Seeking Answers: Strategies for Interpreting Consumption Analyses

Matthew Deslauriers and Kathryn Parlin, West Hill Energy and Computing, Brattleboro, VT
Lina Riveros, Con Edison, New York, NY

ABSTRACT

Feedback from impact evaluations can provide critical advice to program staff for understanding how to improve program performance. While consumption analysis is a common method for impact evaluation, it often does not provide insight into underlying causes for differences between program reported and evaluated savings. Metering and on-site measurement are more likely to be able to identify project-specific issues and trends, but these methods tend to be substantially more expensive.

The question becomes how to use consumption analysis along with readily available data and/or targeted measurement and verification (M&V) to understand how to improve program reported savings while keeping evaluation costs down. We developed and tested various strategies in the impact evaluation of a multifamily energy efficiency program in the Northeast.

We used the following strategies to explore drivers of realization rates:

- Review equations and inputs that were used to calculate *ex ante* savings and identify key inputs
- Develop a list of potential contributors to the overstatement of *ex ante* savings based on review of the inputs
- Perform a reality check of the program *ex ante* inputs using data collected for the impact evaluation
- Conduct sensitivity analyses to isolate the impact of each input on the realization rates
- Conduct field observations and targeted metering to investigate inputs that could not be assessed by consumption analysis or other methods
- Review other relevant studies that provide a direct comparison to the evaluated measures.

This approach was used in the impact evaluation of a multifamily efficiency program and focused on two measure groups, lighting and natural gas energy boiler controls. For lighting, the outcome of this analysis indicated that overstatement of the baseline wattage was the largest contributor to the low realization rate. The boiler control analysis indicated that the energy savings factor (ESF) was substantially overstated. While this evaluation was limited to lighting and boiler controls, the strategy has the potential to be applied to a broad range of end uses and applications.

Introduction

Consumption analysis was the primary method of evaluating savings for an impact evaluation for a multifamily energy efficiency program in New York City. The consumption analysis compared pre- and post-installation electric and natural gas consumption to estimate the actual reduction in energy use. The impact evaluation resulted in realization rates of about 40% for both lighting and boiler controls.

Evaluators used consumption data in combination with other sources of data to investigate the reasons for the discrepancy between the program reported and evaluated savings. This process involved breaking down the savings into the component parts and assessing how the available consumption data could be used to gain additional insight. In addition, 40 exploratory site visits (20 for lighting and 20 for boiler controls) and a review of relevant studies were conducted to support and verify the findings. This process allowed us to define the major contributors to the realization rate and identify strategies for improving program performance going forward.

While consumption analysis is often selected as a relatively low-cost impact evaluation technique when compared to metering, the methods used in this review can provide additional insight into the reasons for discrepancies between *ex ante* and *ex post* savings without the need for extensive metering. The remainder of this paper covers a brief description of the program, methods, results, and conclusions.

Program Description

The program provides incentives for qualifying upgrades in multifamily buildings. Qualifying upgrades include electric and gas energy efficiency measures such as efficient lighting, heating, ventilation, and air conditioning (HVAC) maintenance and control systems, weatherization, and occupancy sensors. The program has four components: electric common area upgrades, natural gas common area upgrades, in-unit upgrades, and custom measures. The common area measures were the focus of this evaluation since they made up 92% of electric and 91% of natural gas savings.

Among the common area measures, lighting efficiency accounted for 99% of the electric savings and boiler controls covered 87% of the natural gas savings. Consequently, these measures were the focus of this evaluation.[1] Lighting measures were primarily lighting efficiency measures where less efficient bulbs were replaced with LEDs. The boiler controls measures were more advanced steam boiler controls with Wi-Fi access, outdoor air reset, and indoor temperature sensors to prevent overheating.

Methods

The initial impact evaluation involved analyzing electric and natural gas billing records from the common area meters of participating buildings to determine the evaluated savings. There were 678 projects with electric bills and 231 with the natural gas bills in the final analysis. This consumption analysis resulted in a low realization rate for the lighting and boiler controls measures.

After the initial impact evaluation was performed, a more in-depth analysis was conducted to understand the drivers of program savings. This analysis included the following steps:

1. Review equations and inputs that were used to calculate *ex ante* savings and identify key inputs
2. Develop a list of potential contributors to the overstatement of *ex ante* savings based on review of the inputs
3. Perform a reality check of the program inputs using data collected for the impact evaluation
4. Conduct sensitivity analyses to isolate the impact of each input on the realization rates
5. Conduct field observations and targeted metering to investigate inputs that could not be assessed by consumption analysis
6. Review other relevant studies that provide a direct comparison to the evaluated measures

The data sources used for this evaluation are described in the Table 1 below.

[1] Weatherization measures were rarely installed and there were no projects with both weatherization and boiler controls during the evaluation period.

Table 1. Data sources

Source	Data Used
Program Data	Measures installed, install date, *ex ante* savings, contractor name, account number, address
Billing Data	Consumption, read date, read code, service class code, part supplied code, account number, address, borough, block, lot
NOAA Weather Data	Hourly outside air temperature, date, hour, station location
On-site Data Collection	Light meter data to estimate hours of use; boiler size collected at site visit
External Studies	Evaluations of similar programs and earlier evaluations of the multifamily program

The methods and results of using the strategies identified above are described in the identifying equations and inputs, developing a list of potential contributors to the overstatement of program savings, reality checks, sensitivity analysis, field observations, and review of relevant studies sections below.

Identifying Equations and Inputs

The New York State Technical Resource Manual[2] (TRM) for the lighting and boiler controls measures was reviewed to identify the inputs into the calculations of the *ex ante* savings. The relevant inputs are highlighted in red.

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

Equation 1: TRM lighting savings

The key inputs from this equation are units, baseline watts, efficient watts, and annual hours of use. The inputs, their source of information, and their likely contribution to program savings are summarized in Table 2.

Table 2. Lighting program inputs and contribution to program savings

Calculation Input	Definition	Source of Input	Contribution to Magnitude of Program Savings
Units	Installed units	Site-specific	High
$W_{baseline}$	Baseline watts	TRM lookup based on site-specific data	High
W_{ee}	Efficient case watts	TRM lookup based on site-specific data	Moderate
$Hrs_{operating}$	Annual hours of use	TRM lookup	High
$HVAC_c$	Interactive effects factor	TRM	Low; eliminated from further review since it was a small adjustment to savings and difficult to evaluate

2 New York Standard Approach for Estimating Energy Savings from Energy Efficiency Programs - Version 4

The boiler controls measure used the equation below with the relevant portions highlighted in red. The important inputs for this measure were boiler size (BTU_{in}), equivalent full load hours for heating ($EFLH_{heating}$), and the energy savings factor (ESF). The BTU_{in} input was from the nameplate rating of the boiler at the site, the $EFLH_{heating}$ was a lookup value in the TRM based on building age and city, and the ESF was a constant of 22%.

$$\Delta therms = \frac{BTUin}{100{,}000} \times EFLH_{heating} \times ESF$$

Equation 2: Savings from boiler controls

The key inputs for boiler controls are the input Btu of the boiler, the heating EFLH and the ESF (or percent reduction as compared to pre-install heating use). The inputs, their source of information, and their likely contribution to program savings are summarized in Table 3.

Table 3. Boiler controls, program inputs, and contribution to program savings

Calculation Input	Definition	Source of Input	Contribution to Magnitude of Program Savings
BTU_{in}	Input Btu of the boiler	Site-specific	Moderate
$EFLH_{heating}$	Equivalent full load hours for heating	TRM Lookup based on site-specific data	Moderate
ESF	Energy Savings Factor	TRM deemed value of 22%	High

Developing a List of Potential Contributors to the Overstatement of Program Savings

Each program input was assessed to understand how errors in the inputs would affect program savings. The list of inputs and their potential contributions to the overstatement of savings are presented in Table 4.

Table 4. Calculation inputs and potential contribution to overstatement of program savings

Measure	Calculation Input	Possible Source of Errors in Program Input	Expected Outcome of Errors in Program Inputs
Lighting	Installed units (Units)	Overstating the number of installed units	Both baseline and efficient lighting use would be too high
	Baseline watts ($W_{baseline}$)	Overstating the baseline W	Only baseline lighting use would be overstated
	Efficient case watts (W_{ee})	Understating the efficient W	Efficient lighting use would be understated
	Hours of use ($hrs_{operating}$)	Overstating the hours of use	Both baseline and efficient lighting use would be too high
Boiler Controls	Boiler Input BTUs (BTU_{in})	Overstating the boiler's size	Pre-install use would be overstated
	Equivalent full load hours ($EFLH_{heating}$)	Overstating hours of use	
	Energy savings factor (ESF)	Overstating the percent of energy savings	Savings would be overstated

Reality Checks

The purpose of the reality check is to use the available information to assess whether the baseline and efficient case from the program *ex ante* inputs are reasonable. For example, Equation 1 shows that lighting savings are calculated using the wattages of the baseline lamps and estimated hours of use; this information can be used to estimate the pre-install lighting kWh as calculated by the program and this number can then be compared to the actual pre-install billed consumption. Each of the program inputs was tested and the estimated consumption was compared to the either pre- or post-consumption data. The discussion below begins with lighting measures followed by the boiler controls.

Lighting

The baseline and efficient wattages were tested by comparing the program inputs to the consumption at the building level, and the results were then aggregated to the program level. The steps are as follows:

1. Equation 3 was used to calculate the baseline from the program inputs on a building-by-building basis.
2. The annual consumption calculated from the program *ex ante* inputs was compared to the pre-install base consumption (nonweather dependent use) determined from the bills.
3. The analysis was aggregated to the program-level by binning the buildings according to the baseline *ex ante* estimated lighting use as a percent of the pre-install base consumption.

The same process was used to calculate the efficient *ex ante* use and compare it to the post-install consumption.

$$kWh_{base} = \sum \frac{Watts_{base}}{1000} \times Qty \times \ Hours\ of\ Use$$

Equation 3: Baseline energy use from the program *ex ante* inputs

The expected outcome is that the total pre-install base consumption from the bills should be *higher* than the estimated lighting use from the program *ex ante* inputs, as the billed consumption includes lighting and other end uses. Figure 1 shows the approach to analyzing the results.

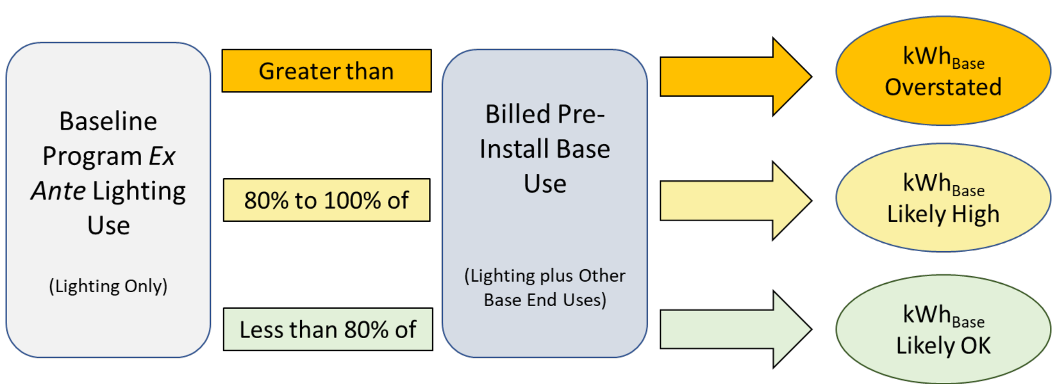

Figure 1: Analysis of comparison of baseline program *ex ante* lighting use to billed pre-install base use

This analysis showed that program *ex ante* lighting estimates exceeded the total base use from the consumption analysis in 43% of buildings, indicating that the program *ex ante* inputs overstate lighting use in *at least* 43% of buildings. Figure 2 shows that the average realization rate is substantially lower in buildings where the program *ex ante* estimated use is greater than 100% of the billed use. This result suggests that the baseline wattage and/or hours of use in the program data are overstated and this overstatement is likely to be a major driver of the low realization rate.

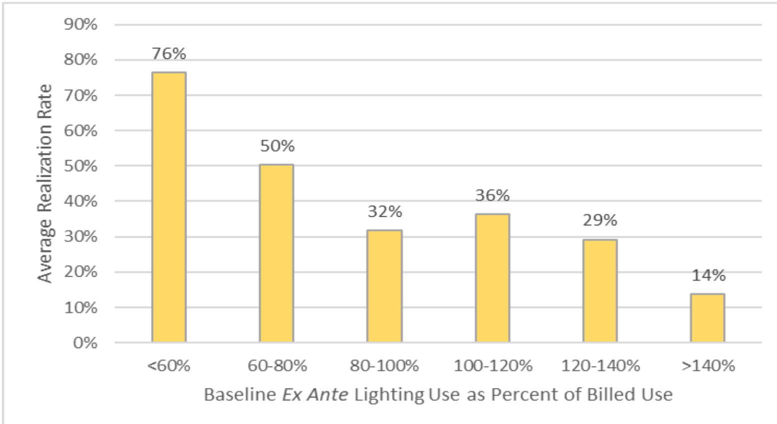

Figure 2: Baseline *ex ante* lighting use relative to billed use and realization rate

A similar analysis was conducted for the efficient case using the post-install base consumption from bills, program *ex ante* inputs for the efficient lighting and the same bins as the baseline analysis. If lighting use was overstated in both the baseline and efficient cases, it would suggest that the hours of use or fixture quantity was substantially overstated.

Keeping the buildings in the same bins as shown in Figure 2 (using the baseline), the efficient case was added in Figure 3. For the efficient case (in green), *no buildings were at or above the pre-installed base use* from the bills. This result shows no indication that the efficient use as calculated from the post-install *ex ante* inputs is overstated, as is illustrated most dramatically in the final group (>140%), where the average program *ex ante* baseline use for the 135 buildings is over three times (342%) the billed use in the baseline and only 72% of the billed use in the efficient case.

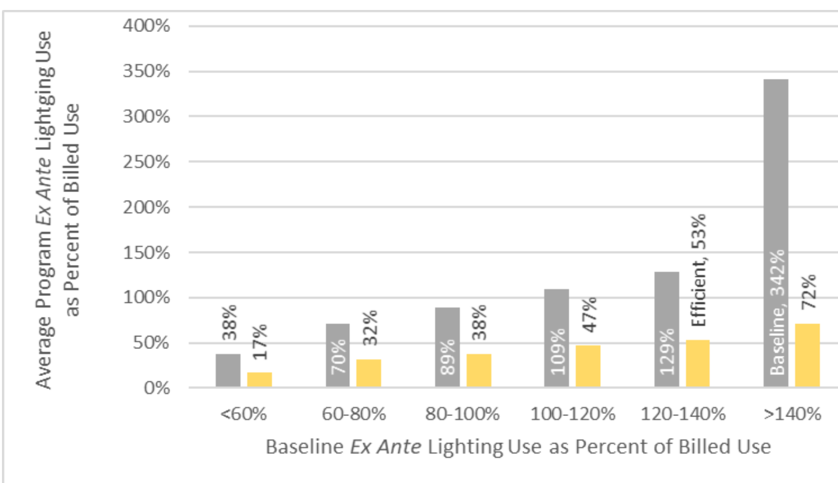

Figure 3: Comparison of baseline and efficient program *ex ante* lighting use to billed consumption

If either the total number of fixtures or hours of use were substantially overstated, both the efficient and baseline lighting use would be overstated. Thus, combining the baseline and efficient case analyses indicates the major source of error is in the characterization of the baseline conditions.

A third analysis was conducted to estimate the non-lighting base use for the baseline and efficient cases. The average non-lighting use is expected to be similar in the pre- and post-period, as the electric component of the Multifamily Program is focused entirely on lighting measures. The yellow in Figure 4 is the program *ex ante* use and the grey is the remaining non-lighting use from the bills. This analysis shows that the non-lighting use is almost *three times higher* in the post-period, on average over all 678 buildings in the analysis. This finding also supports the hypothesis that the baseline wattages are overstated.

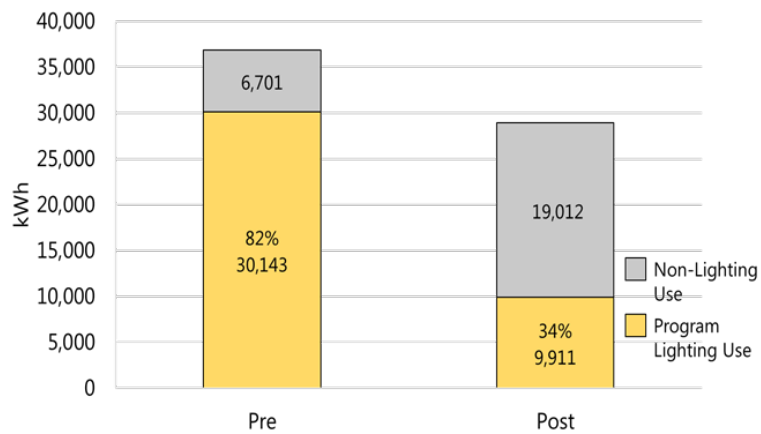

Figure 4. Comparison of average program *ex ante* lighting use to average billed use

From these three analyses, we concluded that baseline watts were most likely a major source of error and efficient watts were a less likely source. These findings are summarized in Table 5 below.

Table 5. Summary of lighting reality check findings by program input

Measure	Calculation Input	Findings
Lighting	Baseline watts ($W_{baseline}$)	Likely to be overstated as baseline energy use (units x $W_{baseline}$ x hours of use) is not in a reasonable range when compared to bills.
	Installed units (Units)	Less likely to be a factor as efficient energy use (units x W_{ee} x HOU) is in a more reasonable range when compared to bills.
	Efficient case watts (W_{ee})	
	Hours of use (HOU)	

Boiler Controls

Savings could be overestimated due to an overestimation of the annual heating use or by overstatement of the ESF. Consequently, this analysis was divided into two parts:

1. Compare the program-estimated and billed annual heating use
2. Determine the ESF from the consumption analysis

The first two terms (BTU_{in} x EFLH) in the TRM formula estimate the annual heating consumption for the building and savings were calculated as a percent of this heating consumption.

The consumption analysis provided us with average estimates of the pre-install consumption and allowed us to conduct a direct comparison between the program *ex ante* consumption (BTU_{in} x EFLH) and the pre-install heating consumption from the bills. The program *ex ante* consumption was found to be about 13% higher than the pre-install heating consumption from the bills as seen in Figure 5.

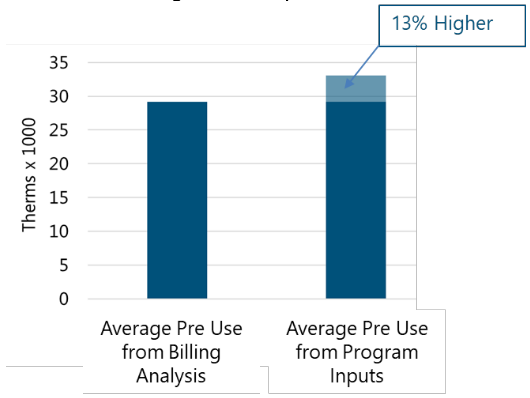

Figure 5. Comparison of pre-install use from program *ex ante* inputs and from the consumption analysis

The consumption analysis also provided the opportunity to compare the ESF of 22% from the TRM and the energy savings as seen in the bills. This analysis indicated the average ESF is about 10% versus the TRM estimate of 22%. These results indicated that the ESF was the primary driver of the low realization rate, while the BTU_{in} x EFLH was the secondary driver. Further analysis of BTU_{in} was conducted through spot checks to understand if field collection of boiler size was problematic or if the EFLH was the only factor causing the 13% increase in the heating use when compared to the bills.

Table 6. Summary of boiler controls reality check findings by program input

Measure	Calculation Input	Findings
Boiler Controls	Boiler input BTUs (BTU_{in})	Not likely to be a substantial contributor to low realization rate, as boiler input x $EFLH_{heating}$ is only 13% higher on average when compared to bills, yet the realization rate is 40%.
	Equivalent full load hours ($EFLH_{heating}$)	
	Energy savings factor (ESF)	Likely to be the primary contributor of the low realization rate. Billing analysis indicates savings of about 10% on average compared to the ESF of 22%.

Sensitivity Analysis

This sensitivity analysis study was specific to lighting and assessed how the change in savings by a consistent discrepancy in the baseline and efficient wattages. The results of the sensitivity analysis are shown in Figure 5. A 30% change in the median baseline wattage caused a 70% change in savings while a 30% change in efficient wattage only caused a 10% change in savings, as shown in Figure 6.

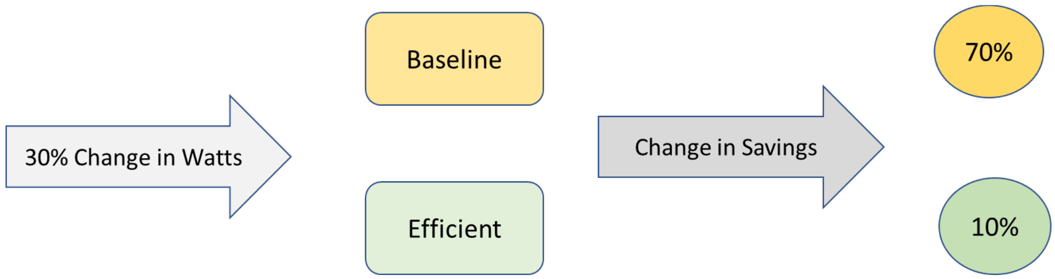

Figure 6: Change in savings for a 30% change in baseline and efficient wattage

Table 7 provides a summary of the lighting sensitivity analysis. The main finding is that an error in recording the baseline watts has a dramatically greater impact on the savings than a similar error for the efficient watts.

Table 7. Summary of lighting sensitivity analysis findings by program input

Measure	Calculation Input	Findings
Lighting	Baseline watts ($W_{baseline}$)	Likely to be overstated as the overstatement of the baseline watts has a substantially larger impact on the savings than overstatement of the efficient watts, further indicating that baseline wattage is the largest driver of the low realization rate.
	Installed units (Units)	Not evaluated in this section
	Efficient case watts (W_{ee})	Less likely to be a factor as changes in the efficient wattage does not change the savings substantially
	Hours of use ($hrs_{operating}$)	Not evaluated in this section

Field Observations

Two sets of on-site visits were conducted, one to meter the lighting hour of use and the other to assess the boiler controls. The lighting on-site visits are discussed first, followed by the boiler controls.

Lighting Field Observations and Metering

Lighting was metered in 20 buildings to obtain the hours of use by space type. This meter data was analyzed and aggregated to develop a rough estimate of the accuracy of the program inputs for hours of use. The two main findings are described below.

- Average operating hours for interior multifamily common area lighting were found to be 6,280 hours per year in the overall meter data, versus 7,665 in the TRM V4.
- TRM V4 provided a single number (7,665) for both interior and exterior lighting. Metering indicated that exterior annual hours of use were closer to 4,100 hours.

A summary of findings from this analysis are shown in Table 8.

Table 8. Summary of lighting field observation findings by program input

Measure	Calculation Input	Findings
Lighting	Baseline watts ($W_{baseline}$)	Not evaluated in this section
	Installed units (Units)	
	Efficient case watts (W_{ee})	
	Hours of use ($hrs_{operating}$)	Interior hours of use were found to be about 22% higher in the TRM than metering suggested. This showed that interior hours of use were overstated but were likely a secondary driver of the low realization rate.

Boiler Control Field Observations

The intent of the boiler controls site visits was to identify characteristics that were contributing to the low or high savings. In addition, boiler nameplate input capacity data was recorded. The field data was compared to the program reported boiler input capacity to get a sense of how often errors in boiler input capacity occurred in the program data. The sample design called for the inclusion of high and low performers rather than a random sample of all buildings; however, the site visits provided a sense of how well the boiler capacity was recorded. A summary of findings for this analysis is provided in Table 9.

Table 9. Summary of boiler controls field observation findings by program input

Measure	Calculation Input	Findings
Boiler Controls	Boiler input BTUs (BTU_{in})	Not likely to be a major source of error, typically recorded accurately.
	Equivalent full load hours ($EFLH_{heating}$)	Not evaluated in this section
	Energy savings factor (ESF)	Not evaluated in this section

Review of Relevant Studies

We reviewed the existing literature to identify studies relevant to the findings of our evaluation. The coincidence factors found in the metering in this study were compared to those found in a previous study of the multifamily sector in the Brooklyn and Queens area. Our study covered sites dispersed throughout Manhattan, Brooklyn, Queens, and the Bronx. The space type definitions also differed slightly between the two studies. A comparison of the space type categories is provided in Table 10.

Table 7. Comparison of space type definitions

Step	Definition in this Study	Definitions in Previous Study
Community Rooms	Community rooms, libraries, gyms, auditoriums, or computer rooms	Bathrooms, laundry rooms, daycare, locker rooms, libraries, conference rooms, classrooms, open areas, pantries, kitchens, and stores
Other	Bathrooms, laundry rooms, staff conference rooms, or staff locker rooms	Bulkheads

A summary of findings from the review is provided in Table 11.

Table 8. Summary of review of similar studies by program input

Measure	Calculation Input	Findings
Lighting	Baseline watts ($W_{baseline}$)	Not evaluated in this section
	Installed units (Units)	
	Efficient case watts (W_{ee})	
	Hours of use ($hrs_{operating}$)	Coincidence factor and hours of use are similar to the results in our evaluation

For boiler controls, the ESF estimated in the billing analysis was compared to a previous study of the multifamily sector, leading to the findings listed below.

- Current evaluation: *ex ante* savings of 7,270 therms per building, *ex post* savings of 2,870 therms per building, and a realization rate of 39%.
- Previous impact evaluation: *ex ante* savings of 2,879 therms per building, *ex post* savings of 2,586 therms per building, and a realization rate of 90%.

Backing out the ESF from the previous study suggests an ESF of about 10%. A summary of findings for this analysis is provided in Table 12.

Table 9. Summary of boiler controls field observation findings by program input

Measure	Calculation Input	Findings
Boiler Controls	Boiler input BTUs (BTU_{in})	Not evaluated in this section
	Equivalent full load hours ($EFLH_{heating}$)	Not evaluated in this section
	Energy Savings Factor (ESF)	While the *ex post* savings are 11% higher in this study than in the previous evaluation, the *ex ante* savings are 153% higher. This further indicates that the ESF has been overstated in the *ex ante* savings and should be reduced.

Conclusions

The results of the lighting inputs analysis indicated that lighting realization rates may be improved by more rigorous determination of baseline wattage. Also, hours of use were somewhat overstated in comparison to the metering results. The high-level results and the evidence that supports these conclusions are summarized in the bullets below.

- The low lighting realization rate was primarily due to overstatement of the baseline wattages. Supporting evidence for this conclusion included the following:
  - 43% of buildings have lighting pre-use higher than the billing analysis
  - Average non-lighting use in the post-period is almost 3 times higher than in the pre-period
  - The sensitivity analysis shows baseline wattage has a much larger impact on savings than efficient wattage
  - Meter data indicated hours of use revisions alone were not enough to explain the 40% realization rate.
- A secondary factor was the hours of use. The meter data indicates hours of use are 22% lower than the interior hours used in the TRM.

These findings suggest that improving the accuracy of the baseline wattages either through more rigorous on-site verification and quality assurance/quality control or establishing a prescriptive baseline would increase the realization rate.

The natural gas consumption analysis ESF is too high and the full load hours were somewhat overstated. The high-level results and the evidence that supports these conclusions are summarized in the bullets below.

- The low boiler controls realization rate is primarily due to the overstatement of the ESF. The billing analysis indicates that the ESF should be about 10%, whereas 22% was used in the TRM.
- A secondary factor is the equivalent full load hours in the TRM are 13% higher than the billing analysis indicates.

The evaluators recommended reducing the ESF and proposed updated numbers for the equivalent full load hours based on the consumption analysis.

This approach of parsing out the savings calculations and assessing the impact of each component on the realization rate provided a solid basis for determining the main drivers of the realization rate and providing actionable recommendations for improving the program savings. The same type of approach could be applied to other retrofit measures in the residential and commercial and industrial (C&I) sectors, either at the program level as discussed above or at the project-level for more complex C&I applications. For example, an industrial process measure evaluated through a consumption analysis could be separated into the pre-periods and post-periods. The program assumptions used to estimate the baseline consumption could be compared to actual baseline bills to determine whether discrepancies are due to incorrect assumptions about the operation of the equipment or due to performance issues with the new equipment. The underlying principles described above can be used to tailor a project-specific analysis as needed.

References

Energy and Resource Services (ERS). 2014. "Con Edison EEPS Programs – Impact Evaluation of Multifamily Electric and Gas Programs."

Energy and Resource Services (ERS). 2016. "BQDM Metering and Market Characterization Effort."

New York State Joint Utilities. New York State Technical Research Manual (TRM) Version 4 for Program Year 2017. Issue Date - April 29, 2016. Effective Date – January 1, 2017

West Hill Energy and Computing, Inc. Multifamily Program – Program Year 2017 Impact Evaluation. Prepared for Consolidated Edison, New York City, New York. January 7, 2020.