Small Business Direct Install Program Evaluation Review

Final Report

January 30, 2015

Prepared for: E^2 Working Group

Prepared by: Evaluation Studies Subcommittee

TABLE OF CONTENTS

Executive Summary

This report documents the work of the Evaluation Studies Subcommittee in reviewing independent evaluations of New York State's Small Business Direct Install programs. Results of the subcommittee's review of these evaluations, the conclusions drawn by the subcommittee and the recommendations that the subcommittee made are described here in detail. Brief summaries of the evaluations are provided in the Appendix.

This subcommittee was formed at the request of the New York State Energy Efficiency (E^2) Working Group and tasked with reviewing completed evaluations of New York State's Small Business Direct Install (SBDI) programs for cycle one (2009-2011) of the Energy Efficiency Portfolio Standard (EEPS 1) .

SBDI program administrator representatives who participated in this review were: Con Edison - Rosanna Jimenez and Steven Mysholowsky, Central Hudson - Amanda DiMaso and Thomas Rizzo, National Grid - Joseph Dolengo and Tamara Prodrick, New York State Electric & Gas (NYSEG) and Rochester Gas and Electric (RG&E) - John Zabliski and Debbie Pickett, Orange & Rockland (O&R) – Charmaine Cigliano and Sandra Eason-Perez. Staff from the New York State Energy Research and Development Authority (NYSERDA) – Tracey DeSimone and staff from the New York State Department of Public Service – Joe Hitt, William Saxonis, Kanchana Paulraj and Pete Sheehan also participated on the subcommittee.

The objectives of the review were to review the completed New York State SBDI impact evaluations, and specify any required changes and/or follow up based on the results of the evaluations, including:

- determine any required changes to the New York Standard Approach for Estimating Energy Savings from Energy Efficiency Programs – Residential, Multi-Family, and Commercial/Industrial Measures (TRM),including when and how such changes should be made,
- determine the need, timeline and process for any additional data collection for SBDI, and
- make a recommendation on the next steps for the currently tabled Statewide SBDI Impact Evaluation RFP as necessary.

Conclusions

The subcommittee found that (refer to section 5.0 below for more detail):

- Differences in annual measure operating hours, and facility type definitions produced the greatest variations in the evaluation results.

- Additional time is required to standardize facility type definitions and deemed operating hour values in the TRM,

- This review process worked well and should be employed for future program evaluation reviews, and that considering the status of the currently tabled statewide SBDI impact evaluation RFP, the subcommittee should also review the SBDI evaluation templates recently submitted by the PAs to determine if any future SBDI program evaluations might be conducted jointly.

1.0 Objectives

The objectives of this subcommittee were:

1. To review the completed New York State program administrator EEPS Cycle 1 (2009-2011) SBDI impact evaluations to compare evaluation results (e.g., compare the site-verified lighting measure savings values to the deemed lighting measure savings values in the TRM). Consider both annual energy savings (kWh) and peak demand reduction (kW).

2. If changes were required in the TRM, to make recommendations for those changes and to specify when and how those changes should be implemented.

3. If additional data collection was necessary to fill gaps in either the evaluation results or the TRM, to recommend a data collection process and a timeline for data collection.

4. To determine whether or not to proceed with the currently tabled Statewide SBDI Impact Evaluation, either in its current form or modified, to address specific aspects of the subcommittee's recommendations, and each Program Administrator's own plans for needed additional SBDI program data collection / evaluation in the 2015 – 2016 program years.

2.0 Methods

2.1 Formation of the subcommittee

The E^2 Working Group, in preparation for future energy efficiency program cycles and in preliminary work addressing New York State's Reforming the Energy Vision proceeding (REV), determined that program administrators (PAs) should have available the best and the most current per-unit savings estimates for planning purposes. To that end, a subcommittee was formed with representatives from each program administrator's evaluation team and with participation by DPS staff. The task given the subcommittee was to review completed New York State SBDI program evaluations for program years 2009 – 2011 (EEPS 1) and the current TRM.

The subcommittee was kept to 15 individuals involved in energy efficiency program evaluation in New York State. This size and familiarity with the material allowed the group to act quickly - scheduling meetings, reviewing data, discussing findings and forming recommendations for the this report.

The E^2 working group gave the subcommittee a three month deadline for completing the review. Six meetings were held over the three month period.

2.2 The Subcommittee Process

The process began with a reading and review of completed New York State SBDI evaluation reports by all committee members. Subcommittee objectives were determined. (See Section 1 above.) Discussions ensued on the differences in reported per-unit savings and per-unit demand reductions between the reports and the savings values in the TRM.

As the work progressed, additional detailed data was provided by program administrators to support the analysis. This allowed the subcommittee to determine the reasons for differences in per-unit savings and to begin to discuss means for resolving those differences.

As the subcommittee reached consensus on issues, decisions were recorded, and the work progressed to the next outstanding issue or the next level of analysis. Near the end of the three month timeframe the subcommittee began to form recommendations for next steps that would be the basis this report.

2.3 Final Report

This preliminary final report was drafted at the close of the third month based on consensus items. It includes the group's recommendations for next steps and further work. Where conclusions were reached, explanations and supporting data are provided. Where there is insufficient data to reach a conclusion, recommendations for additional data collection are made. If changes to the TRM were contemplated, they are outlined and discussed. (See Section 4.0, Conclusions and Recommendations.)

3.0 Review of SBDI Evaluation Reports

3.1 Comments on Reports

The subcommittee began by reviewing and discussing each of the EEPS 1 (2009-2011) SBDI program year evaluation reports. Four evaluation reports, prepared by independent evaluation consultants, were available for review:

1. An evaluation of Consolidated Edison's program prepared by Energy Resource Solutions,
2. An evaluation of National Grid's program prepared by DNV GL (previously known as KEMA),
3. An evaluation of Central Hudson's program prepared by Applied Energy Group, and
4. An evaluation of NYSEG and RG&E's programs prepared by Itron.

During the subcommittee's first meeting an overview of each program administrator's SBDI evaluation report was given by one of the subcommittee members, followed by a discussion of the report. The subcommittee members discussed general conclusions presented by the various evaluations. Following are brief summaries of those discussions for each evaluation report:

1. Consolidated Edison – The executive summary provides the reviewer with a good framework of the evaluation report content and the evaluation results. SBDI is defined as customers with facilities that have an average monthly peak demand of less than or equal to 100 kilowatts (kW).

The majority of the program savings came from lighting measures. The evaluation's on-site light logging operating hours differ from the TRM deemed values, somewhat significantly for some facility types. The logger study period was for 12 months, with an interim report prepared after 3 months of data collection. On-site sampling was stratified by facility type, with 133 total sites selected for metering.

TRM facility type categories with similar operating hours and functions were collapsed for evaluation (For example, 9 TRM "office" related facility types were collapsed into a single evaluation "Office" facility type category.)

The subcommittee discussed categorization of facility types in further detail, and areas for further exploration surrounding: 1) building characteristics and 2) service territory locational effects (cityscape vs. rural, etc.

The subcommittee agreed on the need to dive into the specifics of "customer" at more detailed levels; and to develop changes to the current TRM thinking and/or expand the definitions of buildings and facility types in the C&I section.

2. National Grid – The executive summary provides good detail of the evaluation report content. Net-to-Gross (which includes Free Ridership and Spillover) was quite a bit higher than other PA's evaluation results.

Long term (12 months) on-site light logging was done for a number of sites, with additional site visit data used to inform the evaluation results. As with the other evaluation reports, lighting operating hours from the light loggers were found to be different than the TRM deemed values. On-site light logger sampling was stratified by site savings, with break-outs by facility type (and calculation of resulting confidence intervals) once the logger data had been collected. 70 sites were light-loggered.

The subcommittee further discussed mixed use customers (strip malls as an example) and the difficulty in determining the overall lighting operating hours evaluation results in these cases. The subcommittee determined that it was important to consider this customer "type" when reviewing the results from each evaluation.

3. Central Hudson – The executive summary did not discuss some of the key findings and details important to the overall framework of the evaluation. As examples: Definitions of "small" and "mid-sized" customer types are not addressed in the executive summary [Central Hudson noted that small = 0-100kW accounts, mid-size = 100kW and above], and the allocation of account types (91% small, 9% mid-sized) does not appear to be representative of program participation.

Minimal on-site light logging was done, and the results are not as statistically significant as generally preferred for New York State evaluations. An explanation of that issue would be helpful in the report (reasons for the smaller sample size and/or the shorter length of time the loggers were in place). [CH noted that they had two implementation contractors working during the evaluation timeframe, and that data from the first implementation contractor was sometimes incomplete. The second implementation contractor, however, provided better data.] On-site light logging was done for 20 participants recruited from the Net-to-Gross telephone surveys. While the Net-to-Gross telephone survey sample was stratified by facility type, the number of light loggers installed (20) and the time period installed (1 week) did not make it possible to draw statistically valid conclusions from the results. However, the light-logging results align well with light logging results of the other SBDI evaluations, and as with the other evaluation results, differ significantly from the TRM deemed lighting operating hours.

While the free ridership and spillover reported in the evaluation are essentially the same as those found in the Con Edison evaluation, the various studies we reviewed did not use identical measurement approaches and survey

questions. This could account for some variation in the results. While there are pros and cons of mandating a standardized measurement approach, standardization should be considered.

Additional points: Gaining a sense of specific customer project sizes and locations is important for further data review between studies; also the need to take into consideration the effect that time lag (from a project installation to evaluation survey/on-site work) may have affected the evaluation results obtained.

4. NYSEG and RG&E – As with the Consolidated Edison study, the executive summary provides the reviewer with a good framework of the study content and results. Program savings are essentially all lighting. The Net-to-Gross ratio found in this study is significantly lower than either Central Hudson or Consolidated Edison. This study did assess both kWh and kW savings.

A number of sites received site visits and light loggers to record lighting operating hours. Similar to the other studies, the evaluation identified lighting operating hours differ significantly from the TRM deemed values. The study was done at the measure level (CFL and linear fluorescent) which does provide additional information for each facility type. Similar to the other evaluation reports, the effect of time lag on the evaluation results should be considered. Sampling was stratified by facility type and 50 sites were light-loggered.

5. Orange & Rockland – Orange & Rockland as not yet released its final SBDI evaluation report, however, the Orange & Rockland subcommittee member indicated that much of the discussion of other program administrator evaluation results also applies to their evaluation results.

3.3 SBDI Program Realization Rates

The savings for each program, adjusted by the program realization rate, are shown in Table 1. (Note: Please see the Glossary at the end of the report for the definitions of savings terms.)

TABLE 1: PROGRAM EVALUATION RESULTS SHOWING REALIZATION RATES

Program Administrator	**Gross Savings (kWh)**	**Realization Rate**	**Adjusted Gross Savings (kWh)**
Central Hudson	41,841,892	Not Available[1]	Not Available[1]
Consolidated Edison	107,839,484	76%	81,958,008
National Grid	51,594,070	80%	41,275,256
NYSEG and RG&E	66,174,000	58%	38,381,000

It can be seen from the table that the Realization Rates varied between program administrators, in one case significantly. The realization rate is the ratio of project tracking system savings data (i.e., initial estimates of project savings) to savings adjusted for data errors and incorporating the evaluated or verified results of the tracked savings.

[1] Sample size not statistically valid.

The Adjusted Gross Savings are the product of the reported Gross Savings and the Realization Rate determined by the program evaluator. The Realization Rate includes adjustments for any misreporting of gross savings that may have occurred, adjustments for calculation errors in reporting, verification that program measures were installed by program participants and, in the case of lighting measures, field verification of annual lighting operating hours.

4.0 Results

After review, the subcommittee turned its attention to *annual lighting operating hours* as the factor showing the greatest variability between the evaluation data and the TRM. To begin the comparison of results it was necessary to understand how lighting measure savings are calculated. Annual lighting savings are defined in the TRM as:

Annual Electric Energy Savings

$$\Delta kWh = units \times \left[\frac{(W \times units)_{baseline} - (W \times units)_{ee}}{1{,}000} \right] \times FLH \times (1 + HVAC_c)$$

where FLH are Full Load Hours, W is watts and Units is the number of lamps or fixtures.

$HVAC_c$ is a heating and cooling interaction factor that accounts for the effect of reduced heat from the lighting system on the building heating and cooling system.

Full Load Hours are the *annual operating hours* of the lights. Annual operating hours used for reporting savings may come from the initial SBDI sales call, or from the table of deemed annual operating hours, listed by facility type, in the TRM.

As the subcommittee's review of the SBDI evaluation data continued, it became clear that there were significant variations between evaluation reported annual operating hours and the deemed annual operating hour values in the TRM. A review of TRMs from other states indicated variation in annual operating hours between TRMs for the same or similar facility types. Table 2 provides a sample of some of the data reviewed by the subcommittee, comparing annual operating hours from the evaluations with deemed operating hour values from several current TRMs.

TABLE 2: ANNUAL OPERATING HOURS FROM VARIOUS TRMs

Commercial Lighting Annual Operating Hours Research

	Evaluation Results					North East				
Building Type	Central Hudson	Consolidated Edison	National Grid	NYSEG	RG&E	New York	Connecticut	Mass	Rhode Island	Vermont
Automobile Retail	2,706		2,469	2,810	2,237	4,056	4,056	4,056	4,056	
Non-Enclosed Mall										
Nursing Home/Residential care						5,840	5,840			
Office	2,520	3,013	N.A.	1,476	1,903	3,748	3,748			3,642
Other								3,951	3,951	
Parking Garage		7,717				4,368	4,368			
Restaurant						4,182	4,182	5,110	5,110	4,089
Retail	2,602	3,458	3,463	2,728	3,022	4,057	4,057			4,103
Small Office								3,610	3,610	
Small Retail/Business								4,089	4,089	
Sports Arena						1,954	1,954			
Transportation						6,456	6,456			
University/College						2,586	2,586	3,255	3,255	3,416
Warehouse						2,602	2,602	3,759	3,759	4,009
Waste Water treatment plant						6,631	6,631			
Workshop						3,750	3,750			

All four program evaluations used on-site light loggers to determine actual operating hours. A statistically valid sample (combining facility types, if necessary) of participants was chosen in each program evaluation and multiple loggers were installed at each site. Since light logger deployment can be one of the most expensive and time consuming aspects of an evaluation, light loggers were only installed where the greatest savings were achieved. All four evaluations, however, did place loggers at Office, Retail or Automotive Related facilities and these three facility types became the basis for comparison between the evaluation results and the TRM.

To determine the reasons for the variability of operating hours data between evaluations for the same facility type, as shown in Table 2, individual site visit data was pulled and reviewed for each evaluation. It became apparent that even for a specific evaluation there could be significant variations in operating hours for the same facility type. Table 3 presents a sample of this data, where each row in the table is a unique site; even for rows of the same facility type at the same program administrator, differences in operating hours are noted. (For example, NYSEG shows 3946 hours at one Retail site and 1183 hours at another Retail site.)

TABLE 3: INDIVIDUAL SITE DETAIL FROM EVALUATION REPORTS

Company Name	Facility Type (as defined in TRM)	Sample Group Type	TRM Hours	Study Measured Operating Hours	Delta (TRM - Study)	Hrs used for Savings Calc (i.e. TRM/ actual)
CentHud	Office (General Office Types)	Office/Retail	3,748	2,983	765	actual
CentHud	Retail	Office/Retail	3,748	3,708	40	actual
CentHud	Office/Retail	Office/Retail	3,748	3,517	231	actual
CentHud	Office (General Office Types)	Office/Retail	3748	2,577	1,171	actual
CentHud	Auto Related	Automotive Facility	4056	2,706	1,350	actual
CentHud	Office/Retail	Office/Retail	3748	2,119	1,629	actual
CentHud	Auto Related	Automotive	4056	2,604	1,452	actual
CentHud	Retail	Office/Retail	3748	1,440	2,308	actual
ConEd	Office (General Office Types)	Offices - Downstate	3748	3,013	735	TRM
ConEd	Auto Related	Parking Garages- Downstate	4368	7,717	(3,349)	TRM
ConEd	Retail	Retail - Downstate	4057	3,458	599	TRM
NYSEG	Retail	Retail - Small	4057	111	3,946	TRM
NYSEG	Retail	Retail - Small	4057	2,874	1,183	TRM
NYSEG	Office (General Office Types)	Office - Small	3748	284	3,464	TRM
NYSEG	Office (General Office Types)	Office - Small	3748	56	3,692	TRM
RG&E	Retail	Retail - Small	4057	2,279	1,778	TRM
Ngrid	Other	Retail	4057	2,706	1,351	actual
Ngrid	Automobile	Automotive Facility	4056	2,768	1,288	actual
Ngrid	Other	Office	3748	2,039	1,709	actual
Ngrid	Automobile	Automotive Facility	4056	2,486	1,570	actual
Ngrid	Automobile	Retail	4057	8,760	(4,703)	actual

Note: Program Administrators may use TRM deemed operating hours or site hours, depending on the availability and accuracy of site hours.

It is clear from review of the data in Table 3 that some of the difficulty in comparing operating hours between sources is in making a clear determination of facility type (it is very subjective). Customers with the same facility type may have significantly different operating hours In the New York State TRM (as well as in TRMs from other states) facility types often lack detailed definitions. There is clearly a need to better define facility types in the New York State TRM.

One alternative for eliminating some of the confusion surrounding facility types would be for the TRM to adopt facility type descriptions from the North American Industry Classification System (NAICS). This system is currently used by the United States Census Bureau, and categorizes businesses by a two to six digit numeric code. The greater the number of digits included in the code, the more specific and detailed the classification is. NAICS business types are comparable to TRM facility types. Further, collection of NAICS codes is currently required by the New York State Evaluation Guidelines. Table 4 provides an excerpt from the NAICS classification system showing some of the three digit NAICS classifications from Category 44 Retail Trade that might be found among participants in an SBDI program.

TABLE 4: EXCERPT FROM NAICS CODES, CATEGORY 44 - RETAIL

2012 NAICS US Code	2012 NAICS US Title
441	Motor Vehicle and Parts Dealers
442	Furniture and Home Furnishings Stores
443	Electronics and Appliance Stores
444	Building Material and Garden Equipment and Supplies Dealers
445	Food and Beverage Stores
446	Health and Personal Care Stores
447	Gasoline Stations
448	Clothing and Clothing Accessories Stores
451	Sporting Goods, Hobby, Musical Instrument, and Book Stores
452	General Merchandise Stores
453	Miscellaneous Store Retailers
454	Nonstore Retailers

Looking at the three digit codes in Table 4, we can see a much more detailed breakdown of business type than the list of facility types included in the New York State TRM. More importantly, the NAICS classification system also provides a detailed written description for each code. So, for example, NAICS code 445 in Table 4 is defined as:

> "Industries in the Food and Beverage Stores subsector usually retail food and beverage merchandise from fixed point-of-sale locations. Establishments in this subsector have special equipment (e.g., freezers, refrigerated display cases, refrigerators) for displaying food and beverage goods. They have staff trained in the processing of food products to guarantee the proper storage and sanitary conditions required by regulatory authority."

Use of NAICS codes may be one alternative that the subcommittee reviews when considering changes to the TRM Facility Type definitions.

At this point, further detailed work also needs to be completed to fully resolve the differences in lighting operating hours:

1. Individual site visit sheets must be pulled and reviewed for each light logger site in the evaluations. From these, a specific and detailed description of the facility type should be documented.

2. Individual site descriptions must then be grouped under corresponding TRM facility types.

3. These groupings of field data by TRM facility type must be reviewed for consistency among individual sites and their operating hours within a group.

4. Groupings of field data by TRM description should be used to answer the following questions:

a) Are the TRM facility type groupings specific enough to reliably characterize each individual site within that grouping? (for example, "Office – General") Are the TRM groupings broad enough that it keeps to the minimum necessary the total number of facility types required?

b) Are the specific sites within a TRM grouping homogeneous enough that a single value for operating hours can be used to reliably represent that entire facility type group? (for example, "Retail")

c) Is there a different way of defining the facility types in the TRM that would be more compatible with the field data?

d) How should the value of operating hours be determined for each facility type grouping? A strict numeric average of the field data within a grouping, a weighted average weighted by the frequency of facility sub-types within a grouping, or by other variables such as geographic location within the state or building square footage?

5.0 Conclusions and Recommendations

The subcommittee has made excellent progress in completing the objectives of this project. However, additional detailed work remains to be done and the subcommittee respectfully requests an extension of the project deadline to April 1, 2015 to further investigate Objective 3 (additional data collection) and Objective 4 (Statewide SBDI Impact Evaluation) . The subcommittee provides the following conclusions and recommendations at this time:

Conclusions:

1. From the equation in the TRM, annual operating hours are the variable with the greatest effect on the savings of lighting measures. The subcommittee focused its review on site-verified annual operating hours that were documented in the evaluation reports, and compared those operating hour values with the deemed operating hour values in the TRM. Lack of clear and specific definitions for facility types in the TRM made this process more difficult.

   Note: Free-Ridership and Spillover may also significantly affect program savings. However, Free Ridership and Spillover are not components in the TRM deemed savings equations and so were not considered in this review.

2. Additional time is required to standardize facility type definitions and deemed operating hour values in the TRM.

3. This review process worked well and should be employed for future program evaluation reviews, and that considering the status of the currently tabled statewide SBDI impact evaluation RFP, the subcommittee should also review the SBDI evaluation templates recently submitted by the PAs to determine if any future SBDI program evaluations might be conducted jointly.

Recommendations:

1. Further analysis must be completed to update Facility Type definitions and to determine annual operating hour values for Facility Type in the New York State TRM highlighted on the evaluation reports. More clarity is needed in the TRM facility type definitions to reflect the specific business types in the evaluation data. (Use of NAICS codes may be useful in adjusting TRM definitions.) Changes to TRM operating hour values must be made to account for the differences between territories, for example between New York City (Downstate) and the rest of New York State (Upstate).

2. Changes to the New York State TRM facility type definitions and operating hour values resulting from the work of this subcommittee should not become effective until January 1, 2016 (end of the current EEPS 2 period), and should only be applied to program planning / program results going forward from that date. Changes should be published in the same manner as current changes to the TRM are published by the E^2 Technical Manual subcommittee. The results of this SBDI subcommittee's work, however, may be shared prior to the date of publication and used by program administrators in their Energy Efficiency Transition Implementation Plans (ETIPs).

3. This subcommittee's process of reviewing the results of the completed New York SBDI program evaluations should be used for review of the results of other New York State program evaluations, and for making recommendations for applying or using those results on a statewide basis.

APPENDICES

APPENDIX A: GLOSSARY OF TERMS

Annual Energy Savings	The reduction in electricity consumption (kWh) or in fossil fuel use in thermal unit(s) from the savings associated with an energy saving measure, project, or program in a given year. (Annualized Energy Savings are calculated based on a full year's installation and operation)
Gross kW	Expected demand reduction (kW) based on a comparison of standard or replaced equipment, and equipment installed through an energy efficiency program.
Gross kWh	Expected reduction in energy consumption (kWh) based on a comparison of standard or replaced equipment, and equipment installed through an energy efficiency program.
Gross Savings	The change in energy consumption and/or demand that results directly from program-related actions taken by participants in an efficiency program, regardless of why they participated and unadjusted by any factors.
Adjusted Gross Savings	The change in energy consumption and/or demand that results directly from program-related actions taken by participants in an efficiency program, regardless of why they participated. It adjusts for such factors as data errors, installation and persistence rates, and hours of use, **but does not adjust for free ridership or spillover.** Can be calculated as an annual or lifetime value.
Evaluated Gross Savings	The change in energy consumption and/or demand that results directly from program-related actions taken by participants in an efficiency program, regardless of why they participated, as calculated by program evaluators.
Net Savings	The total change in load that is attributable to an energy efficiency program. This change in load may include, implicitly or explicitly, the effects of free drivers, free riders, energy efficiency standards, changes in the level of energy service, and other causes of changes in energy consumption or demand.
Evaluated Net Savings	The total change in load that is attributable to an energy efficiency program, as calculated by program evaluators. This change in load may include, implicitly or explicitly, the effects of free drivers, free riders, energy efficiency standards, changes in the level of energy service, and other causes of changes in energy consumption or demand.
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.

Participant	A consumer that received a service offered through the subject efficiency program, in a given program year. The term "service" is used in this definition to suggest that the service can be a wide variety of services, including financial rebates, technical assistance, product installations, training, energy-efficiency information or other services, items, or conditions. Each evaluation plan should define "participant" as it applies to the specific evaluation.
Non-Participant	Any consumer who was eligible but did not participate in the subject efficiency program, in a given program year. Each evaluation plan should provide a definition of a non-participant as it applies to a specific evaluation.
Free Rider	A program participant who would have implemented the program measure or practice in the absence of the program. Free riders can be: a total Free Rider, in which the participant's activity would have completely replicated the program measure; a partial Free Rider, in which the participant's activity would have partially replicated the program measure; or a deferred Free Rider, in which the participant's activity would have completely replicated the program measure, but at a future time.
Free Ridership	Refers to the percentage of savings attributed to customers who participate in an energy efficiency program but would have, at least to some degree, installed the same measure(s) on their own if the program had not been available.
Net-to-Gross Ratio (NTGR)	Is represented as a ratio that compares the gross savings of a program to the energy savings actually attributable to the program. Energy savings are estimated after adjusting for factors such as measurement error, measure installation quality, user behavior, and the actions program participants and non-participants would have taken absent the program (e.g., free ridership and spillover). The decision path proposed to arrive at net savings should be discussed. NTG ratio = (1 - Free ridership) + Spillover
Spillover	Refers to the energy savings resulting from action by consumers influenced by an energy efficiency program, but where the consumers do not receive direct financial or technical assistance from the energy program for their spillover savings. Spillover is one of the more difficult-to-measure components of an impact evaluation but, for some programs, it can represent a significant percentage of the energy savings. To help guide evaluators in the measurement of spillover, Appendix F (of the DPS Evaluation Guidance) offers methodological guidance including a discussion of the required levels of rigor.

	Participant spillover is the additional energy savings that occur when a program participant independently installs energy efficiency measures or applies energy saving practices after having participated in the efficiency program as a result of the program's influence. **Non-participant spillover** refers to energy savings that occur when a program non-participant installs energy efficiency measures or applies energy savings practices as a result as a result of a program's influence.
	Inside Spillover occurs when, due to the project, additional actions are taken to reduce energy use at the same home, but these actions are not included as program savings. **Outside Spillover** occurs when an actor participating in the program initiates additional actions that reduce energy use at other sites that are not participating in the program.
	Spillover Rate is the estimate of energy savings attributable to spillover effects expressed as a percent of savings installed by participants through an energy efficiency program.
Realization Rate	It is important to clarify that our definition of "realization rate" reflects adjustments to a program's gross energy savings estimate, but does not reflect the impacts of free ridership and spillover. The realization rate is the ratio of project tracking system savings data (i.e., initial estimates of project savings) to savings adjusted for data errors and incorporating the evaluated or verified results of the tracked savings. Free ridership and spillover are captured in the net to gross ratio to reflect the degree of program induced actions. Specifically, the gross energy savings estimate, refined by the realization rate, is adjusted to reflect the negative impacts of free ridership and the positive impacts of spillover. NTG ratio = (1 - Free ridership) + Spillover

Sources for definitions:

Definitions for the EM&V Terms listed above have been approved by the Evaluation Advisory Group and provided by DPS staff for this report.

Central Hudson Gas & Electric Corp Commercial Lighting: Impact Evaluation Summary

Evaluation Conducted by: Applied Energy Group (AEG), DATE: April 2014

PROGRAM SUMMARY

The objective of the Central Hudson Small & Mid-Size Business Lighting Program is to replace existing inefficient lighting with energy efficient lighting and lighting controls in Central Hudson business customer facilities in order to capture significant energy savings and environmental benefits. Participants receive a free energy audit conducted by Alliance Energy Solutions ("Alliance"), the third-party program implementer. Incentives cover up to 70 percent of the equipment and installation costs. A financing program initiated by Central Hudson and approved by the New York Department of Public Service ("DPS") provides zero percent financing to cover the remaining equipment and installation cost.

EVALUATION OBJECTIVE AND HIGH LEVEL FINDINGS

An independent evaluation contractor Applied Energy Group (AEG) designed the impact evaluation to estimate energy and demand savings impacts of the 2010-2011 program. The evaluation utilizes various methods to calculate savings and other program impacts, including engineering and metering analyses. AEG performed the following tasks to determine the impacts of the program:

- Reviewed program tracking database to verify savings and develop participant samples for the billing analysis.
- Conducted a participant survey to verify program participation and determine the influences of free ridership and spillover.
- Conducted an engineering analysis to identify changes in energy usage as a result of program participation.
- Performed a cost-effectiveness analysis of the program.

In 2010-2011, a total of 2,062 participants accounted for 7,975 projects and 153,318 measures. The majority of participants were Small-Size (91%) with the remaining Mid-Size (9%). All business types were represented with Office/Retail (33%), Automotive (18%), Industrial (17%) and Assembly (11%) making up the majority of building types participating in the program. The overall net energy and demand savings determined by the engineering analysis were 41,841,182 kWh and 12,684 kW, respectively. AEG estimated a net-to-gross factor of approximately 91 percent using the results of a telephone survey of program participants designed to assess the effects of free ridership and spillover.

DETAILED IMPACT EVALUATION FINDINGS

AEG performed engineering analysis to determine program impacts. A metering study was also conducted to estimate lighting usage hours.

Savings were spread across all building types. Office/Retail buildings accounted for the greatest amount of energy savings (31%) followed by Industrial (20%), Assembly (15%), and Automotive (12.5%). The combination of Schools, Healthcare, Other, Food Service, and Hotel/Motel combined to account for significant program savings (20%).

The net-to-gross factor includes estimates of free ridership and spillover. Using the results of a participant telephone survey, AEG estimated that 17 percent of savings would have been achieved due to natural market activity without the influence of the program. However, the program motivated participants to engage in spillover energy saving actions that amounted to 8

percent of additional savings. These included actions such as installing energy appliances, upgrading their HVAC system, and installing a programmable thermostat.

EVALUATION METHODS AND SAMPLING

The impact evaluation includes four major components: an engineering analysis to determine the amount of expected savings, a metering study to determine observed lighting usage by building type, and a participant survey to assess free ridership and spillover..

AEG performed the engineering analysis consistent with the *2010 New York Standard Approach Manual for Estimating Energy Savings from Energy Efficiency Programs* ("Tech Manual").[2] AEG utilized the program tracking data with savings algorithms provided by the Tech Manual for each lighting measure installed through the program. The savings algorithms yielded savings per measure for each type of lighting, which were multiplied by the number of participants to determine total gross savings.

The metering study used light loggers to determine the lighting usage patterns among commercial participants by building type. Only twenty participants, whose representativeness was not able to be determined, volunteered for the study, so the results and the information gained from the study are for information purposes only. Each participant had up to ten light loggers installed in their facility for a period of 60 days. The results of the study provided lighting full load hours by facility type.

In October 2012, AEG conducted a telephone survey of 72 randomly selected program participants to evaluate the impacts of free ridership and spillover. Free ridership was determined through a series of questions designed to predict whether the participant would have installed the lighting without receiving an incentive through the program. Responses to the free ridership questions were weighted based on the probability that the participant was a free rider and to account for potential bias. Similarly, participants were asked if the program motivated them to engage in spillover energy savings actions beyond the program, such as upgrading HVAC, installing a programmable thermostat, or energy efficient appliances. The spillover factor was calculated based on the ratio of spillover savings to total gross savings for each participant.

[2] *New York Standard Approach for Estimating Energy Savings from Energy Efficiency Programs*, Prepared for New York Department of Public Service by TecMarket Works, October 15, 2010.

Consolidated Edison Company of New York: Small Business Direct Install Program Impact Evaluation Summary

Evaluation Conducted by: ERS (Energy & Resource Solutions, Inc.), DNV-GL, Opinion Dynamics Corporation, and Apprise, May 2014

PROGRAM SUMMARY

CECONY designed the SBDI programs for rapid deployment of energy efficiency measures to existing small commercial and industrial (C&I) customers. CECONY defined small C&I customers as those with facilities that have an average monthly peak demand of less than or equal to 100 kilowatts (kW).

SBDI participants are recruited primarily through door-to-door solicitation. The prospective participant is audited, free CFLs are sometimes installed, and additional measures are identified for installation. If the customer agrees to the installation of additional measures, the implementation contractor arranges for an installation contractor to visit the customer site and provide complete pricing and installation services. The customer is required to pay 30% of the installed cost; the program pays the balance. For the 2009 – 2011 period, two implementation contractors served the CECONY SBDI program – Willdan Energy Services (WES) and Free Lighting Corporation (FLC). FLC was responsible for Staten Island while WES covered all of the other CECONY territories, including Westchester County. [1]

Approximately 97% of the savings to date are attributable to the installation of high efficiency lighting. A mix of refrigeration and HVAC control measures constitutes the balance.[2]

EVALUATION OBJECTIVE AND HIGH LEVEL FINDINGS

The intent of this impact evaluation of the SBDI program is threefold:

1. To evaluate the program's recent performance by developing gross savings realization rates (RRs) and a net-to-gross ratio (NTGR) that measures the attribution of SBDI program savings.
2. To provide information to the New York Standard Approach for Estimating Energy Savings from Energy Efficiency Programs (from here on referenced as New York Technical Manual [NYTM]) to help update key deemed savings input parameters affecting the SBDI program based on New York-specific performance data.
3. To provide actionable recommendations for improving the program's implementation including future evaluation activities.

The results of this study show three main conclusions:

1. **The evaluated parking garages operating hours were found to be significantly greater than the NYTM deemed value and the evaluated operating hours for the offices and retail sectors were found to be lower than the NYTM-deemed value.**
2. **The contractors are not consistently using the NYTM hours for all the lighting projects.**

[1] As of 12/31/11, FLC's involvement was curtailed. Subsequently Lockheed Martin was added as an implementation contractor

[2] Data from the CECONY Program Tracking Database.

3. **Parking garages were found to have a high free-ridership (FR) and chains were found to have higher FR than non-chains.**

Savings were calculated for each of four strata – by building type (retail, offices, parking garages, and other). The total verified gross energy savings are shown by stratum in Table 1.

Table 1. Program Gross Energy Impacts

Building Type	Percent Program Savings (Tracking)	Min n*	Gross Savings Realization Rate	Relative Precision at 90% Confidence	Measured Gross Program kWh Savings
Retail	30%	63	80%	14%	25,810,748
Offices	8%	73	75%	12%	6,634,822
Parking Garages	9%	57	83%	8%	8,251,934
Other	53%	38	72%	16%	41,260,504
Entire program**	**100%**	**231**	**76%**	**9%**	**81,958,008**

* The minimum number of sampled measures from w hich the results are based

** Overall program level numbers are savings w eighted totals/averages

The three most significant reasons that engineers identified for the evaluated savings deviating from – and in particular for being less than – the tracking savings were operating hours (17% lower than tracking estimates), quantities (6% lower), and technology (2% lower).

The net-to-gross ratios (NTGR) are shown by stratum in Table 2.

Table 2. Program Net-To-Gross Ratios

Building Type	Min n*	Combined Net-to-Gross Ratio	Relative Precision at 90% Confidence
Retail	136	86%	4%
Offices	75	85%	8%
Parking Garages	21	69%	15%
Other	311	83%	5%
Overall NTGR	**543**	**83%**	**3%**
Overall FR	**543**	**17%**	**3%**

* The minimum number of sampled measures from w hich the results are based

Attribution research found that 17% of the program's savings would have occurred in the absence of the program due to free ridership (FR). The measured participant spillover (SO) was negligible. Initially FR of 17% seemed high for SBDI programs but in a NTGR study of similar SBDI programs in the northeast conducted in 2011, FR was 13% and SO < 1%. So the 17% free ridership level is within the range of results for these other programs.

The net program results are calculated by multiplying the gross program results by the net-to-gross ratio. The total program net energy impacts are shown in Table 3.

Table 3. Program Net Energy Impacts

Building Type	Min n*	Net Realization Rate	Relative Precision at 90% Confidence	Net Program kWh Savings
Retail	63	69%	15%	22,261,770
Offices	73	64%	14%	5,661,715
Parking Garages	21	57%	17%	5,666,991
Other	38	60%	17%	34,383,753
Entire program**	**231**	**63%**	**9%**	**67,974,229**

* The minimum number of sampled measures from w hich the results are based

** Overall program level numbers are savings w eighted totals/averages

EVALUATION RECOMMENDATIONS AND PROGRAM ADMINISTRATOR RESPONSE

The recommendations for the program and the NYTM resulting from the results and conclusions of this evaluation are shown in the following sections. Program administrator responses are shown after each program recommendation.

Program Recommendations

The evaluation team offers seven recommendations that may help to increase program cost-effectiveness. These are listed below and then discussed in detail.

1. **Revise operating hours –** The evaluation team found that parking garages operate more hours than that suggested by NYTM, whereas the offices and retail facilities operate fewer hours than the NYTM-suggested values. If the New York State Public Service (DPS) chooses not to update the NYTM-deemed hour estimates in response to this research, the evaluators recommend that the SBDI program itself apply the operating hours for the various building types identified in the study. This will improve the savings estimation accuracy using data most applicable to the SBDI program population and increase the RR in the next evaluation.

   Response to Program Recommendation: 1 The SBDI team agrees with this recommendation. The change will impact the SBDI program prospectively. Discussions will be initiated with the Tech Manual review committee to include the revised downstate hours of operation for metered building/business types. Once the operating hours data is approved, we propose to use the ex-post verified (revised deemed) hours of operation to calculate energy savings.

2. **Account tracking –** Identify all Con Edison primary usage accounts associated with each facility or part of a facility that participates in the program for customers that possess multiple accounts.

   Response to Program Recommendation: 2 Con Edison is currently reviewing programmatic proceedures in conjunction with customer data extracts provided to its' implementation contractor, with respect to program eligible accounts. Better identification of these program eligible accounts will lead to a higher quality of data housed in the tracking database and mitigate some of the issues identified in the EEPS cycle 1 assessment with respect to participant sites that possess multiple customer accounts. Action will be taken as soon as proceedures are formulated.

3. **Use pre-retrofit billing data for QC –** During the first EEPS cycle, the SBDI installation contractor used pre-retrofit billing data to cross check the energy savings potential for each project. The EM&V group flags sites with a high savings to usage ratio and visits them to verify the measures. However, no action

was taken by the program contractors on such sites with high savings to usage ratios. We recommend that the program staff, or the implementation contractor, inspect all the job sites with savings to usage ratios greater than or equal to 80% and adjust the claimed savings accordingly.

Response to Program Recommendation: 3 The SBDI team will accept this recommendation. The program team feels this would be worthwhile and they will work with the implementation contractor to resolve the problem. The following three items should be noted and resolved prior to implementation of this recommendation:

- The statement "However, no action was taken by program contractors on such sites with high savings to usage ratios," is incorrect. WES, the SBDI implementation contractor, has built into their data management system (SMART) a tool that flags survey reports when the savings to usage ratio is greater than 50%.This recommendation did not impact FLC, as they were not retained after their EEPS Cycle 1 contract expired. Additionally Lockheed Martin's involvement in the SBDI program was curtailed in mid-2014.
- A much more concerted effort should be placed upon management and evaluation of data. Until recently customer data was updated once per quarter. There are many factors that are out of the program team's control, which have a significant impact on the billing analysis approach. For example, reducing the hours of operation will impact the energy savings and subsequently the billing analysis approach. Timeframe and data extract protocols also have an effect; for example annual kWh consumption is dependent upon the time of the year.
- There will be additional costs and time associated with this recommendation that will impact the program negatively.

  An inspection contractor currently does all of the SBDI "In Process" site inspections to meet regulatory criteria. The task of verifying savings associated with SBDI projects flagged by the implementation contractor for high savings to usage ratio can be completed by the inspection contractor as part of their "In Process" inspections. This process can begin as soon as contract changes are in effect.

4. **Standardize the use of NYTM hours –** Standardize the use of NYTM hours as a matter of general procedure instead of site-specific estimates. The program standardized the use of NYTM hours in October of 2010, but evaluators still found some sites where installer reported site specific hours were used for estimating savings.

   Response to Program Recommendation 4: The implementation contractors have been directed to use the NYTM operating hours or be censured. Once revised deemed hours of operation are approved by DPS, they will replace the original NYTM deemed hours where ever appropriate for savings calculations.

5. **Do not rely on savings from CFLs** – CFLs had lower installation rates and higher FR than the rest of the program measures. As CFLs become standard practice, net savings from CFLs will continue to decline. CFLs may still be useful to the program as a "loss leader" measure that can be offered to get auditors in the door, but most of the savings from CFLs will likely not be realized.

   Response to Program Recommendation 5: The SBDI team agrees with this recommendation, in fact, the program has been offering LEDs since 2011. In 2014, the program offers a greater variety of LEDs and has begun to phase CFLs out.

6. **Include install location information in the tracking data –** The tracking data included all the lighting inventory information except the installation locations. The evaluators extracted the installation location

information from the actual work orders, which was a time-consuming process. We recommend that CECONY look into including the installation location information in the tracking data for each measure or group of measures installed, which will then serve as a more complete source of data for impact evaluations and post installation inspections.

Response to Program Recommendation 6: The SBDI team agrees with this recommendation and will take steps to implement it.

7. **Track service addresses –** Ensure that the site address matches the location where the measures were installed and not the billing address for the participant.

Response to Program Recommendation 7: The SBDI team agrees with this recommendation. The program staff has been collecting service addresses as part of the customer extract protocol; implementation contractors provide both service address and mailing address. Additionally, implementation contractor surveyors enter the business address into the field survey tool if the addresses provided do not match. This is very significant, because the differences in addresses become a major complication when the program staff is validating eligibility as part of the invoice review. At the program staff's request, implementation contractors provide the service address in their reporting. This allows program staff to verify reported service addresses with the data Con Edison has on record.

Recommendations for New York Technical Manual

The evaluators recommend that the NYTM be modified to either include updated operating load hours for selected building types based on this study's research, as shown in Table 4, or create new "small business" or "downstate" retail, office, and parking garage categories specifically for SBDI programs that reflect this study's findings. The table presents both the current NYTM deemed hours and the recommended updated ex-post revised values based on this evaluation's research.

Table 4. Recommended Updated Hours

Description	NYTM Current Hours	Study-Measured Hours - CECONY
Banks	3,748	3,013
Court house	3,748	3,013
Library	3,748	3,013
Medical offices	3,748	3,013
Museum	3,748	3,013
Office/retail	3,748	3,013
Post office	3,748	3,013
Town hall	3,748	3,013
Small services	3,750	2,995
Retail	4,057	3,458
Parking garages	4,368	7,717

Response to NYTM Recommendations: We propose to initiate discussions as per response #1.

Evaluation Recommendations

Upon completing this evaluation, the evaluation team has some recommendations for ways to improve future SBDI evaluations.

CECONY: SBDI Program Impact Evaluation Summary

1. The error ratio used for future SBDI evaluations needs to account for the reliability of the tracking data. The final error ratio based on the RR of this study was 0.9, indicating that there was more variation in the sample than anticipated by the typical 0.3 or 0.4 assumption generally used for sampling.
2. Billing analysis – This population is not a good candidate for program-level billing analysis of savings. We recommend that it not be attempted again for Con Edison due to the complex account association challenges found in New York City, unless the program is able to definitively identify all accounts affected by project activities, with emphasis on the primary usage accounts located at the site of record. Based on billing analysis results from studies conducted for regional programs, this appears to be a consistent pattern.
3. Prior to processing the twelve month logged data, the evaluators expected the numbers to remain in the same ball park as the three month logged data. However, the overall twelve month metered operating hours went up by 8% compared to three month metered data. Given the observed increase in operating hours for longer duration metering, the evaluators recommend metering for longer duration whenever possible.

EVALUATION METHODS AND SAMPLING

Two samples were designed for the SBDI evaluation, an on-site sample to evaluate gross energy savings realization rates (RRs) and installation of measures, and a computer-assisted telephone interview (CATI) sample to verify installations and gather data for determining program attribution and SO. Both sample designs used the population of projects completed in 2011 as the frame. The evaluation used a model-based stratified sampling approach that defines strata within segments (utility and building type or utility and technology) based on kWh savings.

Program net energy savings were determined through a sample of 133 on-sites that determined gross program savings and a sample of 310 measures that determined the attribution.

For the on-sites, the evaluation team selected accounts using a model-based stratified sampling approach. This approach defines strata within segments (program and building type) based on account kWh savings and allocated targets. Using this technique, the overall relative precision for each of the programs was optimized.

In contrast to the on-site sample which used the project account as the sampling unit, the attribution sample design used the measure as the sampling unit. For the on-sites, the primary area of uncertainty was lighting operating hours, which were purportedly assigned based on building type, an account-level characteristic. For the telephone survey, the primary area of uncertainty was the influence of the program on the decision to install each measure. The evaluation team chose to sample at a level that would allow respondents to balance other influences (such as measure cost) with the program influence and respond at a level detailed enough to accurately reflect the program's influence. Therefore, the CATI sample design selected customers based on the kWh savings associated with individual measures rather than individual customer accounts.

The evaluation team calculated gross impacts by leveraging the program tracking data, program participants' billing data, data collected via phone surveys, and on-site metered data as well as other data collected on-site. 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 $(1 - FR + SO)$. These components are derived from self-reported information from telephone interviews with program participants. The evaluation team relied on the self-report method to derive both FR and SO estimates. Program participants were interviewed and asked a series of structured and

open-ended questions about the influence of the program and its various components and on the decision to purchase or install energy efficient equipment.

Iberdrola RG&E/NYSEG Small Business Direct Install Program: Impact Evaluation Summary

Prepared by: Itron, July 2013

PROGRAM SUMMARY

The Small Business Direct Install (SBDI) Program provides free energy efficiency lighting assessments and direct installation of measures to nonresidential customers with less than 100 kW of metered demand on a cost-sharing basis. The direct installation of selected measures is subject to 70/30 cost-sharing, in which the utility is responsible for 70% of the total installed measure costs. Measures include linear fixtures, CFLs, LED exit signs and occupancy sensors.

EVALUATION OBJECTIVE AND KEY FINDINGS

The evaluation objective was to estimate the first year gross and net energy and demand impacts resulting from PY2011 activity. The gross impacts evaluation included review of tracking system calculations and consistency with the NY TRM, phone survey verification of measure installations, onsite data collection (including lighting loggers) and an analysis of lighting hours-of-use. Net-to-gross (NTG) estimates, including free ridership and participant spillover, were developed using enhanced self-report surveys.

The evaluated first year annual net savings for the SBDI Program were equal to 36% of the targeted savings of 75,980 MWh. As shown in Table 1, the evaluated energy and demand NTG ratios were 0.72 and 0.71, respectively. The gross evaluation realization rate combined with the evaluated NTG rate resulted in ex post net energy and demand savings (evaluated) that were 41% and 45%, respectively, of the ex ante tracked (utility-reported) savings.

Table 1. Net Program Impact

Parameter	Electric Energy (MWh/yr)	Electric Demand (MW)	Natural Gas (MMBtu/yr)
Ex Ante Tracked Gross Savings	66,174	18.989	NA
Evaluation Realization Rate (RR)	0.58	0.63	NA
Evaluation Net-to-Gross Ratio (NTG)	0.72	0.71	NA
Ex Post Net Impact	27,395	8.589	NA

DETAILED FINDINGS: REALIZATION RATE AND NET-TO-GROSS

Gross Realization Rate

The SBDI gross realization rates reflects three components: 1) onsite survey verification that measures were in service at the time of the site visit; 2) review of tracking system algorithms, assumptions and calculations; and 3) adjustments to self-reported hours of use, which was also used to estimate peak demand coincidence factors.

Evaluated gross energy savings, along with the major adjustments that were made to reported gross savings, are reported in Table 2. The relative precision (at 90% confidence level) associated with each adjustment is reported in parentheses. As shown, the major driver of the low gross realization rates was the adjustment to self-reported hours of use and coincidence factors for CFLs and linear fluorescent lighting.

Table 2. Gross Energy Savings and Adjustment Factors by Measure (with Relative Precision)

Measure Type	Reported Gross Savings MWh	In-Service Rate (ISR)[3]	ISR-Adjusted Gross Savings MWh	HOU Adj Rate[4]	HOU-Adjusted Gross Savings MWh	Evaluated Gross Savings MWh	Gross Energy Realization Rate
CFLs	11,093	91% *(6%)*	10,067	40% *(10%)*	4,478	4,064	37%
Faucet Aerators	293	98%	287	100%	293	287	98%
LED Exit Signs	2,164	98% *(3%)*	2,121	100%	2,164	2,121	98%
LED Lamps	3,255	100% *(0%)*	3,255	100%	3,255	3,255	100%
Occupancy Sensor	14	98%	14	100%	14	14	98%
T-8 Fluorescent Lamps	49,355	99% *(1%)*	48,745	58% *(5%)*	28,807	28,452	58%
Total	**66,174**	**97% *(1%)***	**64,489**	**59% *(4%)***	**39,012**	**38,192**	**58%**

Evaluated gross peak demand savings, along with the major adjustments that were made to reported gross savings, are reported in Table 3. The relative precision (at 90% confidence level) associated with each adjustment is reported in parentheses.

[3] The program-level in-service rates were used for faucet aerators and occupancy sensors, since no on-site verification was performed for these measures

[4] The lighting hours of use analysis did not include occupancy sensors, LED measures or aerators. Tracking system HOU values were accepted for these measures.

Table 3. Gross Demand Savings and Adjustment Factors by Measure

Measure Category	Gross Reported Savings MW	In-Service Rate (ISR)[5]	ISR-Adjusted Gross Savings MW	Coincidence Factor (CF)[6]	CF-Adjusted Gross Savings MW	Evaluated Gross Savings MW	Gross Demand Realization Rate
CFLs	3.38	91% *(6%)*	3.06	42% *(10%)*	1.42	1.29	38%
LED Exit Signs	0.27	98% *(3%)*	0.27	100%	0.27	0.27	98%
LED Lamps	0.97	100%	0.97	100%	0.97	0.97	100%
Occupancy Sensor	0.00	NA	0.00	NA	0.00	0.00	NA
T-8 Fluorescent Lamps	14.36	99% *(1%)*	14.19	67% *(4%)*	9.61	9.49	66%
Aerators[7]	0.00	NA	0.00	NA	0.00	0.00	NA
Total	**18.99**	**97% *(1%)***	**18.49**	**65% *(3%)***	**12.27**	**12.02**	**63%**

Net-to-Gross

The evaluated NTG ratios were 0.72 and 0.71 for kWh and kW, respectively. As shown in Table 4, the program level NTG ratios were driven almost entirely by free ridership. Participant spillover increased the evaluated NTG ratios for both kWh and kW by less than 2%. Non-participant spillover was outside the scope of the evaluation. The program level relative precision of the NTG estimates was 6% with 90% confidence.

Table 4. Free Ridership and Spillover Estimates

Attribution Variable	Energy	Demand
Free ridership	0.30	0.29
Inside spillover	0.02	0.02
Outside spillover	NA	NA
Non participant spillover	NA	NA
Net-to-gross factor (equals 1-FR+SO)	0.72	0.71

Table 5 reports the evaluated net kWh energy savings for the SBDI program. The relative precision (at the 90% confidence level) associated with each measure level NTG estimate is reported in parentheses for each measure and the total.[8]

[5] The program-level in-service rates were used for faucet aerators and occupancy sensors, since no on-site verification was performed for these measures.

[6] The lighting hours of use analysis did not include occupancy sensors, LED measures or aerators. Tracking system CF values were accepted for these measures.

[7] The TRM does not specify demand savings for aerators and tracking data reports zero demand savings for aerators.

[8] The program-level In-Service Rate was used for faucet aerators and occupancy sensors since no on-site verification was performed for these measures.

Table 5. Summary of Reported and Evaluated kWh Impacts by Measure Category

Measure Category	Number of Net Survey Sample Sites with Measure	Evaluated Gross kWh	Evaluated NTG[9]	Evaluated Net kWh
CFLs	121	4,063,833	66% *(17%)*	2,685,921 *(18%)*
Faucet Aerators	0	287,087	90%	258,378
LED Exit Signs	64	2,120,829	77% *(8%)*	1,633,974 *(8%)*
LED Lamps	9	3,254,717	67% *(54%)*	2,181,920 *(54%)*
Occupancy Sensor	1	13,599	90%	12,239
Linear Fluorescent	181	28,451,579	72% *(7%)*	20,622,160 *(7%)*
Total	**376**	**38,191,644**	**72% *(6%)***	**27,394,592 *(6%)***

EVALUATION METHODS AND SAMPLING

Evaluation of Gross Impacts

The gross realization rates reflects three components: 1) verification that measures were installed and operational and conformed to technical specifications as documented in the tracking databases (In-Service Rates); 2) review of tracking system algorithms, assumptions and calculations; and 3) adjustments to self-reported hours of use. The hours-of-use analysis, which adjusted self-reported hours of operation and use to reflect actual observed hours of use, was a significant part of the SBDI program evaluation effort and was the major driver of the low gross realization rates for CFLs and linear fluorescent lighting.
Primary data collection included: 215 telephone surveys, 75 on-site visits, and placement of 730 lighting loggers at 55 sites. For the gross impacts evaluation, the surveys were used to verify measure installation and select operating parameters, as well as recruit participants for the on-site survey. Table 6 provides the phone survey sample dispositions along with the percent of total sample for that disposition.

Table 6: SBDI Participant Survey Sample Disposition

Sample Disposition	Number of Customers	% of Sample Frame
Participants in Sample Frame	1,043	100%
Completes	187	17.9%
Refusals	6	0.6%
Reached Maximum Number of Attempts	10	1.0%
Language Barrier	4	0.4%
Incorrect Phone Number	12	1.2%
Designated Respondent Not Available	21	2.0%
Remaining Sample when Sample Target Achieved	756	72.5%
Appointment Made for Later Date	47	4.5%

[9] The deemed NTG ratio of 90% was used for faucet aerators and occupancy sensors since a NTG survey was not performed for faucet aerators, and only a single measure was included in the sample for occupancy sensors, which was considered not statistically significant.

To increase confidence and precision in the measure level gross savings estimates, the 75 on-site surveys from PY 2011 were combined with 78 on-sites completed for PY 2010. Participant sites were recruited as part of the phone surveys. For each on-site evaluation completed, the evaluation team verified that the lighting equipment was installed and operational. Other essential tracking data were also verified, including the locations, sizes, configurations, makes, and models of the rebated lighting.
Table 7 presents the weighted installed and operable measure-level in-service verification rates as a percent of reported savings, which were developed from the onsite surveys.

Table 7. Population Installed and Operable Verification Rates (% of reported kWh savings)

Utility	Measure	ISR	Verification Measures Surveyed
NYSEG	Linear Fluorescents	99.63%	113
	CFL	98.01%	41
	Occupancy Sensor	--	0
	Faucet Aerator	--	0
	LED Exit Signs	98.38%	12
	LED Lamps	100.00%	1
RG&E	Linear Fluorescents	96.98%	97
	CFL	68.91%	45
	Occupancy Sensor	--	0
	Faucet Aerator	--	0
	LED Exit Signs	97.28%	14

*The weighted average program level verification rate was applied to occupancy sensors and faucet aerators.

Lighting loggers were installed in April 2011 for seven to ten months, which ensured that data were captured prior to the 2011 summer solstice and through the winter solstice. This duration should provide a reasonably accurate and complete representation of the 8760-hour lighting HOU for both the linear fluorescent and CFL lighting measures.
To develop lighting HOU and CF estimates for CFLs and linear fluorescent lighting measures, hourly lighting usage profiles (covering all 8,760 hours in a year) were developed for five building types. The hourly profiles were based on self-reported data on HOU and business hours, which were adjusted using data obtained from an evaluation of small commercial programs for the California Public Utility Commission.[10] Logger data was used to validate and adjust the methods and results. The HOU and coincidence factor estimates were derived from these hourly lighting usage profiles.
As shown in Table 2 and Table 3 (above), the program level relative precision (at a 90% confidence level) of the installation verification was 1%, for the hours of use analysis was 4%, and for the coincidence factor analysis was 3%.

Evaluation of Net Impacts
An enhanced self report approach was used for the NTG analysis, based on data obtained from the 187 PY 2011 participant telephone surveys. The phone survey instrument included a structured battery of questions designed to triangulate participant responses. The distribution of NTG scores calculated for each respondent for the different measures installed is provided in Table 8.

[10] California Public Utility Commission, *Small Commercial Contract Group Direct Impact Evaluation Report*, prepared by Itron, Inc. February 9, 2010

Table 8: SBDI Participant NTG Scores

NTG Sc8ore	T-8 Fluorescent Lamps	CFLs	LED Exit Signs	LED Lamps	Occupancy Sensor
0-0.25	22	12	8	3	0
0.25-0.5	3	2	2	0	0
0.5-0.75	34	23	16	3	1
0.75-1	80	53	28	2	0
1.00 - 1.25	42	31	10	1	0

A representative weight was created for each measure category by dividing the total ex-ante kWh savings of the entire population by the total ex-ante kWh savings of the entire sample. Each measure category weight was then multiplied by the measure level ex-ante kWh savings at the site to obtain a site level weight. This process was used to obtain a weight at the measure category level specific for each site. The weights were then applied to the phone survey results and used in the analysis of the NTGRs.
Table 9 reports the results of the NTG analysis, along with relative precision (at 90% confidence), for each measure.

Table 9: SBDI NTG Relative Precision and Bounds

Measure Category	Number of Survey Sample Respondents with Measure	Evaluated NTG	Relative Precision	Lower Confidence Limits for Means	Upper Confidence Limits for Mean
CFLs	121	66%	17%	55%	77%
Faucet Aerators	0	--	--	--	--
LED Exit Signs	64	77%	8%	71%	83%
LED Lamps	9	67%	54%	31%	103%
Occupancy Sensor	1	--	--	--	--
T-8 Fluorescent Lamps	181	72%	7%	68%	77%
Total	**376**	**72%**	**6%**	**68%**	**76%**

RECOMMENDATIONS AND PROGRAM ADMINISTRATOR RESPONSE

No recommendations were provided as part of the impact evaluation of the PY2011 SBDI programs.

National Grid
Niagara Mohawk

Small Business Services Program:
Impact Evaluation Summary

Prepared by: DNV GL, September 23, 2014

PROGRAM SUMMARY

The SBS Program provides direct install of energy-efficient measures for small commercial and industrial ("C&I") customers. The small business sector (defined as customers with an average monthly demand of less than 100 kW) is a heterogeneous group of end users that represent a sizeable portion of the National Grid commercial electric customer base in their Niagara Mohawk territory.

EVALUATION OBJECTIVE AND KEY FINDINGS

The primary objective of this evaluation is to quantify the gross and net annual energy and summer demand impacts of prescriptive lighting installed through the SBS program. The study design was laid out to provide two primary means of assessing gross impacts: an M&V study of prescriptive lighting installed in the 2010 program year and a billing analysis of the 2010 and 2011 program years. The M&V study provides results by primary discrepancy, including hvac interactive. Net to gross factors were derived from surveys performed on the 2011 and 2012 program years.

The saving values in the table below are for prescriptive lighting installed without lighting controls installed through the program[11] and represent roughly 58% of the program savings in 2010. The values are based on the M&V site results discussed below. The billing analysis results were substantially lower and after consideration of the strengths and weaknesses of each method, we have based our final impact estimate on the site level M&V work. The net to gross ratio and accompanying values are derived from the core NTG approach.

Table 1. Prescriptive Lighting without Controls Net Program Impact

Parameter	Electric Energy (MWh/yr)	Electric Demand (MW)	Natural Gas (MMBtu/yr)
Ex Ante Tracked Savings	64,265,935	18.437	N/A
Evaluation Realization Rate (RR)	80.3%	95.7%	N/A
Evaluation Net-to-Gross Ratio (NTG)	97%	97%	N/A
Ex Post Net Impact	50,057,379	17.11	N/A

DETAILED FINDINGS: REALIZATION RATE AND NET-TO-GROSS

Realization Rate:

- The final energy savings realization rate for prescriptive lighting without controls from the on-site M&V work is 80.3% with a precision of +/- 9.1% at the 90% confidence interval.
- The connected demand realization rate for prescriptive lighting without controls from the on-site M&V work is 95.7% with a precision of +/- 2.0% at the 90% confidence interval.

[11] The lighting measures may have had pre-existing controls installed.

Net-to-Gross:
The table below presents the final NTG results by measure type and attribution variable. In this study, we assessed free ridership[12] and participant like inside spillover[13]. The overall program free ridership rate is 5.8% while the participant inside like spillover rate is 1.7%. The overall NTG comprising these two factors is 95.9%, with a precision of +/-1.4% at the 90% confidence interval.

Table 2. Free Ridership and Spillover Estimates

Attribution Variable	Prescriptive Lighting without Controls Factor	Prescriptive Lighting with Controls Factor	Refrigeration Factor	Non-HVAC Motors and Drives	Overall
Free ridership	5.2%	6.7%	6.4%	21.9%	5.8%
Participant Inside spillover (Like)	2.2%	0.6%	1.2%	0.1%	1.7%
Outside spillover	N/A	N/A	N/A	N/A	N/A
Non participant spillover	N/A	N/A	N/A	N/A	N/A
Net-to-gross factor (equals 1-FR+SO)	97.0%	93.9%	94.8%	78.2%	95.9%

Other Results:

There were several other savings factors of interest to National Grid that were assessed as part of this study. These include:

- Summer kW coincidence factor of 56.6% with a precision of +/-15.3% at the 90% confidence interval.
- HVAC kW Interactive Effect Factor of 110.4% with a precision of +/-2.7% at the 90% confidence interval.
- HVAC kWh Interactive Effect Factor of 104.0% with a precision of +/-1.1% at the 90% confidence interval.
- Average annual hours of use among all sampled sites of 2,708.
- % on peak kWh of 72% with a precision of +/-6.1% at the 90% confidence interval.
- Heating HVAC Interaction Effect (MMBtu/kWh) of -.0023 with a precision of +/-86.9% at the 90% confidence interval.

EVALUATION METHODS AND SAMPLING
There were three primary activities undertaken as part of this study. The bullets below provide a brief description of each.

- On-site assessments at 70 participants with program installed prescriptive lighting without controls from the 2010 program year. These on-site visits were statistically selected, and included comprehensive inventories and time-of-use metering performed for a year. The method for the prescriptive on-sites with metering adheres to IPMVP Option A. A spreadsheet engineering model

[12] Free ridership is the percent of savings attributed to customers who participate in an energy efficiency program but would have, at least to some degree, installed the same measure(s) on their own if the program had not been available.
[13] Participant inside like spillover is the percent energy savings associated with energy efficient equipment that is the same as that installed through the program that is also installed in the same facility by consumers who were influenced by an energy efficiency program, but without direct financial or technical assistance from the program.

was used to develop all savings estimates and factors of interest for each sampled site. This analysis was performed in a manner that allowed the determination of impacts at each site and the primary reason for discrepancies observed between the gross and tracking savings estimates. These site level results were then expanded up to represent the impacts of the full prescriptive lighting without controls population, along with all accompanying precisions.

- Two billing analyses were conducted in 2010 and 2011 participants. One methodological approach employed a pooled, time-series, cross-section approach that weather-normalized measures savings in a single model with multiple months of data for all accounts/sites. A second approach was a site-level modeling approach, which included a site-level weather normalization process followed by savings estimates based on the pre-/post- differences in weather-normalized consumption. As part of exploring the relationship between the billing analysis and on-sites, this study included a billing analysis on the subset of on-sites.
- A self-report net to gross approach to assess free ridership and inside like spillover. This approach utilized a core algorithm consistent with the standardized approach that National Grid has exercised for assessing net to gross in Massachusetts. We also calculated NTG based on an alternative method that we have seen implemented in NY State. This survey was performed with 484 participants from 2011 through mid-year, 2012.

RECOMMENDATIONS AND PROGRAM ADMINISTRATOR RESPONSE

The following recommendations were made by the evaluators conducting this study. National Grid's initial response to these recommendations is also summarized below and will be tracked over time.

Recommendation 1: The reliability of a billing analysis is largely dependent upon the availability and relationship of consumption data for program treated spaces to the estimate of savings in those spaces. Despite a significant effort to acquire all accounts associated with each participant and those spaces in particular in the population as well as the on-site sample, we suspect the inability to do this consistently at either level limited the ability to produce a reliably accurate billing analysis result.

Response to Recommendation 1:

A billing analysis approach offers the advantage of near census inclusion of sites in analysis. It is difficult, however, to control for changes in consumption at the site that occur at the time of the retrofit but are not specifically caused by the retrofit measures. While gathering more comprehensive account information is useful, the possibility and likelihood of these kinds of changes should be explored further as part of any future billing analysis on this customer segment.

We agree with this specific recommendation regarding better collection of account numbers on projects. It is to be noted that this analysis was based on 2010 and 2011 data and there have been many process improvements since this timeframe as a result of process evaluation findings. There have been efforts made to have program vendors more thoroughly collect meters and/or accounts that serve program treated areas in the interest of making a billing analysis a more feasible option for consideration when evaluating the small business program in the future.

Recommendation 2: Currently, the NY TM stipulates using a 0.9 factor as a stand in for NTG in reporting program savings. There are two ways in which one might apply net to gross factors from this report. The first is if the factor is applied to tracking savings, under which case the net realization rate from this report can be used. Alternatively, if changes to the TM are made such that it is believed that the tracking savings become more similar to those estimated as gross savings in this report, then using the combined net-to-gross ratio from this report may be more appropriate for consideration.

Response to Recommendation 2: The results suggest that collective consideration (National Grid and DPS) be made with regard to how to apply the results of this study (to the Tech Manual hours or to a National Grid specific realization rate). It is our recommendation that further analysis be done through statewide evaluations in order to get enough samples by building/facility type to make adjustments to specific values such as hour of use, diversity factors, HVAC interactive factors and NTG factors. Using just one NTG factor would not accurately reflect the current state and processes of the program since improvements were made since 2010 and 2011 as a result of process evaluation recommendations. In addition, it is not clear how a statewide result should be applied for a program that is being implemented by several PAs in slightly different ways or if each PA would continue to apply different results.

Recommendation 3: The lighting hours of use assumptions in the Technical Manual appear to be high in general, although only the automotive facility TM assumed hours falls outside of the 90% confidence interval result. As such, while there is evidence from this study that the TM hours might be overstated, we would encourage National Grid and the DPS to compile added evidence from other New York PAs to be sure this trend holds across other territories before making TM revisions on this matter[14].

Response to Recommendation 3: This recommendation is actionable in that compilations of all PA evaluations of hours of use are needed in order to have a sample size large enough to be statistically applicable by building type/facility type. The current effort in this study provides a start toward this; however, there was not a large enough sample in order to fully assess results by building / facility use type for hour changes to the technical manual. It is our recommendation that further analysis be done through statewide evaluations in order to get sufficient samples by building/facility type to draw more meaningful conclusions before updating the TM.

Recommendation 4: The Technical Manual currently assumes a coincident factor of 1.0 for commercial indoor lighting measures. The estimated coincidence factor from this study is 0.566. We recommend consideration of this value for small business direct installation lighting applications or that minimally, National Grid and the DPS to compile additional data from other New York PA's evaluations.

Response to Recommendation 4: The results suggest that the Technical Manual may be overstated with regard to the coincident factor. But, further discussions would need to take place with DPS in order to compile results from all PAs in order to determine next steps.

Recommendation 5: Currently, National Grid uses HVAC interactive estimates of 1.19 for summer kW and 1.07 for annual energy savings. These values are proxies that represent average factors from the table of HVAC interactive factors by Building Type in the Technical Manual. This study suggests that National Grid might consider the use of a kW HVAC Interactive Effect Factor of 1.11, and a kWh interactive factor of 1.04 or again work with DPS to combine this data with that from other New York PAs.

Response to Recommendation 5: Further discussions would need to take place with DPS in order to compile results from all PAs in order to determine next steps.

[14] National Grid uses hours of use vendor estimates at the site and location level at the time of participation, however such adjustments may help other PAs that rely more heavily on the Technical Manual.

Recommendation 6: We recommend that future studies of energy impacts of the small business program use an error ratio of 0.5 for lighting for-sites metering sample sizing. The error ratio we recommend for targeting demand savings is a 0.2 error ratio.

Response to Recommendation 6: As part of future evaluation planning, National Grid will work with the DPS and other stakeholders to assess the appropriateness of using this recommended error ratio.

Recommendation 7: We recommend that National Grid work with Program vendors to ensure that the location of measures installed can be acquired for evaluation purposes.

Response to Recommendation 7: This recommendation has been acted upon and completed. The documentation for projects has greatly improved since the 2010 and 2011 timeframe as a result of process evaluation recommendations. We believe these improvements will better support future impact studies.