FINAL REPORT

Impact Evaluation of PY2019 Custom Gas Installations in Rhode Island

National Grid

Date: December 8, 2021

Table of contents

List of figures

List of tables

List of acronyms used in this report

CDA	comprehensive design approach
C&I	commercial and industrial
CI	confidence interval
EMS	energy monitoring system
HVAC	heating, ventilation, and air-conditioning
ISP	industry standard practice
M&V	measurement and verification
MA	Massachusetts
MBSS	model-based statistical sampling
PA	program administrator
PY	program year
PY2016	program year 2016
PY2017	program year 2017
PY2018	program year 2018
PY2019	program year 2019
RI	Rhode Island
RR	realization rate
SEMP	strategic energy management partnership
TMY3	typical meteorological year 3

1 EXECUTIVE SUMMARY

This Executive Summary provides a high-level review of the results for the Rhode Island (RI) Commercial and Industrial (C&I) Impact Evaluation of Program Year (PY) 2019 Custom Gas Installations. This section states the study objectives; summarizes the evaluation approach; and presents key findings, conclusions, and recommendations. The scope of work of this impact evaluation covered the PY2019 Custom Gas impact category, which included HVAC, EMS, Steam Trap, Insulation, and Other measures. All the measures are commercial retrofit and new construction projects.

The work was completed between 2020 and 2021. DNV performed a site-based Measurement and Verification (M&V) impact evaluation to quantify the achieved natural gas energy savings for a sample of custom gas sites from completed projects in the Program Year 2019 (PY2019). The ongoing COVID-19 health emergency caused changes to onsite work during the metering period of the study. DNV adapted the evaluation process by continuing the operational and non-operational result split (named two-sample inner-year method in the PY2018 report) used for the PY2018 evaluation. As discussed below, due to the pandemic not all sites received full M&V treatment with assessments of operational savings. Sites that did not have meter or trend data collected had non-operational adjustments quantified through site verification either from a virtual or onsite inspection (one site had only a virtual visit). In response to the reduced operational metering, DNV included an imputed historical operational adjustment for site evaluations that did not have operational discrepancy adjustments. A total of ten sites were evaluated, with four receiving only non-operational discrepancy adjustments and six receiving both non-operational and operational discrepancy adjustments. An additional sampling category was also added to account for strategic energy management partnership (SEMP) projects by including the sampling error for those projects within the precisions of the study. This category is equivalent to the existing categories of steam traps and non-steam traps.

1.1 Study purpose, objectives, and research questions

The objective of this Impact Evaluation of the Program Year 2019 (PY2019) Custom Gas Installations is to provide verification or re-estimation of energy (therms) savings for sampled Custom Gas sites through site-specific inspections, end-use monitoring, and analysis. Site-specific results were aggregated to determine realization rates for National Grid's custom gas installations in RI. Custom gas evaluations for National Grid in RI territory starting from PY2016 are designed to be rolling/staged evaluations. The goal of this approach was to repeat M&V annually as the previous year's tracking data becomes available. This study is considered Year 3 of the rolling/staged evaluation, with PY2017 and PY2018 as Year 1 and Year 2, respectively.

This study:

- Achieved gross natural gas energy savings for RI custom gas projects, with targeted sampling precision of ±20% at 80% confidence when RI PY2019 results are pooled with RI PY2017 and PY2018 results

The evaluation process of PY2019 was adapted to limit the impact on customers selected for evaluation. More information for COVID-19 evaluation adaptations can be found in Section 3.3.2.

1.2 Key findings and results

The site-level evaluation results were aggregated using the final adjusted case weights in operational and non-operational results approach based on non-operational and operational adjustment factors developed for 2019 sites and imputed historical operational discrepancy adjustments for those 2019 sites that did not get the full M&V required to develop operational adjustments. The realization rates were calculated and then applied to total tracking savings to determine their total evaluated savings. DNV developed realization rates (and associated precision levels) for annual therms savings of the program by combining three consecutive custom gas study results (conducted for PYs 2017, 2018, and 2019).

1.2.1 Rolling statewide sample: PY2017, PY2018, and 2019

The Rhode Island Piggybacking Diagnostic Study developed guidance on when it is appropriate to "piggyback" or combine RI evaluations efforts with MA studies or adopt MA results as a proxy for RI versus standalone RI studies. The "piggybacking" study report recommends which approaches National Grid RI should use for C&I measure groups and residential programs. For custom gas, it recommends using a RI Independent Sample approach. Therefore, the rolling statewide evaluation approach was planned to effectively produce independent results for RI by the end of a three-year rolling cycle if reasonable relative precisions could be achieved. The results presented in this report achieved reasonable precisions by combining three program years (PY2017, PY2018, and 2019), as shown in Table 1-1. Overall, the study achieved better than its target relative precision (RP) of ±20% at an 80% confidence interval.

Table 1-1. Yearly RI-specific results and pooled results

Parameter	PY2017	PY2018	PY2019	PYs 2017+2018+2019
Tracking Savings	1,948,383	2,350,739	1,944,204	6,243,326
Sample Size	6	6	10	22
Realization Rate (RR)	92.7%	83.3%	85.3%	86.9%
Relative Precision @ 80% CI (%)	±2.3%	±22.6%	±4.5%	±6.8%

CI = confidence interval

1.2.2 Operational and non-operational estimation

Due to the COVID-19 pandemic, fieldwork remains altered when compared to pre-pandemic evaluation methods. In PY2019, DNV continued the modified evaluation approach of operational and non-operational adjustments (named two-sample ratio estimation in PY2018). Non-operational adjustments (see Table 1-2) were calculated for all sites using the in-depth desk review and the available initial site visit results, which do not include metering or trend data. Operational adjustments were calculated for six of ten site evaluations, as the other four site evaluations did not collect meter or trend data, as explained in Section 5.1.1.2.

Table 1-2. Adjustment factors for evaluation

	Adjustment Factors						
Ratio Name:	Non-Operational Adjustments					Operational Adjustments	
Obtain During:	In-depth desk review			1st onsite or virtual site visit		Logger Installation	
Factor:	Baseline	Methodology	Tracking & Admin	Technology	Quantity	Operational	HVAC Interactive

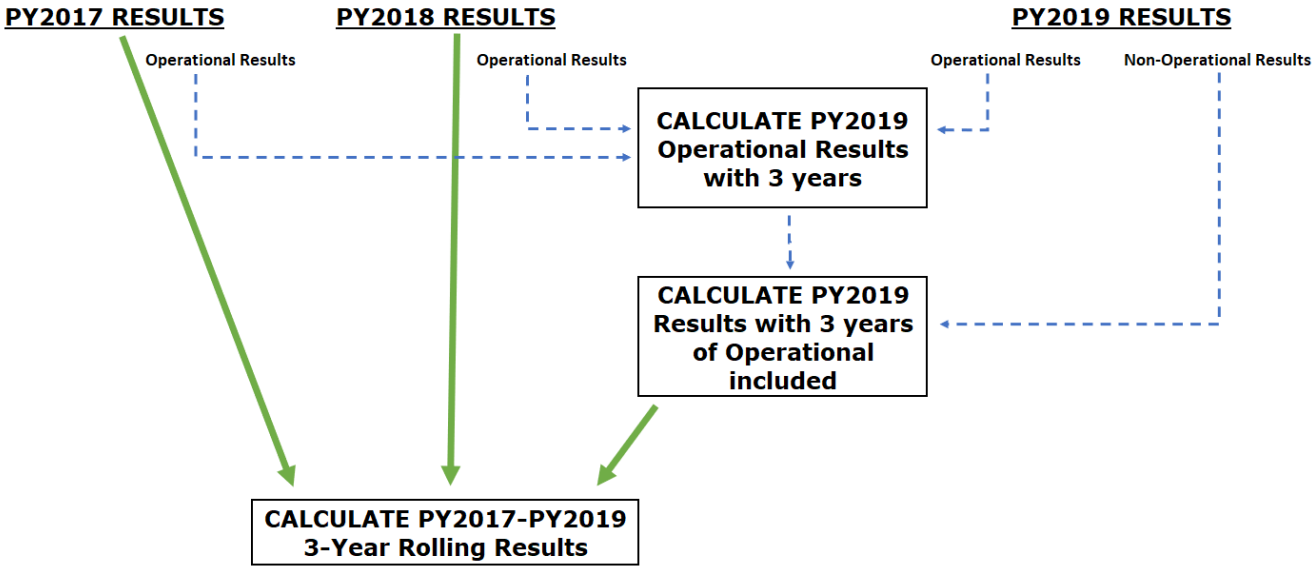
DNV and National Grid agreed to calculate savings realization rates from energy usage and savings that remained unaffected by the pandemic as operational adjustments influenced by changes in usage from the pandemic would unfairly affect future site realization rates. Therefore, operational conditions were monitored or collected and converted to operational adjustment factors only for sites unaffected by the pandemic, while non-operational adjustment factors were

calculated for all sites recruited during the program evaluation period. DNV developed the operational and non-operational sample adjustment factor results (named two-sample ratio estimation method in the PY2018 report) that uses two sets of case weights to calculate the realization rate (RR) for PY2018. These case weights were based on the sites that included non-operational adjustments only (first set – After Non-Operational Adjustments columns in Table 1-3) and sites that included both non-operational and operational adjustments (second set – After Operational Adjustments columns in Table 1-3).

DNV collected operational data for six out of ten sites and non-operational data for all ten sites. Trend data, metered data, or a combination of both were collected for five of the six operationally adjusted sites. Data was not collected for the last operationally adjusted site because the project was never completed. The site was found to provide zero savings resulting from an administrative error where site savings were accidentally claimed even though the project was incomplete. The evaluation team would not have included the findings in the operational adjustment for the calculated PY2019 operational results using the current workflow, since the team could not monitor the system. However, due to the extenuating circumstance that the project was never completed, the site adjustments were included in the operational results, rather than only including them in the non-operational results. This is important because the project never would have realized the operational savings and the overall operational results would miss a key factor from the site evaluations if not included.

Operational data was not collected for four sites as the pandemic impacted site energy usage at those facilities. Sites were affected by the pandemic due to a reduction in personnel at the locations, reduced manufacturing, occupancy restrictions at restaurants, or reduced HVAC needs from pandemic-related logistical changes. Both sets of non-operational and operational adjustment factors in PY2019 were then combined to calculate the overall site RR for operational adjusted sites.

Figure 1-1. Historical operations adjustment inclusion map for PY2019 and PY2017-PY2019 three-year rolling results



The two sets of site level results, non-operational and operational, are combined to calculate the PY2019 annual results with historical operations data included from PY2017 and PY2018, as depicted in **Error! Reference source not found.**. Annual results for PY2017 were broken into components of operational and non-operational site results to match PY2018 and PY2019. Then operational results from the three years are combined to calculate a PY2019 operational result with historical adjustments included. That means, for PY2019, the operational results included the operational components from PY2017 and PY2018. The process then becomes the same as PY2018 where the operational result (newly calculated with historical

data) is combined with the non-operation result to form the PY2019 annual results. Only PY2019 received historical operational adjustments as PY2018 annual results were calculated without considering pre-PY2018 operational results[1]. The PY2019 historically adjusted annual results are then combined with the annual results from PY2017 and PY2018 to form the three-year rolling results.

Table 1-3. PY2019 combined realization rates after discrepancy adjustment factors

		After Non-Operational Adjustments		After Operational Adjustments	
Site ID	**Tracking Savings (therms)**	**Evaluated Savings (therms)**	**Site Level Realization Rate from Tracking (%)**	**Evaluated Savings (therms)**	**Site Level Realization Rate from Non-Operational (%)**
2019RIG002N	26,120	26,025	100%	N/A	N/A
2019RIG019N	4,246	3,981	94%	N/A	N/A
2019RIG026S	2,120	2,120	100%	2,953	139.3%
2019RIG057N	24,296	24,296	100%	26,215	107.9%
2019RIG064S	13,038	13,038	100%	12,546	96.2%
2019RIG065N	8,327	8,327	100%	2,152	25.8%
2019RIG095S[2]	8,530	9,404	110%	7,207	84.5%
2019RIG104S/105N[2]	111,021	112,026	101%	N/A	N/A
2019RIG108N[2]	379,460	374,728	99%	N/A	N/A
2019RIG138N[2]	731	0	0%	0	0.0%

Site-level therm savings and site realization rates are shown in Table 1-3. More information for this table can be found in Section 5.1.1.3. All site IDs are independent site evaluations. The first six site IDs in the table are separate commercial or industrial locations while the last four (2019RIG095S, 2019RIG104S/105N, 2019RIG108N, & 2019RIG138N) are from the same campus but different building locations (sites) at the campus.

[1] PY2018 RI Custom Gas was the first program to calculate operational and non-operational results where historical adjustments were not considered nor the process developed at the time of regulation filing deadlines.

[2] The last four site IDs (2019RIG095S, 2019RIG104S/105N, 2019RIG108N, & 2019RIG138N) are from the same campus but different building locations (sites) at the campus. The campus was part of the strategic energy management partnership (SEMP) portfolio.

1.3 Conclusions, recommendations, considerations, and guidance for future research

1.3.1 Conclusions

1.3.1.1 PY2019 Performance

The program continues to generate significant natural gas savings. In RI, the PY2019 custom gas projects saved an estimated 2.04 million therms (adjusted gross savings) annually, with 85.3% of the program year savings realized based on the program evaluation sample for RI PY2019 sites. The current results are accurate within state and regulatory standards and provide adequate planning and program reporting savings estimates.

A more detailed explanation of the PY2019 performance is found in Section 5.1.1.3. Site-specific details are shown in APPENDIX A. More details on the PY2019 results are presented in the sections below, and each site report included in APPENDIX D.

1.3.1.2 Combined three-year rolling (PY2017, PY2018, and PY2019) performance

Combined over the three-year rolling sampling period, the program realized gross savings of 6.24 million therms, with 86.9% savings realized as shown in Table 1-1. The current three-year rolling results are accurate within state and regulatory standards and provide adequate planning and program reporting savings estimates.

A more detailed explanation of the combined three-year results is found in Section 5.2.

1.3.2 Recommendations

1.3.2.1 R1: Realization rate

DNV recommends National Grid to use the PY2017, PY2018, and PY2019 combined RR of 86.9% for planning and program reporting, starting with PY2022 and continuing to subsequent years until new program impact evaluation study results are available. The applicable RRs are noted in Table 1-1. This recommendation was based on the following factors:

- When PY2017 (92.7%), PY2018 (83.3%), and PY2019 (85.3%) results are pooled, the study produced statewide results that met precision targets of ±20% relative precision at 80% confidence (actual: ±6.8% at 80% confidence level).

Based on the results listed for PY2019, an individual program year sampling Error Ratio Target of 0.55 for non-steam trap and SEMP projects has been recommended for the 2020 RI Custom Gas Impact Evaluation to achieve the next three-year rolling savings program evaluation precision targets. Steam traps are recommended 0.65 for the error ratio target.

1.3.2.2 R2: Memo should accompany savings changes in submitted calculators

DNV recommends that project savings calculators have an accompanying memo that documents changes when savings deviate from submitted calculators or changes during the final stages of compiling claimed savings. Site evaluations depend on final savings calculators to understand assumptions and methodology to construct a site evaluation plan. An accompanying memo for savings changes would provide necessary information on what changed and the reasons for those changes. A specific example of this problem is listed below:

2019RIG108N: The project folder for this measure contained several calculators labeled 'FINAL.' After many attempts by both National Grid and DNV evaluators, the evaluation team identified the proper calculator used for claimed savings in tracking data. The major change between submitted calculators was reduced savings due to gas usage moved from the main boiler heat plant to distributed boilers that were not considered in the original calculations. Due to the multiple calculators, the evaluation team is recommending a memo written for modifications to calculators after submittals or final versions of the calculators are created to document reasons for changes. There should be a clearly documented path to

explaining the final savings claimed in the tracking system when multiple versions of the savings calculations exist in the project files

1.3.2.3 R3: Post-inspection verification for HVAC EMS projects

DNV recommends HVAC EMS project applicants submit EMS screenshots, schedule printouts, or trend data to National Grid representatives at the time of project completion or before incentives are paid to customers for comparison of installed condition with original design documents submitted for initial project approval. The submissions will provide an opportunity for implementation teams to identify differences from original submission and installed condition to determine operation is within original parameters or if modifications must be made to the installation or incentive due to project changes during installation. A specific example of this problem is listed below:

2019RIG065N: This project installed a new dehumidification system to serve a swimming pool at the recreational complex of an elementary school. The installed dehumidification system is based around a hot-gas heat recovery system. The project saves energy when the installed system is in dehumidification mode by utilizing the hot-gas heat recovery system to reheat the supply air to reduce HVAC energy consumption. Several parameters were not as stated in the as-built documents submitted during the project and most major discrepancies would have been avoided with the submissions of EMS screenshots, schedule printouts, or trend data submissions.

1.3.2.4 R4: Evaluate lifetime savings for PY2020 projects

DNV recommends evaluating lifetime savings and reporting them at the site level in all future custom gas site evaluations. This is to prepare for reporting the new lifetime savings goal as National Grid transitioned away from annual savings goals to lifetime savings goals beginning in PY2021. Reporting them in every study is critical to producing three-year rolling-based lifetime savings results. Should lifetime savings results be available beginning with PY2020 evaluated projects, then a standard three-year rolling reporting cycle would be available after the PY2022 program evaluation.

1.3.3 Considerations

1.3.3.1 C1: SEMP sampling in next program year evaluation

DNV proposes a three-stage process to include SEMP projects in the next program year evaluation sampling for the 2020 RI Custom Gas Impact Evaluation. Stage 1 will identify SEMP projects and categorize them as a separate category using the line items as they are claimed in tracking savings and creating site IDs as traditionally completed in recent program evaluations. DNV will identify which SEMP projects are identified for site evaluation, which leads to Stage 2. Stage 2 will require project files to disaggregate sampled SEMP sites to the measure level. DNV will sample again with the disaggregated sampling breakout at the measure level to identify measure-specific projects to evaluate that will provide results within statewide precision targets. Stage 3 is an extra step only if an identified site in Stage 2 proves to be a much larger burden on evaluators and customers than need be. A second disaggregation would then occur at a level needed to reduce the burden on a case-by-case basis. Sub-categories (steam trap or non-stream trap) will be provided for each claimed savings and measure row item as traditionally completed, and stratification (number of sites per stratum) will follow the same methodology as has been completed in previous program evaluations.

1.3.3.2 C2: Dual baseline inclusion for lifetime savings

As a preparation for the coming lifetime savings analysis in future program year evaluations, DNV proposes that the National Grid implementation team consider dual baseline calculations to more accurately reflect the claimed savings over the lifetime of the retrofit measure.

2 OUTLINE

This section presents the objective for DNV's Impact Evaluation of the 2019 Custom Gas Installations for National Grid in Rhode Island (RI) and describes the organizational format of the report.

2.1 Study purpose, objectives, and research questions

The objective of this program report is to document the evaluated achieved savings for the 2019 RI Custom Gas installations using a sample of projects grouped within sites based on location. DNV provides program achieved savings with a precision tolerance for regulatory reporting. The regulatory filing is then used for planning future program cycles. This program evaluation performed site-based M&V impact evaluations to quantify the achieved natural gas energy savings using 10 RI custom gas sites from projects completed in the PY2019 cycle. The results of this study were combined with the results from the PY2017 and PY2018 studies to produce updated, statewide realization rates.

2.2 Organization of report

The remainder of this report is organized as follows:

- Section 3: Methodology and Approach. The methods associated with sampling and the M&V tasks are described in this section.
- Section 4: Data Sources.
- Section 5: Analysis and Results. The results associated with the program evaluation of PY2019 and the latest rolling three-year results are presented in this section.
- Section 6: Conclusions, Recommendations, and Considerations. Conclusions and recommendations from analyzing the M&V findings are presented in this section.

3 METHODOLOGY AND APPROACH

The evaluation team's approach was consistent with the procedures and protocols developed during the previous round of custom gas impact program evaluation conducted for PY2017 and PY2018, with a few adaptations to account for the ongoing health emergency. The evaluation consists of a randomly selected sample of projects at participating facilities. As described in the next subsections, the impact evaluation consisted of virtual and onsite visits. For virtual and some onsite visits, only non-operational observations were collected, and operational metering was performed for the remainder of the onsite visits.

3.1 Description of sampling strategy

DNV designed the PY2019 program impact evaluation sample to pool annual program evaluation results with the PY2017 and PY2018 results to produce a rolling updated result. This allowed the sampling precision to meet and exceed the targets laid out in Table 3-1.

PY2017, PY2018, and PY2019 results were pooled together to use in PY2022. In subsequent years, the realization rate will reflect the pooling of the three most recent impact results.

Based on the results achieved in the previous studies, this sample design assumed the error ratios shown in Table 3-1 for the targets listed. The sample design for this round of the study was developed assuming the results would pool with prior (and future) custom gas results. The general principle for the design is each independent program evaluation year would need to achieve a ±35% precision at an 80% confidence interval to maintain a three-year pooled result of ±20% precision at 80% confidence for gross therms savings RRs. DNV used a Model-Based Statistical Sampling (MBSS) technique to develop the sample design. The sampling unit is the sum of all projects installed in the evaluated program year for an account or location if the account serves multiple locations.

Table 3-1. Sampling targets

Annual Sampling Target	Three-Year Pooled Sampling Target	Error Ratio
±35% expected relative precision - 80% CI	±20% expected relative precision - 80% CI	0.55 (non-steam trap)
		0.65 (steam trap)
		0.55 (SEMP site)

An additional category was added to the sample design after the evaluation team discovered two years' worth of projects (85 measures) was completed by a single site and compiled into three application IDs. The single SEMP randomly selected for a site evaluation is part of a National Grid initiative called the strategic energy management partnership (SEMP[3]). The three application IDs correspond to a parent/child[4] application combo for 2018 (31 measures) and a single application for 2019 (54 measures). In previous program evaluation years, evaluators created a randomly selected sample of measures from the SEMP site after it was selected in the sample design to limit burden on the site contacts and provide a feasible amount of field work for engineers. However, the evaluation team decided, beginning in PY2019, to categorize the initiative in a separate category due to the size of the SEMP program relative to the total portfolio and continuing partnership in the following program years. Details for categorizing the broader initiative are still underway, but National Grid and DNV agreed to evaluate this single SEMP site in a separate category until stakeholders agree on how to continue with the new category.

[3] Strategic Energy Management Partnership (SEMP) is RI National Grid's portfolio partnership program with a few large customers in the state. The program constitutes an assortment of multiple energy efficiency projects that are completed in these facilities in a given program year. The projects included in the portfolio range anywhere from a few measures to over 50 measures installed across multiple buildings on the campus

[4] For some large projects, National Grid typically does not pay out the total incentive upfront but splits the project into two applications as parent and child. The child payment is made after the project is fully commissioned and completed. Savings are split between parent and child applications. In this case, the parent application was claimed in 2018, and the child application was claimed in 2019, so the entire project is classified as being completed in 2019 for site evaluation purposes.

3.1.1 PY2019 sample frame

The initial population for this program impact evaluation was the set of custom gas projects rebated in 2019. Table 3-2 shows the distribution of all tracking records and the associated savings by National Grid.

Table 3-2. PY2019 population distribution of custom gas accounts

Distribution	Number of Accounts	Gas Savings (Therms)	% Savings
C&I Custom General	91	1,944,204	95.1%
Custom Design Approach (CDA)	10	81,282	4.0%
Custom-Prescriptive	2	12,993	0.6%
Less than 1,000 therms savings	11	5,812	0.3%
Grand Total	114	2,044,291	100%

As previously completed in prior program year evaluations, small sites (<1,000 therms) were excluded from the sample frame. These small sites account for less than 1% of total tracking savings and do not warrant the expense of site M&V. There were 11 such gas accounts with annual savings of less than 1,000 therms that were removed from the population frame, with a total savings of 5,812 therms as shown in Table 3-2. There were 10 Comprehensive Design Approach (CDA) sites that were removed from the population frame as CDA projects are evaluated in a separate study. Two sites were considered prescriptive and were removed from the population. Therefore, the final population included 91 unique customer accounts.

The final PY2019 population frame has a total of 91 accounts with savings of 1,944,204 therms. Table 3-3 shows the selected sample frame after dropping the small sites, dropping CDA projects, and removing prescriptive measures.

Table 3-3. PY2018 adjusted (final) project sample frame

Accounts	Tracking Savings (Therms)
91	1,944,204

3.1.2 PY2019 sample design

Table 3-4 shows the selected sample for this project. DNV estimated that ten sampled sites would give reliable precisions to achieve the required annual and three-year rolling program target per Table 3-1. An annual target of ±35% precision at 80% confidence interval was set for the individual program year, and, when combined with the other two program years in the three-year rolling sample, a target of ±17.3% precision at 80% confidence was estimated to result if all planned sites were completed in the current program year. An analysis was performed at the request of the C-Team to determine if a sample reduction was possible due to the adaptations from the ongoing COVID-19 health emergency. Sample targets are typically met or improved when the study includes imputed historical operational adjustments, as large variability exists when introducing new operational adjustments. DNV found that reducing the sample was not possible due to the low number of observations in each stratum, and the original sample was kept. The table also includes the completed 10 of the designed 10 sites regarding non-operational adjustment factors in PY2019 and the completed 5 of the 10 operational adjustment factor sites. The study also achieved reliable statistical precision targets at an 80% confidence interval, as shown in Table 3-4.

Table 3-4. PY2019 project sample

Accounts	Savings	Error Ratio	Sample (n)		Expected Relative Precision @ 80% CI	Achieved Relative Precision @ 80% CI
			Designed	Completed		
91	1,944,204	0.55 (non-ST) 0.65 (ST) 0.55 (SS)	10	10 Non-Op 6 Op	±35.0%	±4.5%

Non-Op: Sites with Non-Operational adjustment factors; Op: Sites with Operational and Non-Operation adjustment factors; Non-ST: non-steam trap; ST: steam trap; SS: SEMP Site.

3.1.2.1 Strategic Energy Management Partnership (SEMP) category addition

Strategic Energy Management Partnership (SEMP[3]) is RI National Grid's portfolio partnership program with a few large customers in the state. The program targets an assortment of multiple energy efficiency projects completed in these facilities in a given program year. SEMP project portfolios range from a few measures to over 50 measures installed across multiple buildings at the facility. The SEMP project included in the PY2019 program evaluation sample is a single SEMP participant portfolio. The total savings from all the measures are entered into National Grid's tracking system under three application IDs: two applications as a parent/child[5] system for 2018 projects and a single application for 2019 projects. The two years fell under the PY2019 program evaluation period due to the inability to complete the full project closeout of 2018 projects in PY2018.

Evaluating these large projects could be very expensive and burdensome for the customers. Therefore, DNV and National Grid chose to select a sub-sample of the SEMP portfolio to reduce the burden and be cost-effective. During sampling, a SEMP site was identified for site evaluation. The one sample unit was found to contain a portfolio of projects. Instead of evaluating the entire project, a smaller, more manageable number of buildings were selected. The one sample point portfolio was disaggregated into measures and recombined at the building level (the chosen sample unit for the SEMP portfolio) using project files provided by National Grid. Traditionally, a sample unit is based on a location. Building names provided in savings summary sheets were used to recombine sample units. The sample units (commonly referred to as a site during program evaluations) range from one to five measures. The SEMP project had a total of 60 sites after recombination. All SEMP building level sample units (sites) were then added back to sampling with the rest of the population to include all statistical errors within the precision for reporting purposes and to choose a more manageable number of buildings. Sampling identified four of the 60 SEMP sites to achieve annual and three-year rolling precision goals discussed in Section 3.1.2. Therefore, the one sampling unit originally identified was broken into four individual sampling points.

The PY2019 study is the first effort to include a standard error from the subsample of large portfolios under one or a few application IDs that were aggregated together into a single site ID or application ID. The inclusion assumed all errors into the overall sample design. In this current study, the SEMP project was disaggregated and evaluated after the site was chosen for site evaluation during sampling. Future efforts for determining how to include the SEMP in the custom gas program impact evaluation are in discussions with National Grid, where DNV's suggested changes are found in Section 1.3.3.1.

The sample for PY2019 is found in **Error! Reference source not found.**.

[5] For some large projects, National Grid typically does not pay out the total incentive upfront but splits the project into two applications as parent and child. The child payment is made after the project is fully commissioned and completed. Savings are split between parent and child applications.

3.1.3 Rolling sample design

The expected precision from the PY2019 sample design was combined with the achieved PY2017 and PY2018 study results to produce a combined precision for the overall program. Table 3-5 provides the combined expected precision based on this sample design.

Table 3-5. PY2017, PY2018, and PY2019 combined expected precision at 80% confidence interval

Program Year	Accounts (N)	Therms Savings	Error Ratio	Sample (n)		RP @80% CI	
				Design	Achieved	Design	Achieved
PY2017	98	1,948,383	0.6	7	6	±30.0 %	±2.3%
PY2018	87	2,350,739	0.40 (non-ST) 0.65 (ST)	8	8 non-OP 6 OP	±21.0 %	±22.6%
PY2019	91	1,944,204	0.55 (non-ST) 0.65 (ST) 0.55 (SS)	10	10 non-OP 6 OP	±35.0 %	±4.5%
PYs (2017, 2018, & 2019)	276	6,243,326	N/A	25	24	±17.3 %	±6.8%

ST = Steam Trap; SS = SEMP Site; OP = Operational; N/A = Did not calculate.

3.1.4 PY2019 final sample disposition

Two primary sites were dropped from the study sample and were replaced by two secondary sites. The two sites were removed and replaced due to the non-responsiveness of the site contact for one site and a refusal for the other. The energy savings measures were a condensing boiler retrofit and a steam trap site. They were replaced by the next sites on the priority list.

The final (achieved) sample includes ten sites, as shown in Table 3-7. **Error! Reference source not found.** summarizes the ten sites for which M&V activities were completed. The summary includes the site ID, the verified measure description, tracking savings, and site RR.

3.2 Site M&V planning

The site evaluation (M&V) plan played an important role in establishing approved field methods and ensuring that the ultimate objectives for each site evaluation were met. The M&V plan for each evaluated site provided detailed information on the procedures for accomplishing those objectives.

DNV submitted full individual M&V plans for each evaluated site. These plans were reviewed by National Grid. Each site plan included the following sections:

- **Project description** – A description of how the project saves energy.
- **Tracking savings** – A short description of how the tracking savings were estimated and their source, including:
  - Analysis method was used.
  - Identification of the key baseline assumptions.

- Identification of the key proposed assumptions.
- Evaluator assessment of tracking savings methods or assumptions, including program-reported baseline.

- **COVID-19 impacts** – A description, if any, of impacts from the current health emergency.
  - Suggested site evaluation method taking into account COVID-19 Impacts.
  - Reasoning for the chosen site evaluation method.
- **Project (site) evaluation** – A short description of the methods to be used to evaluate the project, including, but not limited to:
  - Methods for verifying the measure installation and current operation.
  - Methods for observing and/or assessing building use and occupancy.
  - Identification of the tracking and expected evaluator baseline of each measure.
  - The data to be collected by DNV; where several similar items have been installed or are being controlled, the site evaluation plan described and justified the sampling rate of the equipment to be monitored.
  - Site staff interview questions (to understand the baseline operation and determine if any changes in the operation of the impacted system occurred after the project was installed).
  - The data provided or to be provided by the site (e.g., EMS trends, production, pre-metering) and/or National Grid.
  - The expected site evaluation analysis method to be used, including any deviations from the implementer savings estimation method. In general, the same methodology used to estimate tracking savings was used to estimate evaluated savings. DNV presented an alternative methodology only if the tracking methodology was flawed, unfeasible, or a more accurate methodology that utilized post-installation data was available.
  - Key parameters that are determined through the site evaluation preparation to compare to those used in the original savings estimate.
  - Measurement verification equipment to install on select equipment and quantity of devices intended for installation.

DNV updated the M&V plan, responding to National Grid comments, and in most of the cases, submitted a revised M&V plan before the site visit. For some sites, the initial visit was scheduled within a couple of days or less, and National Grid reviewers did not have the chance to approve the entire M&V plan before the site visit. For those sites, DNV evaluators emailed the plan for a quick review and response specifically for the tasks to be conducted onsite and the metering approach.

3.3 Data collection

DNV performed a site contact interview and scheduled a site visit to perform the tasks described in the site M&V plan. Data collection occurred from December 2020 to May 2021.

3.3.1 Customer outreach

Using the information provided in the project files, project engineers reached out to customer site contacts. During this initial outreach, the engineers discussed the purpose of the site evaluation, the scope of measures installed, availability of onsite trend/EMS/production data, any other applicable parameters relevant to the site evaluation, impact from a COVID-19 health emergency, and confirmed that the site would allow DNV to conduct the site visits. The site-specific M&V planning effort did not commence until the customer site contacted indicated they would accommodate the ex-post virtual or onsite evaluation process. After the customer outreach discussion, if the engineer determined significant barriers were preventing M&V of substantial parts of the completed project, the site was flagged for review and, if warranted, replaced with a backup site.

The agreed-upon communication protocol memo for National Grid is found in APPENDIX C.

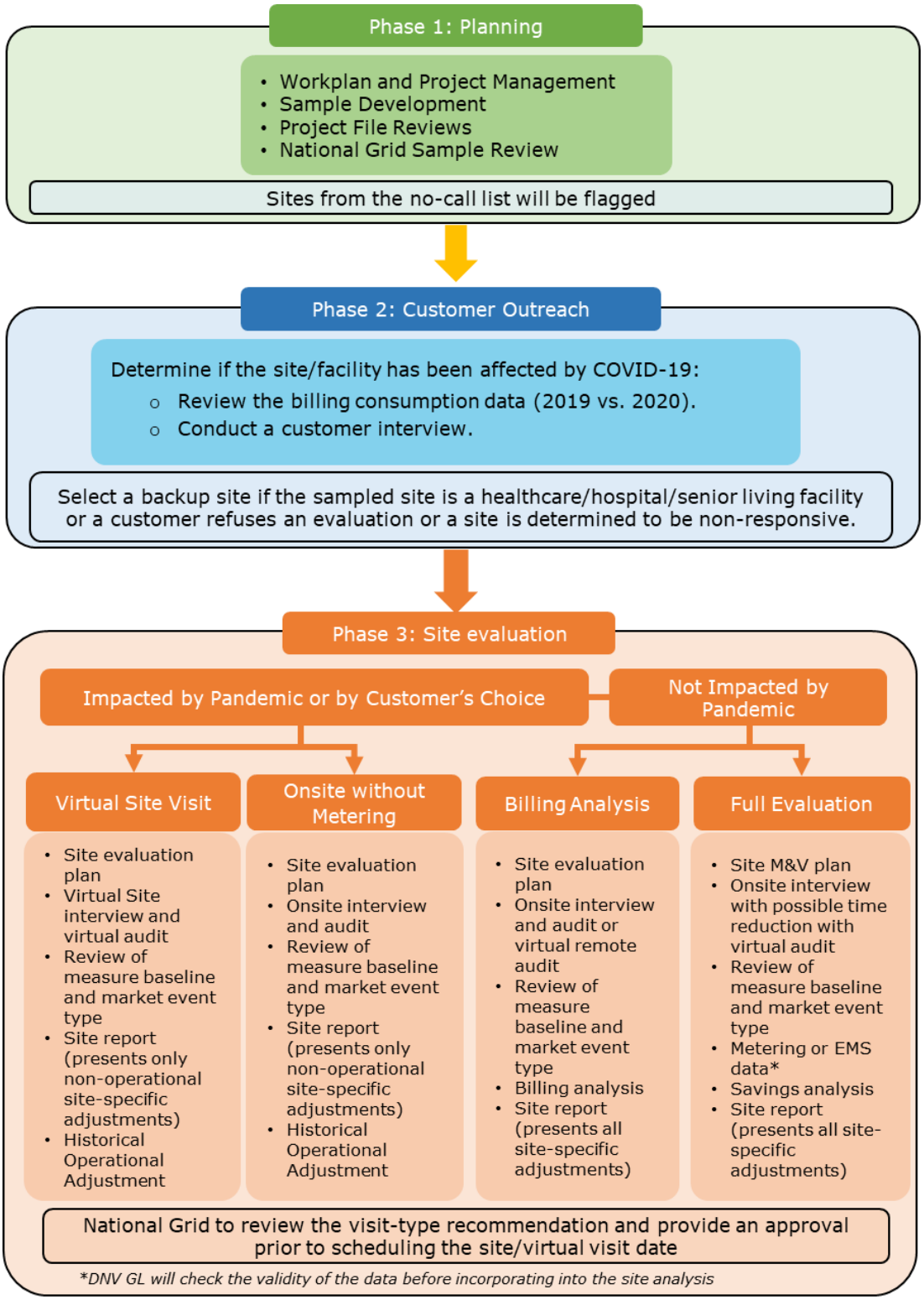
3.3.2 Site evaluation type determination

DNV conducted one of four types of site evaluations for each site. The deciding factors were: COVID-19 gas consumption impact; site contacts long-term absence or physical presence onsite; and site contact willingness for a physical visit from engineers. The following conditions were applied to the site evaluation type:

- Base Method – Virtual Site Visit: The site was recruited for virtual visits due to the restrictions on having an in-person visit. Evaluators collected only non-operational impacts for this option.
- Add-On 1 – Onsite Site Visit with only non-operational impacts: The site was open for an onsite visit, but the COVID-19 pandemic impacts the installed measure's operation, and little meaningful data would be obtained by performing onsite metering. Evaluators collected only non-operational impacts for this option.
- Add-On 2 – Billing Analysis: In this case, the engineer determined the measures installed were best analyzed by billing analysis. This option was possible if the site was open for an onsite or virtual visit and the customer operation was not impacted by the COVID-19 pandemic. Billing analysis does not allow for savings determination across measures and only delivers results at the site level.
- Add-On 3 – Onsite Visit with both non-operational and operational impacts: The site was open to an onsite visit, and the customer was not impacted by the COVID-19 pandemic. Evaluators collected both non-operational and operational impacts using M&V.

Evaluators suggested a virtual site visit depending upon the site contact's willingness for a physical visit to the site. A backup site was selected when the customer refused a site visit or a virtual visit. Figure 3-1 shows the decision tree graphically. A virtual visit was conducted for non-operational sites, and an onsite visit was conducted for both non-operational and operational metering sites if the customer permitted. The onsite site visits were used to verify the non-operational and operational adjustment factors. In the absence of metered or trend data (i.e., non-operational sites), operational adjustment factors are based on historical results (operational data collected in PY2017 and PY2018, and operational data collected in PY2019) as discussed in Section 3.7.2.

Figure 3-1. Summary of M&V flow through sampling to site evaluation to determine site evaluation type



As found in Table 3-6, the site level information of data types collected and the type of site visit is listed.

Table 3-6. Site-level information for the type of visit and data collected

Site ID	Type of Site Visit	Data Collected				Type of Adjustment
		Site Interview	Equipment Verification	Trend Data	M&V Data	
2019RIG002N	Virtual	X	X	N/A	N/A	Non-Operational
2019RIG019N	Onsite	X	X	N/A	N/A	Non-Operational
2019RIG026S	Onsite	X	X	-	X	Operational
2019RIG057N	Onsite	X	X	-	X	Operational
2019RIG064S	Onsite	X	X	-	X	Operational
2019RIG065N	Onsite	X	X	-	X	Operational
2019RIG095S[2]	Onsite	X	X	-	X	Operational
2019RIG104S/105N[2]	Onsite	X	X	NU	N/A	Non-Operational
2019RIG108N[2]	Onsite	X	X	N/A	N/A	Non-Operational
2019RIG138N[2]	Onsite	X	X	-	-	Operational

NU: Not used but collected; CD: Consumption Data; [2]Ibid, page 4.

3.3.3 Site visit

Each initial site visit consisted of the determination of the COVID-19 health emergency impacts; site verification of installed equipment; a discussion with facility personnel regarding the baseline characteristics of the measure; if called for, the installation of measurement equipment; the collection of available trend data; and/or the creation of a plan to gather trend data coinciding with the measurement period. Trend data beyond the measurement period was also requested and used to improve the accuracy of measure savings estimates.

A second site visit to retrieve meters was scheduled for sites where evaluators installed meters during the initial visit.

3.3.4 M&V plan update

DNV submitted an updated site M&V plan to National Grid after the completion of the initial site visit. This updated plan included the following information, based on the site visit:

- Any deviations from the plan that occurred during the visit or were expected to occur; deviations included cases where a portion of the proposed M&V plan was not feasible for unforeseen reasons.
- Provides a summary of the information to be collected, information that will not be available for analysis purposes, and lists tasks to complete on the return for meter pickup.

The update was intended to keep National Grid current on the status of the site evaluation and communicate any anticipated or resulting deviations from the plan.

3.4 Site analysis

As previously discussed, the evaluation team evaluated six projects with operations adjustments (traditionally called full site evaluations). However, the team reviewed all data collected and then utilized the data to complete a site evaluation analysis for five of the ten sampled projects. Even though the project was never completed, the remaining operationally adjusted site is considered operationally adjusted (full discussion found in Sections 1.2.2 and 5.1.1.2). For five operationally adjusted projects, the analysis generated evaluated savings estimates for all measures installed at each sampled site. Results were normalized to typical production or weather data. For weather-dependent measures that result in savings, the site analysis involves normalizing the models to weather data using Typical Meteorological Year 3 (TMY3) data from the closest representative weather station to each site.

For the remaining four sites of ten, engineers did not complete a full site evaluation due to COVID impacts, limiting the program representativeness of the operational aspects from their site evaluations. The four sites had non-operational data collected but no metered or trended data to calculate operational adjustments. All sites that did not have an analysis completed were still included in the final project realization rate using the operational and non-operational adjustments described in detail in Section 3.7.2 due to evaluators and National Grid's wanting to include the non-operational information.

3.5 Site reporting

DNV submitted draft site reports to National Grid, who provided comments or questions to the engineer who led the site analysis. The engineer responded to comments and questions until a final agreement was reached on the analysis approach, the results, and the report itself. Each site report contains the following sections:

- Project summary and results – Provides a brief description of how the evaluated measures at the site save energy and a high-level summary of why the site evaluation results may differ from the tracking estimates. The site results are also presented in this section. A description of COVID impact and site evaluation type is included to describe what data was included and if an M&V based operational analysis was completed during the site evaluation.
- Evaluated measures – Describes the evaluated measures, including, but not limited to:
  - Applicant baseline and proposed conditions
  - Applicant savings calculation methods
  - Evaluator assessment of the applicant savings calculation methods
  - Measure verification results and methods for verifying measures
  - The data collected by DNV, summarized in graphical or tabular form for each data point
  - The data provided by the site and/or National Grid, with key data summarized in graphical or tabular form
  - Site evaluation baseline used
  - The site evaluation analysis method used, identifying any deviations from the original savings estimation method
  - Key savings parameters determined through the site evaluation, and a comparison to those used in the original savings estimate
  - A summary of the evaluated savings calculated and the primary drivers for differences between the tracking savings estimates and site evaluation savings estimates

An internal quality assurance lead reviewed a select amount of site reports. This review determined if the reports complied with the requirements for this deliverable and if the document communicates information clearly and consistently.

3.5.1 Measure event type and baseline review

A review of event measure types and baselines for each measure installed at sites in the sample selected for the program evaluation were completed for this study. DNV selected a measure baseline event type from stakeholder agreed-upon categories based on a preponderance of evidence presented in the project file, the data gathered during the site contact interview, and information gathered during the site visit. National Grid classified measures into two event types: 1) new construction measures, which include both new buildings and replace on failure or planned new measure purchases, and 2) retrofit measures which include single baseline, early replacement, and add-on. New construction and retrofit event types were found in the PY2019 program evaluation. The evaluation team reclassified EMS and insulation measures as add-ons and new construction projects as lost opportunity.

Table 3-7 shows the measure event types used in National Grid tracking information and site evaluations. Sites 2019RIG002N, 2019RIG057N, and 2019RIG095S; 2019RIG104S; and 2019RIG105N have multiple application numbers for different measures. The multiple application numbered sites are considered Parent/Child[6] projects or have multiple application numbers due to claiming measures separately. Sites 2019RIG095S, 2019RIG104S, and 2019RIG105N; 2019RIG108N; and 2019RIG138N are part of the disaggregated SEMP site for site evaluation.

[6] For some large projects, National Grid typically does not pay out the total incentive upfront but splits the project into two applications as parent and child. The child payment is made after the project is fully commissioned and completed. Savings are split between parent and child applications.

Table 3-7. Measure event type in National Grid tracking information and site evaluations

Site ID	Measure Type	National Grid Application#	Tracking Event Type	Site Evaluation Event Type
2019RIG002N	EMS/DCV	8175997	Retrofit	Add-On
	Condensing Water Heater	10722806	New Construction	Lost Opportunity[7]
	Vent Hood w/ Controls	8375229	Retrofit	Add-On
2019RIG019N	DCV	9581345	Retrofit	Retrofit
2019RIG026S	Steam Traps	8884415	Retrofit	Retrofit
2019RIG057N	Insulation	9257490	Retrofit	Add-On
	Process Controls	10426524	Retrofit	Retrofit
2019RIG064S	Steam Traps	8651739	Retrofit	Retrofit
2019RIG065N	Dehumidification	7722897	New Construction	Lost Opportunity[8]
2019RIG095S[2]	Steam Traps	8116687, 9208255, 9494451	Retrofit	Retrofit
2019RIG104S/105N[2]	Heat Recovery, Steam Trap, HVAC Controls	8116687, 9208255, 9494451	Retrofit	Retrofit
2019RIG108N[2]	Heat Plant Improvements and Hot Water Loop Efficiency	9494451	Retrofit	Retrofit
2019RIG138N[2]	HVAC Control Schedules	9494451	Retrofit	Retrofit

[2]Ibid, page 4.

After the measure event type was selected, the evaluator selected the evaluated baseline for the event type. Measures classified as retrofit used pre-existing conditions as a baseline. Measures classified as new construction used ISP or code as the baseline. The evaluation team completed an independent review of the baseline for each sampled project. Using site data project documentation and interviews at the facility, DNV assessed the reasonableness of the baseline for each sampled project.

[7] ISP defined from the International Energy Conservation Code 2015 – Table 404.2: baseline consisted of five 500 MBH gas-fired hot water heaters with an efficiency of 80%

[8] ISP defined from the International Energy Conservation Code 2012 - Table C403.2.3(4): baseline consisted of a dehumidification system with gas-fired reheat with an efficiency of 80%

3.6 Non-operational site collected data

With the ongoing COVID-19 health emergency, the Team decided to implement a two-tiered data collection approach such as the approach used in PY2018. Evaluators and National Grid understood there would be sites incapable of accepting evaluators at the location, sites that may not have full occupancy, sites with reduced production capacity, and sites with reduced or increased energy usage due to the previous scenarios or others. Refer to the site evaluation type decision tree, found in Figure 3-1, to understand when sites received non-operational vs. operational adjustment factors.

Non-operational data was collected for all ten sites in PY2019. From the total of ten sites evaluated, only non-operational data were collected for four of the ten sites. Non-operational and operational data were collected for the remaining six sites.

The evaluation team conducted an in-depth review of the baseline, methodology, administrative tracking/documentation, quantity, and technology adjustment factors for each evaluated measure. The non-operational data collection focused on measure-specific assessments of the following criterion:

- Measure event type classifications (retrofit, add-on, lost opportunity, etc.)
- Applicant baseline source
- Applicant and evaluator measure life
- Evaluator assessment of the baseline (pre-existing single/dual, ISP, unique)
- Assessment of baseline change impact on the measure savings
- Savings calculation method used by the applicant
- Most applicable savings calculation method, per evaluator
- Applicant key assumptions quality
- Assessment of methodology change impact on the measure savings
- Availability of native tracking savings calculations in electronic form
- Tracking savings source (applicant, equipment vendor/contractor, National Grid implementer, independent TA consultant)
- Assessment of quantity of items installed
- Verify the unit(s) is/are installed and if there are any discrepancies for installed quantities
- Does the installed technology match the applicant claimed technology or serve the same function?
- Does the applicant analysis consider interactivity with other end-uses, equipment, or fuel types?
- Were the applicant savings calculations normalized?
- Evaluator assessment of the quality of the applicant's savings estimations

The results of the non-operational data-only sites were documented and reported in a final site report as per full site evaluation standard procedure. Discrepancies were documented in the final site reports and used in the program savings expansion. Operational adjustment factors were imputed for all sites where only non-operational data was collected from prior years of program evaluations and the current operational data collected (PY2017, PY2018, and PY2019 operational collected sites). More information on the imputed historical adjustments can be found in Section 3.7.2.

3.7 Sample expansion

3.7.1 Site weight calculation

Non-operational weights are calculated similarly to previous rounds of custom gas program evaluations and are determined by taking the total number of observations in the stratum and dividing by the number of evaluated observations. Operational adjustments are based on the same methodology; however, the final realization rate and error calculations are based on imputed calculations. Initially, PY2019 operational adjustments are calculated and then combined with PY2017 and PY2018 operational adjustments using the methodology found in APPENDIX B. The methodology is similar to combining results from the three-year rolling sample with only operational adjustment calculations; however, APPENDIX B contains the specific algorithm followed to calculate imputed historical results.

3.7.2 Operational and non-operational sample with imputed historical adjustments

The operational and non-operational sample estimation approach accounts for the difference within the program year of 2019 from two results: operational and non-operational adjustment factors. National Grid and DNV chose to keep the integrity of the randomly selected sample by collecting as much information from each site if a minimum amount of information was collected to verify measure installation. That minimum included a site contact interview and measure verification from either a site visit or virtual visit.

Operational adjustment factors were not collected from a site for two reasons: 1) the location was affected at the time of the site evaluation by COVID-19 restriction measures that reduced occupancy or energy consumption or 2) meter installation, trend data collection, or physical access by evaluators to the installed measure for direct observation was impossible from the COVID-19 restriction regulations. Restrictions resulted from the business itself or another governing entity.

The methodology in APPENDIX B is used to calculate the realization rates for both sample components of the 2019 program year. The overall 2019 program year realization rate is shown and discussed in detail in Section 5.1.

Table 3-8 shows the adjustment factors used by evaluators to categorize discrepancies from tracking data and how those factors are categorized within PY2017, PY2018, and PY2019. Non-operational adjustment factors include factors that are obtained during a desk review, site contact interview, and primary site visit. Operational adjustments require metering or trends collected for analysis which is obtained during logger installation or delivered after the initial site visit.

Table 3-8. Adjustment factors for site evaluation

	Adjustment Factors						
Ratio Name:	Non-Operational Adjustments					Operational Adjustments	
Obtain During:	In-depth desk review			1st site visit (onsite or virtual)		Logger Installation	
Factor:	Baseline	Methodology	Tracking & Admin	Technology	Quantity	Operational	HVAC Interactive

Operational adjustment results were used from PY2017, PY2018, and PY2019. The historical adjustment is essentially extrapolating results from the operational adjustment factors from the last three years to calculate a combined operational realization rate. Table 3-9 details the sites used from each program year that were used to calculate the imputed historical operational adjustment for PY2019. The total number of operational adjusted sites from each program year are included, along with the total number of sites the program year contained. For PY2018, six out of eight sites had operational

adjustments and for PY2019, six out of ten sampled sites had operational adjustments, as shown below. For 2017, all sites sampled sites were able to complete operational adjustments.

Table 3-9. Sites used for imputed historical operational adjustment calculations

Program Year	Site IDs	Number of Sites in Imputed Ops Adjustments	Number of Sites in Program Year
PY2017	2017RIG015, 2017RIG047, 2017RIG053, 2017RIG097, 2017RIG098, 2017RIG107	6	6
PY2018	2018RIG78, 2018RIG26, 2018RIG27, 2018RIG43, 2018RIG19, 2018RIG58	6	8
PY2019	2019RIG065N, 2019RIG057N, 2019RIG026S, 2019RIG064S, 2019RIG095S, 2019RIG138N	6	10

4 DATA SOURCES

To support the findings of the study, the team used the following data sources:

- PY2019 tracking data provided by National Grid
- PY2019 parent/child tracking data provided by National Grid
- PY2017 and PY2018 tracking data
- PY2017 and PY2018 program impact evaluation results
- Project files, which typically include one or more of the following: original applications, offer letter, BCR screenings, invoices, minimum requirements documents, technical assistance studies, applicant savings calculations that match claimed savings, and post-installation reports
- Onsite observations and data collection including inspection and verifications of equipment, nameplate data, staff interviews, vendor interviews, spot measurements of various parameters including kW, longer-term measurements, and combustion efficiency
- Metered and/or EMS trend data from operational adjusted sites that participated in the study

5 ANALYSIS AND RESULTS

The RI PY2019 study achieved the target precisions for that individual year's projects as well as while combining the latest three years (PY2017, PY2018, and PY2019). PY2017 program impact evaluation results were finalized in March 2019, and the PY2018 program impact evaluation results were finalized in June 2020. The following subsections provide more details on the PY2019 results.

5.1 PY2019 results

This section provides an overview of the results from comparing PY2019 tracking and evaluated results.

5.1.1 Site-level results

Figure 5-1 and Table 5-2 illustrate the comparison of reported (x-axis) and evaluated (y-axis) annual natural gas savings for each of the ten sites included in the program evaluation sample for PY2019. Table 5-1 shows all sites, and Table 5-2 shows a magnified version for the smaller savings sites.

Ideally, the evaluated savings would always match the reported savings; each chart shows this ideal as a solid green line. Any evaluated sites above the 100% RR line indicate a RR greater than 100%, and those below the 100% RR line indicate a RR less than 100%. Table 5-1 shows the largest evaluated site with tracking savings of about ~379,000 therms per year. That evaluated site almost lies directly on the 100% RR line at 99% RR. However, the largest site only received non-operational adjustments, so these results from this site did not affect the operational realization rate. Site realization rates in Figure 5-2, though smaller therm savings are weighted higher than the larger savings sites in the expansion, where these figures may underrepresent their impact on the PY2019 overall realization rate. Therefore, the 85.3% realization rate is lower than would be expected from the largest savings site achieving a precise realization rate.

As Figure 5-2 illustrates, one evaluated site has a 0% RR rate and is shown along the bottom of both graphs with 0 evaluated therms savings. APPENDIX A summarizes the ten sites for which M&V activities were completed, with statistics such as the site ID, the verified measure description, tracking savings, and RR.

Figure 5-1. PY2019 reported and evaluated annual natural gas savings (all savings sites)

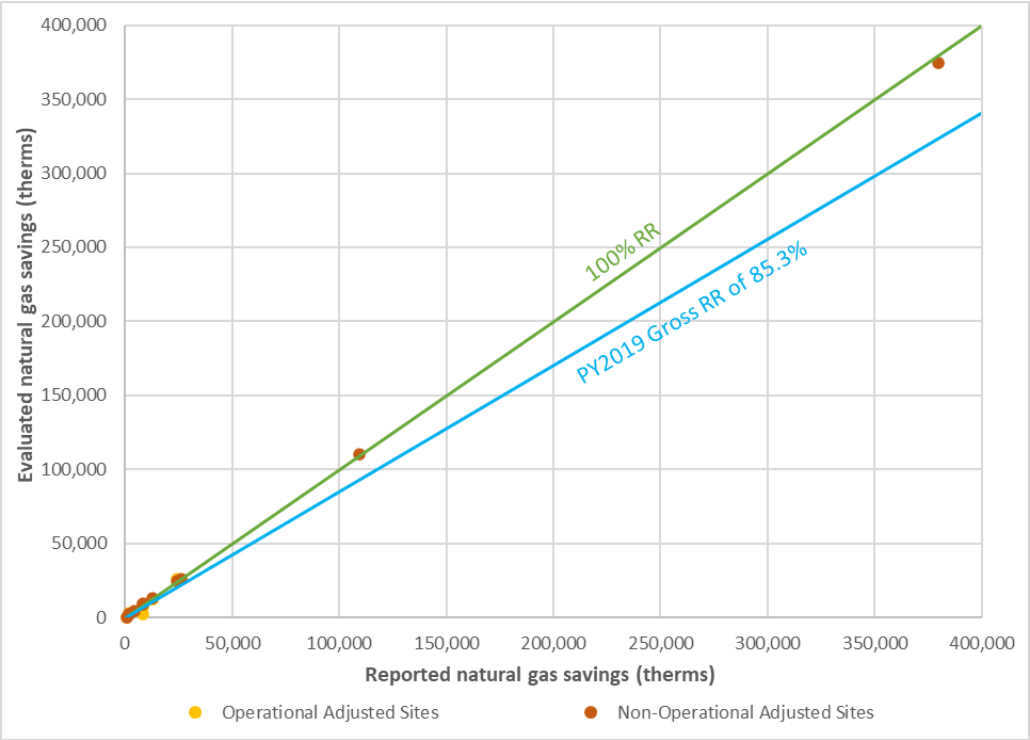
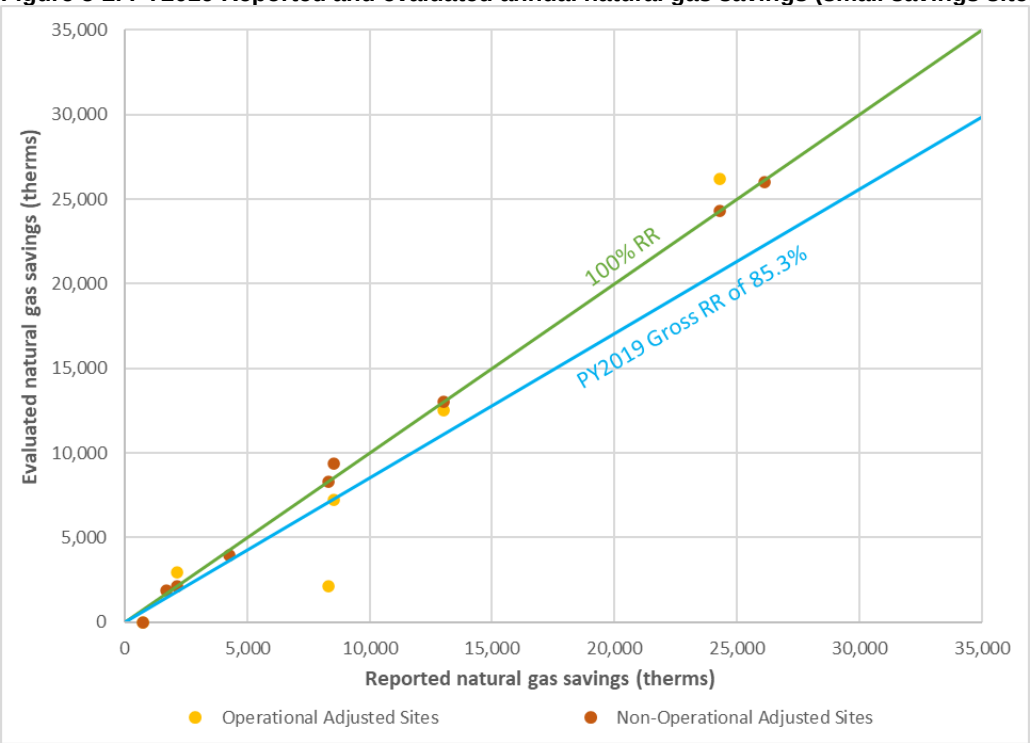


Figure 5-2. PY2019 Reported and evaluated annual natural gas savings (small savings sites only)



5.1.1.1 Non-operational adjustment results

Non-operational adjustments contain baseline, methodology, tracking/admin, technology, and quantity adjustment factors. Baseline, methodology, and tracking/admin factors are completed with project files, tracking data, and preferably a site contact interview. Technology and quantity adjustments are verified and obtained during the first site-visit or through a virtual visit. The team was able to collect information from non-operational adjustments on all 10 of the sampled sites. Table 5-1 shows how the savings change from tracking as adjustments are tabulated; adjustments were applied in the same order for each site. For example, no gas savings (therms) change was made for 2019RIG002N to baseline adjustments from tracking, so the total therm savings remains constant (26,120 therms for tracking and 26,120 therms for baseline). However, a 95 therm savings reduction was made from a methodology adjustment, so the change is found as 26,025 therms in the Methodology column in Table 5-1.

Table 5-1. PY2019 site-level verified savings after non-operational adjustments

Site ID	Tracking Savings (therms)	Verified Savings After Non-Operational Adjustments				
		Baseline	Methodology	Tracking/Admin	Technology	Quantity
2019RIG002N	26,120	26,120	26,025	26,025	26,025	26,025
2019RIG019N	4,246	4,246	4,246	4,246	3,981	3,981
2019RIG026S	2,120	2,120	2,120	2,120	2,120	2,120
2019RIG057N	24,296	24,296	24,296	24,296	24,296	24,296
2019RIG064S	13,038	13,038	13,038	13,038	13,038	13,038
2019RIG065N	8,327	8,327	8,327	8,327	8,327	8,327
2019RIG095S[2]	8,530	8,530	9,404	9,404	9,404	9,404
2019RIG104S/105N[2]	111,021	111,021	111,197	112,026	112,026	112,026
2019RIG108N[2]	379,460	379,460	374,728	374,728	374,728	374,728
2019RIG138N[2]	731	731	731	0	0	0

[2]Ibid, page 4.

The applicant never completed 2019RIG138N, and savings were accidentally claimed in the SEMP program summation table. Savings were removed in tracking and admin adjustment factors due to the administration error. The savings determination was made after the site contact interview. No operational adjustments were made since the project was never completed. However, the 0% realization rate carried through to the operational adjustment analysis.

For 2019RIG019N, evaluators found the installed technology nameplate showed a different heating efficiency than reported within the savings calculator. The efficiency value was updated, and a discrepancy was logged in the technology adjustment factor. Therefore, the savings were reduced after updating the increase in efficiency.

For 2019RIG108N, a formula within the savings calculator did not reference the correct cell. After the cell reference was fixed, the savings were reduced. The discrepancy was logged as a methodology error.

For 2019RIG095S, there was an adjustment factor applied to measure savings where the evaluation engineer did not believe the factor was necessary. The adjustment was made from steam trap survey competed at the facility to determine the number of retrofitted steam traps across the campus based on a sample of 50 steam traps (n=50, N=13,455). Instead of applying the adjustment factor based on the whole campus inspection result the applicant applied, the evaluator used the calculated tracking savings from the statewide 2017 Steam Trap calculator as the evaluated savings. The removal of the adjustment factor discrepancy was logged as a methodology change.

All other sites either do not contain non-operational adjustment factor discrepancies, or the discrepancies affected <1% of tracking savings. Forthcoming sections present the descriptions of the general discrepancies for all sampled sites in operational adjustment factors.

5.1.1.2 Operational adjustment results

The results from a full site evaluation include all adjustment factors found in the non-operational adjustment factors while also including the additional operational adjustments (Operation and HVAC-interactive adjustments). These factors are obtained after logger pickup and after analyzing long-term data (trend data is categorized in these adjustments). Table 5-2 shows the operational adjustments after non-operational adjustment factors are considered from Table 5-1. Of the ten site visits, six were evaluated for operational adjustment factors. The other four sites either had COVID-impacted energy usage or metering was not possible with COVID restrictions from the State of RI or a site facility mandate. The four sites where operational adjustments were not calculated are listed as 'N/A.'

Table 5-2. PY2019 site-level unweighted operational savings after operational adjustments

Site ID	Savings After Non-Operational Adjustments (therms)	Savings After Operational Adjustments	
		Operation	Interactive
2019RIG002N	26,025	N/A	N/A
2019RIG019N	3,981	N/A	N/A
2019RIG026S	2,120	2,953	2,953
2019RIG057N	24,296	26,215	26,215
2019RIG064S	13,038	12,546	12,546
2019RIG065N	8,327	2,152	2,152
2019RIG095S[2]	9,404	7,207	7,207
2019RIG104S/105N[2]	112,026	N/A	N/A
2019RIG108N[2]	374,728	N/A	N/A
2019RIG138N[2]	0	0	0

[5]Ibid, page 18

There were no interactive savings adjustments found for the evaluated sites. All adjustments were found within operation adjustment factors. Steam trap sites were adjusted based on metering data that captured operating hours for the facility or

were adjusted based on operating steam pipe pressure, operation status, calculation methodology input updates from metered or trend data, and boiler efficiency differences. All adjustments are operational that require metered data or observable changes.

2019RIG057N has two measures with individual operational adjustments. Most savings increases are due to more idle hours than estimated from process controls on a production line fan. Savings reductions are from temperature differences for the production line and fixture insulation calculations or operational hour differences for fixture insulation.

2019RIG065N savings were decreased based on decreased operation from the dehumidification unit relative to what was anticipated in applicant estimations. Small discrepancies were found from differences in heating efficiency and supply air temperature.

2019RIG138N, the project that was never completed, does not contain operational adjustments since all adjustments were made from non-operational factors, which effectively reduced the savings to zero. In the site evaluation schema listed in Figure 3-1 for site evaluation type determination, this site would not be included in the operational adjustment factor expansion since meter or trend data was not collected. However, this site was categorized as operational for expansion purposes since the project savings would not have existed whether we completed operational or non-operational site evaluation type.

5.1.1.3 PY2019 combined operational and non-operational results with historical operational adjustments

As previously discussed, a two-sample approach (operational and non-operational adjustment factors) was implemented due to the current health emergency where site operation may be atypical, or site officials may not want external evaluators at their locations to collect meter data (operational adjustment factors), and evaluators would like to achieve, at a minimum, a site interview with measure verification (non-operational adjustment factors). Therefore, two sets of weights are applied in aggregating study results; one set based on the ten sites receiving non-operational adjustments and the second set based on the subset of six sites receiving operational adjustments. Results are tabulated using the methodology described in Section 3.7.2.

Table 5-3. Historical operation adjustment and impact on PY2019

Program Year of Operational Adjustment	# of Sites w/ Ops Adj	Population Savings (therm)	Operational Realization Rate (%)		Historical Ops Adjustment (%)		PY2019 Operational w/ Historical Ops Adjustment
PY2017 OpsAdj	6	1,948,383	104.3%				
PY2018 OpsAdj	6	2,350,739	92.9%	>	94.2%	>	85.6%
PY2019 OpsAdj	6	1,944,204	75.0%				

The historical operational adjustment is calculated after operational/non-operational realization rates and standard errors are calculated in a program year as shown in Table 5-3. The historical operational adjustment is calculated from the prior three years of operational adjustments weighted by the population savings as followed in **Error! Reference source not found.** and discussed in APPENDIX B. The combination of the PY2019 operational adjustment and the PY2017-19 historical

operational adjustment form the PY2019 operational realization rate with the historical operations adjustment found in Table 5-4 under 'Operational RR'. The non-operational and operational adjustments are then combined as explained below.

Table 5-4 presents the discrepancy change percentage of non-operational and operational adjustment factors from tracking and the resulting therms totals for the two adjustment classifications (non-operational and operational). The non-operational realization rate is calculated with tracking savings as the denominator. This realization rate is used to calculate the non-operational realization rate and precisions. In this table, the operational realization rate contains all non-operational adjustment factors for the sites where operational adjustments were collected. This realization rate here is used in the expansion process to compute the overall operations realization rate using the historical adjustment factor. This realization rate is the final site-level realization rate for operational adjustment collected sites. The final site-level realization rate for non-operational only adjustment sites is the realization rate found in the non-operational adjustment column. See Section 5.1.2 for specific non-operational and operational discrepancy percentages when compared with tracking individually that combine to achieve the site level realization rate in Table 5-7.

Table 5-4. PY2019 combined realization rates after discrepancy adjustment factors

		After Non-Operational Adjustments		After Operational Adjustments	
Site ID	**Tracking Savings (therms)**	**Evaluated Savings (therms)**	**Site Level Realization Rate from Tracking (%)**	**Evaluated Savings (therms)**	**Site Level Realization Rate from Non-Operational (%)**
2019RIG002N	26,120	26,025	100%	N/A	N/A
2019RIG019N	4,246	3,981	94%	N/A	N/A
2019RIG026S	2,120	2,120	100%	2,953	139.3%
2019RIG057N	24,296	24,296	100%	26,215	107.9%
2019RIG064S	13,038	13,038	100%	12,546	96.2%
2019RIG065N	8,327	8,327	100%	2,152	25.8%
2019RIG095S[2]	8,530	9,404	110%	7,207	84.5%
2019RIG104S/105N[2]	111,021	112,026	101%	N/A	N/A
2019RIG108N[2]	379,460	374,728	99%	N/A	N/A
2019RIG138N[2]	731	-	0%	-	0.0%
	Non-Operational RR		99.7%	**Operational RR***	85.6%
				Total	85.3%

[2]Ibid, page 4; *Includes imputed historical adjustment

The realization rate for PY2019 RI Custom Gas installations is 85.3%. The realization rate is calculated using previous program year cycle operational adjustments (PY2017 and PY2018) to provide more results, since operational adjustments

were not collected from all PY2019 sites. 2019RIG108N largely impacts the non-operational realization rate as the non-operational weighted savings accounts for 23.3% of weighted savings. The next highest weighted tracking savings is the combination site of 2019RIG104S and 2019RIG105N at 13.3%.

5.1.1.4 Influence tests

The team conducted outlier testing on the results from this study to determine if any sample points could be considered outliers. To do this, the team followed the same methods used in previous rounds of custom gas program impact evaluations to assess leverage and influence.

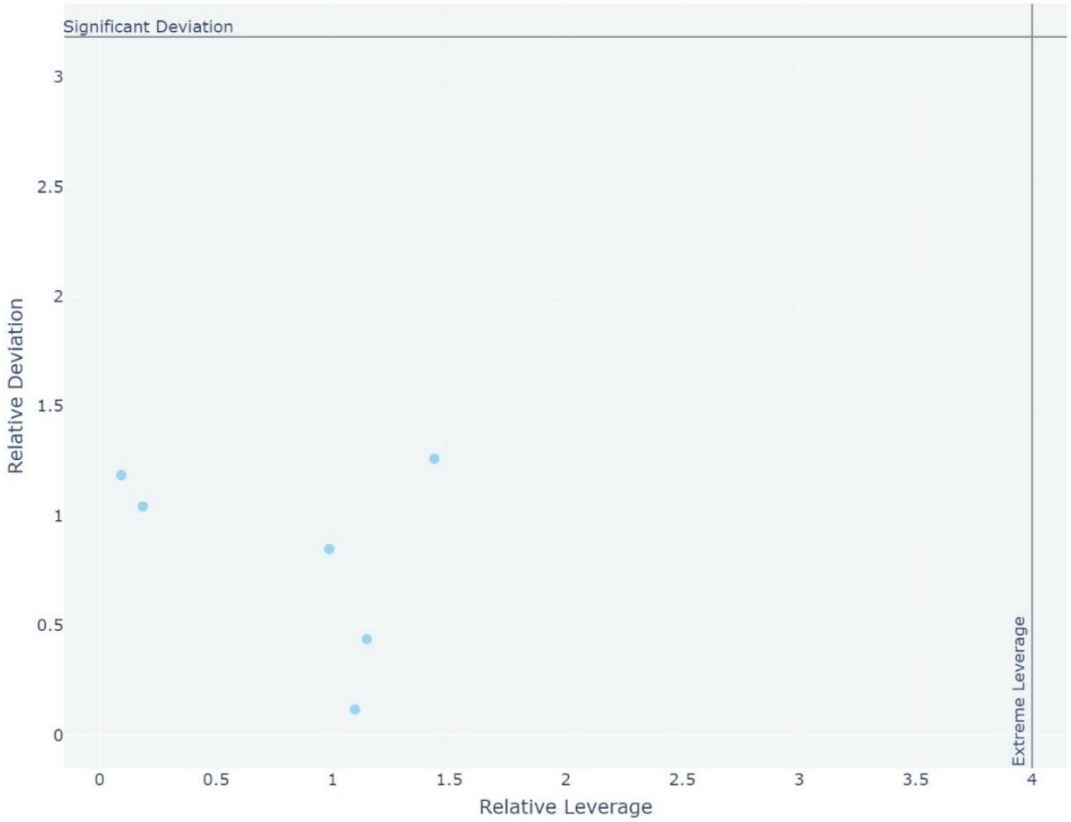
We do not want to exclude or reduce the weight of cases just because they have leverage or influence. We do our best to create uniform leverage, but we can't create a perfectly optimal sample. Moreover, we don't want to automatically reduce the weight of very large cases, especially on a retrospective basis where these cases represent a large fraction of the savings, which is why they are sampled at high rates.

We would want to reduce the weight of a high-leverage case if:

1. It is highly influential, and
2. We have reason to believe its RR is anomalous and somehow much more extreme than is typical of its stratum or the population overall.

Figure 5-3 presents the results of the leverage testing. All sample points were within acceptable levels for deviation and leverage. If any observation points were above the "Significant Deviation" constraint in the plot or greater than the "Extreme Leverage" constraint, then it might have led to action items to reduce the influence. However, all observation points are less than these constraints.

Figure 5-3. Significant deviation and extreme leverage plot of site realization rates



5.1.2 Discrepancy results

For each of the ten sites included in the PY2019 study, the site engineers identified factors that led to differences between the program-reported (tracking) savings and the evaluated savings. The factors are classified into seven categories: baseline, methodology, tracking/administrative, technology, quantity, HVAC interaction, and operational. A more discrete breakdown of possible differences and how they are categorized is presented below in Table 5-5.

Table 5-5. Possible discrepancy factors and their mapping to major categories

Major Discrepancy Category	Discrepancy Definition or Examples
Baseline	Change in the baseline of the post-retrofit condition
Methodology	Accuracy/appropriateness of Analysis Methodology Calculation changes Non-metered data input updates
Tracking/Admin	Accuracy of Tracking Savings Errors during claimed savings input Savings changed but not changed in tracking savings
Technology	Differences in proposed vs. installed technology or measure type
Quantity	Quantity of installed equipment is different
Operational	Boiler combustion efficiency Difference in equipment hours of operation Different equipment load profile Inaccurate pre-project characterization Steam operating pressure difference System optimization or programming not implemented Faulty or improperly installed equipment Operating temperature differences
HVAC Interaction	Interactive effects

The evaluation team used the site-specific, non-operational sampling weights and the sum of site-specific impacts of each discrepancy category to calculate the impact of adjustment factors for differences between the program tracking and evaluated results at the population level. Table 5-6 below presents the discrepancy factors and their impacts. There were no baseline or quantity adjustments discrepancies found in the sample. Most discrepancies are operational, with site-specific comparisons found in Table 5-7. The large methodology discrepancy is from the four sites where methodology discrepancies were found (2019RIG002N, 2019RIG095S, 2019RIG104S, and 2019RIG108N).

Table 5-6. PY2018 weighted discrepancy factors between tracking and evaluated results

Adjustment Factor	Site Counts	Impact on RR	Impact (%)
Baseline	0		0.0%
Methodology	4		-9.4%
Tracking/Admin	1		0.2%
Technology	1		-0.7%
Quantity	0		0.0%
Operational*	5		-15.2%
Interactive*	0		0.0%
Historical Operations Adjustment			10.3%
Total			-14.7%

*Only for the 6 sites with a full site evaluation completed.

Baseline adjustments were not necessary in the two sites where New Construction designation was reclassified as Lost Opportunity or Retrofit was renamed as an Add-On. For example, the change from New Constuction to Lost Opportunity is a term change where the baseline discrepancy would refer to the discrepancy between applicant and evaluator baseline reference (code/ISP) used to calculate savings. Not having baseline adjustments means the applicant used the correct code/ISP when calculating savings.

Adjustment percentages found in Table 5-7 are the magnitude of changes from tracking for each site and are reported at the site level. The combination of non-operational and operational discrepancies sums to the change from tracking to evaluated (realization rate). The percentages are the total adjustments for operational and non-operational adjustments when compared to site-level savings.

Table 5-7. Non-operational and operational weighted discrepancies – PY2019

Site ID	Tracking Savings (therms)	Evaluated Savings (therms)	Site Level Discrepancies		Realization Rate (%)
			Non-Operational	Operational	
2019RIG002N	26,120	26,025	-0.4%	N/A	99.6%
2019RIG019N	4,246	3,981	-6.2%	N/A	93.8%
2019RIG026S	2,120	2,953	0.0%	39.3%	139.3%
2019RIG057N	24,296	26,215	0.0%	7.9%	107.9%
2019RIG064S	13,038	12,546	0.0%	-3.8%	96.2%
2019RIG065N	8,327	2,152	0.0%	-74.2%	25.8%
2019RIG095S[2]	8,530	7,207	10.2%	-25.7%	84.5%
2019RIG104/105N[2]	111,021	112,026	0.9%	N/A	100.9%
2019RIG108N[2]	379,460	374,728	-1.2%	N/A	98.8%
2019RIG138N[2]	731	0	-100.0%	0.0%	0.0%

[2]Ibid, page 4.

The three largest percent of tracking savings sites' discrepancies factors are discussed below:

2019RIG026S discrepancy stems from higher metered operating hours. There was a smaller increase in savings due to a slightly lower boiler efficiency than reported in savings calculations.

2019RIG065S did not dehumidify at the rate expected from applicant reported usage. Savings result from the use of pool dehumidification which operated considerably less than claimed.

2019RIG138N was never completed. The site was accidentally claimed within a large portfolio of projects. The claimed savings were then removed as a tracking/admin discrepancy.

Section 3 of each site report presents detailed information on site-specific differences, which is included in APPENDIX D.

5.2 Combined three-year rolling results (PY2017, PY2018, and PY2019)

The evaluators calculated the gross RR and precisions using the results from PY2017, PY2018, and PY2019. The results are summarized in Table 5-7. The PY2019 relative precision improved greatly when compared to PY2018 using imputed historical adjustment factors (Section 3.7.2). Higher than the usual variance in PY2018 when combined with the inclusion of the two-sample approach (non-operational vs. operational) that introduces a lower than a typical number of operational adjustment observations causes a higher impact on expansion results from each site. It was determined that historical operations data would improve precisions since evaluators could not collect operation data for all sites during the current health emergency. The historical adjustments in this program year are calculated using operational adjustments from PY2017, PY2018, and a portion of the PY2019 sample. The imputed historical adjustments, when combined with the lower variance from PY2019 sites, improved the overall relative precision.

Table 5-8. Three-year rolling plan results and statistics

Parameter	PY2017	PY2018	PY2019	PYs 2017+2018+2019
Tracking Savings (therms)	1,948,383	2,350,739	1,944,204	6,243,326
Non-Operational Sample Size	6	8	10	24
Operational Sample Size[9]	6	6	6	18
Realization Rate (RR)	92.7%	83.3%	85.3%	86.9%
Relative Precision @ 80% CI (%)	±2.3%	±22.6%	±4.5%	±6.8%

The PY2019 individual year results meet the design precision targets proposed and presented above, in Section 3.1.2 and 3.1.3, for a standalone evaluation program year cycle at 85.3% RR and ±4.5% precision. The relative precision of the overall 3 year rolling sample RR (±6.8%) also meets the design precision targets proposed and presented above, in Section 3.1.2 and 3.1.3, after combining the current evaluation program year (2019) and the prior two years of results (2017 and 2018).

The original sample was designed to estimate the overall realization rate of the program by combining results from three program year evaluation studies (PYs 2017, 2018, and 2019) to achieve the agreed-upon precision targets of ±20% relative precision at 80% confidence for a custom gas study. In this case, the precision target was achieved by combining results from PY2017, PY2018, and PY2019 (based on the methodology discussed in Section 3). Table 5-8 shows the individual PY2017, PY2018, and PY2019 results along with the combined three-year rolling program evaluation for PY2017, PY2018, and PY2019. Table 5-9 shows the non-operational and operational realization rates that are used to calculate the three-year rolling realization rate. PY2019 includes imputed historical operation data from PY2017 and the sites that received operational adjustments for PY2018.

[9] The minimum sample size of each of the inner samples (sites with operational adjustments) dictates the overall sample size of the year for combined results.

Table 5-9. RRs used to calculate three-year rolling RR

Program Year	Tracking Savings (therms)	Non-Operational RR	Operational	Combined
PY2017	1,948,383	88.9%	104.3%	92.7%
PY2018	2,350,739	89.7%	92.9%	83.3%
PY2019*	1,944,204	99.7%	85.6%	85.3%
3-Year Rolling	6,243,326	92.6%	94.2%	86.9%

*PY2019 includes imputed historical adjustments (operational results from PY2017, PY2018, and PY2019)

6 CONCLUSIONS, RECOMMENDATIONS, AND CONSIDERATIONS

6.1 Conclusions

6.1.1.1 PY2019 Performance

The program continues to generate significant natural gas savings. In RI, the PY2019 custom gas projects saved an estimated 2.04 million therms (adjusted gross savings) annually, with 85.3% of the program year savings realized based on the program evaluation sample for RI PY2019 sites. The current results are accurate within state and regulatory standards and provide adequate planning and program reporting savings estimates.

A more detailed explanation of the PY2019 performance is found in Section 5.1.1.3. Site-specific details are shown in APPENDIX A. More details on the PY2019 results are presented in the sections below, and each site report included in APPENDIX D.

6.1.1.2 Combined three-year rolling (PY2017, PY2018, & PY2019) Performance

Combined over the three-year rolling sampling period, the program realized gross savings of 6.24 million therms, with 86.9% savings realized as shown in Table 1-1. The current three-year rolling results are accurate within state and regulatory standards and provide adequate planning and program reporting savings estimates.

A more detailed explanation of the combined three-year results is found in Section 5.2.

6.2 Recommendations

6.2.1.1 R1: Realization rate

DNV recommends National Grid to use the PY2017, PY2018, and PY2019 combined RR of 86.9% for planning and program reporting, starting with PY2022 and continuing to subsequent years until new program impact evaluation study results are available. The applicable RRs are noted in Table 1-1. This recommendation was based on the following factors:

- When PY2017 (92.7%), PY2018 (83.3%), and PY2019 (85.3%) results are pooled, the study produced statewide results that met precision targets of ±20% relative precision at 80% confidence (actual: ±6.8% at 80% confidence level).

Based on the results listed for PY2019, an individual program year sampling Error Ratio Target of 0.55 for non-steam trap and SEMP projects has been recommended for the 2020 RI Custom Gas Impact Evaluation to achieve the next three-year rolling savings program evaluation precision targets. Steam traps are recommended 0.65 for the error ratio target.

6.2.1.2 R2: Memo should accompany savings changes in submitted calculators

DNV recommends that project savings calculators have an accompanying memo that documents changes when savings deviate from submitted calculators or changes during the final stages of compiling claimed savings. Site evaluations depend on final savings calculators to understand assumptions and methodology to construct a site evaluation plan. An accompanying memo for savings changes would provide necessary information on what changed and the reasons for those changes. A specific example of this problem is listed below:

2019RIG108N: The project folder for this measure contained several calculators labeled 'FINAL.' After many attempts by both National Grid and DNV evaluators, the evaluation team identified the proper calculator used for claimed savings in tracking data. The major change between submitted calculators was reduced savings due to gas usage moved from the main boiler heat plant to distributed boilers that were not considered in the original calculations. Due to the multiple calculators, the evaluation team is recommending a memo written for modifications to calculators after submittals or final versions of the calculators are created to document reasons for changes. There should be a clearly documented path to

explaining the final savings claimed in the tracking system when multiple versions of the savings calculations exist in the project files

6.2.1.3 R3: Post-inspection verification for HVAC EMS projects

DNV recommends HVAC EMS project applicants submit EMS screenshots, schedule printouts, or trend data to National Grid representatives at the time of project completion or before incentives are paid to customers for comparison of installed condition with original design documents submitted for initial project approval. The submissions will provide an opportunity for implementation teams to identify differences from original submission and installed condition to determine operation is within original parameters or if modifications must be made to the installation or incentive due to project changes during installation. A specific example of this problem is listed below:

2019RIG065N: This project installed a new dehumidification system to serve a swimming pool at the recreational complex of an elementary school. The installed dehumidification system is based around a hot-gas heat recovery system. The project saves energy when the installed system is in dehumidification mode by utilizing the hot-gas heat recovery system to reheat the supply air to reduce HVAC energy consumption. Several parameters were not as stated in the as-built documents submitted during the project and most major discrepancies would have been avoided with the submissions of EMS screenshots, schedule printouts, or trend data submissions.

6.2.1.4 R4: Evaluate lifetime savings for PY2020 projects

DNV recommends evaluating lifetime savings and reporting them at the site level in all future custom gas site evaluations. This is to prepare for reporting the new lifetime savings goal as National Grid transitioned away from annual savings goals to lifetime savings goals beginning in PY2021. Reporting them in every study is critical to producing 3-year rolling-based lifetime savings results. Should lifetime savings results be available beginning with PY2020 evaluated projects, then a standard 3-year rolling reporting cycle would be available after the PY2022 program evaluation.

6.3 Considerations

6.3.1.1 C1: SEMP sampling in next program year evaluation

DNV proposes a three-stage process to include SEMP projects in the next program year evaluation sampling for the 2020 RI Custom Gas Impact Evaluation. Stage 1 will identify SEMP projects and categorize them as a separate category using the line items as they are claimed in tracking savings and creating site IDs as traditionally completed in recent program evaluations. DNV will identify which SEMP projects are identified for site evaluation, which leads to Stage 2. Stage 2 will require project files to disaggregate sampled SEMP sites to the measure level. DNV will sample again with the disaggregated sampling breakout at the measure level to identify measure-specific projects to evaluate that will provide results within statewide precision targets. Stage 3 is an extra step only if an identified site in Stage 2 proves to be a much larger burden on evaluators and customers than need be. A second disaggregation would then occur at a level needed to reduce the burden on a case-by-case basis. Sub-categories (steam trap or non-stream trap) will be provided for each claimed savings and measure row item as traditionally completed, and stratification (number of sites per stratum) will follow the same methodology as has been completed in previous program evaluations.

6.3.1.2 C2: Dual baseline inclusion for lifetime savings

As a preparation for the coming lifetime savings analysis in future program year evaluations, DNV proposes that the National Grid implementation team consider dual baseline calculations to more accurately reflect the claimed savings over the lifetime of the retrofit measure.

APPENDIX A. SITE EVALUATION RESULTS & REALIZATION RATES

This Appendix includes the site ID, the verified measure description, tracking savings and site RR that were used to calculate over realization rates for the program.

Table 6-1. Evaluated site summary

Sample ID	Applications	Measure Description	Site Evaluation Type	Tracking Savings	Evaluated Savings	Realization Rate
2019RIG002N	8375229, 10722806, 8175997	EMS/DCV, Condensing Water Heater, Vent Hood w/ Controls	Non-Operational	26,120	26,025	104.8%
2019RIG019N	9581345	DCV	Non-Operational	4,246	3,981	93.8%
2019RIG026S	8884415	Steam Traps	Operational	2,120	2,120	139.3%
2019RIG057N	10426524, 9257490	Insulation, Process Controls	Operational	24,296	24,296	107.9%
2019RIG064S	8651739	Steam Traps	Operational	13,038	13,038	96.2%
2019RIG065N	7722897	Dehumidification	Operational	8,327	8,327	25.8%
2019RIG095S[2]	8116687, 9208255, 9494451	Steam Traps	Operational	8,530	9,404	73.9%
2019RIG104S/105N[2]	8116687, 9208255, 9494451	Heat Recovery, Steam Trap, HVAC Controls	Non-Operational	111,021	112,026	100.9%
2019RIG108N[2]	9494451	Heat Plant Improvements and Hot Water Loop Efficiency	Operational*	379,460	374,728	99.0%
2019RIG138N[2]	9494451	HVAC Control Schedules	Operational	731	0	0%

[2]Ibid, page 4; *Billing analysis was performed.

APPENDIX B. ADJUSTING GROSS REALIZATION RATE STANDARD ERRORS FOR IMPUTED OPERATING ADJUSTMENT

This appendix explains the process for calculating the current and three-year realization rates incorporating imputed operational adjustment for part of the third-year sample.

Basic structure

We have samples for three successive periods: 1, 2, and 3. In this evaluation these samples are 1) PY2016 (P80), 2) PY2017/18 (P88), and 3) PY2018/19 (PY1819). Samples 1 and 2 were full samples. Sample 3 has non-operational results for all sites and operational results for only a subset of sites.

Full-sample weights for period 3 are calculated in the usual way, as the ratio of population count to sample count within the sampling cell that contains a particular site, where the sample count is for all sites in the sample.

Notation

w_j = full-sample weight for sample site j in the period-3 sample

S_y = population tracked savings of period y

S_T = population tracked savings for all 3 periods combined

$= S_1 + S_2 + S_3$

q_y = period-y savings as a fraction of the 3-period total

$= S_y/S_T$

f_{g3} = fraction of Period-3 savings represented by “good” sites, i.e. those with operational data

= (full-sample-weighted savings of Period-3 sample sites with operational data)/(total full-sample weighted savings for Period 3)

S_{Tg} = total savings for population represented by sites with operational data, across all samples

$= S_1 + S_2 + f_{g3}\, S_3$

RR_{oy} = operational-only realization rate for the period-y sample

RR_{Ny} = non-operational-only realization rate for the period-y sample

RR_{og3} = operational-only realization rate for the population represented by good sites in the period-3 sample, those with operational data

RR_{ob3} = imputed operational-only realization rate for the population represented by bad sites in the period-3 sample, those without operational data

SE(X) = standard error of estimate X

RSE(X) = relative standard error of estimate X

=SE(X)/X

Period 3 operational realization rate RR_{o3}

- For the portion of the population represented by sampled sites with operational adjustments ("good" sites g), RR_{og3} is directly calculated from the sample, using the full sample weights w_j. That is, RR_{og3} is the weighted sum of verified gross savings, divided by the weighted sum of tracked gross savings.
- For sampled sites without operational adjustment ("bad" sites b), RR_{ob3} is imputed as

  $RR_{ob3} = (S_1RR_{o1} + S_2RR_{o2} + f_{g3}\ S_3RR_{og3})/S_{Tg}$

  That is, all available sites with operational data are used to impute the RR for the uncovered portion of the period-3 population, with the RR from different periods weighted by the savings it represented.
- Overall Operational Adjustment for Period 3 is calculated as

  $RR_{o3} = f_{g3}\ RR_{og3} + (1\text{-}f_{g3})RR_{ob3}$.

  That is, the operational adjustment for the directly represented portions of the population and the remainder are combined in proportion to their shares of period-3 tracked savings. This formula can be expanded as

  $RR_{o3} = f_{g3}\ RR_{og3} + (1\text{-}f_{g3})\ (S_1RR_{o1} + S_2RR_{o2} + f_{g3}\ S_3RR_{og3})/S_{Tg}$
  $= (1 + (1\text{-}f_{g3})\ S_3/S_{Tg})f_{g3}RR_{og3} + (1\text{-}f_{g3})(S_1/S_{Tg})RR_{o1} + (1\text{-}f_{g3})(S_2/S_{Tg})RR_{o2})$
  $= a_{og3}\ RR_{og3} + a_1RR_{o1} + a_2RR_{o2}$,

  Where

  $a_{og3} = (1 + (1\text{-}f_{g3})\ S_3/S_{Tg})f_{g3}$
  $a_1 = (1\text{-}f_{g3})(S_1/S_{Tg})$
  $a_2 = (1\text{-}f_{g3})(S_2/S_{Tg})$

  This expansion expresses the overall Period 3 operational realization rate as a weighted average of three independently estimated terms, the directly observed operational realization rate from each period. The factors multiplying the three realization rates have the property that

  $a_{og3} + a_1 + a_2 = 1$.

- Standard error of Period 3 realization rate: The standard error is calculated from the individual standard errors as

  $SE(RR_{o3}) = sqrt[a_{og3}^2\ SE^2(RR_{og3}) + a_1^2\ SE^2(RR_{o1}) + a_2^2\ SE^2(RR_{o2})]$

  This is true because the three RRs at step 3 are from independent samples.

Period 3 combined RR

- **The non-operational realization rate** RR_{N3} is calculated from the full sample using the full sample weights and the non-operational adjusted savings for the sample, via the usual formulas.

- **The Overall RR** is the product of the operational and non-operational RR

  $RR_3 = RR_{o3}\ \ RR_{N3}$

- **Standard error:** First calculate the relative standard error
  a. $RSE(RR_3) = sqrt[RSE^2(RR_{o3}) + RSE^2(RR_{N3})]$

  This formula is approximately correct, assuming that even though RR_o and RR_N are from a common sample, they are essentially unrelated so can be treated as independent.

  The standard error is then calculated from the RSE.

  b. $SE(RR_3) = RR_3\, RSE(RR_3)$

3-year combined RR

Preferred calculation

$RR_{1\text{-}3} = (S_1RR_1 + S_2RR_2 + S_3RR_3)/S_T$
$\quad = q_1RR_1 + q_2RR_2 + q_3RR_3$

That is, the three-year RR is the savings-weighted average of the three separately estimated RRs.

This calculation produces an overall realization rate for each period, then combines these across periods. This approach is the natural one, combining the historical overall results with the most recent, consistent with our general method for three-year rolling realization rate calculation, and is therefore the preferred way to produce the three-year value.

However, because the third term RR_3 is determined in part from the operational portions of RR_1 and RR_2, the three are not independent estimates. Moreover, there's no obvious way to express the calculation as the sum of independent estimates, as would be needed to produce the standard error. We therefore look at an alternative calculation for purposes of standard error calculation only.

SE calculation

We use the standard error of an alternative calculation as an approximate to the standard error of the preferred calculation. The alternative calculation would be to calculate separate operational and non-operational realization rates for the three-year period and multiply them. We calculate this SE. We can check how different the results are, but the SEs or inflation of SE ought to be ballpark the same.

Alternative RR calculation for SE calculation only

- 3-year operational realization rate
  $RR_{o1\text{-}3} = q_1RR_{o1} + q_2RR_{o2} + q_3RR_{o3}$
- 3-year non-operational realization rate
  $RR_{N1\text{-}3} = q_1RR_{N1} + q_2RR_{N2} + q_3RR_{N3}$
- Combined 3-year realization rate
  $RR_{1\text{-}3} = RR_{o1\text{-}3}\ \ RR_{N1\text{-}3}$

Standard error calculations for the alternative RR calculation

Non-operational three-period realization rate SE

The non-operational three-period realization rate is the savings-weighted average of the separate period realization rates. Since these are all independent, we can use the formula for combinations of independent estimates to produce the standard error.

$$SE(RR_{N1\text{-}3}) = sqrt[q_1^2\, SE^2(RR_{N1}) + q_2^2\, SE^2(RR_{N2}) + q_3^2\, SE^2(RR_{N3})]$$

Operational three-period realization rate SE

The operational realization rate is also the savings-weighted average of the three periods' operational realization rates, but these aren't all independent. We rearrange the formula to express the operational realization rate as a combination of independent estimates.

$$RR_{o1\text{-}3} = q_1\, RR_{o1} + q_2\, RR_{o2} + q_3 RR_{o3}$$
$$= (q_1 + a_1\ q_3)\, RR_{o1} + (q_2 + a_2\, q_3)\, RR_{o2} + q_3\, a_{og3}\, RR_{og3}$$

where the factors a_x are as defined above. With this expression of the three-period operational realization rate as a combination of independent estimates, is standard error is calculated as

$$SE(RR_{o1\text{-}3}) = sqrt[(q_1 + a_1\ q_3)^2\, SE^2(RR_{O1}) + (q_2 + a_2\, q_3)^2\, SE^2(RR_{O2}) + (q_3\, a_{og3})^2\, SE^2(RR_{O3})].$$

Relative standard error of overall three-period realization rate

By the same argument as above, the relative standard errors of the two realization rate factors are combined as if they were independent estimates. This is approximately correct, assuming that even though RR_o and RR_N are from a common sample, they are essentially unrelated so can be treated as independent.

$$RSE(RR_{1\text{-}3}) = sqrt[RSE^2(RR_{o1\text{-}3}) + RSE^2(RR_{N1\text{-}3})]$$

Standard error of the three-year realization rate

$$SE(RR_{1\text{-}3}) = RR_3\, RSE(RR_{1\text{-}3})$$

Level of aggregation for applying the formulas

Calculating Period 3 and three-period realization rates

The formulas for calculating the Period 3 operational realization rate RR_{o3}, the Period 3 overall realization rate RR_o, and the preferred three-period overall realization rate $RR_{1\text{-}3}$ are applied separately for each reporting category of realization rate. Typically, each reporting category includes sample points from multiple sampling cells.

For reporting categories with no Period 3 sample that has operational data the same formulas are used, with Period 3 contributing nothing to the three-period operational realization rate. For this study all of the reporting categories used had at least one sample point with operational data.

APPENDIX C. CUSTOMER COMMUNICATION PROTOCOL

Memo to: National Grid

Prepared by: Srikar Kaligotla, Chad Telarico and Olav Hegland
DNV

Date: 10/5/2020

Copied to:
Wendy Todd, Glenn Gavi, Jeffrey Zynda,
DNV

RI – Customer Outreach Communications Protocol

Introduction

In response to the COVID-19 pandemic, National Grid has provided guidelines for customer communications as it relates to all studies that are underway or in the planning stages. National Grid recently provided additional guidelines for impact studies, specifically related to short- and long-term planning activities. This memo represents our response and implementation of these guidelines, specifically as they relate to RI PY2018 Custom Electric study applications.

The following summarizes key takeaways from National Grid guidelines[10]:

Customer Outreach

- Evaluation will not conduct any outreach of any kind to health care facilities, such as hospitals, nursing homes, or assisted living facilities.
- Evaluation will not ask anyone to go into their facility if they are not already there or create a situation where a customer feels compelled to provide data during this difficult time.
- Evaluation will be mindful that response rates could differ from normal circumstances and any data we collect may be influenced by the current COVID-19 situation. We will use our best judgment about what types of data would be meaningful to try to collect.
- Evaluation may contact internal implementation staff (i.e., National Grid employees) regarding evaluation topics, but any communication will be sensitive to the fact that National Grid staff may be overwhelmed with other requests or immediate issues related to COVID19. We will be patient and understanding in communications.
- DNV will present all the results from the outreach to National Grid and seek their approval to visit the site.
- DNV will provide a 1-week notice to the QA vendor about the upcoming site visit.
- A site will be replaced with a backup site when there is a customer refusal.

[10] National Grid may change these guidelines over time, which may warrant revisions to this protocol.

Methodology

In response to the guidelines, we have prepared the following protocols to be used by all evaluators involved with the study for conducting customer outreach and for setting appropriate customer expectations. This protocol includes all planned customer interaction steps in the recruiting phase ONLY of the study, including:

- Initial outreach email
- Phone script for initial customer phone call

The evaluation team will first attempt to reach out to the site contact via email, followed by a phone call to solidify the customer's willingness to participate in the study, and an outline of the customer's involvement in the next steps. The protocol consists of the following steps:

- **Step 1**: Send an initial outreach email to the site contact describing the reason and objectives for reaching out.
- **Step 2:** If the contact does not reply within 48 hours after the initial contact, we will follow-up with a phone call. If the contact answers the call, we will use the same script. If the contact does not answer the call, we will leave a brief voicemail referring to the initial email and requesting a call-back. If we cannot make contact after Step 2, we will notify National Grid to discuss options.
- **Step 3:** We will report all responses gathered through communications described in steps 1 and 2 above, to National Grid.
- **Step 4:** We will monitor the situation and, based on recommendations provided by ***National Grid,*** we will adjust the language and/or approach for future follow up.

Step 1: Initial outreach email

Group 1: New Sites (no previous calls were made or emailed)

APPLICABLE PROJECTS:

- RI Custom Electric PY2019
- Upstream Lighting 2019
- RI Custom Gas PY2019

Email the customer for recruiting purposes as soon as National Grid approves the script provided below. The purpose of this step is to reach the site contact and gather site-specific details on the measure(s) installed and possibly collect some intelligence on the status of the Business right now.

Script:

Hello, my name is _____, and I am an Engineer at DNV. I'm contacting you on behalf of National Grid.

I would like to talk to you about the energy efficiency measure(s) installed at your facility through National Grid's [Program Name with program year]. National Grid has contracted independent consultants from my firm, DNV/ERS/DMI, to contact a sample of program participants and evaluate the performance of this energy-efficiency equipment. Our records indicate the following:

National Grid Application Number	*Installed Equipment Description*	*Estimated Annual Energy Savings (kWh)*	*Year of Completion*
1234567	*Variable Speed Drives on 2 Rooftop unit fans*	*45,678*	*2018*

*I am mainly interested in gathering more details on this equipment, its operation, and any other relevant details over the phone or by email. I would like to set up a phone call with you to go over these details at a time that is convenient for you. Please let me know if [**Date & Time**] would work for you. If not, can you please provide a time that works best for you?*

If you are not the right person, kindly forward the contact information of the person who is most familiar with this equipment. If you have any questions or concerns, please don't hesitate to contact me @email, ph.no).

I really appreciate your time and thank you for participating in National Grid's Energy Efficiency Program.

Best regards,

[Evaluation Engineer's Name]

Group 2: Previously contacted sites (initial calls made in March but not recruited) & Previously Recruited but not visited sites.

APPLICABLE PROJECTS:

- RI Custom Electric PY2018

Script:

Hello, my name is _____, and I am an Engineer at DNV. I'm contacting you on behalf of National Grid.

We spoke [or left a voicemail or emailed] a few months ago about an evaluation of the energy efficiency equipment XXXXX [specify the equipment details] that was installed at your facility through RI National Grid's Custom Electric Energy Efficiency Program in 2018. I would like to talk to you about that equipment at a time that is convenient for you.

*I am mainly interested in gathering more details on this equipment, its operation, and any other relevant details over the phone or by email. I would like to set up a phone call with you to go over these details. Please let me know if [**Date & Time**] would work for you. If not, can you please provide a time that works best for you?*

If you are not the right person, kindly forward the contact information of the person who is most familiar with this equipment. If you have any questions or concerns, please don't hesitate to contact me @email, ph.no).

I really appreciate your time and thank you for participating in National Grid's Energy Efficiency Program.

Best regards,

[Evaluation Engineer's Name]

Step 2: Follow-up phone call

The purpose of this call is to follow up on the objectives listed in the initial email, including:

a) Confirm that the contact person listed in the tracking data is the correct person to speak with
b) If not, identify the person most familiar with the installed measures and request an introduction and/or contact information for that person
c) The call should be brief and should aim at confirming the site's willingness to participate in the study and to briefly go over what's involved and briefly outline the next steps
d) If the site contact has time, also ask a few questions regarding facility operation during the COVID pandemic (Current/Planned facility operations)

Note for all engineers: continuously monitor conversation tone and acceptance to talk. If the site contact does not have time to speak, gently see if there is a time more convenient for them. Thank and terminate the call if the site contact is audibly frustrated.

If the site contact answers the phone:

Hello, my name is _____, and I am an Engineer at DNV/ERS/DMI. I'm contacting you on behalf of National Grid who has contracted with my firm, DNV/ERS/DMI, to contact a sample of program participants and evaluate the performance of energy-efficiency measures installed at your facility as part of National Grid's [Program Name with program year].

*Screener Question: Based on our records, these measures are installed at a [Facility Type]. Can you please confirm the same? [**Thank and terminate** if it's a* hospital, nursing home, or an assisted living facility]

*[**Q1**]: National Grid had listed you as the contact person and I want to confirm that you are the person I should speak to about these measures. [if NO, go to **Q2**, if YES, go to **Q3**]*

*[**Q2**] I would appreciate it if you could provide the name and contact information of the person I should speak with, or make an introduction [using the email I sent as background]?*

*[**Q3**]: Great! We don't have to go into details in this call, however if you have a few minutes, I'd like to go over the steps in the evaluation process. It should take no more than 5 minutes. Is that ok with you? [If NO, go to **Q4**, if **YES**, go to the **Scope Summary**]*

*[**Scope Summary**] The objective of the evaluation is to 1) verify the installed measures (listed in my email) and 2) to evaluate the associated kWh savings listed in the National Grid application(s).*

The evaluation process typically involves us reviewing all the project or application materials and conduct a site-visit either virtually or in-person or both, to verify and measure the actual energy usage of the installed equipment. If available, we will take account of any data sources available at the site. These sources include: Energy Management systems, sub-meters, trend data, etc. I'd appreciate it if you could describe any data sources available at your facility.

*I should mention that this process does not affect any incentives you have already received for the installed measures. Instead, it provides National Grid with general and independent feedback on how these measures perform and is mainly used by National Grid to improve the design of future energy efficiency programs. [When done, go to **Q5**]*

*[**Q4**]: Ok, that's no problem. Let's instead schedule a time that works better for you. Does [**date, time**] work for you? If not, can you suggest a time that works for you?*

*[**Q5**] Do you have any questions about this and can we schedule a time to go over the specifics of the measures installed? [if YES, go to **Q6**. If NO, go to **Refusal Wrap-up**]*

*[**Q6**] Great, thank you! Before we wrap up, I'd like to ask a few questions about your facility's operations. Is that ok? [If YES, ask **Q7-Q10**, if NO go to **Next Steps Wrap-up** and cover Q7-Q10 in the next call]*

Q7. Are you currently onsite or working remotely?

01 [Yes] *How often do you come onsite?*	*When done recording the answers Go to **[Q8]***
02 [No]	
97 [Don't know]	
98 [Refused]	

Q8. Are there other staff currently at the building?

01 [Yes] *How many on average and how does this relate to normal or historical staffing levels?*	*When done recording the answers Go to **[Q9]***
02 [No]	
03 [Don't know]	
04 [Refused]	

Q9. Has there been a significant change in building operations due to the ongoing health crisis?

01 [Business running as normal] – ***Skip Q10***	*When done recording the answers Go to **[Q10]***
02 [Business operations have changed] *How has it impacted the Business and how has that affected the measure(s)*	
03 [Don't know]	
04 [Refused]	

Q10. Can you describe any short or near-term plans for reopening?

	*When done recording the answers Go to **[Next Steps Wrap Up]***

[*Answer to **Q5** was NO, **Refusal Wrap up]***

Thank you for taking the time to speak with me today and thank you for participating in National Grid's energy efficiency programs.

*[**Next Steps Wrap up**]*

Thank you so much for your time, today. I will review these responses with my Team and get back to you about an onsite visit shortly. Is there a good time for me to call you back this week or would you prefer an email?

On behalf of National Grid, we would like to thank you for taking the time to speak with me today and for participating in their energy efficiency programs.

[if follow up call date and time is confirmed]

*I look forward to speaking with you on **[date, time**].*

[if follow up call date and time is <u>not</u> confirmed]

I'd appreciate it if you can get back to me with a date and time for a follow-up conversation.

*In the meantime, don't hesitate to contact me or [**PA lead**] if you have any questions.*

Thank you!

APPENDIX D. SITE REPORTS

About DNV

DNV is a global quality assurance and risk management company. Driven by our purpose of safeguarding life, property and the environment, we enable our customers to advance the safety and sustainability of their Business. We provide classification, technical assurance, software and independent expert advisory services to the maritime, oil & gas, power and renewables industries. We also provide certification, supply chain and data management services to customers across a wide range of industries. Operating in more than 100 countries, our experts are dedicated to helping customers make the world safer, smarter and greener.