NAVIGANT

Itron

230 Horizon Drive,
Suite 101-B
Verona, WI 53593

Memorandum

To: Erin Daughton, ComEd

From: Jeff Erickson, Randy Gunn, Rob Neumann - Navigant Consulting
Jennifer Fagan, Kumar Chittory – Itron

CC: Milos Stefanovic, ComEd
Jennifer Morris, ICC Staff

Date: July 26, 2018

Re: PY8 and PY9 ComEd Custom Rebates Program Recommended NTGR Updates

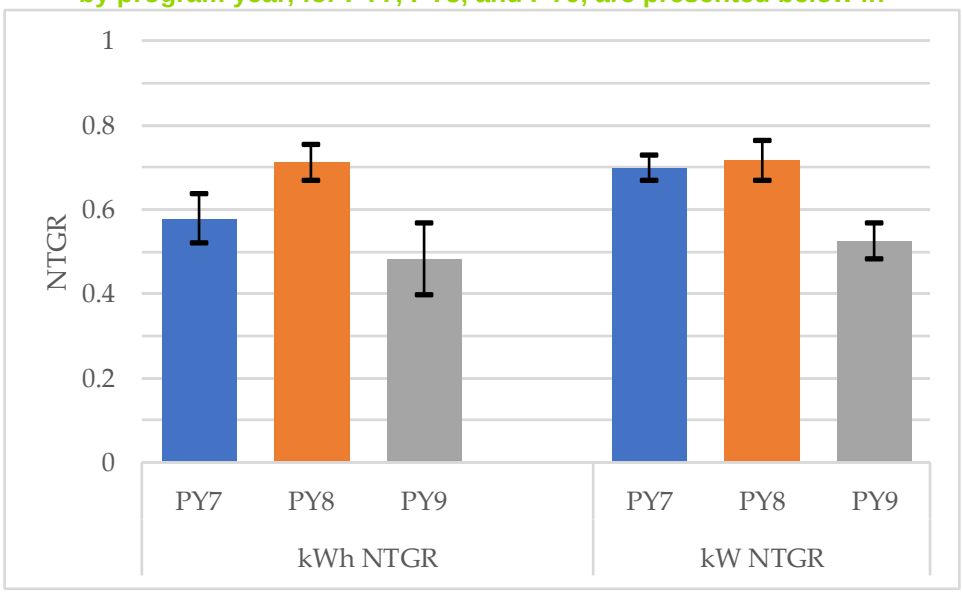
1. INTRODUCTION

This memorandum presents the *Evaluation Research*[1] PY8 and PY9 net-to-gross ratio (NTGR) estimates for ComEd's Custom Rebates program. Regarding PY8, note that net-to-gross (NTG) interviews were completed immediately following the end of the program year, but analysis of the PY8 data was postponed until the conclusion of the PY9 evaluation. Thus, this memo reports findings for PY8, PY9 and combined PY8/PY9 NTGR results.

[1] It should be noted that the NTGR estimates presented here are the evaluation verified estimates (based on the PY9 participating customer and non-participating retailer surveys).

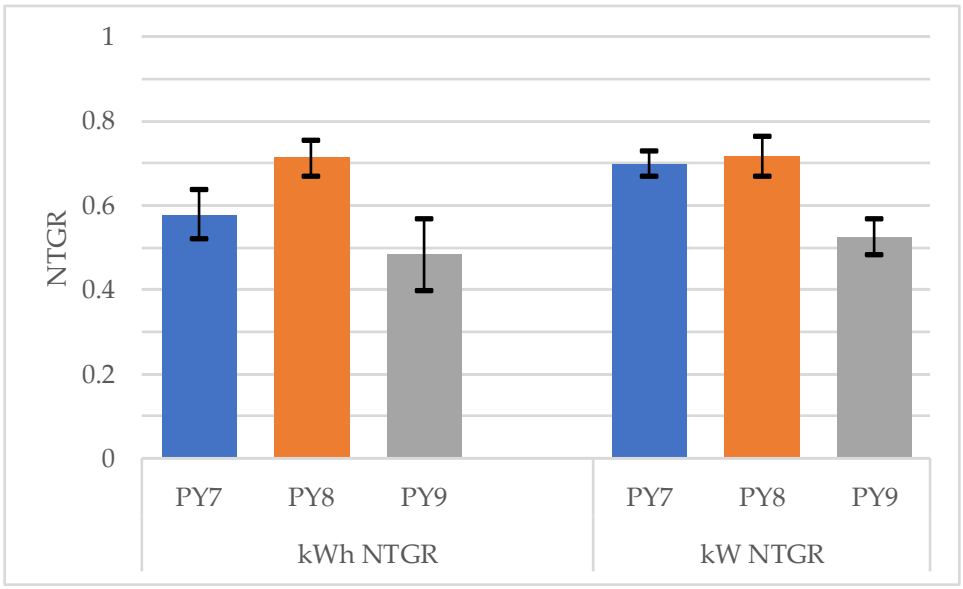
2. SUMMARY OF FINDINGS

The evaluation research findings energy and demand-weighted net-to-gross ratio (NTGR) by program year, for PY7, PY8, and PY9, are presented below in



. The PY8 evaluated kWh NTGR for Custom projects of 0.71 is higher relative to the PY7 NTGR of 0.58; however, the PY9 evaluated kWh NTGR of 0.45 is lower.

Figure 2-1. Evaluated Custom NTGR by Program Year with 90% Confidence Intervals



A combined PY8 and PY9 NTGR was also calculated. This value was determined using savings weighted NTGRs from PY8 and PY9 and computing a weighted average value. The combined PY8 and PY9 value of 0.559 is similar to the PY7 NTGR of 0.57. *It is recommended that the combined PY8 and PY9 value of 0.559 be used to compute*

program-verified savings for CY2018 projects going forward. This combined value is being recommended because it is based on a larger and more robust sample representing two-years' worth of projects, and it reflects the latest available information from the evaluation effort.

3. EVALUATION RESEARCH NET IMPACT FINDINGS

3.1 Free-Ridership

The program's net-to-gross ratio is equal to one minus the free ridership rate plus the spillover rate. The EM&V team calculated PY8, PY9 and combined PY8/9 net-to-gross ratios for future consideration using a self-report method which relies on the results of surveys with PY8 and PY9 participants.

3.1.1 NTG Algorithm Specifications

The PY8 and PY9 NTGR calculations were based on the NTG algorithms specified in the Illinois TRM version 6.0. Approval to use version 6.0 was provided by the Illinois Stakeholder Advisory Group and Illinois Commerce Commission staff via an email seeking permission dated April 2, 2018 and their lack of objections by April 16, 2018, which was interpreted as consensus. The NTG protocols in version 6.0 were developed by the Illinois Net-to-Gross Working Group, in their deliberations during the summer and fall of 2017.

The protocols provide two options for combining the three scores. These two options use different specifications to account for the impact of the program on project timing (referred to as "deferred free ridership). Evaluators are to calculate free ridership using both options, and to select one option for purposes of calculating the annual incremental energy savings for comparing to the legislated goal.

The evaluation team's preferred algorithm specification **is Core Free Ridership Algorithm 1**, shown graphically below (Figure 3-1). The majority of NTG findings discussed below are based on this version. The second option, Core Free Ridership Algorithm 2 (Figure 3-2) has also been analyzed, and those findings will be presented as a sensitivity case later in this memo. The rationale for selecting Algorithm 1 over Algorithm 2 is that Algorithm 1 provides for equal weighting of each of the three scores, which represent different ways of determining program influence. In contrast, Algorithm 2 applies a 50% weight to the program's effect on the timing of the project, which we believe is too high. Such a high weighting essentially discounts the effect of the other factors influencing program influence, which in our view is inappropriate.

Figure 3-1. Core Free Ridership Algorithm 1

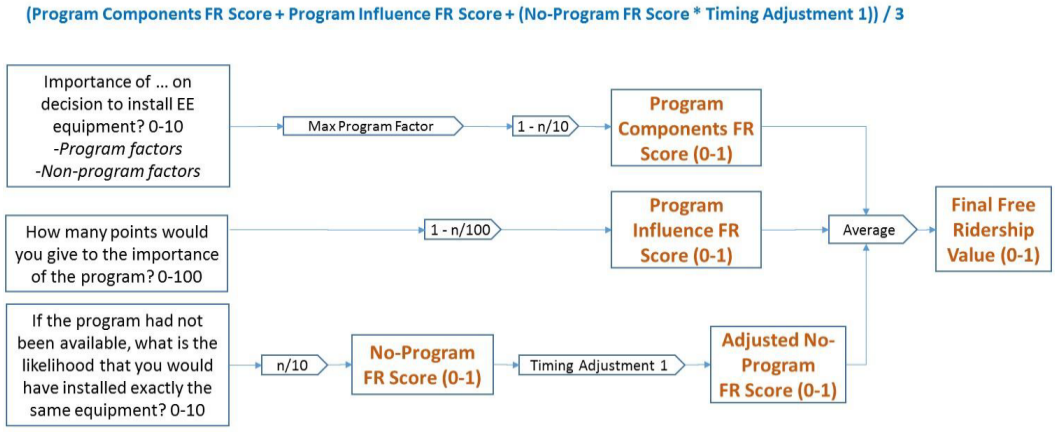
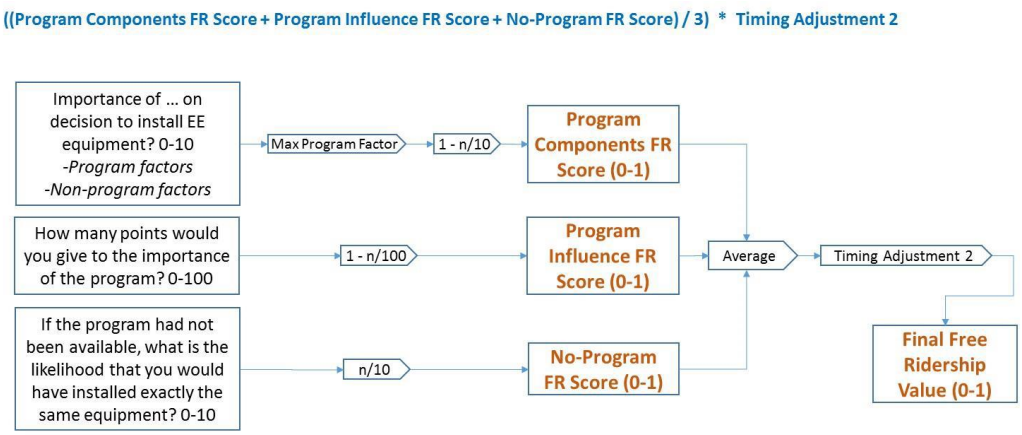


Figure 3-2: Core Free Ridership Algorithm 2



3.1.2 NTGR Calculation

The calculation of both the free ridership rate and each project's net-to-gross ratio (NTGR) is a multi-step process. Responses from the telephone survey are used directly to calculate a timing and selection score, a program influence score and a no-program score for each project (as outlined in Table 3-1 below for both versions of the NTGR algorithm). These three scores can take values of 0 to 10 where a lower score indicates a higher level of free-ridership. The calculation then averages those three scores and incorporates spillover findings to come up with a project-level net-to-gross ratio.

Table 3-1. Net-to-Gross Scoring Algorithms for the PY8 and PY9 Custom Program[2]

Scoring Element	Algorithm 1 Calculation	Algorithm 2 Calculation
Timing and Selection Score. The maximum self-reported score (on a 0 to 10 scale of importance) for the following program elements: A. Availability of the program incentive B. Technical assistance from program staff C. Recommendation from program staff D. Information from program marketing materials E. Endorsement or recommendation by account rep F. Recommendation from an equipment vendor	Maximum of A, B, C, D, E, and F	Maximum of A, B, C, D, E, and F
Program Influence score. From a Total of 10 points, the self-reported number of points assigned to the importance of the Program in their decision to implement the <PROJECT> (as versus other non-program factors.	Points awarded to the program. Reduce by half if decision made BEFORE learning about rebate eligibility	Points awarded to the program. Reduce by half if decision made BEFORE learning about rebate eligibility
No-Program score. If the Program had not been available, the self-reported likelihood (on a 0 to 10 scale, where 0 is "Not at all likely" and 10 is "Extremely likely") that they would have installed the same PROJECT.	Linear adjustment to self-reported No Program Likelihood Score and 10 (maximum score based on deferred installation 48 months or more later).	Self-reported No Program Likelihood.Score.
Timing Adjustment. Timing credit provided for deferred installation absent the Program. Linear adjustment with gradually increasing credit value for each year of deferral of 25% for one year,50% for two years, 75% for three years and 100% for four years or more.	Incorporated into No Program score.	Applied to the average of the Timing and Selection, Program Influence and No-Program scores
Project-level Free-ridership (ranges from 0.00 to 1.00)	1 minus Sum of scores (Timing and Selection, Program Influence, No-Program)/30	1 minus the average of the Timing and Selection, Program Influence and No-Program scores, adjusted for Timing
PY8 and PY9 Project level Net-to-Gross Ratio (ranges from 0.00 to 1.00)	1 minus Project level Free-ridership	1 minus Project level Free-ridership

3.1.3 NTG Sample Design and Completed Surveys

During both PY8 and PY9, the original sample design consisted of 20 sample points that corresponded to and overlapped with the gross impact M&V sample. However, given customer willingness to participate and other factors, the final net samples did not fully

[2] Based on the NTG algorithm specifications in TRM v.6.0 Attachment A (Illinois Statewide Net-to-Gross Methodologies)

match the gross sample. During PY8, telephone surveys were conducted for two waves of sample, yielding a total of 16 completed interviews. In PY9, surveys were completed for three waves of sample, and 19 interviews were completed. However, one of the PY9 survey completes was dropped from the analysis frame because the project was not completed. Therefore, the PY9 findings are based on a total of 18 completed interviews to support the calculation of the net-to-gross ratio calculation. The 16 PY8 and 18 PY9 NTG completes represent a subset of the 20 gross M&V sample points in each year (i.e. they are completely overlapping).

Table 3-2 and Table 3-3 below summarize the number of completed telephone surveys in each year, and the percent of ex ante kWh claims represented. The surveys completed represent 38 percent and 48 percent of ex ante kWh claims in PY8 and PY9, respectively.

Table 3-2: Profile of the PY8 Participant Survey Net-to-Gross Sample by Strata

Program Population Summary				NTG Interviews Completed		
Sampling Strata	Number of Records (N)	Ex Ante kWh Impact Claimed	kWh Weights by Strata	N	% of kWh	% of Population Impacts Surveyed
1	4	9,412,193	0.28	3	5,655,362	60%
2	14	12,928,025	0.38	7	6,317,952	49%
3	57	11,608,685	0.34	6	841,698	7%
TOTAL CUSTOM	75	33,948,903	-	16	12,815,012	38%

Table 3-3. Profile of the PY9 Participant Survey Net-to-Gross Sample by Strata

Program Population Summary				NTG Interviews Completed		
Sampling Strata	Number of Records (N)	Ex Ante kWh Impact Claimed	kWh Weights by Strata	N	% of kWh	% of Population Impacts Surveyed
1	4	11,309,438	0.23	4	11,309,438	100%
2	15	18,863,237	0.39	8	10,043,736	53%
3	98	18,274,524	0.38	6	1,672,332	9%
TOTAL CUSTOM	117	48,447,199	-	18	23,025,506	48%

3.1.4 Weighted NTG Results Based on Core Free Ridership Algorithm 1 (Preferred specification)

Weighted results are presented in this section for each sampling size stratum, and for the program overall. To produce an estimate of the net-to-gross ratio (NTGR), the individual NTGRs for each of the projects in the sample were weighted by the size of the ex ante

savings estimates (savings) associated with the project, and the proportion of the total sampling domain savings represented by each sampling stratum. NTGR results are weighted by ex ante kWh.

The separate ratio estimation technique was used to estimate NTGR for the program. The separate ratio estimation technique follows the steps outlined in the California Evaluation Framework. The standard error was used to estimate the error bound around the estimate of verified evaluation NTGR.

Spillover effects were examined in this evaluation and their magnitude was found to be zero, as discussed below in the spillover section.

3.1.4.1 PY8 NTG Results

The PY8 program level NTGR, along with precision estimates, is shown below in Table 3-4. The overall program NTGR for PY8 is 0.71, which represents a significant improvement over PY7. By strata, the mean energy NTGR values are 0.81 for stratum 1 (large sized projects), 0.73 for stratum 2 (medium sized projects), and 0.62 for stratum 3 (small sized projects) which indicates the free-ridership level for the largest sized projects (stratum 1) is lower than the free-ridership of the smaller project sizes. The strong results for stratum 1 projects are a key factor in the improved PY8 NTG results.

Table 3-4. PY8 MWh NTG Ratio and Relative Precision at 90% Confidence Level

Sampling Strata	Relative Precision ± %	Low	Mean	High
1	1%	0.80	0.81	0.82
2	12%	0.64	0.73	0.82
3	13%	0.54	0.62	0.70
Custom PY8	**6%**	**0.67**	**0.71**	**0.76**

By stratum, highlights include the following:

- For all three of the stratum 1 interviews completed, the NTGRs ranged from 0.80 to 0.83, indicating strong program influence. In all cases, the customer knew about the program well ahead of their decision, the program rebate helped them to meet key economic metrics for investment, and their equipment had significant remaining life, giving them discretion over whether to install the rebated measure or to retain their old equipment.
- For stratum 2 projects, NTGRs ranged from 0.10 to 1.00 indicating wide variation. For those projects with the highest NTGRs, the program rebate was a key factor in enabling them to meet their company's required economic threshold for making this type of investment. One customer noted that energy prices in Illinois are not as high as other states, and therefore, energy efficiency projects have longer payback periods, which makes it harder for them to justify the investment absent any incentive. Projects with the lowest NTGRs had the following common characteristics – measures had significant non-energy benefits, were routinely installed at all their other locations, or were selected to replace aging equipment.

- NTGRs for stratum 3 projects ranged from 0.37 to 0.75, indicating a medium level of free ridership. The program rebate was rated highly for many projects, for moving the project payback to an acceptable level, and/or helping to pay for some of the up-front costs for more expensive energy efficient equipment.

3.1.4.2 PY9 NTG Results

The PY9 program level NTGR, along with precision estimates, is shown below in in Table 3-5. The program-level PY9 mean energy NTGR averaged 0.45. In general, PY9 mean energy NTGR values are much lower than in PY8 and somewhat lower than PY7. NTGR values for the three Custom sampling strata are 0.50 for stratum 1 (large sized projects), 0.37 for stratum 2 (medium sized projects), and 0.51 for stratum 3 (small sized projects) which generally indicates the free-ridership level for the largest sized projects (stratums 1 and 2) is somewhat higher than the free-ridership of the smaller project sizes.

Table 3-5. PY9 kWh NTG Ratio and Relative Precision at 90% Confidence Level

Sampling Strata	Relative Precision ± %	Low	Mean	High
1	0%	0.50	0.50	0.50
2	27%	0.27	0.37	0.47
3	44%	0.29	0.51	0.73
Custom PY9	**21%**	**0.36**	**0.45**	**0.54**

Stratum-level highlights include the following:

- For the largest stratum 1 projects, NTGRs varied widely and ranged from 0.27 to 0.73. The NTGR for this stratum averaged 0.50. Circumstances surrounding these customers' decisions to install energy efficient equipment were very different.
    - For those projects with the lowest NTGRs, non-energy benefits were a strong motivation. One customer was highly motivated to pursue LEED Silver certification to support their firm's marketing strategy. Energy cost reduction was only a secondary objective. This is a good example of a project motivated by a non-energy benefit. Another customer needed to improve the lighting in their work environment, and LEDs provided the perfect solution. Absent the program they would have installed the same measure at the same time they did.
    - For the project with the highest NTGR, the program audit and feasibility study and rebate played a large role in their decision to do the project at this time. Without the program they would have eventually installed the same equipment, some 24 months later. Thus, the program had a strong acceleration effect.
- For stratum 2 projects, NTGRs ranged from 0.00 to 0.77 with a mean value of 0.41. For those projects with the highest NTGRs, the program features, including the audit and feasibility study, the rebate and the assistance provided by program staff were key decision influences. Projects with lower NTGRs were characterized by decisions that were largely dictated by their standard corporate policies or practices such as one firm's decision to incorporate LEDs into their

prototype store designs several years ago and another firm's policy to use on-load transformers at their other facilities around the U.S.

- Across the smallest stratum 3 projects, NTGRs ranged from 0.05 to 0.90, and averaged 0.51, indicating a medium level of free ridership. It is interesting to note that project NTGRs were clustered around medium-high values (0.63 to 0.90, four projects), and extremely low values (0.05 to 0.19, two projects). The program rebate was the number one driver for the high NTGR projects. For the low NTGR projects, prime influences included the urgent need to replace old equipment, and the need to meet corporate sustainability goals. These firms would have installed the same equipment at about the same time absent the program.

The PY8 and PY9 project-specific NTGRs are plotted in Figure 3-3 and Figure 3-4, respectively. Each plot point in the figure represents a sampled project. The plot points are grouped by strata. The blue and lavender horizontal lines denote the stratum-level energy and demand weighted NTGRs, respectively. Note that strata 1 and 2 were combined for the demand weighted NTGR, as there was only a single stratum 1 project with demand savings.

Figure 3-3: PY8 Sample NTGR by Stratum

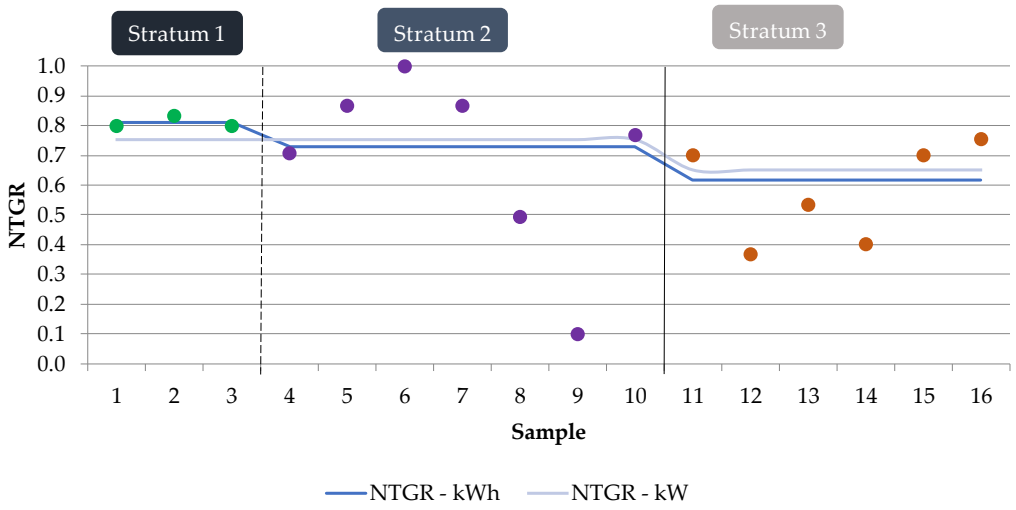
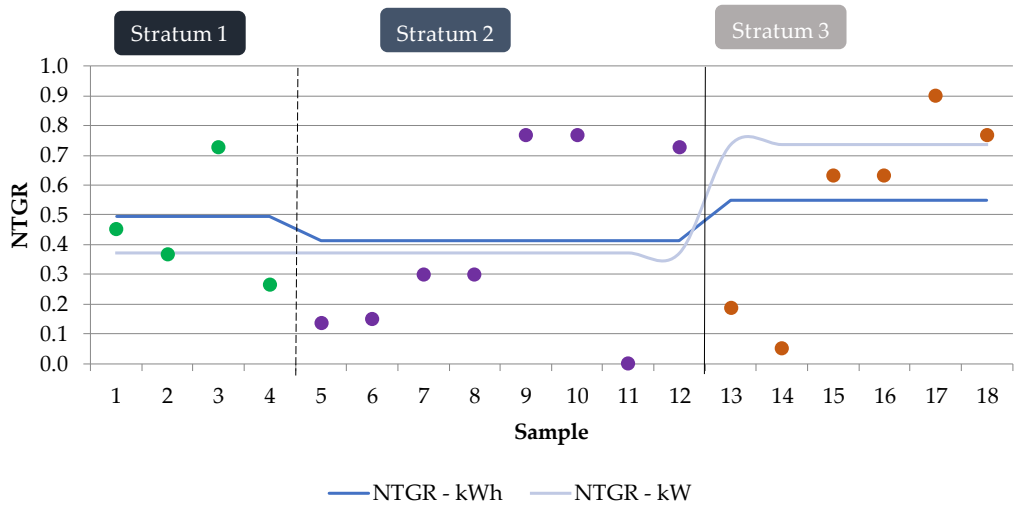
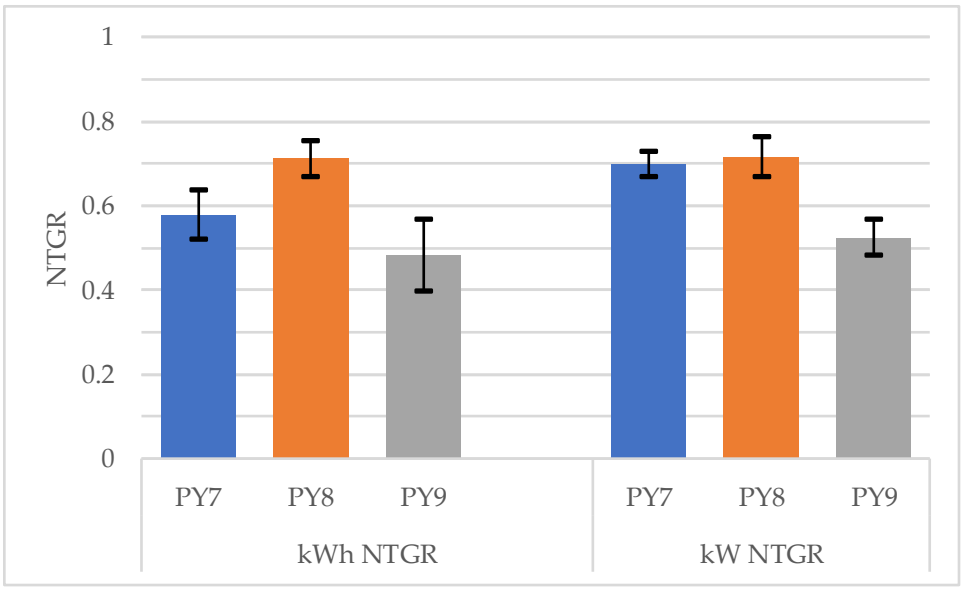


Figure 3-4: PY9 Sample NTGR by Stratum



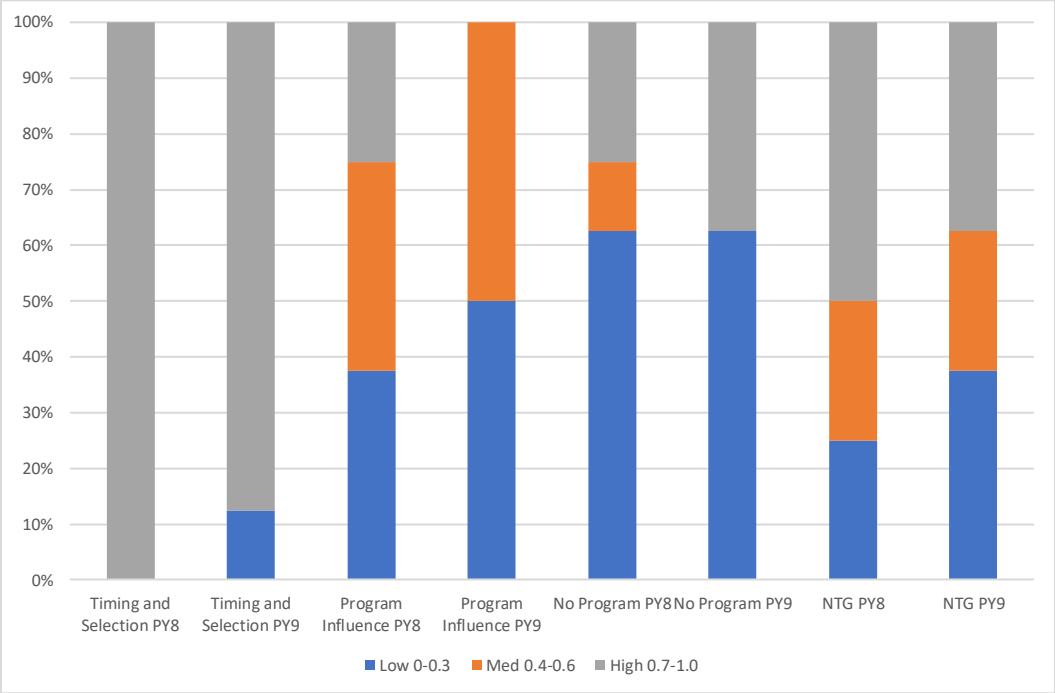
The evaluation research findings energy and demand-weighted NTGR by program year, for PY7, PY8, and PY9, are presented below in Figure 3-5. The PY8 evaluated kWh NTGR for Custom projects of 0.71 is improved relative to the PY7 NTGR of 0.58; however, the PY9 evaluated kWh NTGR of 0.45 is lower. Note that the 90 percent confidence interval (CI) of the PY9 kWh NTGR does overlap with the CI of the PY7 kWh NTGR, indicating that the PY9 kWh NTGR is not statistically different from the PY7 value.

Figure 3-5. Evaluated NTGR by Program Year with 90% Confidence Intervals



A breakdown of the NTGR by the three component scores is shown in Figure 3-6. The timing and selection score reflects the importance of various program and program-related elements in the customer's decision and timing of the decision in selecting specific program measures. The program influence score reflects the relative degree of influence the program had on the customer's decision to install the specified measures versus non-program factors. The no-program score captures the likelihood of various actions the customer might have taken at this time and in the future if the program had not been available.

Figure 3-6. NTGR Level by Component Scores



A scan of the PY8 vs. PY9 bars provides additional insight into a key causal factor for the drop in the NTGR value between PY8 and PY9. For all three scores, the concentration of High values is moderately to significantly higher in PY8 than PY9. As a result, for the overall NTGR, the share of High scores in PY8 exceeds that in PY9 by a wide margin. Stratum-level causal factors leading to these results were discussed previously. In general, PY9 projects were characterized by program-related factors that were either less important or not applicable to the final decisions to do the project.

3.1.4.3 Combined PY8 and PY9 Results

A combined PY8 and PY9 NTGR was also calculated. This value was determined using savings weighted NTGRs from PY8 and PY9 and computing a weighted average value. The combined PY8 and PY9 value of 0.559 is similar to the PY7 NTGR of 0.57. *It is recommended that the combined PY8 and PY9 value of 0.559 be used to compute program-verified savings for CY2018 projects going forward.*

Table 3-6. Combined PY8 and PY9 MWh NTG Ratio

Year	N	kWh	Weight	NTGR	NTG SE
PY8	75	33,948,903	41%	0.713	3%
PY9	117	48,447,199	59%	0.451	7%
Custom PY8/PY9	192	82,396,102	**100%**	**0.559**	**6%**

3.1.5 Sensitivity Case - Weighted NTG Results Based on Core Free Ridership Algorithm 2

The evaluation team also performed a sensitivity analysis based on Core Free Ridership Algorithm 2. NTG results are slightly higher due to the greater weight given to the acceleration (i.e. timing) effect of the program. This algorithm varies from Algorithm 1 with respect to how it treats the effect of timing in the calculation of the NTGR. Algorithm 1 adjusts for Timing within the No-Program score, then averages the 3 scores. Algorithm 2 determines the No-Program Score without a Timing adjustment, averages the 3 scores, then applies a Timing adjustment factor to the 3-score average, based on the formula below:

> **Timing Adjustment Factor (Free Ridership Score)** as equal to:
> 1 - ((Number of Months Expedited - 6)/42) *((10 - Likelihood of Implementing within One Year)/10)

3.1.5.1 NTG Algorithm 2 –PY8 Weighted NTG Results

The PY8 program-level NTGR for version 2 of the algorithm, along with precision estimates, is shown below in Table 3-7 . The overall program NTGR for PY8 is 0.79, which is somewhat higher than the Algorithm 1 value of 0.71. This timing “bump” is due to reports by several decisionmakers reported that the program accelerated the installation of their installed project compared to if there had been no program and incentive.

Table 3-7. Algorithm 2 PY8 MWh NTG Ratio and Relative Precision at 90% Confidence Level

Sampling Strata	Relative Precision ± %	Low	Mean	High
1	1%	0.95	0.96	0.97
2	13%	0.68	0.78	0.88
3	13%	0.58	0.67	0.76
Custom PY8 Alg 2	**6%**	**0.74**	**0.79**	**0.84**

3.1.5.2 NTG Algorithm 2 –PY9 Weighted NTG Results

For this second version of the NTG algorithm, the PY9 program level NTGR, along with precision estimates, is shown below in Table 3-8. The program-level PY9 mean energy NTGR average of 0.48 is slightly higher than the NTGR of 0.45 under NTG Algorithm 1. Again, this small bump is due to the heavier weight given to the acceleration (timing) effect under Algorithm 2 as versus Algorithm 1.

Table 3-8. Algorithm 2 PY9 kWh NTG Ratio and Relative Precision at 90% Confidence Level

Sampling Strata	Relative Precision ± %	Low	Mean	High
1	0%	0.50	0.50	0.50
2	38%	0.21	0.34	0.47
3	46%	0.34	0.62	0.91

Sampling Strata	Relative Precision ± %	Low	Mean	High
Custom PY9 – Alg 2	**25%**	**0.36**	**0.48**	**0.60**

Figure 3-7 (PY8) and Figure 3-9 (PY9) below compare the evaluated NTGRs for Algorithms 1 and 2 for each sampling stratum. For PY8, when compared to Algorithm 1, the mean energy NTGR values are 0.96 vs. 0.81 for stratum 1 (large sized projects), 0.76 vs. 0.73 for stratum 2 (medium sized projects), and 0.67 vs. 0.62 for stratum 3 (small sized projects. The improved results for stratum 1 projects are the sole reason for the increase in the average program NTGR.

In PY9, when compared to Algorithm 1, NTGR values for the three Custom sampling strata are the same 0.50 for stratum 1, slightly lower (0.34) for stratum 2, and much higher (0.62 vs. 0.51) for stratum 3. For both PY8 and PY9, these results indicate that the free-ridership level based on Algorithm 2 for the largest sized projects (stratums 1 and 2) is higher than the free-ridership of the smaller project sizes.

Figure 3-7. Comparison of PY8 Evaluated NTGR by NTG Algorithm and Stratum

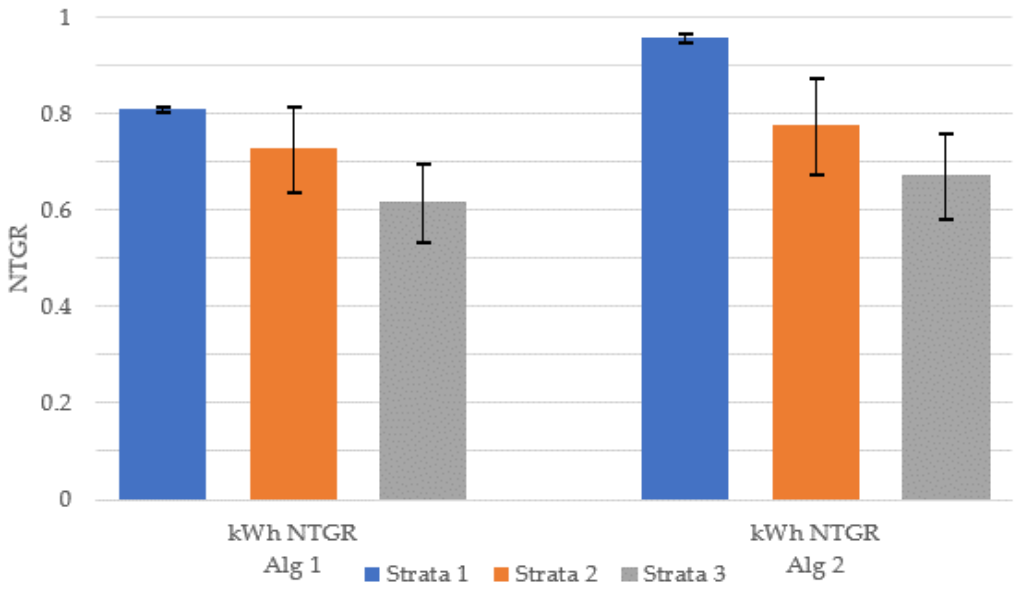
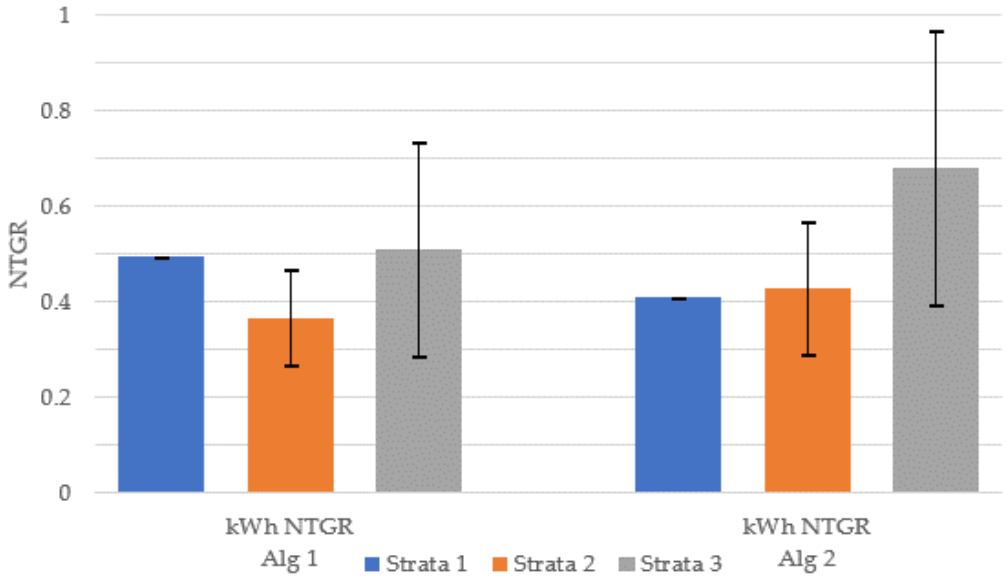


Figure 3-8. Comparison of PY9 Evaluated NTGR by NTG Algorithm and Stratum



3.1.6 Procedures to Reduce Free Ridership

Without a doubt, the large non-residential market is perhaps the most challenging to address in terms of the size and sophistication of end-use customers and suppliers, and the complexity of end-user projects. As a result, a certain amount of free ridership is to be expected in this market. Despite these challenges, there are several different strategies available to ComEd to adjust program design elements and implementation procedures to reduce free ridership. These recommendations are as follows:

Recommendation: Adopt procedures to limit or exclude known free riders.

The best way to accomplish this is to conduct screening for high free ridership on a project-by-project basis. In cases where it is found, the program implementer should continue and expand their current pre-approval process to provide more explicit consideration and re-formulation of projects already planned for completion by the customer. The NTGRs for the Custom Program have fluctuated between 0.56 and 0.72 since the program began, and are in line with similar programs offered elsewhere in the U.S. However, the decline in the PY9 NTGR to 0.45 suggests that a more aggressive approach is warranted since the NTG ratios indicate significant free ridership is still present.

Another path is for the program to set the standard for incentive eligibility higher across-the-board so that all such projects will need to meet a higher standard to qualify. Note that **none** of these options equates to rejecting a customer for energy efficiency funding. Instead, the concept is to “upsell” the customer to an energy efficiency project that they weren’t already planning to do on their own.

3.1.6.1 Screening out Free Riders

One way to assess the rate of free ridership likely on a given project is to critically examine the key reasons behind the project **before** the incentive is approved. For example:

- Has the project already been included in the capital or operating budget? Has the equipment already been ordered or installed?
- Is the measure one that the company or other comparable companies in the same industry or segment routinely installs as a standard practice? Is the measure installed in other locations, without co-funding by incentives? Is the measure potentially Industry Standard Practice?
- Is the project being done, in part, to comply with regulatory mandates (such as environmental regulations)?
- Are the project economics already compelling without incentives? Is the rebate large enough to make a difference in whether the project is implemented?
- Is the company in a market segment that is ahead of the curve on energy efficiency technology installations? Is it part of a national chain that already has a corporate policy to install the proposed technology?
- Does the proposed measure have substantial non-energy impacts? Is it largely being considered for non-energy reasons (such as improved quality or increased production)?
- Is the project payback quite short even without the incentive?

By conducting a brief interview regarding these issues before the incentive is approved, ComEd can better assess the likely degree of free ridership and may be able to then decide if the project should be excluded or substantially re-scoped to a higher efficiency level.

3.1.7 Spillover

Spillover effects were addressed in the PY8 and PY9 evaluations, based on responses to a battery of spillover questions in the telephone survey. Detailed spillover-related findings from the surveys are reported in Table 3-9 below.

Table 3-9. Detailed Spillover-Related Findings for PY8 and PY9

	Evidence of Spillover	
Spillover Question	**PY8**	**PY9**
Since receiving an incentive for the project we just discussed, did you implement any ADDITIONAL energy efficiency measures at this facility or at your other facilities within ComEd's service territory that did NOT receive incentives through any utility or government program?	Of the 16 surveyed customers that responded, one (6%) implemented an additional measure without receiving an incentive. This respondent implemented one energy efficiency measure.	Of the 18 surveyed customers that responded, one implemented an additional measure without receiving an incentive.
What type of energy efficiency measure was installed without an incentive?	Energy efficient chiller, rating of 130,000 Btu/hr, efficiency rating 0.96. (did not provide tons of capacity or COP)	Energy efficient 60-watt lamps to replace 150-watt lamps (n=90)
On a scale of 0 to 10, where 0 means "not at all significant" and 10 means "extremely significant," how significant was your experience in the ComEd program in your decision to implement this energy efficiency measure?	Significance rating of 10	Don't know
If you had not participated in the ComEd program, how likely is it that your organization would still have implemented this measure? Use a 0 to 10, scale where 0 means you definitely would NOT have implemented this measure and 10 means you definitely WOULD have implemented this measure?	Likelihood rating of 0 – would have installed on their own 2 years later	Don't know
Why did you purchase this energy efficiency measure without the financial assistance available through the ComEd's program?	They applied too late for an incentive	The energy savings and payback were sufficient. Also, they needed to address immediate issues.

Only one respondent each in PY8 and PY9 installed a measure with potential savings that could be attributed to calculation of the spillover ratio. Unfortunately, the PY8 respondent did not know the tonnage of the installed equipment, which is critical to support a savings estimate calculation. The PY9 respondent did not provide an importance score to link (attribute) their decision back to the ComEd program. Therefore, no spillover is attributable to either the PY8 or PY9 program.

3.1.8 Cronbach's Alpha Results

Cronbach's Alpha is a measure of internal consistency or reliability. It is used to assess how closely related a set of items are as a group. In this memo, Cronbach's Alpha is used to assess how closely related the items going into the NTG score are to each other. In general, the higher the measured Cronbach's Alpha value, the more consistent and

reliable are the results. However, given the small number of items (i.e., the three scores) being considered in this application of Cronbach's Alpha, a high alpha value is not expected. Realistically, Alpha values ranging from 0.4 to 0.6 are considered an acceptable measure of reliability for this analysis given the small number of items being analyzed.

We used the Standardized Cronbach's Alpha calculation as specified below:

$$\alpha = \frac{N \cdot \bar{r}}{1 + (N - 1) \cdot \bar{r}}$$

Where:

N = the number of items
$\bar{r}$ = the average correlation

We calculated the Cronbach Alpha for each program year, for the two algorithm variations discussed previously.

Figure 3-9 and Figure 3-10 below present the Cronbach's Alpha and the 90% confidence intervals for the two NTGR algorithm variations for the PY8 and PY9 Custom Program, respectively. Overall Cronbach's Alpha values for PY8 were quite low, 0.37 (Algorithm 1) and 0.50 (Algorithm 2). In PY9, Alpha values were significantly higher, 0.88 for both Algorithms.

Note that the confidence intervals around Alpha are expected to be quite large due to the relatively small sample sizes. In PY8, the results show wide confidence bands and low Alpha values, particularly for Algorithm 1, due to the relatively small sample size and diverse project-level NTGR results implying a lack of inter-item correlations. In PY9 however, the Alpha results for the two algorithm variants are identical. Most likely this is because the formula leads to higher values when the inter-item correlations are higher (as was the case in PY9).

Figure 3-9: PY8 Custom Program Cronbach's Alpha and 90% Confidence Intervals for the Two Algorithm Variations (N=16)

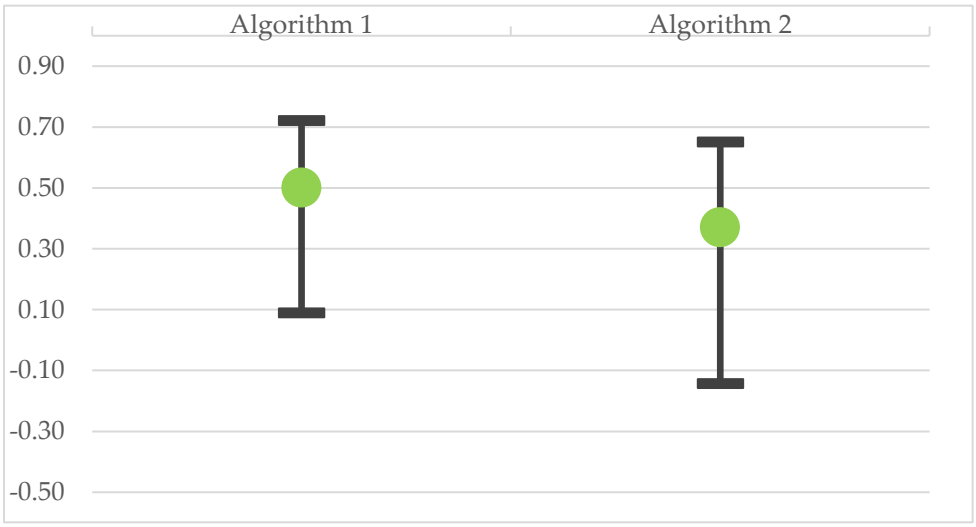


Figure 3-10: PY9 Custom Program Cronbach's Alpha and 90% Confidence Intervals for the Two Algorithm Variations (N=18)