic transit time flow meter	GE sensing ultrasonic flow meter	PT878	±1% of rate >1 fps>6" pipe ID. ±2% of rate >1 fps <6" pipe ID.
Combustion analyzer	UEI combustion analyzer	UEI C75	Temp: ±(0.3% rdg+1F(1C)); Gas: 02 ±0.2%, CO ± 10 ppm < 100 ppm ±5% of rdg, CO2 ±0.3% reading, Efficiency ±1% of reading, excess air ±0.2% reading
4-channel data logger	Onset (HOBO)	H08-006-04	Temp accuracy: ±0.7° at 21°C (±1.27°F at 70°F)
Data logger RH/temp	Onset (HOBO)	H08-032-08	Temp accuracy: ±0.2° at 21°C (±0.33° at 70°F)
4-external channel data logger	Onset (HOBO)	U12-006	Time accuracy: ±1 minute per month at +25°C (+77°C)
Temp/RH/light external data logger	Onset (HOBO)	U12-012	Temp: ±0.35C from 0 to 50C; RH: ±2.5% from 10 to 90%; External input channel: ±2 mV ±2.5% of absolute reading
Temp/RH/ 2 external data logger	Onset (HOBO)	U12-013	Temp: ±0.35C from 0 to 50C; RH: ±2.5% from 10 to 90% to a max of ±3.5%; External input channel: ±2 mV ±2.5% of absolute reading

Light on/off data logger	Onset (HOBO)	U9-002	Time accuracy: approximately ±1 minute per month at 25°C (77°F)
Motor on/off data logger	Onset (HOBO)	U9-004	Time accuracy: approximately ±1 minute per month at 25°C (77°F)
Transport flow meter	Panametric ultrasonic flow meter	Transport PT868	Transit-time mode: 1% of reading typical for calibrated systems.; TransFlection mode: 2% of reading typical for calibrated systems.
Current transformer	Sentran CTs	4DS2-50A:2.5VDC	1% full scale to 50% of scale, 1.5% below 10%
Current transformer	Sentran CTs	4DS2-100A:2.5VDC	1% full scale to 50% of scale, 1.5% below 10%
Current transformer	Sentran CTs	4DS2-200A:2.5VDC	1% full scale to 50% of scale, 1.5% below 10%
Current transformer	Sentran CTs	4DS2-20a:2.5VDC	1% full scale to 50% of scale, 1.5% below 10%
Current transformer	Sentran CTs	4DS2-600A:2.5VDC	1% full scale to 50% of scale, 1.5% below 10%
Current transformer	Sentran CTs	4DS2-1000A:2.5VDC	1% full scale to 50% of scale, 1.5% below 10%
Current transformer	Sentran CTs	4DS2-1500A:2.5VDC	1% full scale to 50% of scale, 1.5% below 10%
10 to 3000 amps AC current probe	Summit Technology (PowerSight) CTs	FX-3000A	2% of reading ±0.2 amps (for currents below 400ARMS) or ±2 Arms (for currents from 400- 3000 ARMS)
Power logger	Summit Technology (PowerSight)	PS2500	1% plus accuracy of current probe
Combustion analyzer	TPI	A714	Gases: O2 ±0.3%, CO ±5 ppm or 5%, NO ±5 ppm to 5%, CO2 calculated, NOX calculated, efficiency calculated
Combustion analyzer	TPI	A712	Gases: O2 ±0.3%, CO ±5 ppm or 5%, CO2 calculated, NOX calculated, efficiency calculated

APPENDIX C – SITE SPECIFIC M&V REPORTS

Due to the large number of the site specific M&V reports, this appendix is provided as a separate document from the main report.

APPENDIX D – ATTRIBUTION ANALYSIS METHODS

Attribution Calculations

Program attribution accounts for the portion of the gross energy savings associated with a program-supported measure or behavior change that would not have been realized in the absence of the program. The program-induced savings, indicated as a net-to-gross ratio (NTGR), is made up of free ridership (FR) and spillover (SO) and is calculated as $(1 - FR + SO)$. Free ridership is the portion of the program-achieved verified gross savings that would have been realized absent the program and its interventions. Spillover is generally classified into participant and nonparticipant spillover. Participant spillover (PSO) occurs when participants take additional energy-saving actions that are influenced by the program interventions but did not receive program support. Nonparticipant spillover is the reduction in energy consumption and/or demand by nonparticipants because of the influence of the program.

As part of this evaluation, the evaluation team focused on the estimation of free ridership and participant spillover. Through interviews with trade allies, the evaluation team also explored the presence of nonparticipant spillover and whether additional research is justified to accurately quantify nonparticipant spillover. Quantifying savings from nonparticipant spillover activities is a challenging task that warrants a separate study and was outside of the scope of this evaluation effort.

FR component of the NTGR was derived from self-reported information from telephone interviews with program participants and further adjusted through the interviews with participating trade allies.

Below is a detailed overview of the method for developing free ridership and spillover estimates, and assessing the presence of nonparticipant spillover.

Free Ridership Estimation through Participant Survey

Free riders are program participants who would have implemented the incented energy efficient measure(s) even without the program. In other words, free ridership represents the percent of savings that would have been achieved in the absence of the program.

The goal of most incentive-based energy efficiency programs is to influence customer decision-making regarding energy efficient improvements. Programs can do this by changing *what* customers install, *when* they install it, and *how much* they install. In other words, programs influence the *efficiency*, *timing*, and *quantity* of customers' energy-using equipment installations.

The bulk of program savings is typically achieved by encouraging customers to install higher efficiency equipment that they would have installed on their own. Programs may also encourage early replacement of still functioning equipment that is less efficient, thus impacting the timing of the installation so that savings can be realized earlier. The incentive may also make it more affordable for customers to install a greater number of high efficiency measures.

The free ridership algorithm outlined here combines estimates of each of these concepts:

- Program influence on the efficiency level of the installed equipment **(FRE)**
- Program influence on the timing of the installation of high-efficiency equipment **(FRT)**
- Program influence on the quantity of the high-efficiency equipment installed **(FRQ)**

Each concept takes a value between 0 and 1. The values are expressed in FR terms, with 0 meaning no FR and 1 meaning full FR. To calculate an overall estimate of program influence, the algorithm first multiplies the estimates of efficiency (FRE) and quantity (FRQ) then averages the resulting estimate with the estimate of timing (FRT), but ONLY in cases where the FRT value is lower than the product of FRE by FRQ multiplication. In cases where the FRT value is higher, FRT, the timing component is not a part of the algorithm. This is done in order not to penalize the program for influencing the timing of the high efficiency project to a lesser degree than influencing the efficiency and/or quantity of the project. Furthermore, since the concepts of timing and quantity are conditional on efficiency (or the probability of the high efficiency installation taking place), the FRT and FRQ elements are only incorporated in the FR estimate when the FRE component is 0.5 or higher (50% or higher probability of a high efficiency installation taking place).

Thus, the formula to calculate FR can be expressed as:

IF FRE<0.5, FR=FRE

IF FRE ≥0.5 AND FRT<(FRE X FRQ), FR=AVERAGE((FRE X FRQ); FRT)

ALL OTHER CASES

FR=FRE X FRQ

Below is further detail on the how each influence score was calculated, why this logic was used, and the survey questions measuring each area of influence.

Program Influence on Equipment Efficiency (FRE)

Large C&I programs promote the purchase and installation of energy efficient measures by commercial and industrial customers in existing facilities. These programs offer customers subsidized energy studies, marketing and support, as well as financial incentives to offset the higher purchase cost of specific energy efficient equipment.

Since influencing the efficiency level of the installed equipment is at the core of the Large C&I program theory, the evaluation team asked participants multiple questions aimed at assessing program influence on the decision-making process and ultimately developed three independent measurements of the concept or three efficiency component scores:

1. Likelihood of adopting the same efficiency in absence of the program **(EI1**)
2. Influence of the program components on adopting the same efficiency **(EI2)**
3. Influence of the program in comparison to other components **(EI3)**

Since each one of these measurements is an independent measure of efficiency, the final score for program influence on efficiency (FRE) is the average of the three efficiency component scores:

$$FRE = Average(EI1;\ EI2;\ EI3)$$

Below the evaluation team describe the survey questions associated with each efficiency component score and the score calculation method.

Likelihood of Adopting the Same Efficiency in Absence of the Program (EI1)

This measurement was calculated using a single question asking participants to rank, on a 1–7-point scale, where 1 is not at all likely and 7 is very likely, their likelihood of making energy efficient improvements in the absence of the program. Industry opinions on the use of various rating scales differ. However, there is evidence that a 1–7-point scale yields more reliable and valid results.[29]

Survey Question

N5 Now, using a likelihood scale from 1 to 7, where 1 is "not at all likely" and 7 is "extremely likely," if <PROGRAM>, including incentives, and other program factors that the evaluation team have just discussed had not been available, what is the likelihood that you would STILL have completed the HIGH EFFICIENCY project? [RECORD 1 TO 7; 98=DON'T KNOW; 99=REFUSED]

Survey responses were converted from the 1–7-point scale to a value between 0 and 1 using linear transformation.[30] A score of 0 means no FR, while a score of 1 means full FR.

---

[29]M. Lozano, E. García-Cueto, J. Muñiz. 2008. "Effect of the Number of Response Categories on the Reliability and Validity of Rating Scales," *European Journal of Research Methods for the Behavioral and Social Sciences* 4(2): 73-79.

[30]We do not have any reason to believe that linear transformations would yield results that are less reliable or valid than if we were to use non-linear transformations of the scale responses. The linear transformation approach also seems intuitive given the use of the scalars. We therefore selected to use it in our calculations.

Calculation

$$EI1 = (QN5 - 1)/6$$

Influence of the Program Components on Adopting the Same Efficiency (EI2)

This measurement was calculated through multiple questions measuring the influence of individual program components. Based on our knowledge of the program theory, the following could influence the decision-making process and are program-induced:

- Program incentives (EI2A)
- Information and recommendations provided as the result of the energy study (EI2B)
- Interactions with program staff, including implementation partner staff (EI2C)
- Utility account executive endorsement (EI2D)
- Program marketing and outreach (EI2E)
- Previous experience with the program (EI2F)

A participant could be influenced by more than one program component (e.g., energy study and incentives), but the influence of various components could be different (e.g., incentives can be more influential than the energy study). To give the program due credit, EI2 was based on the score given to the most influential program component.

We asked respondents the following questions to measure each of the components:

Survey Questions

N3 *I'm going to ask you to rate the influence of a variety of factors that might have played a role in your decision to complete the project at the efficiency level you selected. Specifically, I am interested in factors that influenced you to choose HIGH EFFICIENCY as opposed to LESS EFFICIENT options. When answering, please use a scale of 1 to 7, where 1 is not at all influential and 7 is very influential.*

N3b *Availability of the <PROGRAM> incentive.*

N3c *Information and equipment recommendations provided through the energy study that was at least partially sponsored by <UTILITY>.*

N3f *Information and equipment recommendations provided through the interactions with <IF UTILITY=1 READ "PROGRAM REPRESENTATIVES AT CON EDISON OR LOCKHEED MARTIN," IF UTILITY=2, READ "PROGRAM REPRESENTATIVES AT ORANGE AND ROCKLAND">.*

N3g *Endorsement or recommendation by <UTILITY> account manager or account executive.*

N3h *Information from <PROGRAM> marketing and outreach activities. This might include training workshops, webinars, brochures, case studies, fact sheets, or information on the program website about energy savings opportunities.*

N3m *Previous experience participating in <PROGRAM>.*

Similar to EI1, the evaluation team used a 1–7-point scale, where 1 means not at all influential and 7 means very influential, to measure each of the program components and convert survey responses from the 1–7-point scale to a value between 0 and 1 using linear transformation. The score of 0 means no FR, while the score of 1 means full FR. Note that because of the direction of the scale (higher rating means lower free-ridership), the evaluation team performed additional transformation to "flip the scale."

Calculation

$$EI2 = MINIMUM(EI2A, EI2B, EI2C, EI2D, EI2E, EI2F)$$
$$EI2A = 1 - ((QN3B - 1)/6)$$
$$EI2B = 1 - ((QN3C - 1)/6)$$
$$EI2C = 1 - ((QN3F - 1)/6)$$
$$EI2D = 1 - ((QN3G - 1)/6)$$
$$EI2E = 1 - ((QN3H - 1)/6)$$
$$EI2F = 1 - ((QN3M - 1)/6)$$

Alongside program-related components described above, as part of the survey instrument the evaluation team explored the influence on the various non-program factors that might have driven the installation of high efficiency equipment. We asked participants to rate the influence of such factors as contractor and vendor recommendations, recommendations from a design or consulting engineer, standard practice, corporate policy, and previous experience with the installed equipment.[31] We interweaved the questions asking about the influence of program factors and non-program factors through question rotation/randomization tool.

In addition, the evaluation team gave respondents the option to indicate any additional factors that might have been of influence in an open ended question. While responses to the questions measuring the influence of the non-program factors did not contribute to the calculation of FR, they served as a point of validation of other responses.

The following questions explored the influence of the non-program components:

[31] We used a scale of 1 to 7, where 1 is not at all influential and 7 is very influential.

N3 I'm going to ask you to rate the influence of a variety of factors that might have played a role in your decision to complete the project at the efficiency level you selected. Specifically, I am interested in factors that influenced you to choose HIGH EFFICIENCY as opposed to LESS EFFICIENT options. When answering, please use a scale of 1 to 7, where 1 is not at all influential and 7 is very influential.

N3d Recommendation from an equipment vendor or contractor that helped you with the choice of the equipment.

N3e Previous experience with this type of equipment

N3i A recommendation from a design or consulting engineer

N3j Standard practice in your business/industry

N3k Corporate policy or guidelines

N3n Were there any other factors the evaluation team hasn't discussed that were influential in your decision to complete a HIGH EFFICIENCY AS OPPOSED TO STANDARD EFFICIENCY project?

N3nn Using the same 1 to 7 scale, where 1 means not at all influential and 7 means very influential, how would you rate the influence of this factor?

Influence of the Program in Comparison to Other Components (EI3)

This measurement was calculated using a single question asking participants to divide 100 points that represent all possible factors in their decision to install high efficiency equipment between program factors and other factors not directly related to the program. Prior to asking the question, the evaluation team reminded respondents of the factors, both program and non-program that they ranked as influential to their decision to install high efficiency equipment.[32]

We flipped the score so that 0 means no FR, while 1 means full FR.

Survey Question

You just mentioned that the following <PROGRAM> related factors were influential in your decision to complete a HIGH EFFICIENCY AS OPPOSED TO STANDARD EFFICIENCY project.

[READ A LIST OF PROGRAM FACTORS WITH THE INFLUENCE RATING OF 5, 6, 7]

You also just mentioned that the following other factors were influential in your decision to complete a high efficiency project.

[32] We will list both program and non-program factors that respondents ranked as 5, 6, or 7 on a scale of 1 to 7, where 1 is not at all influential and 7 is very influential.

[READ A LIST OF NON-PROGRAM FACTORS WITH THE INFLUENCE RATING OF 5, 6, 7]

N3p If you were given a TOTAL of 100 points that reflect the influence of all possible factors on your decision to install high efficiency equipment as opposed to less efficient equipment, and you had to divide those 100 points between: (1) <PROGRAM> factors and (2) other factors not directly related to <PROGRAM> how many points would you give to the <PROGRAM>?

Calculation

$$EI3 = (100 - QN3P)/100$$

Program Influence on Timing (FR_T)

Program influence on timing was measured by asking participants if the installation would have happened later in the absence of the program, and, if so, how much later, with the resulting score taking a value between 0 and 1.

We asked respondents who said that the program had sped up the installation if they would have installed the equipment within 6 months of when they did, 6 months to a year later, 1 to 2 years later, 2 to 3 years later, 3 to 4 years later, or 4 or more years later. We further probed respondents who said that that the program sped up the installation by 4 years or more on why they thought the installation would have happened so far out in the future. The choice to use the 4-year time frame stems from the nature of the decision-making dynamics in the commercial and industrial sector (bigger budget projects that require a great degree of planning) as well as the decision-making process that commercial properties undergo when it comes to capital improvements (longer-term budget planning, etc.).

As mentioned before, the timing question is conditional on at least some probability of the *high efficiency* installation taking place absent the program. We defined *high efficiency* as the EI score of 0.5 or higher. In order to reduce respondent bias, when asking the question, the evaluation team emphasized that the evaluation team are referring to the high efficiency purchase and did *not* ask the timing question if at least one of the following parameters was true:

- Program component rating is 5, 6, or 7 across any relevant component for that respondent[33] (meaning that any one of the program components was influential in the participant’s decision to install high efficiency equipment).
- Likelihood of installing high efficiency equipment is either 1, 2, or 3[34] (meaning that there was low likelihood that a high efficiency installation would have happened in the absence of the program).

[33] On a scale of 1 to 7, where 1 is not at all influential and 7 is very influential (the higher the rating, the lower the free-ridership)

- ❑ Program influence on a 100-point scale is more than 50 (meaning that program factors had considerable influence on the participant's decision to install high efficiency equipment).

Essentially, the evaluation team asked the timing question *only* of program participants who had considerable probability of installing *high efficiency* equipment in the absence of the program (thus making timing conditional on efficiency).

When asking timing questions, the evaluation team will give participants an option to volunteer the response "would not have installed the equipment at all without the program." When such a response is volunteered, the evaluation team followed up to confirm and classify participants as non-free-riders.

Survey Questions

We talked quite a bit about the influence of <PROGRAM> on the efficiency level of the equipment. I would now like to talk about how <PROGRAM> might have influenced the timing of your project. Remember, when I say <PROGRAM>, I mean all of the components that the evaluation team talked about before, such as incentives, support and recommendations from program staff, as well as marketing and outreach activities.

N7 Did <PROGRAM> cause you to complete your project EARLIER than you otherwise would have, or did <PROGRAM> have no influence on when you completed the project?

1. *Caused to install earlier*
2. *Did not influence when installed*
3. *(Would not have installed the equipment at all without <PROGRAM>) [DO NOT READ]*
8. *(Don't know) [DO NOT READ]*
9. *(Refused) [DO NOT READ]*

[34] On a scale of 1 to 7, where 1 is not at all likely, and 7 is very likely (the higher the rating, the higher the free-ridership)

[ASK N7A IF N7=3]

N7a Just to confirm, if <PROGRAM> had not been available, you would NOT have completed the project at all, is that correct?

1. Yes [SKIP TO N9a]
2. No
8. (Don't know) [DO NOT READ]
9. (Refused) [DO NOT READ]

[ASK N7B IF N7=1]

N7b If <PROGRAM> had not been available, when would you have completed the project? Would you say . . .

1. Within 6 months of when you did
2. 6 months to 1 year later
3. 1 - 2 years later
4. 2 - 3 years later
5. 3 - 4 years later
6. 4 or more years later
8. (Don't know) [DO NOT READ]
9. (Refused) [DO NOT READ]

[ASK N7C IF N7B=6]

N7c Why do you think it would have been 4 or more years later?

00. [RECORD VERBATIM]
98. (Don't know) [DO NOT READ]
99. (Refused) [DO NOT READ]

For each response to the series of questions measuring program influence on timing, the evaluation team developed a timing score that varied between 0 and 1. Zero means no FR, while 1 will mean full FR.

Participants who said that the high efficiency project would have happened within six months of when it did were assigned a score of one (full free-riders), as six month is too close to the actual installation date to give program the timing credit. On the other hand, participants who said that the high efficiency project would have happened four or more years later were assigned a score of zero, as this time is so removed from the actual installation date that the program deserves full credit for the timing component. All other responses were converted using linear transformation function that begins in month six (the starting point at which the program can start getting partial credit) and ends in month 48 (the last month after which the program gets full credit), decreasing 0.024 for each month of deferred installation. We used midpoints for each period to calculate the number of months that program should be getting credit for. For example, if the respondent would have completed the project six months to a year later, the linear transformation process looked like:

$$1 - (((6 + 12)/2) - 6) * 0.024) = 0.93$$

For ease of interpretation, the evaluation team calculated FR_T values for each response category.

<u>Calculations</u>

$$IF\ QN7 = 2, \boldsymbol{FRT} = 1$$

$$IFQN7 = 3\ AND\ QN7A = 1, \boldsymbol{FRT} = 0$$

$$IF\ QN7 = 1\ AND\ QN7B = 1, \boldsymbol{FRT} = 1$$

$$IF\ QN7 = 1\ AND\ QN7B = 2, \boldsymbol{FRT} = 0.93$$

$$IF\ QN7 = 1\ AND\ QN7B = 3, \boldsymbol{FRT} = 0.71$$

$$IF\ QN7 = 1\ AND\ QN7B = 4, \boldsymbol{FRT} = 0.42$$

$$IF\ QN7 = 1\ AND\ QN7B = 5, \boldsymbol{FRT} = 0.14$$

$$IF\ QN7 = 1\ AND\ QN7B = 6, \boldsymbol{FRT} = 0$$

To validate the timing scores provided by respondents the evaluation team first asked them to explain, in an open ended fashion, the specific ways in which the program accelerated the installation process. Furthermore, the evaluation team asked respondents who indicated that program influenced the timing of their project the following questions to validate the timing credit given to the program.

[SKIP TO Q16 IF END USE=1 OR END USE=5 AND VMEASD=23,24,28]

Q12 *Approximately how old was the equipment that was removed as part of this project? Was it..?*

1. *Less than five years old*
2. *Between 5 and 10 years old*
3. *Between 10 and 15 years old*
4. *More than 15 years old*

8 *(Don't know) [DO NOT READ]*

9. *(Refused) [DO NOT READ]*

Q13 *How would you describe the condition of your old equipment? Was it in good condition, fair condition, poor condition or was it not working?*

1. *Good condition*
2. *Fair condition*
3. *Poor condition*

4. *Not working*
8 *(Don't know) [DO NOT READ]*
9. *(Refused) [DO NOT READ]*

Q14 *How much longer do you think it would have lasted? [NUMERIC OPEN END. RECORD RESPONSE IN YEARS]*

Q15 *In your opinion, based on the economics of operating this equipment, for how many more years could you have kept this equipment functioning? [NUMERIC OPEN END]*

Q16 *Over the last five years, have maintenance costs been increasing, decreasing or staying about the same?*

1. *Increasing*
2. *Decreasing*
3. *Staying the same*
8. *(Don't know) [DO NOT READ]*
9. *(Refused) [DO NOT READ]*

Program Influence on Quantity (FR_Q)

Program influence on *quantity* was measured by asking participants whether the scope of the project would have been the same in the absence of the program and estimating what share of the project (in percentage terms or in units depending on the project and the installed equipment) would have happened anyway. Those participants whose projects only contained one piece of equipment were not asked questions related to program influence on quantity/scope.

Similar to the assessment of program influence on timing, the quantity question is conditional on at least some probability of the *high efficiency* installation taking place absent the program. We defined *high efficiency* as the EI score of 0.5 or higher. In order to reduce respondent bias, when asking the question, the evaluation team emphasized that the evaluation team are referring to the high efficiency purchase and did *not* ask the quantity question if at least one of the following parameters was true:

- ❑ Program component rating is 5, 6, or 7 across any relevant component for that respondent[35] (meaning that any one of the program components was influential in the participant's decision to install high efficiency equipment).

[35] On a scale of 1 to 7, where 1 is not at all influential and 7 is very influential (the higher the rating, the lower the free-ridership)

- ❑ Likelihood of installing high efficiency equipment is either 1, 2, or 3[36] (meaning that there was low likelihood that a high efficiency installation would have happened in the absence of the program).
- ❑ Program influence on a 100-point scale is more than 50 (meaning that program factors had considerable influence on the participant's decision to install high efficiency equipment).

As with the timing questions, the evaluation team allowed respondents to volunteer the response "would not have installed any equipment without the program" and confirmed that that was indeed true.

<u>Survey Questions</u>

N8 If <PROGRAM> had not been available, would the scope or size of your HIGH EFFICIENCY project have been larger, the same, or smaller? [RECORD 1 TO 7; 98=DON'T KNOW; 99=REFUSED]

1. Larger
2. Same
3. Smaller
4. (Would not have installed any equipment without <PROGRAM>)[DO NOT READ]
8. (Don't know) [DO NOT READ]
9. (Refused) [DO NOT READ]

[ASK N8A IF N8=4]

N8a Just to confirm, if <PROGRAM> had not been available, you would NOT have completed the project at all, is that correct?

1. Yes [SKIP TO N10A]
2. No
8. (Don't know) [DO NOT READ]
9. (Refused) [DO NOT READ]

[ASK N8B IF N8=3 OR N8A=2,8,9 OR N8B=2,8,9]

N8b In percentage terms, what percent of the HIGH EFFICIENCY project would have happened anyway in the absence of <PROGRAM>? [1-99, 998=DK; 999=REF] [PROBE: YOUR BEST ESTIMATE IS FINE]

---

36 On a scale of 1 to 7, where 1 is not at all likely, and 7 is very likely (the higher the rating, the higher the free-ridership)

[ASK IF N8B=998 OR 999]

N8c Would you say you would have done 25%, 50%, or 75% of the project if the program had not been available?

1. *25%*
2. *50%*
3. *75%*
4. *(0%)*
5. *(100%)*
8. *(Don't know) [DO NOT READ]*
9. *(Refused) [DO NOT READ]*

For each response to the series of questions measuring program influence on timing, the evaluation team developed a quantity score that varied between 0 and 1. Zero means no FR, while 1 means full FR.

Calculations

$$IF\ QN8 = 1, \boldsymbol{FRQ} = 1$$

$$IF\ QN8 = 2, \boldsymbol{FRQ} = 1$$

$$IF\ QN8 = 3\ AND\ QN8B\ IS\ NOT\ DON'T\ KNOW\ OR\ REFUSED, \boldsymbol{FRQ} = QN8B$$

$$IF\ QN8 = 3\ AND\ QN8B\ IS\ DON'T\ KNOW\ OR\ REFUSED, \boldsymbol{FRQ} = QN8C$$

$$IF\ QN8 = 4\ AND\ QN8A = 1, \boldsymbol{FRQ} = 0$$

$$IF\ QN8 = 4\ AND\ QN8A\ IS\ DON'T\ KNOW\ OR\ REFUSED$$

$$AND\ QN8D\ IS\ NOT\ DON'T\ KNOW\ OR\ REFUSED,$$

$$\boldsymbol{FRQ} = QN8B$$

$$IF\ QN8 = 4\ AND\ QN8A\ IS\ DON'T\ KNOW\ OR\ REFUSED$$

$$AND\ QN8D\ IS\ DON'T\ KNOW\ OR\ REFUSED,$$

$$\boldsymbol{FRQ} = QN8C$$

Consistency Checks

The scoring algorithm relied on responses from multiple questions to develop FR rates. Because respondents can sometimes give inconsistent answers, the survey instrument included several consistency checks to clarify these responses.

Individual Components vs. Program Score

N4 *Earlier, you gave <N3P RESPONSE> points to the influence of <PROGRAM>. I would interpret that to mean that the program was quite influential in your decision to complete the HIGH EFFICIENCY project. However, when I asked you about the influence of individual elements of the program I recorded some answers that would imply that they were not that influential. Just to make sure I have recorded this properly will you explain the role <PROGRAM> had in your decision to complete the high efficiency project? [OPEN END; 98=DON'T KNOW; 99=REFUSED]*

N4aa *Earlier, you gave <N3P RESPONSE> points to the influence of <PROGRAM>. I would interpret that to mean that the program was not very influential in your decision to complete the HIGH EFFICIENCY project. However, when I asked about the influence of individual elements of the program I recorded some answers that would imply that they were very influential. Just to make sure I have recorded this properly will you explain the role <PROGRAM> had in your decision to complete the HIGH EFFICIENCY project?*

<u>Individual Components vs. Likelihood of Installing without Program</u>

N5a *Earlier, you said that there is <N5 RESPONSE> in 7 likelihood that you would have installed the same efficiency equipment absent <PROGRAM>. I would interpret that to mean that the program was not that influential on your decision to complete the HIGH EFFICIENCY project. However, when I asked you about the influence of individual elements of the program I recorded some answers that would imply that they were quite influential. Just to make sure I have recorded this properly will you explain the role <PROGRAM> had in your decision to complete the HIGH EFFICIENCY project? [OPEN END; 98=DON'T KNOW; 99=REFUSED]*

N5aa *Earlier, you said that there is <N5 RESPONSE> in 7 likelihood that you would have installed the same efficiency equipment absent <PROGRAM>. I would interpret that to mean that the program was quite influential on your decision to complete the HIGH EFFICIENCY project. However, when I asked you about the influence of individual elements of the program I recorded some answers that would imply that they were not that influential. Just to make sure I have recorded this properly will you explain the role <PROGRAM> had in your decision to complete the HIGH EFFICIENCY project? [OPEN END; 98=DON'T KNOW; 99=REFUSED]*

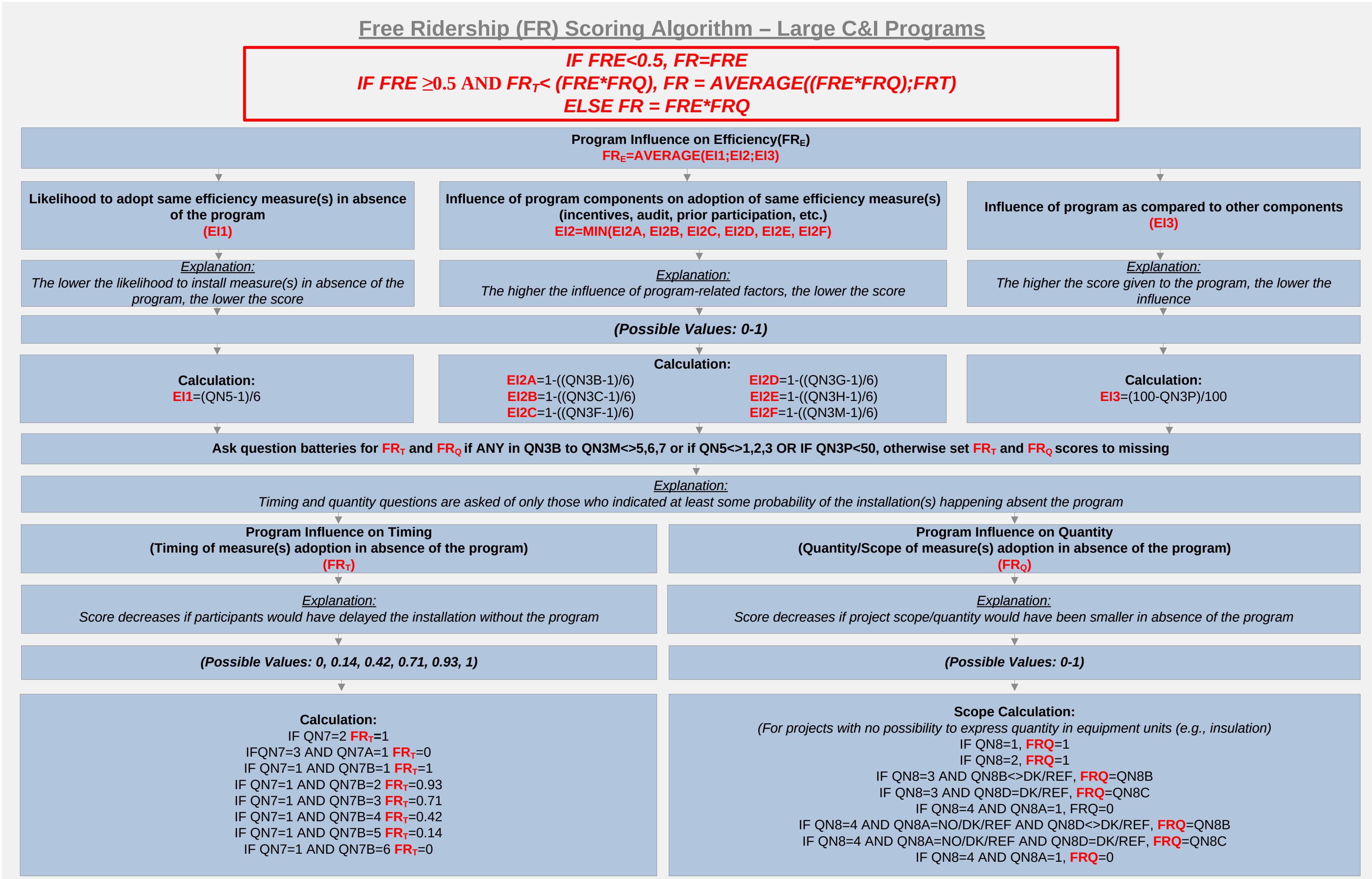
<u>Individual Components vs. Program Knowledge</u>

N6a *Earlier, you said that you learned about the program AFTER you made the decision to complete the HIGH EFFICIENCY project. However, when I asked you about the influence of individual elements of the program I recorded some answers that would imply that they were quite influential. Just to make sure I have recorded this properly will you explain the role*

<PROGRAM> had in your decision to complete the HIGH EFFICIENCY project? [OPEN END; 998=DON'T KNOW; 999=REFUSED]

Figure A-3 provides graphical summary of the FR algorithm.

Figure A-3. Proposed FR Scoring Algorithm



Free Ridership Adjustment through Trade Ally Research

As previously mentioned, FR estimation for the Large C&I program relies primarily on survey research with program participants. It is our understanding, however, that the Large C&I program used various tactics to influence the way trade allies (equipment vendors, installation contractors, designers, engineers, etc.) specify projects to customers. This included marketing and outreach to trade allies, trade ally training, etc. As such, there was a concern that assessment of FR solely through participant research would not credit the program for influences not visible to customers, primarily program outreach to trade allies. To address this concern, the final FR estimates incorporated the results of trade ally interviews, in addition to the research with program participants.

The interviews were completed with a subset of participating trade allies who influenced end-user decisions to install high efficiency equipment that received incentives through the programs. The interviews determined the influence of the Large C&I program on:

- Trade ally sales
- Recommendations
- Installation practices

Using the results of the interviews, the evaluation team adjusted participant-derived FR rates downward to account for program influence on trade allies.

We also used the trade ally interviews to support a limited assessment of whether the program is having a larger impact on the overall market. If so, future research may be warranted to quantify nonparticipant spillover (NPSO). We did not use trade ally interview results to quantify NPSO as part of this study. To do so accurately requires a larger study that is beyond the scope of this evaluation.

Trade ally interviews were triggered for projects where participant-reported trade ally influence is high – a rating of 6 or 7 to either of the following questions[37].

[IF WORKED WITH A TRADE ALLY OR VENDOR]

N3d. Recommendation from an equipment vendor or contractor that helped identify the project or helped with the choice of the equipment.

[IF WORKED WITH A DESIGN OR CONSULTING ENGINEER]

[37] On a scale of 1 to 7, where 1 means not at all influential and 7 means very influential.

N3i. A recommendation from a design or consulting engineer.

Similar to the participant interviews, trade ally interviews were be end use - and project-focused. That is, trade allies were asked about their involvement in and program influence on their equipment recommendations for a specific project and a specific end use within the project.[38] If a single trade ally influenced the decision-making process for more than one project, as part of the trade ally interview, the evaluation team inquired with trade allies separately about their involvement in multiple projects. Similarly, if multiple trade allies were involved in the decision-making process for a single project (e.g., for a project with various end uses), the evaluation team attempted to reach out and complete interviews with multiple trade allies.

While it is possible that in addition to influencing the efficiency level of the equipment trade ally recommendations might have influenced the timing of the installation of the high efficiency equipment or increased the scope of the high efficiency project, the evaluation team feels these influences are visible to the participant and would have been captured through the participant survey. The participant is also better positioned to assess what the timing and scope of the project would have been if the program were not available.

The program's influence on the frequency and manner in which contractors recommend high efficiency equipment is less visible to participants. For participants who rely on their contractor's recommendation for equipment selection, the evaluation team needs to determine the program's influence on those recommendations. Therefore, the core area of trade ally influence, and by extension, potential program influence, is equipment efficiency. The results from trade ally interviews therefore were used to adjust only the efficiency (FRE) component of the participant-derived FR score downward to account for program influence on trade allies not visible to participants. Trade ally interviews supported the development of the trade ally program influence score (FRE-C), which reflects the degree of program influence on trade ally recommendations of high efficiency equipment.

To adjust the participant efficiency score (*FRE-ADJ*) and arrive at the final FR score for the projects or end uses that were heavily influenced by trade allies, the evaluation team took the minimum score (i.e., the score that reflects the highest influence of the program).

$$FRE\text{-}ADJ = Min(FRE; FRE\text{-}C)$$

We believe that taking the minimum is appropriate because it accounts for the greatest source of program influence, thus giving the program the deserved credit. Other possible ways of adjusting the participant scores would be to average or multiply the two scores. We believe that

[38] Consistent with the participant survey effort, we will group all measures into five core end uses – lighting, HVAC, motors, EMS, and other.

averaging the two scores penalizes the program for influencing one actor (e.g., trade allies) to a lesser degree than the other (e.g., participant). Multiplying, on the other hand, might double-count the influence of the program incentives by capturing their indirect influence through trade ally recommendations of high efficiency equipment and their direct influence on participant decisions to install high efficiency equipment.

The final FR score calculation for projects or end uses influenced by trade ally recommendations, as a result, was calculated using the following formula:

If FRT<(FRE × FRQ), FR=Average((FRE-ADJ × FRQ); FRT)

Else

FR=FRE-ADJ × FRQ

Calculating Trade Ally Program Influence Score *(FRE-C)*

We used the following questions as the basis for the (*FRE-C*) score:

Vendor/Trade Ally Questions

FR3. *Using a 1 to 7 point scale, where 1 means no influence and 7 means a great deal of influence, how influential was <UTILITY>'s program, including incentives as well as program services and information, on your recommendation to install* [*DO NOT READ FOR END USES WITH NO LESS-EFFICIENT OPTIONS, E.G., EMS;* HIGH EFFICIENCY] *<END USE> as part of this project?*

FR6. *Thinking about it differently, using a 1 to 7 likelihood scale, where 1 means not at all likely and 7 means very likely, if the PROGRAM, including incentives as well as program services and information, had not been available, what is the likelihood that you would have recommended this specific* [*DO NOT READ FOR END USES WITH NO LESS-EFFICIENT OPTIONS, E.G., EMS;* HIGH EFFICIENCY] *<END USE> to <COMPANY>?*

**Please note that these questions will not be directly following each other during the interview. Please refer to the discussion guide in the Appendix of this document to see the questions' ordering.*

We converted responses to each of the two questions above from a 1 to 7 point scale to a value between 0 and 1 using linear transformation.[39] Each score represents program influence (I) on trade ally recommendations and is expressed in terms of free-ridership, where 0 means no FR

[39] We do not have any reason to believe that linear transformations would yield results that are less reliable or valid than if we were to use non-linear transformations of the scale responses. The linear transformation approach also seems intuitive given the use of the scalars. We therefore selected to use it in our calculations.

and 1 means full FR. Table A-10 below provides an example of the linear transformation of a 1 to 7 point scale to a value between 0 and 1.

Table A-10. Example of Linear Transformation of the Scale Responses

Scale	Equation for Linear Transformation of the Scale Responses	Result of the Linear Transformation of the Scale Responses
1	(1-1)/6	0
2	(2-1)/6	.167
3	(3-1)/6	.333
4	(4-1)/6	.5
5	(5-1)/6	.667
6	(6-1)/6	.833
7	(7-1)/6	1

The values reflect program influence on trade ally recommendations and are expressed in FR terms, with 0 meaning no FR and 1 meaning full FR. Table A-11 provides an overview of how, using the linear transformation approach described above, the ratings from each question was transformed into the trade ally influence scores I1 and I2. Note that in order to translate respondent ratings into trade ally influence scores expressed as FR, question FR3 required an additional transformation step to reverse the direction of the rating scale (higher rating in question FR3 represents higher program influence and as such lower FR, which means that the evaluation team should "flip" the result of the linear transformation for this question). Such additional transformation was not needed for question FR6 because the rating scale was already aligned to express FR (higher likelihood to recommend absent the program means lower program influence and higher FR).

Table A-11. Conversion of Program Influence Responses from Trade Ally Interviews

Trade Ally Influence Score	Question	Conversion Formula
I1	FR3	1-((FR3-1)/6)
I2	FR6	(FR6-1)/6

Because each of these questions is conditionally similar, the evaluation team averaged the scores from the questions above to arrive at the FR adjustment score ($FR_{E\text{-}C}$).

$$FRE\text{-}C=Average(I1;I2)$$

This overall influence score is initial and **served as the basis** for the final adjustment score. Trade ally interviews are qualitative and in-depth in nature, providing an opportunity to obtain rich additional information.

We used a variety of additional open-ended questions and consistency checks to validate the ratings that trade allies gave. When needed, the evaluation team revised the initial adjustment score to encompass the qualitative information gathered during the interview and accurately reflect program influence on trade ally recommendations. For example, if a trade ally rated the influence of the program as high on the project-specific recommendations, but then struggled to provide a reasoning for the rating, the evaluation team adjusted the *FRE-C* upward (i.e., toward higher FR). And vice versa, if the initial rating of the program influence on trade ally recommendation was low, yet the trade ally was deeply engaged with the program and mentions, in an open-ended fashion, that a variety of program-related factors played a role in their decision to recommend and specify high efficiency equipment, the evaluation team adjusted the *FRE-C* downward (i.e., toward lower FR).

As part of the validation and final scoring process, the evaluation team used the inter-rate reliability approach to the analysis of qualitative results. Two experts at Opinion Dynamics analyzed trade ally responses and arrived at an agreement on the final adjustment score (*FRE-C*) for each trade ally.

As the first validation step, the evaluation team asked trade allies to explain, in an open ended fashion, why they gave the rating that they did, as well as explored the specific avenues of program and non-program influence.

Follow-Up Questions

[THESE QUESTIONS ARE USED AS FOLLOW-UPS TO QUESTIONS FR3 AND FR6 RESPECTIVELY]

FR3a. *Why did you give this rating?*

FR6a. *Why do you say that? [IF THE SCORE IS CONTRADICTORY WITH THE ONE GIVEN TO FR3, PROBE FOR REASONS FOR THE INCONSISTENT RESPONSE]*

FR4. *How specifically did the program influence your recommendations for* [*DO NOT READ FOR END USES WITH NO LESS-EFFICIENT OPTIONS, E.G., EMS;* HIGH EFFICIENCY] *<END USE> for this specific project? [IF NEEDED, PROBE FOR SPECIFIC INFLUENCE OF MARKETING, TRAINING, INCENTIVES]*

FR5. *What other factors, if any, influenced your decision to recommend that <COMPANY> install the* [*DO NOT READ FOR END USES WITH NO LESS-EFFICIENT OPTIONS, E.G., EMS;* HIGH EFFICIENCY] *<END USE>? [IF NEEDED, PROBE FOR CHANGING CODES AND STANDARDS, STANDARD PRACTICE OF THE COMPANY, MANUFACTURER INCENTIVES, ETC.]*

a. *Which would you say was more influential – the influence of the program or the other, non-program factors?*

We also explored the trade ally's involvement and role in the project and the factors that trade allies took into account when developing equipment specifications and recommendations for the project in question.

FR1. *Can you please describe your firm's role in the selection and installation of <END USE> at <COMPANY>'s facility? [PROBE IF FIRM MERELY SUPPLIED OR INSTALLED EQUIPMENT OR IF THEY HAD A ROLE IN SELECTING IT. PROBE ABOUT PERCEIVED LEVEL OF INFLUENCE FIRM'S RECOMMENDATION HAD ON COMPANY'S CHOICE.]*

a. *Did you initiate the contact with the customer about this project or did the customer reach out to you with a request for a quote? [PROBE FOR DETAIL]*

FR2. *When developing equipment specifications and recommendations for this project, what factors went into the decision-making process? [PROBE FOR: COST, ENERGY SAVINGS, PERFORMANCE, NON-ENERGY BENEFITS, SUCH AS DECREASED NEED FOR MAINTENANCE, AESTHETICS, ETC.]*

a. *What factors, if any, did <COMPANY> discuss with you when deciding on the equipment for an upgrade project? [PROBE FOR: COST, ENERGY SAVINGS, PERFORMANCE, NON-ENERGY BENEFITS, SUCH AS DECREASED NEED FOR MAINTENANCE, AESTHETICS, ETC.]*

b. *Did the customer request specific equipment or did the customer choose a recommended option?*

Finally, as part of the trade ally interviews, the evaluation team got an understanding of the trade allies' knowledge of and the frequency of their interactions with the program.

Questions

PI1. *How long have you known about <UTILITY>'s program? How long have you been active within the program? [IF NEEDED: WHEN DID YOU COMPLETE YOUR FIRST PROJECT THAT RECEIVED INCENTIVES THROUGH <UTILITY>?]*

PI1A. *[ASK IF PROGRAM TRACKING DATA ARE UNAVAILABLE] Thinking about the timeframe between 2009 and 2011, how many projects did you complete that received incentives through the program? [PROBE FOR BEST ESTIMATE.]*

PI2. *How knowledgeable do you consider yourself about <UTILITY>'s program and its offerings? Would you say very knowledgeable, somewhat knowledgeable, not very knowledgeable, or not at all knowledgeable? [PROBE FOR KNOWLEDGE OF PROGRAM ELIGIBILITY REQUIREMENTS, INCENTIVIZED EQUIPMENT OPTIONS, INCENTIVE LEVELS, PARTICIPATION PROCESS, ETC.]*

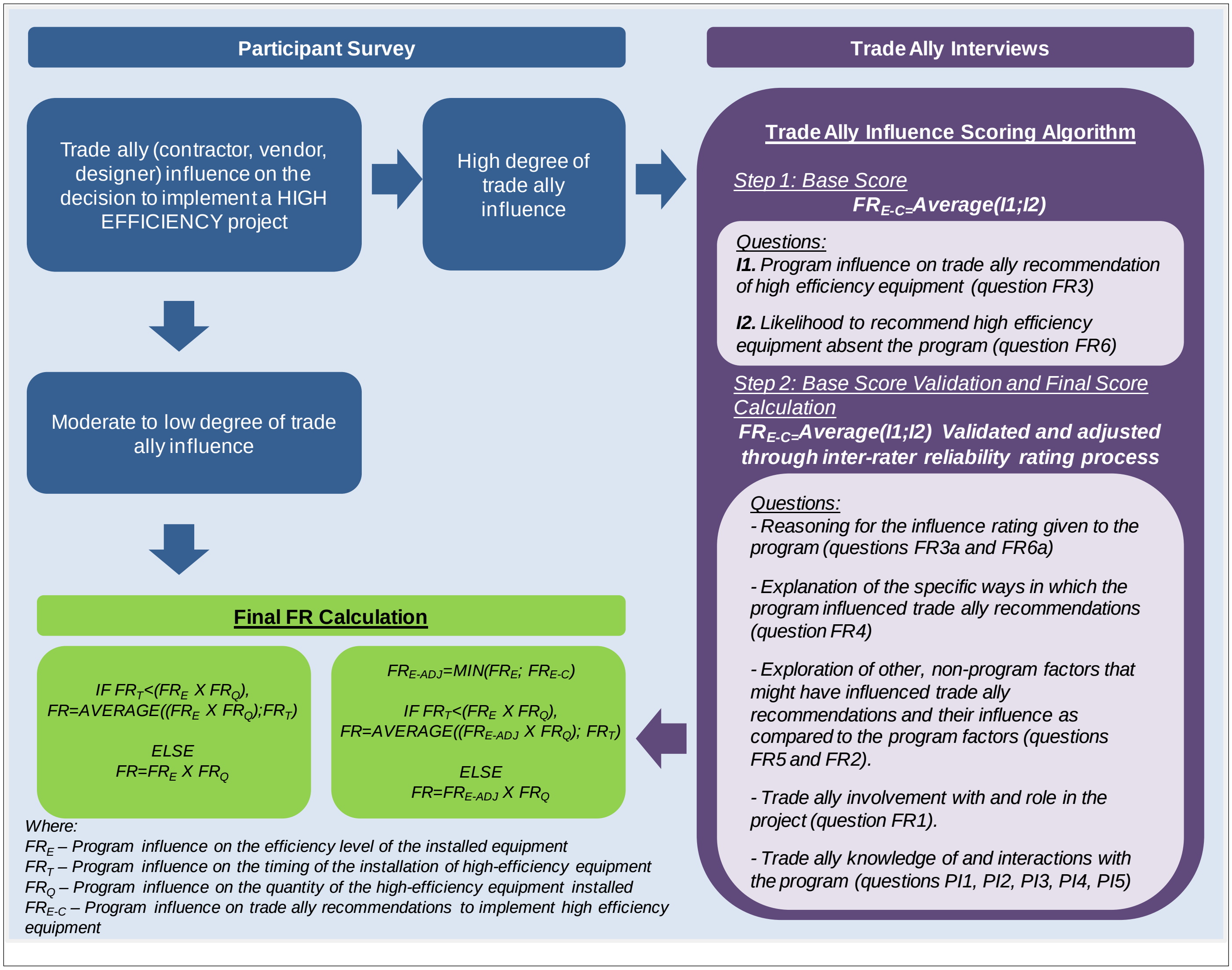
PI3. *Did you attend any training sessions, meetings or events, both formal and informal, facilitated by <UTILITY>'s program between 2009 and 2011? What kind of training did you receive? [PROBE FOR TRAINING AND GUIDANCE ON HOW TO PREPARE CUSTOMER APPLICATIONS, SPECIFICS ON QUALIFIED MEASURE TYPES, SALES TRAINING, ETC.]*

PI4. *Do you remember receiving any marketing or promotional materials or any ongoing communications <UTILITY> between 2009 and 2011? [PROBE FOR TYPES OF MATERIALS – BROCHURES, CASE STUDIES, NEWSLETTERS, CO-BRANDED MARKETING, ETC. – AS WELL AS THEIR FREQUENCY]*

PI5. *Did you receive ANY OTHER information or support from the program that either improved your ability to sell energy efficiency to customers or improved your overall knowledge of energy efficient equipment options? If so, what support did you receive?*

Figure A-4 displays the calculation of the trade ally influence scores and integration of the scores in the final FR score.

Figure A-4. Trade Ally Influence Score Calculation and Final FR Calculation



Participant Spillover

While the program has not had a large marketing component that would promote energy efficiency in general or the installation of other measures aside from the measures rebated through the program, our past experiences conducting evaluation research suggest that once a person purchases energy efficient equipment, they often look for other ways to increase the energy efficiency of their businesses. If program-induced, those additional improvements can result in the participant spillover (PSO) savings that the program could claim.

As part of the participant survey the evaluation team estimated both the presence and magnitude of PSO. While PSO can result from a variety of measures, survey length did not allow for estimation of PSO across all possible scenarios. Given the type of businesses that participated in the program, the evaluation team included measures that could reasonably be expected to be influenced by program participation and would be more likely to have been implemented without program support. PSO was measured for the following equipment:

- ❑ Lighting
- ❑ Cooling
- ❑ Refrigeration
- ❑ Kitchen
- ❑ Motors
- ❑ Heating and water heating
- ❑ Other

Participants were asked if they made any of the above-listed improvements.

Survey Questions

SP1a SINCE you participated in the <PROGRAM>, did you install any additional ENERGY EFFICIENT equipment at THIS facility that did NOT receive incentives through <UTILITY>?

SP1b And, since you participated in the <PROGRAM>, did you install any additional ENERGY EFFICIENT equipment at OTHER facilities in <UTILITY>'s service territory that did NOT receive incentives through <UTILITY>?

[IF SP1A OR SP1B=YES]

SP1c Just to confirm, you made energy efficient improvements AFTER you participated in the <PROGRAM>, is that correct?

[IF SP1A OR SP1B=YES]

SP1d Just to confirm, you made energy efficient improvements that did NOT receive incentives through the <PROGRAM>, is that correct?

Those who did were asked about specific improvements they made.

<u>Survey Question</u>

SP2A/SP3A/SP4A/SP5A/QP6A/SP7A/SP8A
Did you install any energy efficient <END USE> equipment without getting an incentive from <UTILITY>?

For each of the end use categories, respondents were asked to explain why they had purchased the equipment without applying for incentives, as well as to provide a rating of the program influence on their decision to make additional improvements.

<u>Survey Questions</u>

SP2B/SP3B/SP4B/SP5B/SP6B/SP7B/SP8B
Why did you purchase this <END USE> equipment without applying for incentives through <UTILITY>?
SP2C/SP3C/SP4C/SP5C/SP6C/SP7C/SP8C
How much did your experience with <UTILITY>'s <PROGRAM> influence your decision to install the energy efficient <END USE> equipment that you ended up installing? Please use a scale from 1 to 7 where 1 means no influence and 7 means a great deal of influence.

If any energy efficient improvements were heavily influenced by the program,[40] respondents were asked in an open-ended fashion to explain how the Program influenced their decision to install the energy efficient equipment.

<u>Survey Question</u>

SP2M/SP3M/SP4M/SP5M/SP6M/SP7M/SP8M
How did the <PROGRAM>influence your decision to install the energy efficient <END USE> equipment that you ended up installing?

Respondents were also asked a few equipment-specific questions that allowed for the calculation of savings associated with the installed equipment. The equipment details explored as part of the survey effort were limited by the survey length as well as by what the evaluation team believe respondents can provide reliable responses to.

[40] A rating of 6 or 7 on a scale of 1 to 7 where 1 is no influence and 7 is a great deal of influence.

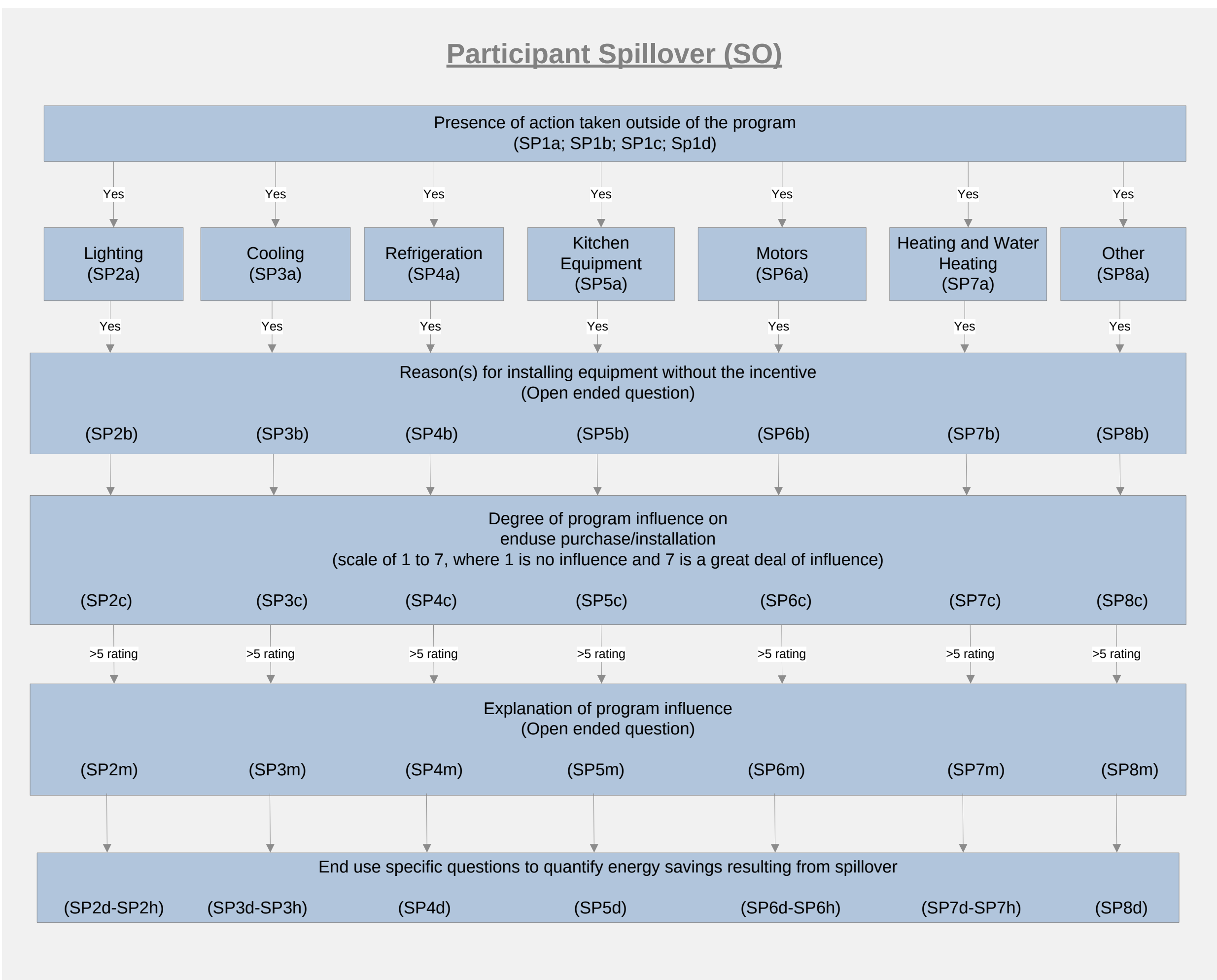
As part of the SO calculation, the evaluation team applied savings values to the measures installed outside of the program. We estimated savings for each measure using the most recent TRM values supplemented by engineering assumptions. We determined the program-level SO factor by dividing the estimated savings of the measures installed by survey respondents outside of the program (but influenced by the program) by the savings the survey respondents realized through the program.

$$\text{SO Rate} = \frac{REx\ Post\ net\ energy\ savings\ from\ measures\ installed\ outside\ the\ program}{Ex\ Post\ gross\ energy\ savings\ from\ measures\ installed\ through\ the\ program}$$

As indicated earlier, SO is included in the overall NTGR as $(1 - FR + SO\ Rate)$.

Figure A-5 provides a graphical depiction of the spillover diagram.

Figure A-5. Participant SO Diagram



APPENDIX E – PARTICIPANT SURVEY INSTRUMENT

The following is the customer survey instrument that was used by Opinion Dynamics for the participant surveys.

Con Edison and Orange & Rockland
Large C&I Program Group

June 2013

SAMPLE VARIABLES

<UTILITY> (1=Con Edison; 2=Orange and Rockland)
<YEAR> (Program participation year)
<CONTACT> (1=Contact name is present; 0=contact name is missing)
<PROGRAM> (1=The Con Edison Commercial and Industrial Electric Energy Efficiency Program; 2= The Con Edison Commercial and Industrial Gas Energy Efficiency Program; 3=O&R Big Energy Solutions Program)
<PROGRAM_TYPE> (1=Prescriptive; 2=Custom)
<FUEL_TYPE> (1=Electric; 2=Gas)
<PROGRAM CONTACT> (Name of the contact person)
<COMPANY> (Name of participating company)
<ADDRESS> (Facility address)
<END USE> (Project end use: 1=Energy Management System and Controls; 2=Lighting; 3=HVAC; 4=Motors; 5=Other)
<MEASQ> (Measure quantity)
<MEASD> (Measure description)

1=Boiler

2=CFL

3=Chiller

4=Cold cathode lamps

5=Compressed air system

6=Custom refrigeration system

7=Energy management system

8=Fluorescent fixtures

9=Furnace
10=Gas system tune-up
11=HVAC system
12=Heat pump
13=Induction lighting fixtures
14=Insulation
15=LED case lighting
16=LED exit sign
17=Lighting controls or occupancy sensors
18=Lighting improvements
19=Low flow pre-rinse sprayer
20=Metal halide fixtures
21=Motors
22=Other
23=Process upgrades
24=Pump fan piping and duct improvements
25=Delamping
26=T5
27=T8
28=VFD
29=Air sealing
30=Automatic doors
31=Cooler control unit
32=Door heater control
33=LED
34=Night setback units
35=T8 and LEDs
36=T5 and T8s
37=VSD
38 = Boiler reset
39 = Thermostat

<MSAME> (A flag for presence of other projects with the same end use)
<NSAME> (Number of projects with the same end use)
<MDIF> (Description of a project with different end uses)
<NDIF> (Number of projects with a different end use)

INTRODUCTION

[READ IF CONTACT=1]

Hello, this is _____ from Opinion Dynamics calling on behalf of <UTILITY>. This is not a sales call. May I please speak with <PROGRAM CONTACT>?

Our records show that <COMPANY> participated in <UTILITY>'s <PROGRAM> in <YEAR> and received an incentive for HIGH EFFICIENCY upgrades made at <ADDRESS>. We are calling to do a follow-up study about <COMPANY>'s participation in this program. I was told that you are the person most knowledgeable about this project. Is this correct? [IF NOT, ASK TO BE TRANSFERRED TO MOST KNOWLEDGEABLE PERSON OR RECORD NAME & NUMBER]

[READ IF CONTACT=0]

Hello, this is _____ from Opinion Dynamics calling on behalf of <UTILITY>. This is not a sales call. May I please speak with the person most knowledgeable about recent changes in cooling, lighting, or other energy-related equipment <COMPANY> installed at <ADDRESS> in <YEAR>?

[ONCE CONNECTED WITH CONTACT] Our records show that <COMPANY> participated in <UTILITY>'s <PROGRAM> in <YEAR> and received an incentive for HIGH EFFICIENCY upgrades made at <ADDRESS>. We are calling to do a follow-up study about <COMPANY>'s participation in this program. I was told that you are the person most knowledgeable about this project and the installation of the related measures. Is this correct? [IF NOT, ASK TO BE TRANSFERRED TO MOST KNOWLEDGEABLE PERSON OR RECORD NAME & NUMBER]

This survey will take about 15-20 minutes. Is now a good time? [IF NO, SCHEDULE CALL-BACK]

SCREENER

S1 Which of the following statements best characterizes your relationship to <COMPANY>?

01. I am an employee of <COMPANY> [INTERVIEWER NOTE: THIS CATEGORY SHOULD INCLUDE THE OWNER/PRESIDENT/PARTNER ETC. OF THE COMPANY]
02. My company provides energy-related services to <COMPANY>
03. I am a contractor and was involved in the installation of HIGH EFFICIENCY equipment for this project
00. (Other, specify) [PUT OWNER/PRESIDENT/PARTNER ETC. OF THE COMPANY IN 1] [DO NOT READ]
98. (Don't know) [DO NOT READ]
99. (Refused) [DO NOT READ]

[IF S1=2 OR 3 – THANK AND TERMINATE, RECORD DISPO AS ENERGY SERVICE PROVIDER IF S1=2 AND CONTRACTOR IF S1=3]

S2 Are you the person who was most involved in making the decision to install the equipment that received the incentives through <PROGRAM>?

1. Yes
2. No
8. (Don't know) [DO NOT READ]
9. (Refused) [DO NOT READ]

[ASK IF S2<>1]

S3 Who was primarily responsible for making the decision to install the equipment purchased through <PROGRAM>?
[RECORD NAME, TITLE AND CONTACT INFORMATION. THANK AND TERMINATE OR ASKED TO BE TRANSFERRED TO THAT PERSON]

[READ IF S1<>1] This survey asks questions about HIGH EFFICIENCY improvements for which <COMPANY> received an incentive in <YEAR>. Please answer the questions from the perspective of <COMPANY>. For example, when I refer to "YOUR COMPANY," I am referring to <COMPANY>. If you are not familiar with certain aspects of the project, please just say so and I will skip to the next question.

INSTALLATION VERIFICATION

To start with, I would like to confirm some information in <UTILITY>'s program tracking database.

V00. Our records show that your company completed a high efficiency project at <ADDRESS> in <YEAR>. Is that address correct?

1. (Yes)
2. (No, at a different address)
3. (Did not make any improvements) - TERMINATE
8. (Don't know) [DO NOT READ]
9. (Refused) [DO NOT READ]

[ASK IF V00=2]
V00a. Can you please provide me with the correct address? [RECORD ADDRESS]

[CREATE VERIFIED ADDRESS VARIABLE]

[TERMINATE IF ADDRESS = DK/REF]

V0. Was the project completed as part of a new construction project or existing building upgrade project?

1. New construction
2. Existing building
8. (Don't know) [DO NOT READ]
9. (Refused) [DO NOT READ]

[IF V0=1, THANK AND TERMINATE]
[REPEAT AS A LOOP FOR ALL MEASURES ASSOCIATED WITH THE PARTICIPANTS IN THE SAMPLE FILE]

[ASK IF MEASD<>18, 14,23,10, 24, 25, 29, 22]

V1A Our program records indicate that your company installed <MEASQ> <MEASD>. Is that correct? [DO NOT READ RESPONSES]

1. (Yes)
2. (Yes, but different number) [RECORD NUMBER]
4. (No, did not make improvement)
8. (Don't know)
9. (Refused)

[ASK IF MEASD=18, 14, 23, 24, 29,22]

V1B Our program records indicate that your company made <MEASD> at <ADDRESS>. Is that correct? [DO NOT READ RESPONSES]

1. (Yes)
3. (No, did not make improvement)
8. (Don't know)
9. (Refused)

[ASK IF MEASD=10]

V1C Our program records indicate that your company performed <MEASQ> gas system tune-up. Is that correct? [DO NOT READ RESPONSES]

1. (Yes)
3. (No, did not make improvement)
8. (Don't know)
9. (Refused)

[ASK IF MEASD=25]

V1D Our program records indicate that your company removed <MEASQ> lighting fixtures. Is that correct? [DO NOT READ RESPONSES]

1. (Yes)
2. (Yes, but different number) [RECORD NUMBER]
4. (No, did not make improvement)
8. (Don't know)
9. (Refused)

[ASK IF V1A OR V1B OR V1C OR V1D=4,8,9]

V2 Is it possible that someone else dealt with this project?

1. Yes
2. No

8. (Don't know) [DO NOT READ]
9. (Refused) [DO NOT READ]

[IF V2=1, RECORD NAME & NUMBER]
[IF V2=1,2,8,9, THANK AND TERMINATE. RECORD DISPO AS "COULD NOT CONFIRM PARTICIPATION"]

[COMPUTE MEASQ_REV=RESPONSE #2 IN V1A, V1D]

[COMPUTE VMEASD=RESPONSE #1 IN V1A, V1D]

[ASK IF MEASD<>10, 14, 18, 23, 24,25, 29]

V3 Are all/Is the <MEASQ_REV> <VMEASD> still installed and in use?

1. Yes
2. No
8. (Don't know) [DO NOT READ]
9. (Refused) [DO NOT READ]

[ASK IF V3=2] [SKIP IF < MEASQ_REV > =1] [SKIP IF VMEAS=22]

V4 How many of the < MEASQ_REV> <VMEASD> have been removed? [NUMERIC OPEN END. ALLOW RESPONSES FROM 0 TO 999; 9998=DON'T KNOW; 9999=REFUSED]

[ASK IF V3=2]

V5 Why were some or all of the <VMEASD> removed? [OPEN END; 998=DON'T KNOW; 999=REFUSED]

PROJECT DEFINITION

[READ IF END USE=1,2,3,4] For the rest of the survey, I would like for you to think about the <END USE> project that included the following improvements:

<VMEASD LIST>

I will refer to this simply as "the project" from now on.

[READ IF END USE=5] For the rest of the survey, I would like for you to think about the high efficiency project that included the following improvements:

<VMEASD LIST>

I will refer to this simply as "the project" from now on.

WARM-UP QUESTIONS

[SKIP TO INTRO BEFORE Q1 IF END USE=1 OR (VMEASD=10, 14, 23,24,29, 38 AND MEAS_SUM=1)]

I would like to start with some questions related to the equipment you replaced as part of this project, as well as about your facility.

Q12 Approximately how old was the equipment that was removed as part of this project? Was it..? [IF RESPONDENT SAYS IT DEPENDS ON THE EQUIPMENT TYPE, PROBE FOR AVERAGE ACROSS ALL MAJOR EQUIPMENT TYPES THAT WERE A PART OF THE PROJECT]

1. Less than five years old
2. Between 5 and 10 years old
3. Between 11 and 15 years old
4. More than 15 years old

8 (Don't know) [DO NOT READ]

9. (Refused) [DO NOT READ]

Q13 How would you describe the condition of your old equipment? Was it in good condition, fair condition, poor condition or was it not working?

1. Good condition
2. Fair condition
3. Poor condition
4. Not working

8 (Don't know) [DO NOT READ]

9. (Refused) [DO NOT READ]

Q14 How much longer do you think it would have lasted? [NUMERIC OPEN END. RECORD RESPONSE IN YEARS]

Q15 In your opinion, based on the economics of operating this equipment, for how many more years could you have kept this equipment functioning? [NUMERIC OPEN END. RECORD RESPONSE IN YEARS]

Q16 Over the last five years, have maintenance costs for the equipment you replaced been increasing, decreasing or staying about the same?

1. Increasing
2. Decreasing
3. Staying the same

8. (Don't know) [DO NOT READ]
9. (Refused) [DO NOT READ]

I'd now like to ask you a few general questions about your participation in <UTILITY>'s <PROGRAM>.

Q1 How did you learn about <UTILITY>'s <PROGRAM>? [MULTIPLE RESPONSE. ACCEPT UP TO 3] [DO NOT READ RESPONSES]

01. (Account Executive/Account Representative)
02. (Lockheed Martin)
03. (<UTILITY> representative)
04. (<UTILITY> website)
05. (Friend/colleague/word of mouth)
06. (Vendor)
07. (Distributor)
08. (Consultant)
09. (Supplier)
10. (Engineer)
11. (Contractor)
12. (Conference)
13. (Participated in the program before)
00. (Other, specify)
98. (Don't know)
99. (Refused)

[SKIP IF Q1=13]

Q1aa Had you participated in <PROGRAM> before completing the project?

1. Yes
2. No
8. (Don't know) [DO NOT READ]
9. (Refused) [DO NOT READ]

[ASK IF Q1AA=1]

Q1aaa When did you participate in <PROGRAM>? [RECORD MONTH AND YEAR]

[SKIP IF Q1AA=1]

Q2 When did you first learn about <UTILITY>'s <PROGRAM>? Was it BEFORE or AFTER you decided to make HIGH EFFICIENCY AS OPPOSED TO STANDARD EFFICIENCY improvements as part of this project at <ADDRESS> in <YEAR>?

1. Before
2. After
3. (At about the same time) [DO NOT READ]
8. (Don't know) [DO NOT READ]
9. (Refused) [DO NOT READ]

Q2a Why did you decide to make HIGH EFFICIENCY AS OPPOSED TO STANDARD EFFICIENCY improvements? [MULTIPLE RESPONSE. ACCEPT UP TO 3] [DO NOT READ RESPONSES]

01. (Funds, incentives, programs were available for new equipment)
02. (Part of a larger renovation/remodel project)
03. (Wanted to reduce energy bills)
04. (Wanted to save energy)
05. (Past experience with the program)
00. (Other, specify)
98. (Don't know)
99. (Refused)

Q3 Who was the most influential in identifying and recommending specific equipment options for this project? [DO NOT READ RESPONSES]

01. (Me/respondent)
02. (Contractor)
03. (Engineer)
04. (Architect)
05. (Manufacturer)
06. (Distributor)
07. (Owner)
08. (<UTILITY> representative/program staff)
00. (Other, specify)
98. (Don't know)
99. (Refused)

[SKIP IF Q3=2]

Q4 Did you work with a contractor or vendor that helped you with the choice of [IF ALL OF THE END USES<>1 READ "HIGH EFFICIENCY EQUIPMENT"; IF AT LEAST ONE END USE=1 READ "THIS EQUIPMENT"] for this project?

1. Yes
2. No
8. (Don't know) [DO NOT READ]
9. (Refused) [DO NOT READ]

Q5 Did you use a design or consulting engineer for this project?

1. Yes
2. No
8. (Don't know) [DO NOT READ]
9. (Refused) [DO NOT READ]

[ASK IF Q3=2, OR Q4=1 ELSE SKIP TO Q9]

I would now like to ask you a few questions about your interactions with your contractor on this project.

Q6 Did you have a prior working relationship with the contractor who assisted you with this project?

1. Yes
2. No
8. (Don't know) [DO NOT READ]
9. (Refused) [DO NOT READ]

[SKIP IF Q6=1]

Q7 How did you find the contractor for this project? [DO NOT READ RESPONSES]

01. (Have worked with this contractor before)
02. (Referral from colleague/friend)
03. (<UTILITY> website)
04. (Phone book / white pages)
05. (Internet search)
00. (Other, specify)
98. (Don't know)
99. (Refused)

Q8 Using a 1-to-7 point scale where 1 is "not at all knowledgeable" and 7 is "very knowledgeable," how would you rate the contractor's knowledge of <UTILITY>'s <PROGRAM> requirements? [SCALE 1–7; 8=DON'T KNOW, 9=REFUSED]

Q9 Did your <UTILITY> account manager or account executive assist you with this project? [DO NOT READ RESPONSES]

(INTERVIEWER NOTE: A utility account manager/account executive is a utility employee that manages your electric or gas account. It is not an internal company employee.)

1. (Yes)
2. (No, don't have a utility account manager)
3. (No, have a utility account manager but they weren't involved)
8. (Don't know)
9. (Refused)

Q10 Prior to starting this project, did you have an energy study performed at this facility that was at least partially paid for by <UTILITY>? An energy study would involve analyzing potential energy savings at this facility and identifying incentives for equipment upgrades. <UTILITY> would have shared the cost of this study with you.

1. Yes
2. No
8. (Don't know) [DO NOT READ]
9. (Refused) [DO NOT READ]

Q11 Did you have any interactions with <IF UTILITY=1 READ "PROGRAM REPRESENTATIVES AT CON EDISON OR LOCKHEED MARTIN"; IF UTILITY=2, READ "PROGRAM REPRESENTATIVES AT ORANGE AND ROCKLAND"> related to equipment selection?

1. Yes
2. No
8. (Don't know) [DO NOT READ]
9. (Refused) [DO NOT READ]

FREE RIDERSHIP

N3 I'm going to ask you to rate the influence of a variety of factors that might have played a role in your decision to complete the project (READ UNLESS MEAS IS ONLY 7, 10, 14, 17, 23, 25, 28, 29, 34, 37, 38 OR 39: at the efficiency level you selected. Specifically, I am interested in factors that influenced you to choose HIGH EFFICIENCY as opposed to LESS EFFICIENT options.) When answering, please use a scale of 1 to 7, where 1 is not at all influential and 7 is very influential.

[IF NEEDED: ON A SCALE OF 1 TO 7, HOW INFLUENTIAL ON YOUR DECISION TO INSTALL THE HIGH EFFICIENCY EQUIPMENT WAS . . .]
[1–7; 96=NOT APPLICABLE; 98=DON'T KNOW; 99=REFUSED]
[RANDOMIZE N3B THROUGH N3K, ALWAYS ASK N3N LAST]

N3b Availability of the <PROGRAM> incentive.

N3bb How did the incentive influence your decision to complete a HIGH EFFICIENCY AS OPPOSED TO STANDARD EFFICIENCY project? [OPEN END; 98=DON'T KNOW; 99=REFUSED]

[ASK IF Q10=1, ELSE SKIP TO N3f]

N3c Information and equipment recommendations provided through the energy study that was at least partially sponsored by <UTILITY>.

[ASK IF N3C=5,6,7]

N3cc Why do you give it this rating? [OPEN END; 98=DON'T KNOW; 99=REFUSED]

[ASK IF Q11=1, ELSE SKIP TO N3G]

N3f Information and equipment recommendations provided through the interactions with <IF UTILITY=1 READ "PROGRAM REPRESENTATIVES AT CON EDISON OR LOCKHEED MARTIN"; IF UTILITY=2, READ "PROGRAM REPRESENTATIVES AT ORANGE AND ROCKLAND">.

[ASK IF N3F=5,6,7]

N3ff Why do you give it this rating? [OPEN END; 98=DON'T KNOW; 99=REFUSED]

[ASK IF Q9=1]

N3g Endorsement or recommendation by <UTILITY> account manager or account executive.
(INTERVIEWER NOTE: A utility account manager/account executive is a utility employee that manages your electric or gas account. It is not an internal company employee.)

[ASK IF N3G=5,6,7]

N3gg Why do you give it this rating?

N3h Information from <PROGRAM> marketing and outreach activities. This might include training workshops, webinars, brochures, case studies, fact sheets, or information on the program website about energy savings opportunities.

[ASK N3HH IF N3H=5,6,7]

N3hh Why do you give it this rating?

N3m Previous experience participating in <PROGRAM>.

[ASK IF N3M=5,6,7]

N3mm How did your previous experience participating in <PROGRAM> influence your decision to complete a HIGH EFFICIENCY AS OPPOSED TO STANDARD EFFICIENCY project? [OPEN END; 98=DON'T KNOW; 99=REFUSED]

[ASK N3D IF Q4=1 OR Q3=2]

N3d Recommendation from an equipment vendor or contractor that helped you with the choice of the equipment.

N3dd Please describe the influence the equipment vendor or contractor had on your decision to complete a HIGH EFFICIENCY AS OPPOSED TO STANDARD EFFICIENCY project. [OPEN END; 98=DON'T KNOW; 99=REFUSED]

N3e Previous experience with this type of equipment

[ASK N3I IF Q5=1]

N3i A recommendation from a design or consulting engineer

[ASK N3II IF Q5=1]

N3ii: Please describe the influence the design and consulting engineer had on your decision to complete A HIGH EFFICIENCY AS OPPOSED TO STANDARD EFFICIENCY project [OPEN END; 98=Don't know; 99=Refused]

N3j Standard practice in your business/industry

N3k Corporate policy or guidelines

N3n Were there any other factors the evaluation team haven't discussed that were influential in your decision to complete a HIGH EFFICIENCY AS OPPOSED TO STANDARD EFFICIENCY project?

00. [RECORD VERBATIM]
96. (Nothing else influential) [DO NOT READ]
98. (Don't Know) [DO NOT READ]
99. (Refused) [DO NOT READ]

[ASK N3NN IF N3N=00]

N3nn Using the same 1 to 7 scale, where 1 means not at all influential and 7 means very influential, how would you rate the influence of this factor? [1–7; 96=NOT APPLICABLE; 98=DON'T KNOW; 99=REFUSED]

[READ IF (N3B, N3C, N3F, N3G, N3H, N3M =5,6,7)]

You just mentioned that the following <PROGRAM> related factors were influential in your decision to complete a HIGH EFFICIENCY AS OPPOSED TO STANDARD EFFICIENCY project.

[READ IF N3B=5,6,7]

N3b Availability of the <PROGRAM> incentive

[READ IF N3C=5,6,7]

N3c Information and equipment recommendations provided through the energy study that was at least partially sponsored by <UTILITY>

[READ IF N3F=5,6,7]

N3f Information and equipment recommendations provided through the interactions with <IF UTILITY=1 READ "PROGRAM REPRESENTATIVES AT CON EDISON OR LOCKHEED MARTIN"; IF UTILITY=2, READ "PROGRAM REPRESENTATIVES AT ORANGE AND ROCKLAND">

[READ IF N3G=5,6,7]

N3g Endorsement or recommendation by <UTILITY> account manager or account executive

[READ IF N3H=5,6,7]

N3h Information from <PROGRAM> marketing and outreach activities.

[READ IF N3M=5,6,7]

N3m Previous experience with participating in <PROGRAM>

[READ IF (N3D, N3E, N3J, N3K =5,6,7)]

You also just mentioned that the following other factors were influential in your decision to complete a high efficiency project.

[READ IF N3D=5,6,7]

N3d Recommendation from an equipment vendor or contractor that helped you with the choice of the equipment

[READ IF N3E=5,6,7]

N3e Previous experience with this type of equipment

[READ IF N3J=5,6,7]

N3j Standard practice in your business/industry

[READ IF N3K=5,6,7]

N3k Corporate policy or guidelines

N3p If you were given a TOTAL of 100 points that reflect the influence of all possible factors on your decision to install high efficiency equipment as opposed to less efficient equipment,, and you had to divide those 100 points between: (1) <PROGRAM> factors and (2) other factors not directly related to <PROGRAM> how many points would you give to the <PROGRAM>?

Points given to program: [RECORD 0 TO 100; 998=DON'T KNOW; 999=REFUSED]

[CALCULATE VARIABLE "OTHERPTS" AS: 100 MINUS N3p RESPONSE; IF N3p=998, 999, SET OTHERPTS=BLANK]

N3o And how many points would you give to other factors not directly related to <PROGRAM>? [RECORD 0 TO 100; 998=DON'T KNOW; 999=REFUSED]

[ASK IF N3O<>OTHERPTS]

INC1 The last question asked you to divide a TOTAL of 100 points between the program and other factors. You just noted that you would give <N3p RESPONSE> points to the program. Does that mean you would give <OTHERPTS> points to other factors?

1. Yes
2. No
8. (Don't know) [DO NOT READ]
9. (Refused) [DO NOT READ]

[IF INC1=2, GO BACK TO N3P]

N5 Now, using a likelihood scale from 1 to 7, where 1 is "not at all likely" and 7 is "extremely likely," if <PROGRAM>, including incentives, and other program factors that the evaluation team have just discussed had not been available, what is the likelihood that you would STILL have completed the HIGH EFFICIENCY project? [RECORD 1 TO 7; 98=DON'T KNOW; 99=REFUSED]

N6 We have talked about program factors and other factors that have been influential in your decision to complete the project at the efficiency level that you did. Please summarize in one sentence what influenced the decision to complete a HIGH EFFICIENCY as opposed to STANDARD EFFICIENCY project. [OPEN END; 98=DON'T KNOW; 99=REFUSED]

Consistency Checks – Individual Components vs. Program Score

[ASK IF (N3P>69 AND ALL OF (N3B, N3C, N3F, N3G, N3H, N3M =1,2,3), ELSE SKIP TO N4AA]

N4 Earlier, you gave <N3p RESPONSE> points to the influence of <PROGRAM>. I would interpret that to mean that the program was quite influential in your decision to complete the HIGH EFFICIENCY project. However, when I asked you about the influence of individual elements of the program I recorded some answers that would imply that they were not that influential. Just to make sure I have recorded this properly will you explain the role <PROGRAM> had in your decision to complete the high efficiency project? [OPEN END; 98=DON'T KNOW; 99=REFUSED]

[ASK IF N3P<31 AND ANY ONE OF (N3B, N3C, N3F, N3G, N3H=5,6,7) ELSE SKIP TO N5]

N4aa Earlier, you gave <N3P RESPONSE> points to the influence of <PROGRAM>. I would interpret that to mean that the program was not very influential in your decision to complete the HIGH EFFICIENCY project. However, when I asked about the influence of individual elements of the program I recorded some answers that would imply that they were very influential. Just to make sure I have recorded this properly will you explain the role <PROGRAM> had in your decision to complete the HIGH EFFICIENCY project?

Consistency Checks – Individual Components vs. Likelihood of Installing without Program

[SKIP IF ASKED N4AA, OTHERWISE ASK IF ANY IN N3B, N3C, N3F, N3G, N3H, N3M=5,6,7 AND N5=5,6,7]

N5a Earlier, you said that there is <N5 RESPONSE> in 7 likelihood that you would have installed the same efficiency equipment absent <PROGRAM>. I would interpret that to mean that the program was not that influential on your decision to complete the HIGH EFFICIENCY project. However, when I asked you about the influence of individual elements of the program I recorded some answers that would imply that they were quite influential. Just to make sure I have recorded this properly will you explain the role <PROGRAM> had in your decision to complete the HIGH EFFICIENCY project? [OPEN END; 98=DON'T KNOW; 99=REFUSED]

[SKIP IF ASKED N4, OTHERWISE ASK IF ALL IN N3B, N3C, N3F, N3G, N3H, N3M=1,2,3 AND N5=1,2,3]

N5aa Earlier, you said that there is <N5 RESPONSE> in 7 likelihood that you would have installed the same efficiency equipment absent <PROGRAM>. I would interpret that to mean that the program was quite influential on your decision to complete the HIGH EFFICIENCY project. However, when I asked you about the influence of individual elements of the program I recorded some answers that would imply that they were not that influential. Just to make sure I have recorded this properly will you explain the role <PROGRAM> had in your decision to complete the HIGH EFFICIENCY project? [OPEN END; 98=DON'T KNOW; 99=REFUSED]

Consistency Checks – Individual Components vs. Program Knowledge

[ASK IF ANY IN N3B, N3C, N3F, N3G, N3H, N3M=5,6,7 AND Q2=2]

N6a Earlier, you said that you learned about the program AFTER you made the decision to complete the HIGH EFFICIENCY project. However, when I asked you about the influence of individual elements of the program I recorded some answers that would imply that they were quite influential. Just to make sure I have recorded this properly will you explain the role <PROGRAM> had in your decision to complete the HIGH EFFICIENCY project? [OPEN END; 998=DON'T KNOW; 999=REFUSED]

[READ IF ANY IN QN3B-QN3M<>5, 6, OR 7, OR IF QN5<>1,2, 3, OR IF QN3P<50]

We talked quite a bit about the influence of <PROGRAM> on the efficiency level of the equipment. I would now like to talk about how <PROGRAM> might have influenced the timing of your project. Remember, when I say <PROGRAM>, I mean all of the components that the evaluation team talked about before, such as incentives, support and recommendations from program staff, as well as marketing and outreach activities.

[ASK IF ANY IN QN3B-QN3M<>5, 6, OR 7, OR IF QN5<>1,2, 3, OR IF QN3P<50]

N7 Did the <PROGRAM> cause you to complete your project EARLIER than you otherwise would, or did the <PROGRAM> have no influence on when you completed the project?

1. Caused to install earlier
2. Did not influence when installed
3. (Would not have installed the equipment at all without <PROGRAM>) [DO NOT READ]
8. (Don't know) [DO NOT READ]
9. (Refused) [DO NOT READ]

[ASK N7A IF N7=3]

N7a Just to confirm, if <PROGRAM> had not been available, you would NOT have completed the project at all, is that correct?

1. Yes [SKIP TO N8a]
2. No
8. (Don't know) [DO NOT READ]
9. (Refused) [DO NOT READ]

[ASK N7B IF N7=1]

N7b If <PROGRAM> had not been available, when would you have completed the project? Would you say . . .

1. Within 6 months of when you did
2. 6 months to 1 year later
3. 1 – 2 years later
4. 2 – 3 years later
5. 3 – 4 years later
6. 4 or more years later
8. (Don't know) [DO NOT READ]
9. (Refused) [DO NOT READ]

[ASK N7C IF N7B=2-6]

N7c How did the program influence your decision to complete the project earlier than you would otherwise? [RECORD VERBATIM]

[ASK N7C IF N7B=6]

N7d Why do you think the installation would have happened 4 or more years later?

00. [RECORD VERBATIM]
98. (Don't know) [DO NOT READ]
99. (Refused) [DO NOT READ]

[READ IF ANY IN QN3B-QN3M<>5, 6, OR 7, OR IF QN5<>1,2, 3, OR IF QN3P<50]

And now, I want to understand if the <PROGRAM> influenced the scope or the size of the the project. When answering please ONLY think about how the program affected the size and scope of [IF END USES<>1 READ "HIGH EFFICIENCY"] project that you installed.

[ASK IF ANY IN QN3B-QN3M<>5, 6, OR 7, OR IF QN5<>1,2, 3, OR IF QN3P<50]

N8 If <PROGRAM> had not been available, would the scope or size of your HIGH EFFICIENCY project have been larger, the same, or smaller?

1. Larger
2. Same
3. Smaller
4. (Would not have installed any equipment without <PROGRAM>) [DO NOT READ]
8. (Don't know) [DO NOT READ]
9. (Refused) [DO NOT READ]

[ASK N8A IF N8=4]

N8a Just to confirm, if <PROGRAM> had not been available, you would NOT have completed the project at all, is that correct?

1. Yes [SKIP TO N10]
2. No
8. (Don't know) [DO NOT READ]
9. (Refused) [DO NOT READ]

[ASKIF N8=3, OR N8A=2,8,9]

N8b In percentage terms, what percentage of the HIGH EFFICIENCY project would have happened anyway in the absence of <PROGRAM>? [1-99, 998=DK; 999=REF] [PROBE: YOUR BEST ESTIMATE IS FINE]

[ASK IF N8B=998 OR 999]

N8c Would you say you would have done 25%, 50%, or 75% of the project if the program had not been available?

1. 25%
2. 50%
3. 75%
4. (0%)
5. (100%)
8. (Don't know) [DO NOT READ]
9. (Refused) [DO NOT READ]

Payback Battery

I'd like to find out more about the payback criteria <COMPANY> may use for its investments in facility improvements.

N10 Does <COMPANY> require that payback criteria be adhered to when making decisions about capital improvement investments?

1. Yes
2. No
8. (Don't know) [DO NOT READ]
9. (Refused) [DO NOT READ]

[ASK IF N10=1]

N10a What payback criteria does your company use? [OPEN END]

[ASK IF N10=1]

N10b Did the program incentives make a difference in meeting your payback criteria or would your payback criteria have been achieved anyway, even without the incentives?

1. Program incentives helped meet payback criteria
2. Payback criteria would have been met without the incentives
8. (Don't know) [DO NOT READ]
9. (Refused) [DO NOT READ]

Corporate Policy Battery

[ASK N11 AND N11A IF N3K=5,6,7]

N11 Does your company have a corporate environmental policy to reduce environmental emissions or energy use? This could include having annual energy savings goals as part of a corporate energy management system such as ISO 50001. Other examples would be to "buy green" or use sustainable approaches to business investments.

1. Yes
2. No
8. (Don't know) [DO NOT READ]
9. (Refused) [DO NOT READ]

[ASK N11A IF N11=1]

N11a What specific corporate policy influenced your decision to complete this HIGH EFFICIENCY project?

00. [RECORD VERBATIM]
98. (Don't know) [DO NOT READ]
99. (Refused) [DO NOT READ]

Standard Practice Battery

[ASK N12 AND N12A IF N3J=5,6,7]

In an earlier question, you rated the influence of STANDARD PRACTICE in your industry very highly in your decision-making.

N12 What industry group or trade organization do you look to in order to establish standard practice for your industry?

00. [RECORD VERBATIM]
98. (Don't know) [DO NOT READ]
99. (Refused) [DO NOT READ]

N12A Could you please rate the influence of the PROGRAM, relative to this standard industry practice, on your decision to complete this high efficiency project? Would you say the program was much more influential, somewhat more influential, equally influential, somewhat less influential, or much less influential than the standard practice or policy?

1. Much more influential
2. Somewhat more influential
3. Equally influential
4. Somewhat less influential
5. Much less influential
8. (Don't know) [DO NOT READ]
9. (Refused) [DO NOT READ]

Additional Projects – Other End Uses ThatAre a Part of the Same Project

[ASK OF PROJECTS WITH MULTIPLE END USES]

Our records show that as part of the same project, <COMPANY> also received an incentive for the following improvements in addition to the improvements the evaluation team just discussed in detail.

[SAMEPROJ_MEASD]

N13 Was it a single decision to complete all of improvements or did these additional improvements go through its own decision process?

01. Single decision
02. Each improvement went through its own decision process
00. (Other, specify) [DO NOT READ]
98. (Don't know) [DO NOT READ]
99. (Refused) [DO NOT READ]

Additional Projects – Other Projects

[ASK OF CONTACTS WITH MULTIPLE PROJECTS]

Our records show that <COMPANY> also received an incentive for <PROJCOUNT> additional projects. [IF NEEDED: ADDITIONAL PROJECTS INCLUDES IMPROVEMENTS MADE AT OTHER SITES, OR OTHER TYPES OF IMPROVEMENTS AT THIS SITE THAT WE HAVEN'T FOCUSED ON.]

N13a Was it a single decision to complete all of those projects for which you received an incentive or did each project go through its own decision process?

01. Single decision
02. Each project went through its own decision process
00. (Other, specify) [DO NOT READ]
98. (Don't know) [DO NOT READ]
99. (Refused) [DO NOT READ]

SPILLOVER

Next, I would like to discuss any HIGH EFFICIENCY equipment you might have installed AFTER participating in the <PROGRAM> that did not receive incentives from <UTILITY>.

SP1a SINCE you participated in the <PROGRAM>, did you install any additional HIGH EFFICIENCY equipment at THIS facility that did NOT receive incentives through <UTILITY>?

1. Yes
2. No
8. (Don't know) [DO NOT READ]
9. (Refused) [DO NOT READ]

SP1b And, since you participated in the <PROGRAM>, did you install any additional HIGH EFFICIENCY equipment at OTHER facilities in <UTILITY>'s service territory that did NOT receive incentives through <UTILITY>?

1. Yes
2. No
8. (Don't know) [DO NOT READ]
9. (Refused) [DO NOT READ]

[ASK IF SP1A=1 OR SP1B=1, ELSE TO FIRMOGRAPHICS]

SP1c Just to confirm, you made HIGH EFFICIENCY improvements AFTER you participated in the <PROGRAM>, is that correct?

1. Yes
2. No
8. (Don't know) [DO NOT READ]
9. (Refused) [DO NOT READ]

[ASK IF SP1A=1 OR SP1B=1]

SP1d Just to confirm, you made HIGH EFFICIENCY improvements that did NOT receive incentives through the <PROGRAM>, is that correct?

1. Yes
2. No
8. (Don't know) [DO NOT READ]
9. (Refused) [DO NOT READ]

[ASK IF SP1C=1 AND SP1D=1]

SP1e What is the primary heating fuel type at the facility or facilities where you installed the high efficiency equipment?

1. Gas
2. Electric
00. (Other, specify) [DO NOT READ]
3. (A mix depending on facility) [DO NOT READ]
8. (Don't know) [DO NOT READ]
9. (Refused) [DO NOT READ]

Lighting

SP2a Did you install any HIGH EFFICIENCY lighting equipment without receiving an incentive from <UTILITY>?

1. Yes
2. No
8. (Don't know) [DO NOT READ]
9. (Refused) [DO NOT READ]

[ASK IF SP2A=1, ELSE SKIP TO SP3A]

SP2b Why did you purchase this HIGH EFFICIENCY lighting equipment without applying for incentives through <UTILITY>? [DO NOT READ RESPONSES]

01. (Did not know how to get incentive after declining originally)
02. (No time to participate/needed equipment immediately)
03. (No incentive was offered)
04. (The amount of the incentive was not large enough)
00. (Other, specify)
96. (I did get an incentive)
98. (Don't know)
99. (Refused)

[ASK IF SP2B<>96, ELSE SKIP TO SP3A]

SP2c1 Did the <PROGRAM> IN ANY WAY influence your decision to install the HIGH EFFICIENCY lighting equipment that you ended up installing afterward?

1. Yes
2. No
8. (Don't know) [DO NOT READ]
9. (Refused) [DO NOT READ]

[ASK IF SP2C1=1]

SP2c2 How much did your experience with <UTILITY>'s <PROGRAM> influence your decision to install the HIGH EFFICIENCY lighting equipment that you ended up installing afterward? Please use a scale from 1 to 7 where 1 means no influence and 7 means great influence. [SCALE 1–7; 98=DON'T KNOW; 99=REFUSED]

[ASK IF SP2C2>5, ELSE SKIP TO SP3A]

SP2m_2 How did the <PROGRAM> influence your decision to install the HIGH EFFICIENCY lighting equipment that you ended up installing afterward? [OPEN END, UP TO 240; 998=DON'T KNOW; 999=REFUSED]

SP2d What types of HIGH EFFICIENCY lighting equipment did you install without receiving an incentive through <UTILITY>? [MULTIPLE RESPONSE. ACCEPT UP TO THREE] [DO NOT READ RESPONSES]

01. (Linear fluorescent light lamps – T8)
02. (Linear fluorescent light lamps – T5)
03. (High-Intensity Discharge or HID lamps)
04. (Compact fluorescent lamps or CFLs)
05. (LED lamps)
06 . (LED exit signs)
07. (Halogen light lamps)
08. (Lighting Occupancy Sensors)
09. (Daylighting Controls)
96. (Did not install any equipment) [SKIP TO SP3A]
00. (Other, specify)
98. (Don't know)
99. (Refused)

[ASK ABOUT FIRST MEASURE; REPEAT SP2E-SP2H FOR EACH MEASURE MENTIONED. DO NOT LOOP IF SP2D=00, AND SKIP TO SP3A]

SP2e How many [READ IN RESPONSE FROM SP2D] did you install? [NUMERIC OPEN END: 1–5000; 9998=DON'T KNOW; 9999=REFUSED]

[SKIP IF SP2D=6 OR 8 OR 9]

SP2f What is the average wattage of the [READ IN RESPONSE FROM SP2D] you installed? [NUMERIC OPEN END: 1–5000; 9998=DON'T KNOW; 9999=REFUSED]

SP2g What equipment did these [READ IN RESPONSE FROM SP2D] replace? [OPEN END; 98=DON'T KNOW; 99=REFUSED]

[SKIP IF SP2D=6 OR 8 OR 9]SP2h What is the average wattage of the removed equipment? [OPEN END]

[SKIP IF SP2D=6 OR 8 OR 9]

SP2i And were these [READ IN RESPONSE FROM SP2D] installed inside, outside, or in a refrigerated space?

01. Inside (in heated or cooled space)
02. Outside (in non-heated and non-cooled space)
03. Refrigerated space (in a cooler or freezer)
00. (Other, specify) [DO NOT READ]
98. (Don't know) [DO NOT READ]

99. (Refused) [DO NOT READ]

[ASK IF SP2I = 1]

SP2j Is the inside space heated, cooled, heated and cooled, or unconditioned?

01. Heated AND cooled
02. Heated only
03. Cooled only
04. Not conditioned
00. (Other, specify) [DO NOT READ]
98. (Don't Know) [DO NOT READ]
99. (Refused) [DO NOT READ]

[ASK IF SP2ID=6]

SP2j Which of the following statements describes the LED exit signs that you installed?

1. Most are single-sided
2. Most are double-sided
8. (Don't know) [DO NOT READ]
9. (Refused) [DO NOT READ]

[ASK IF SP2ID=6]

SP2k What type of equipment did these [READ IN RESPONSE FROM SP2D] replace?

01. Single Sided Incandescent Exit Sign
02. Double Sided Incandescent Exit Sign
03. Single Sided CFL Exit Sign
04. Double Sided CFL Exit Sign
05. Or was the equipment newly installed and did not replace anything?
00. (Other, specify) [DO NOT READ]
98. (Don't know) [DO NOT READ]
99. (Refused) [DO NOT READ]

[ASK IF SP2D=08 OR 09]

SP2l What is the total average wattage controlled by each occupancy or daylighting sensor? [OPEN END]

Cooling

SP3a Did you install any HIGH EFFICIENCY cooling equipment without receiving an incentive from <UTILITY>?

1. Yes
2. No
8. (Don't know) [DO NOT READ]
9. (Refused) [DO NOT READ]

[ASK IF SP3A=1, ELSE SKIP TO SP4A]

SP3b Why did you purchase this cooling equipment without receiving an incentive through <UTILITY>? [DO NOT READ RESPONSES]

01. (Did not know how to get incentive after declining originally)
02. (No time to participate/needed equipment immediately)
03. (No incentive was offered)
04. (The amount of the incentive was not large enough)
00. (Other, specify)
96. (I did get an incentive)
98. (Don't know)
99. (Refused)

[ASK IF SP3B<>96, ELSE SKIP TO SP4A]

SP3c1 Did the <PROGRAM> IN ANY WAY influence your decision to install the HIGH EFFICIENCY cooling equipment that you ended up installing afterward?

1. Yes
2. No
8. (Don't know) [DO NOT READ]
9. (Refused) [DO NOT READ]

[ASK IF SP3C1=1]

SP3c2 How much did your experience with <UTILITY>'s <PROGRAM> influence your decision to install the HIGH EFFICIENCY cooling equipment that you ended up installing afterward? Please use a scale from 1 to 7 where 1 means no influence and 7 means great influence. [SCALE 1–7; 98=DON'T KNOW; 99=REFUSED]

[ASK IF SP3C2>5, ELSE SKIP TO SP4A]

SP3m How did the <PROGRAM> influence your decision to install the HIGH EFFICIENCY cooling equipment that you ended up installing? [OPEN END; 98=DON'T KNOW; 99=REFUSED]

SP3d What types of HIGH EFFICIENCY cooling equipment did you install without receiving an incentive through <UTILITY>? [MULTIPLE RESPONSE. ACCEPT UP TO THREE] [DO NOT READ RESPONSES]

01. (Split air conditioning system)
02. (Packaged air conditioning system)
03. (Heat pump)
04. (Chiller – air cooled)
05. (Chiller – water cooled)
06. (Compressor)
07. (Condenser)
08. (Cooling tower)
09. (Air handler)
10. (Window air conditioner)
11. (Through the wall air conditioner or packaged terminal A/C (PTAC) unit)
12. (Through the wall heat pump or packaged terminal heat pump (PTHP) unit)
00. (Other, specify)
96. (Didn't install any equipment) [SKIP TO SP4A]
98. (Don't know)
99. (Refused)

SP3i How many cooling units of this type were installed? [NUMERIC OPEN END; 1–500; 996=NOT APPLICABLE, 998=DON'T KNOW, 999=REFUSED]

SP3e How many tons of cooling does this new equipment provide? [NUMERIC OPEN END; 1–500; 996=NOT APPLICABLE, 998=DON'T KNOW, 999=REFUSED]

[ASK SC3F IF SP3E=998,999]

SP3f Approximately, how many square feet of floor space does this cooling equipment serve? [NUMERIC OPEN END; 1–500,000; 999,996=NOT APPLICABLE; 999,998=DON'T KNOW; 999,999=REFUSED]

SP3g Did this new equipment replace old equipment?

1. Yes
2. No
8. (Don't know) [DO NOT READ]
9. (Refused) [DO NOT READ]

[SKIP TO SP4A IF SP3G=2,8,9]

SP3j Was the old equipment working at the time of replacement?

1. Yes
2. No
8. (Don't know) [DO NOT READ]
9. (Refused) [DO NOT READ]

SP3h How old was the replaced equipment? [DO NOT READ RESPONSES]

1. (0–4 years)
2. (5–9 years)
3. (10–14 years)
4. (15–19 years)
5. (20 years or older)
8. (Don't know)
9. (Refused)

Refrigeration

SP4a Did you install any HIGH EFFICIENCY refrigeration equipment without receiving an incentive from <UTILITY>?

1. Yes
2. No
8. (Don't know) [DO NOT READ]
9. (Refused) [DO NOT READ]

[ASK IF SP4A=1, ELSE SKIP TO SP5A]

SP4b Why did you purchase HIGH EFFICIENCY refrigeration equipment without receiving an incentive through <UTILITY>? [DO NOT READ RESPONSES]

01. (Did not know how to get incentive after declining originally)
02. (No time to participate/needed equipment immediately)
03. (No incentive was offered)
04. (The amount of the incentive was not large enough)
00. (Other, specify)

96. (I did get an incentive)
98. (Don’t know)
99. (Refused)

[ASK IF SP4B<>96, ELSE SKIP TO SP5A]

SP4c1 Did the <PROGRAM> IN ANY WAY influence your decision to install the HIGH EFFICIENCY refrigeration equipment that you ended up installing afterward?

1. Yes
2. No
8. (Don’t know) [DO NOT READ]
9. (Refused) [DO NOT READ]

[ASK IF SP4C1=1]

SP4c2 How much did your experience with <UTILITY>‘s <PROGRAM> influence your decision to install HIGH EFFICIENCY refrigeration equipment that you ended up installing afterward? Please use a scale from 1 to 7 where 1 means no influence and 7 means great influence. [SCALE 1–7; 98=DON’T KNOW; 99=REFUSED]

[ASK IF SP4C2>5, ELSE SKIP TO SP5A]

SP4m How did the <PROGRAM> influence your decision to install the HIGH EFFICIENCY refrigeration equipment that you ended up installing afterward? [OPEN END; 98=DON’T KNOW; 99=REFUSED]

SP4d What type of HIGH EFFICIENCY refrigeration equipment did you install without receiving an incentive through <UTILITY>? [MULTIPLE RESPONSE. ACCEPT UP TO THREE] [DO NOT READ RESPONSES]

01. (Strip curtains)
02. (Anti-sweat controls)
03. (EC motor for cooler or freezer)
04. (Solid door cooler)
05. (Solid door freezer)
06. (Glass door cooler)
07. (Glass foor freezer)
08. (Case cooler)
09. (Case freezer)
10. (Condenser)
11. (LED case lights with motion controls)
12. (LED case lights without motion controls)

13. (Door gaskets)
14. (Refrigerated case night covers)
15. (Walk-in cooler fan)
16. (Vending machine controls)
17. (Automatic door closer for cooler/freezer)
18. (Refrigerated cooler controls)
19. (Ice machine)
96. (Didn't install any equipment)
98. (Don't know)
99. (Refused)

Kitchen Equipment

SP5a Did you install any HIGH EFFICIENCY kitchen or vending equipment without receiving an incentive from <UTILITY>, such as fryers, griddles, commercial ovens, pre-rinse spray valves, hot food holding cabinets, or vending machine controls?

1. Yes
2. No
8. (Don't know) [DO NOT READ]
9. (Refused) [DO NOT READ]

[ASK IF SP5A=1, ELSE SKIP TO SP6A]

SP5b Why did you purchase HIGH EFFICIENCY kitchen equipment without receiving an incentive through <UTILITY>? [DO NOT READ RESPONSES]

01. (Did not know how to get incentive after declining originally)
02. (No time to participate/needed equipment immediately)
03. (No incentive was offered)
04. (The amount of the incentive was not large enough)
00. (Other, specify)
96. (I did get an incentive)
98. (Don't know)
99. (Refused)

[ASK IF SP5B<>96, ELSE SKIP TO SP6A]

SP5c1 Did the <PROGRAM> IN ANY WAY influence your decision to install the HIGH EFFICIENCY kitchen equipment that you ended up installing afterward?

1. Yes
2. No
8. (Don't know) [DO NOT READ]
9. (Refused) [DO NOT READ]

[ASK IF SP5C1=1]

SP5c2 How much did your experience with <UTILITY>'s <PROGRAM> influence your decision to install HIGH EFFICIENCY kitchen equipment that you ended up installing afterward? Please use a scale from 1 to 7 where 1 means no influence and 7 means great influence. [SCALE 1–7; 98=DON'T KNOW; 99=REFUSED]

[ASK IF SP5C2>5, ELSE SKIP TO SP6A]

SP5m How did the <PROGRAM> influence your decision to install the HIGH EFFICIENCY kitchen equipment that you ended up installing afterward? [OPEN END; 98=DON'T KNOW; 99=REFUSED]

SP5d What types of HIGH EFFICIENCY kitchen equipment did you install without receiving an incentive through <UTILITY>? [MULTIPLE RESPONSE. ACCEPT UP TO THREE] [IF NEEDED: IS THIS EQUIPMENT GAS OR ELECTRIC?] [DO NOT READ RESPONSES]

01. (Electric steamers)
02. (Gas steamers)
03. (Convection ovens)
04. (Combination ovens) [IF NEEDED: THIS IS A COMBINATION CONVECTION AND STEAM OVEN]
05. (Electric griddles)
06. (Gas griddles)
07. (Electric fryers)
08. (Gas fryers)
09. (Pre-rinse spray valves)
10. (Food holding cabinets)
11. (Snack vending machine controls)
12. (Refrigerated cooler controls)
00. (Other, specify)
96. (Didn't install any equipment)
98. (Don't know)

99. (Refused)

[ASK IF SP5D=05]

SP5e Is the kitchen's water heated by a gas or electric water heater?

1. Gas
2. Electric
0. (Other, specify) [DO NOT READ]
8. (Don't know) [DO NOT READ]
9. (Refused) [DO NOT READ]

Motors

SP6a Did you install any HIGH EFFICIENCY motors without receiving an incentive from <UTILITY>?

1. Yes
2. No
8. (Don't know) [DO NOT READ]
9. (Refused) [DO NOT READ]

[ASK IF SP6A=1, ELSE SKIP TO SP7A]

SP6b Why did you purchase HIGH EFFICIENCY motors without receiving an incentive through <UTILITY>? [DO NOT READ RESPONSES]

01. (Did not know how to get incentive after declining originally)
02. (No time to participate/needed equipment immediately)
03. (No incentive was offered)
04. (The amount of the incentive was not large enough)
00. (Other, specify)
96. (I did get an incentive)
98. (Don't know)
99. (Refused)

[ASK IF SP6B<>96, ELSE SKIP TO SP7A]

SP6c1 Did the <PROGRAM> IN ANY WAY influence your decision to install the HIGH EFFICIENCY motors that you ended up installing afterward?

1. Yes
2. No
8. (Don't know) [DO NOT READ]
9. (Refused) [DO NOT READ]

[ASK IF SP6C1=1]

SP6c2 How much did your experience with <UTILITY>'s <PROGRAM> influence your decision to install HIGH EFFICIENCY motors that you ended up installing afterward? Please use a scale from 1 to 7 where 1 means no influence and 7 means great influence. [SCALE 1–7; 98=DON'T KNOW; 99=REFUSED]

[ASK IF SP6C2>5, ELSE SKIP TO SP7A]

SP6m How did the <PROGRAM> influence your decision to install the HIGH EFFICIENCY motors that you ended up installing afterward? [OPEN END; 98=DON'T KNOW; 99=REFUSED]

SP6d For what types of equipment did you install the HIGH EFFICIENCY motors? [MULTIPLE RESPONSE. ACCEPT UP TO THREE] [DO NOT READ RESPONSES]

01. (HVAC system fans)
02. (HVAC system pumps)
03. (Non-HVAC system fans)
04. (Non-HVAC system pumps)
00. (Other, specify)
96. (Didn't install any motors) [SKIP TO SP7A]
98. (Don't know)
99. (Refused)

SP6e How many motors did you install? [NUMERIC OPEN END, 1 TO 500; 998=DON'T KNOW, 999=REFUSED]

SP6f What was the total horsepower of these motors? [NUMERIC OPEN END, 1 TO 5000; 9998=DON'T KNOW, 9999=REFUSED]

SP6g Did the new motors also have new variable frequency drives (VFDs)?

1. Yes
2. No
8. (Don't know) [DO NOT READ]
9. (Refused) [DO NOT READ]

SP6h Did this new equipment replace old equipment?

1. Yes
2. No
8. (Don't know) [DO NOT READ]
9. (Refused) [DO NOT READ]

[ASK IF SP6H =1, ELSE TO SP7A]

SP6i How old was the replaced equipment? [DO NOT READ RESPONSES]

1. (0–4 years)
2. (5–9 years)
3. (10–14 years)
4. (15–19 years)
5. (20 years or older)
8. (Don't know)
9. (Refused)

SP6k Was the old motor working at the time of replacement?

1. Yes
2. No
8. (Don't know) [DO NOT READ]
9. (Refused) [DO NOT READ]

Heating and Water Heating

SP7a Did you install any HIGH EFFICIENCY heating or water heating equipment without receiving an incentive from <UTILITY>?

1. Yes
2. No
8. (Don't know) [DO NOT READ]
9. (Refused) [DO NOT READ]

[ASK IF SP7A=1, ELSE SKIP TO SP8A]

SP7b Why did you purchase this heating or water heating equipment without receiving an incentive through <UTILITY>? [DO NOT READ RESPONSES]

01. (Did not know how to get incentive after declining originally)
02. (No time to participate/needed equipment immediately)
03. (No incentive was offered)
04. (The amount of the incentive was not large enough)
00. (Other, specify)
96. (I did get an incentive)
98. (Don't know)
99. (Refused)

[ASK IF SP7B<>96, ELSE SKIP TO SP8A]

SP7c1 Did the <PROGRAM> IN ANY WAY influence your decision to install the heating or water heating equipment that you ended up installing afterward?

1. Yes
2. No
8. (Don't know) [DO NOT READ]
9. (Refused) [DO NOT READ]

[ASK IF SP7C1=1]

SP7c2 How much did your experience with <UTILITY>'s <PROGRAM> influence your decision to install the HIGH EFFICIENCY heating or water heating equipment that you ended up installing afterward? Please use a scale from 1 to 7 where 1 means no influence and 7 means great influence. [SCALE 1–7; 98=DON'T KNOW; 99=REFUSED]

[ASK IF SP7C2>5, ELSE SKIP TO SP8A]

SP7m How did the <PROGRAM> influence your decision to install the HIGH EFFICIENCY heating or water heating equipment that you ended up installing afterward? [OPEN END; 98=DON'T KNOW; 99=REFUSED]

SP7d What types of HIGH EFFICIENCY heating or water heating equipment did you install without receiving an incentive through <UTILITY>? [MULTIPLE RESPONSE. ACCEPT UP TO THREE] [DO NOT READ RESPONSES]

01. (Gas furnace)
02. (Electric furnace)
03. (Gas boiler)
04. (Condensing unit heater)
05. (Electric heat pump)
06. (Packaged terminal heat pump (PTHP))
07. (Tankless gas water heater)
08. (Tankless electric water heater)
09. (Solar water heater)
10. (Gas storage water heater)
11. (Electric storage water heater)
12. (Geothermal water heater)
00. (Other, specify)
96. (Didn't implement any heating measures) [SKIP TO SP8A]
98. (Don't know)
99. (Refused)

[ASK ABOUT FIRST MEASURE; REPEAT SP7E–SP7H FOR EACH MEASURE MENTIONED. DO NOT LOOP IF SP7D>7 AND <13, AND SKIP TO SP8A]

SP7j How many <SP7d> were installed? [NUMERIC OPEN END; 1–500,000; 999,996=NOT APPLICABLE; 999,998=DON'T KNOW; 999,999=REFUSED]

SP7e Approximately, how many square feet of floor space does this heating equipment serve? [NUMERIC OPEN END; 1–500,000; 999,996=NOT APPLICABLE; 999,998=DON'T KNOW; 999,999=REFUSED]

SP7f Did this new equipment replace old equipment?

1. Yes
2. No
8. (Don't know) [DO NOT READ]
9. (Refused) [DO NOT READ]

[SKIP TO NEXT LOOP OR SP8A IF SP7F=2,8,9]

SP7k Was the old equipment working at the time of replacement?

1. Yes
2. No
8. (Don't know) [DO NOT READ]
9. (Refused) [DO NOT READ]

SP7g What type of old equipment was replaced? [DO NOT READ RESPONSES]

01. (Gas furnace)
02. (Gas boiler)
03. (Electric resistance heater)
04. (Electric heat pump)
00. (Other, specify)
98. (Don't know)
99. (Refused)

SP7h How old was the replaced equipment? [DO NOT READ RESPONSES]

1. (0–4 years)
2. (5–9 years)
3. (10–14 years)
4. (15–19 years)
5. (20 years or older)
8. (Don't know)
9 . (Refused)

[ASK IF SP7D=10 OR 11]

SP7i Approximately how large is the water heater's storage tank, in gallons? [NUMERIC OPEN END, 1 TO 997; 998=DON'T KNOW, 999=REFUSED]

Other Equipment

SP8a Did you install any other HIGH EFFICIENCY equipment without receiving an incentive from <UTILITY>?

1. Yes
2. No
8. (Don't know) [DO NOT READ]
9. (Refused) [DO NOT READ]

[ASK IF SP8A=1, ELSE SKIP TO END]

SP8b Why did you purchase this HIGH EFFICIENCY equipment without receiving an incentive through <UTILITY>? [DO NOT READ RESPONSES]

01. (Did not know how to get incentive after declining originally)
02. (No time to participate/needed equipment immediately)
03. (No incentive was offered)
04. (The amount of the incentive was not large enough)
00. (Other, specify)
96. (I did get an incentive)
98. (Don't know)
99. (Refused)

[ASK IF SP8B<>96, ELSE SKIP TO END]

SP8c1 Did the <PROGRAM> IN ANY WAY influence your decision to install this HIGH EFFICIENCY equipment that you ended up installing afterward?

1. Yes
2. No
8. (Don't know) [DO NOT READ]
9. (Refused) [DO NOT READ]

[ASK IF SP8C1=1]

SP8c2 How much did your experience with <UTILITY>'s <PROGRAM> influence your decision to install this HIGH EFFICIENCY equipment that you ended up installing? Please use a scale from 1 to 7 where 1 means no influence and 7 means great influence. [SCALE 1–7; 98=DON'T KNOW; 99=REFUSED]

[ASK IF SP8C2>5]

SP8d What type or types of HIGH EFFICIENCY equipment did you install?

00. (Other, specify) [DO NOT READ]
98. (Don't know) [DO NOT READ]
99. (Refused) [DO NOT READ]

FIRMOGRAPHICS

I only have a few general questions left.

F6 What is the business type of the facility at <ADDRESS>? [READ LIST IF NECESSARY]

01. (Grocery)
02. (Medical/health)
03. (Hotel/motel)
04. (Office)
05. (Restaurant)
06. (Retail/service)
07. (Warehouse/distribution)
08. (Church/other religious)
09. (Non-profit organization)
10. (Manufacturing)
11. (K-12 school)
12. (College/university)
00. (Other, specify)
98. (Don't know)
99. (Refused)

F5 How old is this facility? [READ LIST IF NECESSARY]

01. (Less than 5 years)
02. (5–9 years)
03. (10–19 years)
04. (20–29 years)
05. (30–39 years)
06. (40–49 years)
07. (50–59 years)
08. (60 years or more)
98. (Don't know)
99. (Refused)

F2 Is the facility at <ADDRESS>?

1. Your company's only location
2. One of several locations of your company
8. (Don't know) DO NOT READ]
9. (Refused) DO NOT READ]

F10 How many employees, full plus part-time, are employed at this facility? [NUMERIC OPEN END, 0 TO 2000; 9998=DON'T KNOW, 9999=REFUSED]

[ASK IF F10=9998]

F11 Do you know the approximate number of employees? Would you say it is . . .?

1. Less than 10
2. 10–49
3. 50–99
4. 100–249
5. 250–499
6. 500 or more
8. (Don't know) [DO NOT READ]
9. (Refused) [DO NOT READ]

F1 Which of the following best describes your company's ownership of this facility?

01. My company leases the facility
02. My company owns and occupies the facility
03. My company owns the facility but does not occupy it
00. (Other, specify) [DO NOT READ]
98. (Don't know) [DO NOT READ]
99. (Refused) [DO NOT READ]

F12a Does your company pay the electric bill at this facility?

1. Yes
2. No
8. (Don't know) [DO NOT READ]
9. (Refused) [DO NOT READ]

F3 In comparison to other companies in your industry, would you describe your company as . . .?

1. A small company
2. A medium-sized company
3. A large company
4. (Not applicable) [DO NOT READ]
8. (Don't know) [DO NOT READ]
9. (Refused) [DO NOT READ]

[ASK IF Q3=2 OR Q4=1]

F4. You mentioned earlier that you worked with a vendor or contractor for this project. What is the name and contact information of the contractor or contractors that you worked with?

[RECORD CONTRACTOR COMPANY NAME, CONTACT NAME, PHONE NUMBER]
[ALLOW TO RECORD UP TO THREE CONTRACTOR CONTACTS]

These are all the questions that I have for you. Thank you very much for your time. It is greatly appreciated. Have a nice day.

Appendix F – Trade Ally Discussion Guide

The following is the trade ally interview guide that was used by Opinion Dynamics for the trade ally interviews.

<PROGRAM CONTACT> (Name of the contact person)

<COMPANY> (Name of participating company)

<ADDRESS> (Facility address)

<PROGRAM> (Con Edison's Large Commercial energy efficiency program OR Orange & Rockland's Big Energy Solutions Program)

<PROGRAM YEAR> (Year of program participation)

<UTILITY> (Con Edison or Orange & Rockland)

<END USE> The efficiency measure/equipment the customer installed and received an incentive for.

Introduction

Hello, this is _____ from Opinion Dynamics calling on behalf of <UTILITY>. THIS IS NOT A SALES CALL. We are doing a brief survey with program trade allies who have been involved in projects supported by the <PROGRAM>.

We are interested in your experience with the program and any feedback you may have received from your customers about the program. <UTILITY> plans to use the information to improve the energy efficiency programs and services it offers to its business customers.

[IF <PROGRAM CONTACT IS NOT MISSING] May I please speak with <PROGRAM CONTACT>?

[IF <PROGRAM CONTACT IS MISSING] Who might be the best person to speak with about the high efficiency project completed for <COMPANY> at <ADDRESS> through <PROGRAM>?

Would you be willing to speak with me for about 15 minutes? Is now a good time or is there a more convenient time when I could call back?

Alert interviewee that the call will be recorded.

Note that responses will remain confidential and will only be reported in aggregate with other responses.

Screener

S1. <COMPANY> has indicated that your firm was involved in the implementation of a high efficiency project at <ADDRESS> back in <MONTH AND YEAR>. Is this correct? [IF NEEDED, PROVIDE DETAIL ABOUT END USE EQUIPMENT THAT WAS A PART OF THE PROJECT] Are you the person who is most knowledgeable about your firm's involvement in this project?

[IF NOT, ASK TO BE TRANSFERRED TO THE PERSON WHO WAS INVOLVED AND/OR IS MOST KNOWLEDGEABLE ABOUT THE PROJECT]

Firmographics

The goal of this section is to gather more information into the type of trade ally's business.

I would now like to learn more about your company.

F1. What is your company's business category? [PROBE FOR: CONTRACTOR, ENGINEER, ESCO, EQUIPMENT VENDOR, ARCHITECT]

F5. What type of equipment, if any, would you say is your company's area of expertise? [PROBE, IF NECESSARY: LIGHTING, HVAC, REFRIGERATION, MOTORS, FOOD SERVICE]

a. [IF MULTIPLE AREAS] What is the MAIN area?

F4. Does your company provide services to commercial customers only or both commercial and residential customers in <UTILITY>'s service area?

a. [IF BOTH, PROBE] What percentage of your installations in <UTILITY>'s service area is in the commercial sector?

Project-Specific Influence

The goal of this section is to develop FR adjustment factors.

I have a few specific questions about your firm's recent involvement in <COMPANY>'s installation of <END USE> through <UTILITY>'s program at <ADDRESS> back in <MONTH AND YEAR>.

Thinking about the project you completed for <COMPANY>…

FR1. Can you please describe your firm's role in the selection and installation of <END USE> at <COMPANY>'s facility? [PROBE IF FIRM MERELY SUPPLIED OR INSTALLED EQUIPMENT OR IF THEY HAD A ROLE IN SELECTING IT.

PROBE ABOUT PERCEIVED LEVEL OF INFLUENCE FIRM'S RECOMMENDATION HAD ON COMPANY'S CHOICE.]

a. Did you initiate the contact with the customer about this project or did the customer reach out to you with a request for a quote? [PROBE FOR DETAIL]

FR2. When developing equipment specifications and recommendations for this project, what factors went into the decision-making process? [PROBE FOR: COST, ENERGY SAVINGS, PERFORMANCE, NON-ENERGY BENEFITS, SUCH AS DECREASED NEED FOR MAINTENANCE, AESTHETICS, ETC.]

c. What factors, if any, did <COMPANY> discuss with you when deciding on the equipment for an upgrade project? [PROBE FOR: COST, ENERGY SAVINGS, PERFORMANCE, NON-ENERGY BENEFITS, SUCH AS DECREASED NEED FOR MAINTENANCE, AESTHETICS, ETC.]

d. Did the customer request specific equipment or did the customer rely on your recommendations?

There are a variety of factors that might impact the efficiency of the equipment that you specify and recommend to your customers. I'm going to ask you to rate the influence of <UTILITY>'s program on your recommendation to go with HIGH EFFICIENCY EQUIPMENT. This influence could be through incentives, program services, and information, as well as other avenues.

FR3. Using a 1 to 7 point scale, where 1 means no influence and 7 means a great deal of influence, how influential was <UTILITY>'s program, on your recommendation to install [DO NOT READ FOR END USES WITH NO LESS-EFFICIENT OPTIONS, E.G., EMS; <u>HIGH EFFICIENCY</u>] <END USE> as part of this project?

a. Why did you give this rating?

[ASK IF FR3>3]

FR4. How specifically did the program influence your recommendations for [DO NOT READ FOR END USES WITH NO LESS-EFFICIENT OPTIONS, E.G., EMS; <u>HIGH EFFICIENCY</u>] <END USE> for this specific project? [IF NEEDED, PROBE FOR SPECIFIC INFLUENCE OF MARKETING, TRAINING, INCENTIVES]

FR5. What other factors, if any, influenced your decision to recommend that <COMPANY> install the [DO NOT READ FOR END USES WITH NO LESS-EFFICIENT OPTIONS, E.G., EMS; <u>HIGH EFFICIENCY</u>] <END USE>? [IF

NEEDED, PROBE FOR CHANGING CODES AND STANDARDS, STANDARD PRACTICE OF THE COMPANY, MANUFACTURER INCENTIVES, ETC.]

a. Which would you say was more influential – the influence of the program or the other, non-program factors?

FR6. Thinking about it differently, using a 1 to 7 likelihood scale, where 1 means not at all likely and 7 means very likely, if the PROGRAM had not been available, what is the likelihood that you would have recommended this specific [DO NOT READ FOR END USES WITH NO LESS-EFFICIENT OPTIONS, E.G., EMS; HIGH EFFICIENCY] <END USE> to <COMPANY>?

a. Why do you say that? [IF THE SCORE IS CONTRADICTORY WITH THE ONE GIVEN TO FR3, PROBE FOR REASONS FOR THE INCONSISTENT RESPONSE]

FR7. Do you know of any other vendors that worked with <COMPANY> during their implementation and/or installation of <END USE>, for example engineers or designers? If so, do you have their name and phone number?

Market Trends and Program Influence

The goal of this section is to understand, and, to the degree possible, quantify influence that <UTILITY>'s program has on the trade ally's ability and decision to recommend energy efficient equipment to their customers.

I now have a few more general questions about the market for <END USE> equipment. Before we proceed, I wanted to ask you if you are familiar with the <UTILITY> service territory.

F2. Do you know the areas where <UTILITY> provides electric and/or gas services? [IF FAMILIAR, ASK ABOUT <UTILITY>'S SERVICE TERRITORY. IF NOT, FOR CON EDISON: DESCRIBE THE AREAS AS DOWNSTATE NEW YORK, AND MENTION MANHATTAN, QUEENS, STATEN ISLAND, BROOKLYN, THE BRONX, AND WESTCHESTER.
O&R: DESCRIBE THE AREA AS ORANGE, ROCKLAND, AND SULLIVAN COUNTIES]

F3. What percentage of your company's business is in <UTILITY>'s service area?

Please answer these questions thinking about your business operations in <UTILITY>'s service area.

M1. What factors go into the decision-making process when making <END USE> equipment recommendations to potential customers? [PROBE FOR: COST, ENERGY SAVINGS, PERFORMANCE, AVAILABILITY, NON-ENERGY BENEFITS, SUCH AS DECREASED NEED FOR MAINTENANCE, COMFORT, ETC.]

M2. In what percent of sales situations do you recommend high efficiency <END USE> equipment to your customers? [IF NEEDED: WHEN I SAY HIGH EFFICIENCY EQUIPMENT, I MEAN ENERGY EFFICIENT OPTIONS THAT QUALIFY FOR PROGRAM INCENTIVES]

a. Of that number, what percent of sales situations do you also recommend less-efficient options?

[ASK IF M2 IS NOT 100%]

b. When you don't recommend high efficiency equipment options, what are the reasons?

M3. Has the frequency with which you recommend high efficiency equipment changed since you became active with the program? [IF NEEDED, CLARIFY: BECOMING ACTIVE WITH THE PROGRAM MEANS STARTING TO INTERACT WITH PROGRAM STAFF AND/OR WORK ON PROJECTS THAT APPLIED FOR <UTILITY> INCENTIVES]

If change noted:

a. What are you doing more of – presenting only high efficiency options to customers instead of presenting both high efficiency and less-efficient, or adding a high efficiency option to less-efficient options in more sales situations?

b. How influential was <UTILITY>'s program in this change? [PROBE FOR SPECIFIC PROGRAM COMPONENTS: INCENTIVES, TRAINING, PROGRAM WEBSITE, OTHER PROGRAM COMPONENTS.]

M4. Has the program had any influence at all on your business? If so has it been positive or negative influence? What influence has the program had? [PROBE FOR PROGRAM INFLUENCE ON THE NUMBER OF COMPLETED PROJECTS, SALES, REVENUES, COMPANY MARKETABILITY AND POSITIONING AMONG THE COMPETITION, ETC.]

M5. How aware, would you say, are your customers of energy efficiency and options available to make their facilities more energy efficient? How interested would you say they are? [PROBE FOR VERY, SOMEWHAT, NOT VERY, NOT AT ALL AWARE/INTERESTED]

a. Has this (awareness/interest) changed as compared to 2009? If so, how did it change?

M6. What do you view as the main barriers to the installation of energy efficient equipment for your customers? Does this vary by customer type or size? Anything else? What could be done to overcome these barriers?

M7. Thinking back to 2009 through 2011, did any of program qualifying high efficiency <END USE> equipment you installed, NOT receive incentives through the program?

a. [ASK IF M7=YES] Approximately what percentage of the program-qualifying <END USE> equipment you installed during the course of 2009 through 2011 did NOT receive incentives through the program? [PROBE FOR BEST ESTIMATE]

b. What are the reasons for NOT applying for incentives? [PROBE FOR PROGRAM APPLICATIONS BEING BURDENSOME, GENERAL RELUCTANCE TO PARTICIPATE IN THE PROGRAM, ETC.]

Program Interactions

The goal of this section is to determine whether <UTILITY> provides trade allies with any information, training, or other support to help them sell energy efficient equipment.

I now would like to ask you about your knowledge of and interactions with <UTILITY>'s program.

PI1. How long have you known about <UTILITY>'s program? How long have you been active within the program? [IF NEEDED: WHEN DID YOU COMPLETE YOUR FIRST PROJECT THAT RECEIVED INCENTIVES THROUGH <UTILITY>?]

PI1A. [ASK IF PROGRAM TRACKING DATA ARE UNAVAILABLE] Thinking about the timeframe between 2009 and 2011, how many projects did you complete that received incentives through the program? [PROBE FOR BEST ESTIMATE.]

PI2. How knowledgeable do you consider yourself about <UTILITY>'s program and its offerings? Would you say very knowledgeable, somewhat knowledgeable, not very knowledgeable, or not at all knowledgeable? [PROBE FOR KNOWLEDGE OF PROGRAM ELIGIBILITY REQUIREMENTS, INCENTED EQUIPMENT OPTIONS, INCENTIVE LEVELS, PARTICIPATION PROCESS, ETC.]

PI3. Did you attend any training sessions, meetings or events, both formal and informal, facilitated by <UTILITY>'s program between 2009 and 2011? What kind of training did you receive? [PROBE FOR TRAINING AND GUIDANCE ON HOW TO PREPARE CUSTOMER APPLICATIONS, SPECIFICS ON QUALIFIED MEASURE TYPES, SALES TRAINING, ETC.]

PI4. Do you remember receiving any marketing or promotional materials or any ongoing communications <UTILITY> between 2009 and 2011? [PROBE FOR TYPES OF MATERIALS – BROCHURES, CASE STUDIES, NEWSLETTERS, CO-BRANDED MARKETING, ETC. – AS WELL AS THEIR FREQUENCY]

PI5. Did you receive ANY OTHER information or support from the program that either improved your ability to sell energy efficiency to customers or improved your overall knowledge of energy efficient equipment options? If so, what support did you receive?

This concludes our survey. On behalf of <UTILITY>, thank you very much for your time today!

APPENDIX G – VARIABLE FREQUENCY DRIVE BASELINE INVESTIGATION

After reviewing project application materials and visiting several sites that received incentives for variable frequency drives (VFDs), the evaluation team noted recurrence of instances of customers installing the VFDs in place of existing drives that had failed. In the course of researching baseline, the evaluation team asked the site contacts of their alternatives to installing a replacement VFD. In multiple instances, the contacts responded that there was no plausible alternative; they had to replace the VFD with another. This effectively made the baseline characterization the same as the measure characterization and reduced the savings to zero. This prompted the evaluation team to establish a formal protocol to identify when VFD was the baseline for VFD measure installations.

The impact team created a battery of questions to ask every participant that installed a VFD in place of another. The goal of the formal set of questions was to determine if the baseline was a VFD or whether the customer could have regressed to a non-VFD baseline for an extended period.

This set of questions was posed both to all remaining VFD participants yet to be visited and also to those that already had been visited, via follow-up phone calls. Each interviewer was trained on the intent of the questions and how to pose them objectively prior to administration.

The questions asked were as follows:

1. How old were the preexisting VFDs?
2. Did the old VFDs fail?
3. Were you operating the VFDs in bypass?
4. How long were you operating the pumps/fans with VFDs in bypass?
5. Could you have operated the fans / pumps with their VFDs in bypass indefinitely?

Also, as an added review, for these specific sites, the evaluation team also noted the simple paybacks for these projects before and after incentives along with their associated benefit-to-cost ratio (BCR) for context. Table G-1 below shows the specific responses the evaluation team received to their interview questions from the site contacts at these sites.

Table G-1. VFD Replacing VFD Site Interview Question Responses Summary

Site ID	Age of old VFDs	Did the old VFDs fail?	Were you operating in Bypass?	How long were you operating the pumps/fans in bypass?	Could you have operated the fans/pumps in bypass indefinitely?	Simple payback before incentive (from XACT tool)	Simple payback after incentive (from XACT tool	B/C Ratio (from XACT tool)
L077	11-12 years	yes	No, would use redundant capacity of backup pumps first.	If back-up pumps were not sufficient, would run in bypass in an emergency until drive could be repaired. (on the order of days to weeks)	No	1.28	0.58	No XACT file
L094	15-20 years	Yes	No, would use redundant capacity of backup fans/pumps first.	If back-up pumps were not sufficient, would run in bypass in an emergency until drive could be repaired. (on the order of days to weeks).	No	1.02	0.7	3.51
L080	> 10 years	No	No	Never operated in bypass.	No	0.56	0.3	4
LR24	Approx. 15 yrs	Yes	No, would use redundant capacity of other pumps first.	Several weeks while waiting for repairs if backup pumps were not sufficient.	No	Approx. 0.3 per site contact	Approx. 0.18 per contact	No XACT
LR22	11-12 years	yes	No, would use redundant capacity of backup pumps first.	If back-up pumps were not sufficient, would run in bypass in an emergency until drive could be repaired. (on the order of days to weeks).	No	0.4 / 1.28 / 1.6 / 1.47	0.2 / 0.58 / 0.8 / 0.7	Locked XACT files
L053	>28 years	1 of 12 unit failed.	1 of 12 operated in bypass for a few weeks. The other 11 operate in "Auto" with the speed modulated based on static pressure.	A few weeks.	No	0.58	0.38	3
LR17	15-20 years	Yes	Yes	2 months.	No	1.5	0.4	3.83
L093	10 years	Yes	Yes	Periodically while getting replacement parts.	No	2.7	1.6	1.86

The most important question posed was whether or the applicant could have regressed and operated the fans/pumps in bypass mode indefinitely. As Table G-1 shows, none of the VFD replacement project applicants answered affirmatively.

The issue of baseline in this instance could easily be confused with free ridership in a standard attribution survey. If this happened the program would be unfairly double-penalized with a low realization rate and a high free ridership for the same phenomenon. For this reason the evaluation team excluded all and any VFD-VFD replacement measures from the free ridership analysis.

It should be noted that if the engineers had not executed the special baseline protocol established for this measure and allowed the full savings in the realization rates, the social scientists' free ridership research would have included tailored questions instead, and almost certainly would have captured this same effect of lost savings in the net impact. For this reason, the evaluation team generally recommends that VFDs replacing VFDs not be allowed funding as a matter of program policy.[1]

[1] The sole possible exception being if the applicant can document that the replaced VFD also had been program-funded.

APPENDIX H – GLOSSARY OF TERMS

CECONY or CE (in some tables) – Consolidated Edison Company of New York.

census stratum – In a stratified sample design, the stratum with those participants with the largest savings may have a calculated sample size that exceeds the population of the stratum. A stratum that meets this condition is referred to as a census stratum.

coefficient of variance (CV) – A normalized measure of dispersion of a probability distribution and defined as the ratio of the standard deviation, σ,to the mean, μ:

$$c_v = \frac{\sigma}{\mu}$$

DPS – New York Department of Public Service.

Energy Efficiency Portfolio Standard (EEPS) – The state-mandated utility-administered programs.

energy management system (EMS) – A system used by building operators to monitor, measure, control, and schedule their building loads.

error ratio – In energy efficiency evaluation, the error ratio is a measure of the degree of variance between the reported savings estimates and the evaluated estimates. For a sample, the error ratio is:

$$er = \frac{\sqrt{\sum_{i=1}^{n} w_i \frac{e_i^2}{x_i^\gamma} \sum_{i=1}^{n} w_i x_i^\gamma}}{\sum_{i=1}^{n} w_i x_i}$$

where n is the sample size, w_i is the population expansion weight associated with each sample point i, x_i is the program-reported savings for each sample point i, y_i is the evaluated gross savings for each sample point i, error for each sample point $e_i = y_i - bx_i$, and γ= 0.8.

ex ante savings estimate – Forecasted savings used for program and portfolio planning purposes.

ex post savings estimate – Savings estimate reported by an evaluator after the energy impact evaluation has been completed.

free rider, free ridership (FR) – A program participant who would have implemented the program measure or practice in the absence of the program.

HVAC – Heating, ventilation, and air conditioning.

interval meter – An electric utility meter that measures and stores energy use and demand in 15-minute intervals. Interval meters are required for New York customers to participate in Independent System Operator demand response programs.

measurement and verification (M&V) – A subset of program impact evaluation that is associated with the documentation of energy savings at individual sites or projects using one or more methods that can involve measurements, engineering calculations, statistical analyses, and/or computer simulation modeling.

National Action Plan for Energy Efficiency – Model energy efficiency program impact evaluation guide abbreviated as NAPEE. This is the DPS-recommended reference guide for impact evaluations.

New York Technical Manual (NYTM) – The DPS-mandated reference document for calculating EEPS program savings.

net to gross, net-to-gross ratio (NTG, NTGR) – The relationship between net energy or net demand savings, where net is measured as what would have occurred without the program, what would have occurred naturally, and gross savings (often evaluated savings). The NTGR is the ratio of net savings to gross savings.

O&R – Orange & Rockland Utilities.

relative precision – Measures the expected error bound of an estimate on a normalized basis. It must be expressed for a specified confidence level. The relative precision (*RP*) of an estimate at 90% confidence is:

$$RP = 1.645 \frac{CV}{\sqrt{n}} \sqrt{1 - \frac{n}{N}}$$

where n is the sample size, N is the population size, and the coefficient of variance is CV = standard deviation / estimate mean value. The square root expression at the end of the equation is the finite population correction factor, which becomes inconsequential and unnecessary for large populations.

realization rate – The term is used in several contexts in the development of reported program savings. The primary applications include the ratio of project tracking system savings data (e.g., initial estimates of project savings) to savings that (1) are adjusted for data errors and (2) incorporate evaluated or verified results of the tracked savings. In the Updated Guidelines, the realization rate does not include program attribution.

self-reported approach (SRO) – A method for gathering program attribution data through direct interviews with participants.

snapback – Snapback occurs when customers actually increase their energy consumption due to reductions in the cost of energy.

spillover (SO) – Includes **participant spillover (PSO)** and **nonparticipant spillover (NPSO)** – Reductions in energy consumption and/or demand caused by the presence of the energy efficiency program, beyond program-related gross savings of participants.

PSO occurs when additional actions are taken to reduce energy use at the same site, but these actions are not included as program savings.

NPSO is the reduction in energy consumption and/or demand from measures installed and actions taken or encouraged by nonparticipating vendors or trade allies because of the influence of the program.

stratified ratio estimator (SRE) – An efficient sampling design combining stratified sample design with a ratio estimator. It's most advantageous when the population has a large coefficient of variation (CV). (A large CV occurs, for example, when a substantial portion of the projects have small savings and a small number of projects have very large savings.) The ratio estimator uses supporting information for each unit of the population when this information is highly correlated with the desired estimate to be derived from the evaluation, such as the tracking savings and the evaluated savings.

The RR calculation for electric energy is shown below:

$$RR = \frac{kWh_{Evaluation}}{kWh_{Program}}$$

Where, RR is the savings realization rate, $kWh_{Evaluation}$ is the evaluation M&V kWh savings (by evaluation M&V contractor), and $kWh_{Program}$ is the kWh savings claimed by program.

TMY3 – Typical meteorological year weather data.