FINAL REPORT

Impact Evaluation of Large Energy User Program (LEUP) for Projects Completed in Years 2021 - 2022

New Jersey Board of Public Utilities

Date: September 12, 2024

Table of contents

List of figures

List of tables

ABSTRACT

DNV conducted an evaluation of the New Jersey (NJ) Large Energy User Program (LEUP) offered by the 8/22/of Public Utilities (NJBPU) for projects initiated in PY2016 through PY2020 and completed in calendar year (CY) CY2021 and CY2022. This study included gross and net impact evaluation research as well as a process evaluation. This evaluation was completed in collaboration with our partner firm Warren Engineering.

The LEUP program offers New Jersey's largest energy-consuming customers (greater than $5,000,000 in energy utility bills) incentives for scoping and installing energy efficiency measures. Participants must prepare and have approved energy efficiency plans that describe the identified measures, cost, and schedule of implementation, as well as the measurement and verification methods that will be used to determine the actual achieved savings. The impact evaluation leveraged the extensive record of M&V provided by the customers to support their savings claims for a sample of 16 projects from a population of 86 sites coupled with in-depth interviews to determine gross and net savings. Two site visits were completed to verify the information provided in the project files aligned with site conditions.

Table AB-1 and Table AB-2 present the program-level evaluation metrics of the retrospective first year and lifetime gross savings**,** respectively, for electric and natural gas energy and electric demand for LEUP projects Initiated in PY2016-PY2020 and completed CY2021-CY2022. These findings are based on confirmation of site conditions through an in-depth interview (IDI) and a review for a statically selected sample of project sites. While the electric relative precision falls short of the target of ±10%, the results are representative of the program sites.

Table AB-1. Summary of LEUP retrospective first year savings

	First year							
Fuel	Tracking savings	Evaluated savings	GRR	Relative precision	Confidence interval	Lower limit	Upper limit	Average tracking savings per location
Electric energy (MWh)	28,576	22,283	78%	23%	90%	17,064	27,502	370,253
Electric demand (kW)	3,881	2,938	76%	25%	80%	2,201	3,675	50
Natural gas (x1,000 therms)	3,828	2,361	62%	9%	90%	2,145	2,577	43,207

GRR – Gross realization rate

Table AB-2. Summary of LEUP retrospective lifetime savings

Fuel	Tracking savings	Evaluated savings	GRR
Electric energy (MWh)	489,224	355,281	73%
Natural gas (x1,000 therms)	75,615	44,351	59%

GRR – Gross realization rate

Our recommendations are consolidated and summarized in Table AB-3.

Table AB-3. Consolidation and summary of recommendations

Evaluation area	Recommendation
Program impacts	Apply the retrospective evaluation metrics when reporting evaluated savings following the process defined in the "New Jersey Energy Efficiency Triennium 2 Evaluation Framework".
	Follow New Jersey Technical Reference Manual (TRM) guidance to adjust measure savings for interaction between fuels or between other measures.
	Review the applicant metering sample and analysis plan to ensure the meter data is representative of the population of installed equipment and representative of operation across the year.
	Include metered data and savings calculations referenced in the M&V report in their stand-alone native files in the project folder.
	Improve the net-to-gross (NTG) ratio by engaging with customers earlier in the project discovery and development phase and screen out projects that are at or near standard practice, particularly lighting measures.
Improve program delivery	Simplify or clarify the application process.
	Offer audit subsidies for less technical commercial properties (versus institutional or industrial customers) to spur activity and to mitigate free-ridership risk.
	Use the decarbonization framework and New Jersey Cost-Effectiveness Test (CET) to formulate appropriate decarb incentives and an initial set of cost-effective decarb measures.
Future program delivery	Reconsider the program measure mix that will provide the best value under a dual baseline framework.
	Lay the groundwork for the application of dual baseline savings and identify a pathway for modifying current systems for the new calculations and establish an internal process for defining second period baselines.
	Either revise ex-ante savings estimation method to reflect the findings of this study or apply the prospective realization rates. Alternatively, apply an appropriate combination of both recommendations.

EXECUTIVE SUMMARY

DNV conducted an impact evaluation, process evaluation, and net-to-gross study of the New Jersey Large Energy User Program (LEUP) for projects initiated in PY2016 through PY2020 and completed in CY2021 and CY2022. This evaluation was completed in collaboration with our partner firm Warren Engineering.

The LEUP offers New Jersey's largest energy-consuming customers (greater than $5,000,000 in energy utility bills) incentives for scoping and installing energy efficiency measures at their facilities. The one to two hundred eligible customers in the state include manufacturing sites, data centers, labs, large commercial office complexes, transit facilities, and water treatment facilities. Participants must prepare and have approved energy efficiency plans that describe the identified measures, cost, and schedule of implementation, as well as the measurement and verification methods that will be used to determine the actual achieved savings. The evaluation team reviewed 16 facilities from a population of 86 sites representing 9 of the 13 participating customers in the evaluation period. The program was evaluated at a state and not a regional level due to the limited number of participants.

Summary of evaluation methods

This evaluation study included gross and net impact evaluation research as well as a process evaluation. The impact evaluation leveraged the extensive record of M&V provided by the customers to support their savings claims, which allowed for a rigorous evaluation while minimizing the amount of data collection required by the evaluator. Interviews were conducted with knowledgeable staff at each of the sites selected for evaluation to support the gross, net, and process objectives of this study. The impact evaluation assessed the claimed savings and the methods used to produce them. The process evaluation explored program objectives, participant recruitment practices, program delivery activities, and participant satisfaction and recommends improvements to program design and delivery. The NTG component of the evaluation closely followed New Jersey's Statewide Evaluator EM&V Guidelines Net-to-Gross (NTG) Guidance for Downstream Rebate Programs.

This program has not been previously evaluated.

Key findings, conclusions, and recommendations

The study research produced findings, conclusions, and recommendations in the two topic domains of impact and process evaluations. Some of the recommendations are directed to the program, while others address broader evaluation methods and TRM guidance.

Program impacts

Table ES-1 and Table ES-2 present the program-level evaluation metrics of the retrospective first year and lifetime gross savings for electric and natural gas energy and electric demand for LEUP projects initiated in PY2016-PY2020 and completed CY2021-CY2022. The metrics include the realization rate (GRR) and the net-to-gross ratio (NTGR summary of savings per participant are available in APPENDIX E).

Table ES-1. Summary of LEUP retrospective first year savings

Fuel	First year							
	Tracking savings	Evaluated savings	GRR	Relative precision	Confidence interval	Lower limit	Upper limit	Average tracking savings per location
Electric energy (MWh)	28,576	22,283	78%	23%	90%	17,064	27,502	370,253
Electric demand (kW)	3,881	2,938	76%	25%	80%	2,201	3,675	50
Natural gas (x1,000 therms)	3,828	2,361	62%	9%	90%	2,145	2,577	43,207

Table ES-2. Summary of LEUP retrospective lifetime savings

Fuel	Tracking savings	Evaluated savings	GRR
Electric energy (MWh)	489,224	355,281	73%
Natural gas (x1,000 therms)	75,615	44,351	59%

Recommendations

R1: Apply the retrospective evaluation metrics when reporting evaluated savings following the process defined in the "New Jersey Energy Efficiency Triennium 2 Evaluation Framework".

Improve ex ante savings. The evaluated electric first year and lifetime savings were less than reported primarily because the estimates did not account for interactive effects. For example, the savings from lighting projects did not include the additional savings from reduced cooling use. The under-reporting of interactive effects created large changes in savings at some sites, hence the relative precision did not meet the design target (±23% versus the target of ±10%). Other primary sources of discrepancies included administrative errors and incorrect analysis of MV data. See Section 4.1.2 for more details.

Recommendations

R2: Follow TRM instruction to adjust measure savings for interaction between fuels or between other measures.

R3: Review the applicant metering sample and analysis plan to ensure the meter data is representative of the population of installed equipment and representative of operation across the year.

R4: Include metered data and savings calculations referenced in the M&V report in their stand-alone native files in the project folder.

Improve NTGR. The LEUP NTGR rates were similar to the rates for a large custom C&I program in other jurisdictions. Research in California shows that participants engaged earlier with the program and placed a high value on the ROI have higher NTGRs.

NTGR research generally shows that lighting projects have lower NTGRs than other EE measures, since LEDs are becoming standard technologies. In dual baseline environment, lighting projects also take a bigger gross savings penalty due to the AML.

Recommendations

R5: Improve the NTG ratio by engaging with customers earlier in the project discovery and development phase and screen out projects that are at or near standard practice, particularly lighting measures.

Improve program delivery

LEUP program satisfaction was very high, with participants giving satisfaction ratings of 4.5 or higher (on a 5-point scale) for the program overall as well as for most program aspects. However, some customers noted the application process could be onerous and some of the requirements were unclear.

One customer shared that commercial properties (compared to industrial and institutional properties) did not have in-house the technical capabilities to identify viable projects and to quantify savings and thought that technical audit subsidies would fill in this gap. Using audits to identify viable projects has the added benefit of reducing free-ridership, since the program has a role in measure identification and advancing a project through the decision makers.

Program staff requested that the policy makers provide an incentive pathway for decarbonization projects in an interview about a year ago. Since then, New Jersey has introduced a decarbonization framework and has monetized decarb benefits in the New Jersey Cost-Effectiveness Test (CET) analysis.

Recommendations

R6: Simplify or clarify the application process.

R7: Offer audit subsidies for less technical commercial properties (versus institutional or industrial customers) to spur activity and to mitigate free-ridership risk.

R8: Use the decarbonization framework and CET to formulate appropriate decarb incentives and an initial set of cost-effective decarb measures.

Future program delivery

New Jersey program administrators have been directed by the New Jersey 2023 Triennial TRM to use dual baseline methods for estimating lifetime savings of early replacement measures. Dual baseline methods adjust the lifetime savings to account for the fact that, for early replacement projects, the old equipment would have failed within a few years and would have been replaced with more efficient standard practice equipment.

Table ES-3 presents prospective first year savings and Table ES-4 presents prospective lifetime savings.

Table ES-3. Prospective first year savings

Fuel	Tracking savings	Evaluated savings	GRR	Relative precision
Electric energy (MWh)	28,576	22,280	78%	23%
Electric demand (kW)	3,881	2,949	76%	25%
Natural gas (x1,000 therms)	3,828	2,361	62%	9%

GRR – Gross realization rate

Table ES-4. Prospective lifetime savings

Fuel	Tracking savings	Evaluated savings	GRR
Electric energy (MWh)	489,224	104,121	21%
Natural gas (x1,000 therms)	75,615	38,377	51%

GRR – Gross realization rate

Dual baseline treatment reduces the cost-effectiveness of a measure compared to a single baseline treatment. These measures include longer lives and substantially better-than-standard practice to ensure savings through the measure life. Lighting measures are becoming increasingly less cost-effective as LED technology becomes standard practice which will be embodied in poorer NTGR and reduced lifetime savings.

Recommendations

R9: Reconsider the program measure mix that will provide the best value under a dual baseline framework.

Implementing dual baseline requires tracking system upgrades to record parameters associated with the calculations, such as first and second period savings. Additional research is required to define and document the Second Period Factor in ex ante savings.

Recommendations

R10: Lay the groundwork for the application of dual baseline savings and identify a pathway for modifying current systems for the new calculations and establish an internal process for defining second period baselines.

Direct application of retrospective evaluation metrics to the prospective evaluation metrics may not be appropriate depending upon the degree to which the recommendations of this study are implemented.

Recommendations

R11: Either revise ex-ante savings estimation method to reflect the findings of this study (for example by better accounting for interactive effects) or apply the prospective realization rates. Alternatively, apply an appropriate combination of both recommendations.

1 INTRODUCTION

DNV conducted an impact evaluation, process evaluation, and net-to-gross study of the New Jersey Large Energy User Program (LEUP) for projects initiated in PY2016 through PY2020 and completed in CY2021 and CY2022. This study was conducted under the guidance of the New Jersey BPU and its statewide evaluators (SWE). This evaluation was completed in collaboration with our partner firm Warren Engineering.

1.1 Program background

The LEUP offers New Jersey's largest energy-consuming customers incentives for scoping and installing energy efficiency measures at their facilities. The program started back in 2013 and has seen a steady, albeit modest, level of participation. Despite the small number of participating entities, the program achieves sizable savings. To qualify for the program, interested entities must have paid a total of $5,000,000 in energy utility bills (combined supply and delivery charges for both electric and gas service) over the prior year across the portfolio of facilities they intend on participating. In addition, the average billed peak demand across that portfolio must meet or exceed 400 kW and/or 4,000 DTh.

Upon approval, the participants must submit a draft energy efficiency plan (DEEP) that describes the identified measures, cost, and schedule of implementation, as well as the measurement and verification methods that will be used to determine the actual achieved savings. After plan approval, a pre-installation site inspection documents the existing condition of the participating facility or facilities, after which the applicant may proceed with implementation. Finally, after the measures have been installed, post-inspection is conducted and M&V-supported analysis is approved. The incentive payment is paid out as the lesser of the following:

- $4 million incentive cap
- 75% of total project cost
- 2.7% of total annual energy cost (as calculated during enrolment)
- $0.33 per kwh and $3.75 per therm saved

Program eligibility and incentive structures described above are in effect through Program Year 2022.

1.2 Evaluation approach and objectives

Table 1-1 provides a summary of the program objectives and approaches described in the work plan and notes where objectives were not met.

Table 1-1. Summary of achieved objectives and approaches to each evaluation

Task	Objective	Approach
Impact Evaluation	• Determine program source energy realization rates meeting 90/10 precision • Determine electric and natural gas energy savings realization rates with 90/10-15 precision or better *The electric GRR was ±23%* • Determine demand savings realization rates at 80/10 *The demand GRR was ±25%*	• Sample of projects meeting the precision targets • Leverage M&V implementation reports where the M&V is provided by the customer • Conduct desk reviews and supplement with various degrees of M&V where necessary to yield rigorous results without additional evaluator data collection

Task	Objective	Approach
	• Gross energy and demand savings for electricity and energy savings for gas • Consider impacts of evolving evaluation standards	• Develop alternate savings estimates using more progressive baseline assumptions
Process Evaluation	• Understand program goals and objectives • Understand participant recruitment processes and barriers to participation. Identify opportunities to increase program participation • Understand program delivery and opportunities to streamline delivery without sacrificing EM&V rigor • Measure participant satisfaction with the program and capture their recommendations for program improvements • Understand the impact of new IOU program offerings on program participation	• In-depth interviews with program managers • In-depth interviews with participating customers. • In-depth interviews with partial/stalled participants *Partial participants were not interviewed in the current evaluation cycle*[1]
Attribution (NTG)	• Savings attributable to the program, adjusting for both free-ridership and spillover	• In-depth interviews with participating customers. • Adherence to New Jersey's Statewide Evaluator EM&V Guidelines Net-to-Gross (NTG) Guidance for Downstream Rebate Programs for calculating NTGRs.

[1] Communications with the Program Implementers indicated the number of partial participants was very small. The evaluation team chose to delay these interviews until the next evaluation cycle when there would be a larger sample to survey.

2 EVALUATION APPROACH

This section provides an overview of the evaluation methods for the impact evaluation, process evaluation, and net-to-gross study.

2.1 Impact evaluation approach

For each evaluated project, DNV calculated the evaluated gross savings based on an independent analysis of the installed measures. The Impact Evaluation Team reviewed and relied on project files provided by each site through the LEUP program. Project files included LEUP applications, FEEP reports, M&V data, and M&V reports. The M&V data provided by each site was reviewed and leveraged for evaluation efforts. In some cases, the M&V data was reanalyzed using methods that differed from the applicant's methodology.

DNV developed an M&V workbook that detailed the proposed evaluation approach for the site. The M&V workbook outlined the measures, baseline conditions, savings calculation methodology, data gaps, and areas of uncertainty in the approach. The M&V workbook also contained guidance on baseline determinations and relevant questions for the in-depth interviews (IDIs).

DNV began outreach to site contacts in parallel with the review of the project files. In most cases, DNV was able to identify and contact a staff member who oversaw the project. An IDI was scheduled with the site contact to verify project information, including equipment, operating profile, and baseline condition.

In addition to the desk reviews, site visits were conducted for two of the projects. The site visits were completed to verify the information provided in the project files aligned with site conditions.

As the program implementers, TRC reviewed all M&V workbooks for the evaluated sites and provided comments to DNV. DNV reviewed and resolved all comments with TRC.

2.2 Net-to-gross approach

The net-to-gross methodology used for this LEUP evaluation follows New Jersey's Statewide Evaluator EM&V Guidelines Net-to-Gross (NTG) Guidance for Downstream Rebate Programs. However, due to changes in the TRM to require the application of dual baseline methods and likely more scrutiny of baselines going forward, a baseline battery of questions was added to both the net survey instrument and to the gross evaluation protocol.

APPENDIX D includes an in-depth analysis of the methodology along with the full survey instrument.

The net-to-gross (NTG) methodology for the impact evaluation of the LEUP primarily relies on responses from in-depth interviews with LEUP participants. The relevant NTG questions in the participant interview guide fall into three categories.

1. *Baseline/market event questions:* This series of questions asks the interviewee about the baseline conditions and market event scenarios for the projects. In earlier discussions with SWE, the study team proposed that the gross savings data collection instrument would primarily be used to collect information about baselines/market events. However, it was deemed prudent to also collect some baseline/market event information in the net interview guide because it was possible that the interviewees for the gross and net interviews might be two different people.
2. *NTG setup and framing questions:* These setup/framing questions are useful for helping the interviewees recall the projects and the decision-making that went into them that later NTG scoring questions will be asking about. This is important because some time has passed since the implementation of the evaluated projects. In addition, many LEUP participants have implemented multiple energy efficiency projects in recent years and therefore it is important for the evaluators to help the participant distinguish the evaluated project from other recent energy efficiency projects. Finally,

these questions can also help the evaluators know whether there are other key decision-makers they should be talking to estimate net savings.

3. *NTG scoring questions:* These questions closely follow the recommended survey questions for nonresidential programs that appear in the New Jersey NTG guidelines and are the primary source of the NTG scoring.

2.3 Process evaluation approach

The first step in the process evaluation was to review the program documents and website. In this review, DNV examined the length and clarity of the program documents and the "user friendliness" of the program website. After this review, DNV completed in-depth interviews with representatives of TRC who had been implementing the LEUP program. DNV completed a group interview with three of the TRC staff who were involved in program design or delivery. Topics covered in the interview included:

- The program goals and objectives
- How the introduction of new C&I energy efficiency program by NJ IOUs has impacted the LEUP program
- The program's participant recruitment process
- Program delivery

After completing this group interview with TRC representatives, DNV developed a participant interview guide that included both process evaluation and NTG questions. The process evaluation questions mostly focused on participant satisfaction with the components and stages of the LEUP program as well as with the program overall. These LEUP program components/stages included:

- The program information
- The enrolment process
- The draft energy efficiency plan (DEEP) stage
- The final energy efficiency plan (FEEP) stage
- The project implementation stage
- The confirmation of installation/M&V stage

To aid recall, the interviewers provided the participants with a brief description of the relevant program component/stage.

Besides these program satisfaction questions, the interviewers also asked the participants about two other topics of interest:

- *Alternative program options.* The interviewers asked the participants whether they had considered implementing the project through other New Jersey energy efficiency programs and why they had chosen the LEUP program for the project. The SWE had expressed interest in this topic.
- *The absence of financing as a LEUP program offering:* The interviewers asked the participants how significant a barrier it was that the LEUP program did not offer financing. The TRC representatives had expressed interest in this topic.

Besides these data collection efforts, DNV also did a quick analysis of consumption data for New Jersey commercial customers to estimate the approximate size of the market of program-eligible commercial customers.

2.4 Sample design

DNV employed a stratified ratio estimation (SRE) sample design targeting 90/10 precision for each fuel. At the highest level, the population is segmented by fuel (electric and natural gas) and then further segmented by sites with single fuel and mixed fuel savings which optimizes the selection of dual savings sites to support electric and natural gas savings. The electric

population has a dedicated segment for the Telcom customer due to its numerous locations with a single measure installation (lighting). The populations are further stratified by the magnitude of savings at the location.

Table 2-1 summarizes the key sampling parameters.

Table 2-1. Summary of sample design

Parameter	Description
Population	Tracking data for projects completed in the 2021 and 2022 calendar years.
Explicit sampling strata	By fuel (electric only, gas only, and both) By Telecom customer/non-Telecom customer By size
Gross sample allocation	Targeted 90/10 by fuel
NTGR sample allocation	Same as the gross sample
Sample design approach	Stratified ratio estimation
Target parameters	GRR
Analysis domains	By fuel
Error ratios	DNV observed error ratio for similar programs of 0.4
Projected precision at 90% confidence (80% confidence for electric demand)	±10% relative precision at the 90% confidence interval for electric and natural gas energy ±10% relative precision at the 80% confidence interval for electric demand
Savings size stratification	Up to 3 levels based on savings

NTGR – Net-to-gross ratio
GRR – Gross Realization Rate

3 EVALUABILITY ASSESSMENT

DNV reviewed project documentation for 16 completed LEUP projects to perform the evaluability assessment. The evaluation assessed:

- Data that are collected by the program to estimate tracking savings
- Completeness of program data collection by reviewing the inventory of project documentation available
- Tracking savings estimation methodologies

Based on project documentation reviewed for each of the 16 evaluated LEUP projects, DNV determined that generally, sufficient data was collected to perform a detailed evaluation in the future. However, dual baseline treatment of early replacement measures will expand the types of data and documentation required in the future.

Deficits in the data provided in the project files and the analysis approach were observed, although the majority of project files incorporated robust methods and documentation.

Native files. Data gaps for the impact evaluation included instances where data was not available in its native form (i.e., building simulation executable files, excel or CSV formats). Metered data was referenced in the M&V report in some cases, while the raw data was not provided in the project files. Additionally, savings calculations that were referenced in the M&V report were not always accompanied by the native savings calculation spreadsheets. Metered data and calculations referenced in the M&V report should be supported by their stand-alone native files.

Representative data. Energy savings for common measures, such as LED lighting, were often estimated using savings algorithms and assumptions that align with guidance from the NJ TRM. Energy savings for more unique measures were calculated using custom savings algorithms that were appropriate based on measure scope, site conditions, and available data. However, there were cases where the savings estimates relied on metering data from a small unrepresentative sample of equipment, which produced untenable savings estimates. For example, at one site, pre- and post- metering of a small sample of circuits in a lighting project showed an increase in usage, which is unreasonable for a fluorescent to LED conversion.

Exclusion of valid data. At other sites, the ex ante analysis did not incorporate all of the available data. For example, for a cooling tower measure, the ex ante savings estimate extrapolated the summer operation to a full year, while the complete M&V data showed a clear operational difference in the summer and winter savings performance which correlated with outdoor temperature. The ex ante analysis, relying on a subset of the available data, overstated the savings significantly.

Dual baseline methods. Incorporating dual baseline methods requires a reconfiguration of universal reporting fields (i.e., first period savings, second period savings, AML) in program tracking data structures and revisions in the logic and equations for calculating lifetime savings. These revisions are fundamental enough, that it is likely implementers should be considering now, how to adapt their systems and methods to this framework. Dual baseline methods require assessing standard practices two to eight years in the future, depending on the EUL of the installed measure, which will require documentation in the project files.

4 KEY FINDINGS

Our key findings of the evaluation of the LEUP program are detailed in this section.

4.1 Gross program impact findings

DNV evaluated the program performance in two ways:

- **Retrospective evaluation:** using the New Jersey protocols or technical reference manual (TRM) in effect at the initial submission of the energy efficiency plan
- **Prospective evaluation:** using the NJ 2023 Triennial TRM which is TRM in effect for planning purposes at the time of this evaluation.

4.1.1 Retrospective gross impact program findings

Table 4-1 shows the program-level **retrospective** savings and gross realization rates (GRR) for electric and gas energy and electric demand for projects Initiated in PY2016-PY2020 and completed CY2021-CY2022. The realization rates are calculated relative to the reported energy and peak demand savings; thus, the evaluated savings are the product of program-reported savings and the GRR.

Table 4-1. Summary of first year LEUP retrospective savings

	First year							
Fuel	Tracking savings	Evaluated savings	GRR	Relative precision	Confidence interval	Lower limit	Upper limit	Average tracking savings per location
Electric energy (MWh)	28,576	22,283	78%	23%	90%	17,064	27,502	370,253
Electric demand (kW)	3,881	2,938	76%	25%	80%	2,201	3,675	50
Natural gas (x1,000 therms)	3,828	2,361	62%	9%	90%	2,145	2,577	43,207

GRR – Gross realization rate

Table 4-2 presents the lifetime savings GRR for electric and gas energy.

Table 4-2. Summary of lifetime LEUP retrospective gross lifetime savings results

Fuel	Tracking savings	Evaluated savings	GRR
Electric energy (MWh)	489,224	355,281	73%
Natural gas (x1,000 therms)	75,615	44,351	59%

The evaluated electric first year and lifetime savings were less than reported primarily because the estimates did not account for interactive effects (described further in Section 6.2). One site accounted for a 16% reduction in program GRR because of an increase in electric grid purchases due to reduced electrical output from a thermal following CHP plant. This site passed

outlier screening. The under-reporting of interactive effects created large changes in savings at some sites, hence the relative precision did not meet the design target (±23% versus the target of ±10%).

The evaluated natural gas first year and lifetime savings were less than reported because of a mix of factors. One site accounted for a 21% reduction in program GRR because of an administrative error where the savings reported in tracking differed from the savings estimates in the project file.

An analysis of the sources of savings discrepancies is discussed in some detail in Section 4.1.2.

4.1.2 Retrospective discrepancy analysis

Discrepancies in reported savings and evaluated savings were categorized into one of seven sub-categories: administrative, baseline, quantity, technology, operating profile, interactivity, and other. The impact of each discrepancy sub-category was determined by taking the ratio of the sum of the discrepancies (positive or negative) to the total weighted reported savings.

- Administrative savings discrepancy included a difference in savings between savings presented in the M&V report and savings reported in LEUP program tracking data. In the case of DNV-12, there is no reference to the reported gas savings in the M&V report, as-built documentation, or savings calculations. The only reference to the reported gas savings is in the Close Out Incentive file.
- Baseline savings discrepancy was classified as evaluators adjusted the equipment baseline and efficiency to align with the NJ TRM. This adjustment was applied to a cooling tower motor replacement measure for DNV-16. Only minor differences were found.
- No discrepancies in savings due to a discrepancy in equipment quantity were identified. Quantities for all equipment were found to be correct.
- Technology savings discrepancies are all associated with LED lighting measures for six evaluated sites. Savings were adjusted to use fixture wattages outlined in the NJ TRM.
- Operating profile savings discrepancies were identified for 15 of the 16 sampled sites. These differences were typically determined through a re-analysis of the M&V data provided in the project files. For example, metered data for a cooling tower measure was available for summer months and winter months. The ex ante estimate excluded the metered data for the winter months. Instead, it extrapolated meter data from the summer months to the full year.
- Interactive savings were rarely captured in the estimated savings for the evaluated sites. Large adjustments to savings were made due to measure interactive effects from implemented measures. Measure interactivity includes the impact of lighting on HVAC systems and the impact on grid purchases from differential operation of CHP output. Section 6.2 provides additional information on the impact of measure interactivity on the savings.

Savings discrepancies classified as "Other" include the application of metered data to calculate savings in a way that was not representative of the site conditions and operations. In some cases, reported savings for LED lighting measures were calculated by applying a savings percentage value from metered data from a small sample. Table 4-3 breaks down the magnitude of factors by discrepancy categories for electric first-year savings in the sample population. The table contains a visual representation of the magnitude of impact on savings.

Table 4-3. Electric first-year savings results discrepancy by category

Discrepancy sub-category	Measure count	Relative impact on energy savings	Net impact (kWh)	Change in GRR
Administrative	1		8,234	0.0%
Baseline	1		3,424	0.0%
Quantity	0		0	0.0%
Technology	6		(66,764)	-0.2%
Operating profile	15		(1,637,704)	-5.7%
Interactivity	9		(3,155,289)	-11.0%
Other	12		(1,444,625)	-5.1%
Total			(6,292,725)	-22%

Table 4-4 breaks down the magnitude of factors by discrepancy categories for natural gas first-year savings in the sample population.

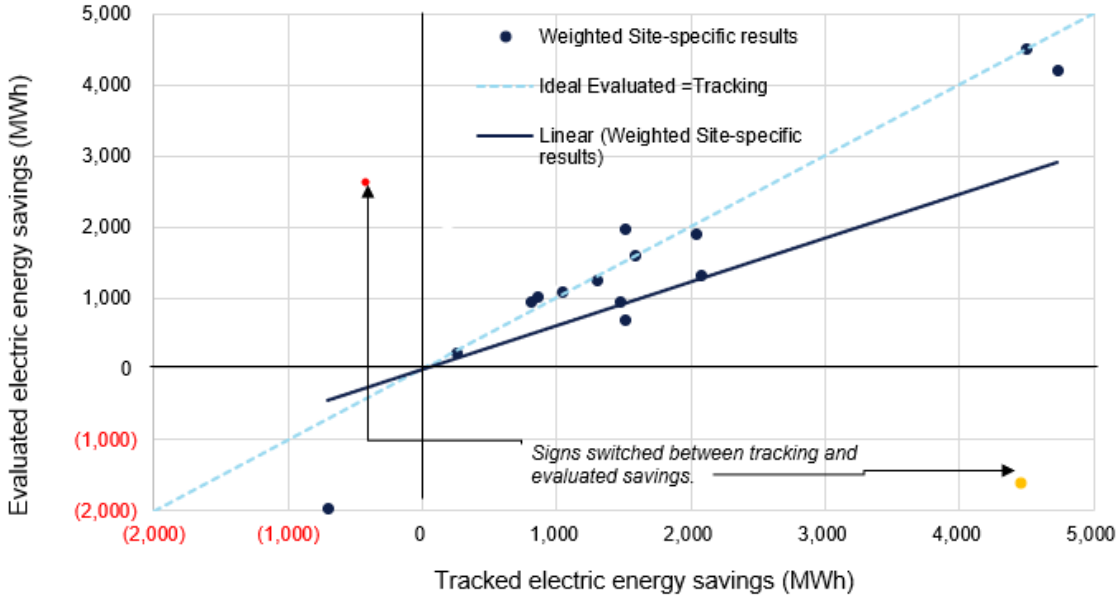
Table 4-4. Natural gas first-year savings results discrepancy by category

Natural gas discrepancy sub-category	Measure count	Relative impact on energy savings	Net impact (therms)	Change in GRR
Administrative	1		(789,488)	-21%
Baseline	0		0	0%
Quantity	0		0	0%
Technology	0		0	0%
Operating profile	3		(364,748)	-10%
Interactivity	2		43,245	1%
Other	3		(355,341)	-9%
Total			(1,466,332)	-38%

4.1.3 Retrospective site-specific gross results

Figure 4-1 presents a graph of the weighted evaluated site first-year gross savings as a function of weighted tracking first-year savings. The gold line represents the ideal relationship where the tracking and evaluated savings are equal and the GRR is 1.0. The blue line represents the actual relationship of tracking and evaluated savings demonstrating a GRR less than 1.0 by the lower slope. The red and yellow points represent two sites where the savings changed from negative to positive in one case (red) and from positive to negative in the second case (yellow). While most of the site data points hew close to the ideal line, the red and yellow marked sites reduced the program GRR and the wide divergence from tracking savings reduced the precision.

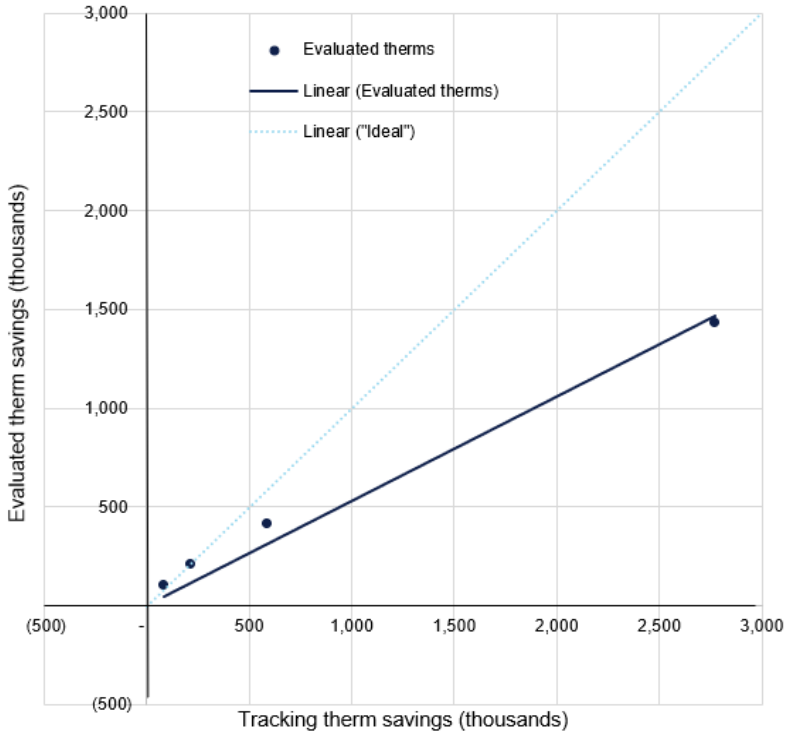
Figure 4-1. Retrospective first-year electric tracking versus evaluated savings (weighted))



Red point indicates case where savings changed from negative to positive.
Yellow point indicates case where savings changed from positive to negative.

Figure 4-2 shows a similar graph of the evaluated site weighted evaluated first year gross gas savings.

Figure 4-2. Retrospective first year natural gas tracking versus evaluated savings



APPENDIX E includes a table of results by site.

4.1.4 Prospective results and dual baseline impacts

Table 4-5 summarizes the program **prospective** savings which are calculated in accordance with the New Jersey 2023 Triennial TRM, compared to tracking savings, which do not include the new features of the 2023 TRM. The first-year prospective and retrospective savings are almost identical, with <1% change in the electric GRR due to a TRM-dictated revision in baseline efficiency.

Table 4-5. Prospective first year savings

Fuel	Tracking savings	Evaluated savings	GRR	Relative precision
Electric energy (MWh)	28,576	22,280	78%	23%
Electric demand (kW)	3,881	2,949	76%	25%
Natural gas (x1,000 therms)	3,828	2,361	62%	9%

GRR – Gross realization rate

Table 4-6 presents prospective lifetime savings.

Table 4-6. Prospective lifetime savings

Fuel	Tracking savings	Evaluated savings	GRR
Electric energy (MWh)	489,224	104,121	21%
Natural gas (x1,000 therms)	75,615	38,377	51%

Discrepancies in lifetime savings between tracking and evaluated savings were categorized into one of five sub-categories. The impact of each discrepancy sub-category was determined by taking the ratio of the sum of the discrepancies (positive or negative) to the total weighted reported savings. The sub-categories are:

- **First-year GRR.** The first-year gross realization rate also impacts lifetime savings and accounts for about one-third of the reduction in expected lifetime savings.
- **Measure life adjustments accounts for differences in non-lighting EUL claimed by the applicant and confirmed by the evaluator.** These changes are relatively minor.
- **Application of lighting AML.** The NJ TRM specifies an “Adjusted Measure Life” (AML) for calculating the dual baseline effective measure live for lighting measures (and lighting measures only). This subcategory accounts for the replacement of the lighting EUL (typically 15 years) with the AML. This accounts for about one-third of the reduction in expected lifetime savings.
- **Non-lighting second period adjustments.** This sub-category accounts for dual baseline adjustments to the second period savings of non-lighting measures.

Table 4-7 breaks down the magnitude of factors by discrepancy categories for electric lifetime savings in the sample population. The table contains a visual representation of the magnitude of impact on savings.

Table 4-7. Electric lifetime savings results discrepancy by category

Electric discrepancy sub-category	Measure counts	Relative impact on energy savings	Net lifetime impact (kWh)	Change in GRR
First year GRR	18		(129,202,877)	-26.4%
Measure life adjustments	10		(7,521,503)	-1.5%
Application of lighting AML	11		(107,105,816)	-21.9%
Non-lighting 2nd period adjustments	21		(141,272,417)	-28.9%
Total			(385,102,613)	-79%

GRR – Gross Realization Rate
AML – Adjusted Measure Life

Table 4-8 presents the magnitude of factors by discrepancy categories for natural gas lifetime savings in the sample population. The table contains a visual representation of the magnitude of impact on savings.

Table 4-8. Natural gas lifetime savings results discrepancy by category

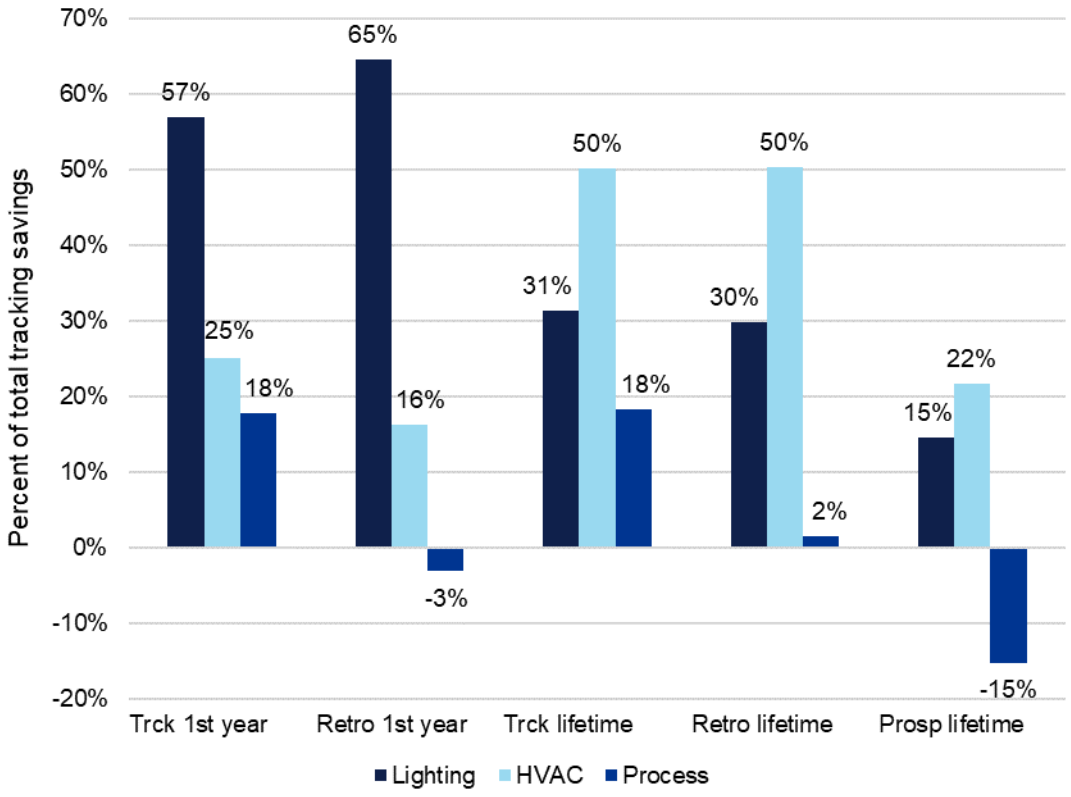
Natural gas discrepancy sub-category	Measure counts	Relative impact on energy savings	Net lifetime impact (therms)	Change in GRR
First year GRR	4		(29,805,900)	-39%
Measure life adjustments	1		(4,440,240)	-6%
Non-lighting 2nd period adjustments	3		(5,192,757)	-7%
Total			(39,438,897)	-52%

GRR-Gross realization rate

These final set of figures shows the relative contribution to savings by the end-uses of Lighting, HVAC, and Process. The figures show the distribution of savings by the three end-uses for first year tracking and evaluated savings and by tracking, retrospective, and prospective lifetime savings. The figures present savings as a percent of total tracked first year or lifetime savings, thus the evaluated savings for each savings type will sum to less than 100% since the GRR is less than 100%.

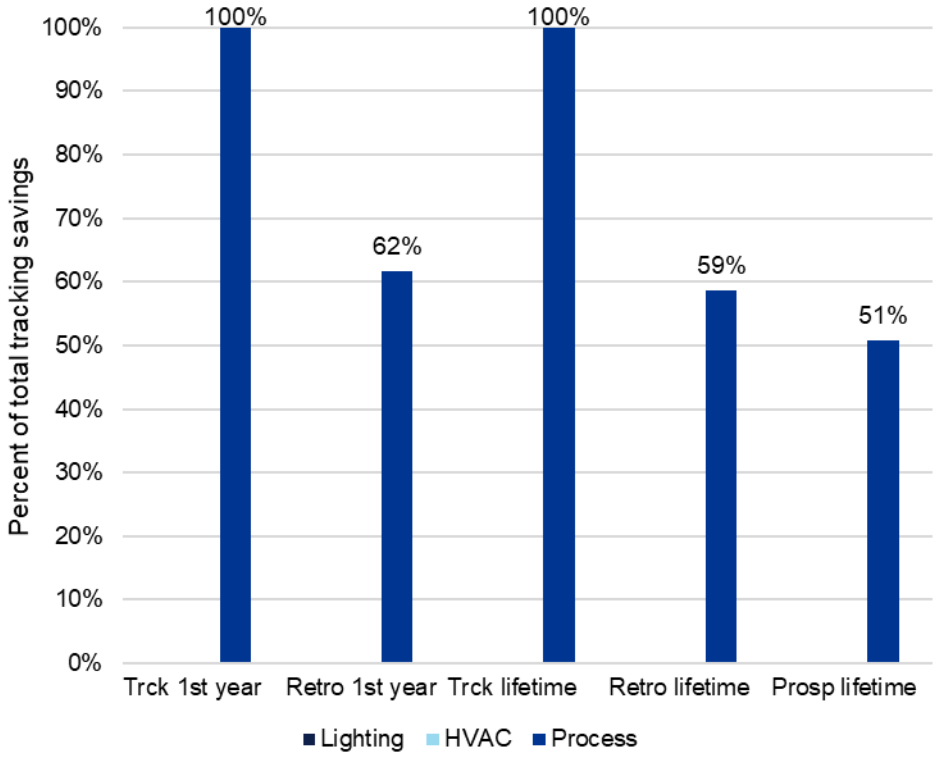
Figure 4-3 presents the end-use distribution for electricity. Lighting accounts for almost two-thirds of the first-year tracking savings, but about one-third of the expected lifetime savings due to its relatively shorter EUL compared to process and HVAC EULs. Lighting faired relatively well in the evaluation with an uptick in first year evaluated savings, although in the long run, lighting lifetime savings will diminish as LED technology becomes standard practice – which is reflected in the prospective lifetime savings. HVAC measures constituted about one-quarter of tracking first year savings, but 50% of lifetime savings due to the longer EULs.

Figure 4-3. Electric savings by end-use as a percent of tracking savings



All the natural gas measures are process measures, thus Figure 4-4 is simpler.

Figure 4-4. Natural gas savings by end-use as a percent of tracking savings



4.1.5 Measure lives

The effective useful life (EUL) for each measure was evaluated through the IDIs with the site contacts in addition to the NJ TRM. Table 4-9 summarizes a list of evaluated industrial measures along with the Ex-Ante EUL and TRM EUL for each measure. Many TRMs provide an EUL for commercial measures that may not always be appropriate for direct application to industrial measures. Equipment, such as motors and pumps, in industrial environments are typically larger than equipment used for similar applications in commercial environments. The commercial "Motors" measure in the NJ 2023 Triennial TRM is only applicable to motors of 200hp or less. Similarly, the commercial "VSD Air Compressors" measure is only applicable to compressors of 100hp or less. Thus, DNV considered both the Ex-Ante EUL and 2023 Triennial TRM EUL for each measure to determine the appropriate EUL to apply to the savings.

Table 4-9. Summary of evaluated industrial measures with associated EULs

Site ID	Measure ID	Measure description	Ex ante EUL	TRM EUL	Evaluated EUL
DNV-04	DNV-04.1	Heat of Compression Dryer: Replace 50+ year old desiccant dryer replaced with electric, ~450 tons	18	13	18
DNV-07	DNV-07.2	Removable Steam Blankets	18	11	11
DNV-12	DNV-12.1	Replace Steam Turbine with More Efficient Unit: 2850 HP gas compressor, replaced 50 year old unit	20	15	20

Site ID	Measure ID	Measure description	Ex ante EUL	TRM EUL	Evaluated EUL
DNV-12	DNV-12.2	Cold Crude Pump Upgrade: 350 HP electric motor, replaced 50+ year old gas driven unit	20	15	20
DNV-17	DNV-17.1	Conversion of a motor drive train from DC to AC with variable speed on an injection molding machine	15	20	20
DNV-17	DNV-17.2	Conversion of a motor drive train from DC to AC with variable speed on an injection molding machine (50 to 200 HP, average 125 HP)	15	18	20

EUL – Effective Useful Life
TRM – Technical Reference Manual

As an example, DNV-04.1 involves the replacement of a compressed air dryer. While the TRM EUL for a commercial building compressed air dryer is 13 years, this equipment is larger than typically found in a commercial building, built to industrial grade standards, and installed in facilities where equipment remains in place and operational for decades. DNV searched the NJ, MA, and CA TRMs, including versions up to 10 years old, for industrial equipment EULs and found none. The MA TRM offers a range of EULs for some equipment but does not offer guidance for the selection of an EUL within that range. DNV agrees with the longer EUL of 18 years used by the applicant for this measure.

4.1.6 Application of dual baseline impact metrics

Whether these GRR values will be applied in the future will depend in part on when and how the PA revises tracking savings methods to incorporate dual baseline methods. Table 4-10 illustrates how the evaluation metrics would change as the PA savings methods begin to align with the TRM. The first column presents the degree of dual baseline methods adoption in the PA tracking systems, with further explanation in the Description column. The last column indicates the evaluation metric that would apply to produce evaluated lifetime savings for that tracking system adoption.

Table 4-10. Changes in lifetime savings due to prospective evaluation metrics for electricity

Status of dual baseline Implementation in PA tracking system	Description	Evaluation metric applied
No dual baseline	Ex ante estimates do not apply either AMLs or dual baseline treatment of non-lighting measures	Apply 21% L-GRR to tracking first year savings × project claimed EUL
Lighting dual only	Ex ante estimates apply lighting AMLs, but no dual baseline treatment of non-lighting measures	Apply 43% L-GRR to tracking first year savings × project claimed EUL
Full dual	Ex ante estimates apply lighting AMLs and dual baseline treatment of non-lighting measures	Apply 72% L-GRR to tracking first year savings × project claimed EUL

AML - Adjusted Measure Life
EUL – Effective Useful Life

4.2 Net-to-gross findings

These sections summarize the net-to-gross impact and methodological findings.

4.2.1 NTGR results

DNV estimated NTGRs based on responses to NTG questions from in-depth phone interviews with LEUP participants. DNV completed seven in-depth NTG interviews with nine individuals responsible for 13 of the 16 sites in the target sample.[2] These sites accounted for 91% of the total program energy savings.

Table 4-11 shows the program-level NTGRs for both electric and gas savings. It is important to note that the program-level NTGR for gas savings is based on a single large project. DNV's review of the site-specific NTGRs revealed an interesting bifurcation of results with five of the sites receiving NTGRs of 100% and eight receiving NTGRs ranging from 50% to 56% with no NTGRs below or between these levels.

Table 4-11. NTG results

Fuel type	NTGR
Electric	60%
Natural gas	56%

NTGR – Net-to-gross ratio

The participant interviews found no evidence of spillover energy savings. While several interviewees identified completed or planned energy efficiency projects which postdated the projects in this evaluation sample, all these subsequent projects were either already in the LEUP program or had been targeted for LEUP participation. As noted, many of the LEUP participants have participated in the program for many years, so it is likely that their growing familiarity with the LEUP program requirements made them more comfortable sending future projects through the program (the unwillingness to deal with program "hassle costs" is a common source of participant spillover).

4.2.2 NTGR program findings

To put the LEUP NTG results in context, DNV compared them to recent NTGRs from C&I programs, especially custom programs, in other jurisdictions. Table 4-12 shows that the LEUP NTGRs were in a similar range as NTGRs from comparable programs in California and Illinois. It also shows that the LEUP custom electric NTGR of 60% was very close to the median custom lighting NTGR (58%) from a NJ/BPU jurisdictional scan study which collected 75 NTGR values. The LEUP NTGRs are much lower than those from a 2022 Atlantic City Electric (ACE) study (102%) but it is important to note that this study only had a sample of five sites.

Table 4-12. Jurisdictional comparison of NTGRs

Source	NTGRs
2024 NJ Large Energy User Program	Custom electric: 60% Custom gas: 56%
2024 CA statewide custom evaluation report (n=72)	Custom electric: 61% 1st year, 56% lifetime Custom gas: 76% 1st year, 75% lifetime
2023 NJ BPU/Rutgers NTG jurisdictional scan study	Custom lighting: 58% (median of 75 values) Custom HVAC electric: 77% (median of 9 values)

[2] The number of interviewees was different than the number of interviews because two of the interviews included multiple individuals. In addition, there was one case where the same interviewee completed two interviews for two different sites.

Source	NTGRs
2022 IL NTG summary memo	Custom electric: 53% (3-year avg.)
2022 NJ ACE prescriptive & custom evaluation report (n=5)	Custom electric: 102%

The in-depth participant interviews identified several drivers of free ridership for the participating LEUP sites including:

- *Non-program project drivers:* The interviews revealed that corporate sustainability policies advocating for carbon reduction as well as opportunities to increase manufacturing productivity, were motivations to undertake some of the energy efficiency projects.
- *Projects greenlighted before LEUP program intervention:* In a few cases, the energy efficiency projects had already received management approval before the LEUP program got involved. Most commonly, the expected energy savings from these projects were large enough to economically justify them without the need for the incentives.
- *Lighting projects:* Many energy efficiency projects in the LEUP program were lighting projects. The 2023 New Jersey BPU/Rutgers jurisdictional scan NTG study found that the median NTGR for custom lighting projects was much lower (58%) than the median NTGR for custom HVAC electric projects (77%). There are several plausible explanations for these lower NTGRs for lighting projects including the growing market adoption and acceptance of LED lighting technologies, the relatively short payback periods of many lighting retrofit projects, and the widespread availability and activity of contractors specializing in lighting retrofits.

These findings point to ways that the LEUP program could reduce the prevalence of free ridership going forward including:

- *Asking screening questions to filter out projects with free ridership risk:* One way for the LEUP program to improve its NTGRs is to ask screening questions of end users who are proposing energy efficiency projects for LEUP program participation. One important screening question would try to determine whether the proposed project has already been approved by company/organization management. DNV's 2024 evaluation of the California statewide custom programs found that 88% of the projects with the lowest NTGRs reported making the decision to install the energy efficient measures before they began discussing incentives with the California custom energy efficiency programs.[3]
- Another possible screening question would ask how important financial criteria such as payback periods or returns-on-investment (ROI) were as project drivers. The 2024 California custom program evaluation found that 82% of the participants with projects in the top NTGR quartile said that payback/ROI considerations were important for their project, while only 13% of those in the bottom NTGR quartile did. If a customer says that payback/ROI considerations are not important for a given energy efficiency project, this indicates that something else is driving the project besides energy savings considerations (e.g., corporate sustainability mandates or manufacturing productivity gains) and that program incentives are unlikely to be very influential. As evidence of the effectiveness of such screening procedures, the California utility with the highest NTGRs for its custom programs – Pacific Gas and Electric (PG&E) – was the only utility to employ a NTG screening tool to weed out custom energy efficiency projects with high free ridership risk.[4]

[3] For this evaluation, DNV conducted a "quartile analysis" where it compared the self-reported project drivers of the 25% of projects which had highest NTGRs (top NTGR quartile) with the self-reported project drivers of the 25% of projects with the lowest NTGRs (bottom NTGR quartile). The evaluation asked the program participants to rate the relative importance of various program- and non-program project drivers using a 0-10 scale where 0 meant "not at all important" and 10 meant "extremely important."

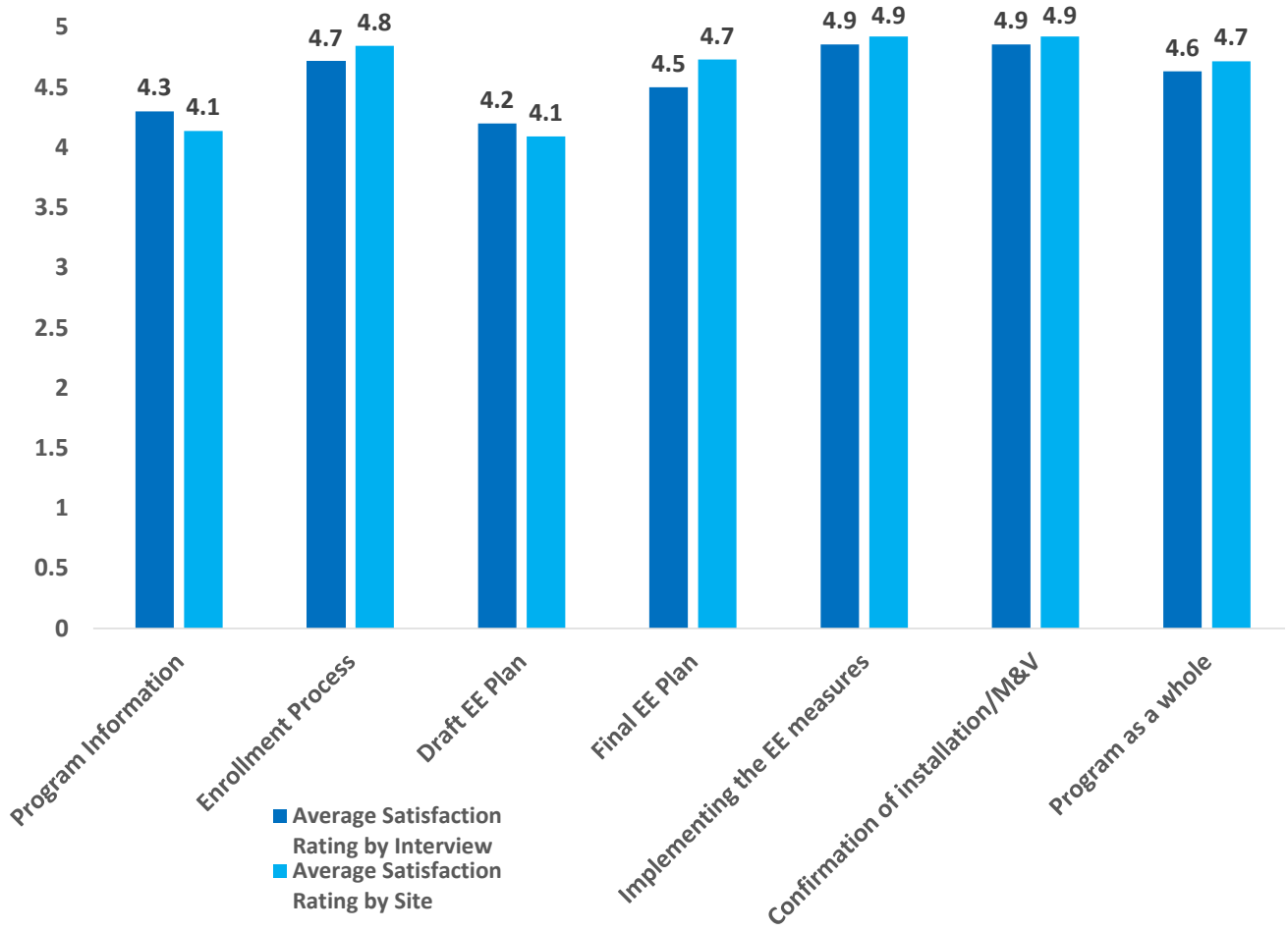
[4] This NTG tool has a list of "showstopper" questions which cover such topics as how likely the customer would have implemented the energy efficiency project with or without the program incentives and whether the customer had already selected or purchased the energy efficiency technology before interacting with the program. If a customer provides responses to these showstopper questions that indicated a high risk of free ridership, PG&E would be unlikely to incentivize this customer's energy efficiency projects.

- *Further reducing the volume of lighting projects:* As noted, recent cross-jurisdictional studies have found that custom lighting projects have much lower NTGRs on average than other custom electric measures such as HVAC. Likely reasons for these lower NTGRs include LEDs becoming standard technologies and payback periods for lighting projects typically being short. Therefore, reducing the volume of lighting projects should increase the overall LEUP program NTGRs. While the LEUP program recently introduced a 50% cap on lighting projects, further reductions in this cap are warranted. Besides the net savings benefits, reducing the volume of lighting projects would also have gross savings benefits because these projects are subject to adjusted measure life (AML) calculations which significantly reduce the second period gross energy savings in the dual baseline scenario.
- *Getting involved earlier in projects:* Earlier program involvement in energy efficiency project development and approval processes should also improve NTGRs. The earlier the LEUP program can be involved in energy efficiency projects, the greater the chance the program can influence project scoping. And even in cases where the LEUP program does not significantly impact project scope, the earlier involvement of the program in the discussion of the project financing should also increase program attribution. For example, the program incentives might accelerate the timing of an energy efficiency project by reducing its payback period enough to move it up in the queue of capital expenditure projects that the company is considering. Early program involvement in projects can also reduce the incidence of the program providing incentives to projects that have already been greenlighted. Finally, the LEUP program's possible subsidization of energy audits, which one program participant recommended as important for encouraging energy efficiency projects in the commercial real estate subsector, would be another mechanism for early project involvement.

4.3 Improving program delivery and outreach

DNV asked customers about their satisfaction with program information, requirements, and processes as well as the program overall. The satisfaction questions used a five-point Likert scale where 5 equaled "very satisfied" and 1 equaled "very dissatisfied." DNV calculated the average satisfaction levels using both the number of interviews and the number of sites. Figure 4-5 shows that overall program satisfaction was very high (4.6-4.7 average satisfaction levels) and satisfaction with most of the program attributes was also very high.

Figure 4-5. Program satisfaction



However, satisfaction with the program information and the requirements for the draft energy efficiency plan were both at lower (4.1-4.3 average scores) satisfaction levels. When DNV asked program participants to explain their lower satisfaction levels with these program components, they mentioned either that program information requirements were unclear or that program requirements were onerous.[5] DNV's review of the LEUP program documents and website also identified opportunities to make information clearer. For example, the program's website landing page had only very high-level information about program eligibility requirements and incentive levels in a single compressed page. If prospective participants wanted to learn more about the program, they were given a link to a program guide over 20 pages long. Therefore, there was no program description available that was shorter than the 20+-page guide but longer than the single webpage. The program guide also contained some information that was hidden or confusing.[6]

Competition with other PAs. One topic of interest for SWE and the BPU was whether the LEUP program faced competition from other C&I energy efficiency programs that New Jersey utilities had recently introduced. Interviews with

[5] Some representative quotes concerning the lack of program clarity included: "When it comes to submitting the nitty-gritty details of the paperwork, some of the aspects could have been clearer" and "There was perhaps some more back and forth [concerning reporting requirements for project labor and material costs] that isn't necessarily as clear as day in the program documentation." One quote concerning onerous program requirements was: "I feel like there's a lot of [requested] information that's not applicable. … We're a site with 40-some-odd buildings. They wanted to know the total square footage of the site. We were doing a project in this one building. So, who cares what the size of the building over there is? It has nothing to do with my project."

[6] For example, important information about the program's new 50% savings cap on lighting projects was not mentioned until page 9 of the program guide where it was buried in a sub-bullet. In addition, while one part of the program guide indicated that the DEEP was optional, a latter part of the guide read: 'the FEEP will build upon the contents of the DEEP submittal," which implies, conversely, that the DEEP is a prerequisite for the FEEP.

LEUP program participants revealed that they had never considered these utility programs as viable alternative channels for their energy efficiency projects. Some participants chose the LEUP program because they had positive previous experiences with this program or because the utility programs could not offer as sizeable a financial incentive package as the LEUP program could. Other LEUP participants were simply unaware of the utility programs.

Value of financing. In-depth interviews with LEUP program implementers revealed a concern that the program's failure to offer project financing options might be a barrier to participation. However, the participant interviews revealed that this lack of a financing option was not a concern for them. The most common reason the participants cited why the lack of a financing option was not a barrier was that their companies/organizations had sufficient sources of internal capital to fund worthwhile projects and therefore outside financing was not needed.

Energy audit support. Besides these process evaluation findings that were in response to direct survey questions, the participants offered other suggestions for improving the LEUP program on a more *ad hoc* basis. One participant, with extensive experience promoting energy efficiency projects in the commercial real estate sector, recommended that the LEUP program subsidize the costs of energy audits. "The real solution to [increasing LEUP EE project volume] would be to frontend the audit process [costs]," he said, "You'd see projects happening all over the place."

This participant explained that these audits were very expensive and that commercial real estate energy managers usually can only get energy efficiency projects approved if they have an analysis demonstrating future energy savings. He elaborated on this predicament:

> *"It's a constant battle. I've got to convince the operations VP to spend the money to do the audit. Then you do the audit. … The results, hopefully, are interesting enough to where the investment is compelling. Then you've got to fight for the money to do the analytics to be able to do the FEEP [Final Energy Efficiency Plan]. And then when you do the FEEP, you've got to fight to get the project approved with the builder."*

This participant also said that this barrier – the high cost of energy audits – was likely more significant for the commercial real estate sector than other sectors such as manufacturing where companies have more routine processes for examining opportunities for process improvements and can justify projects for other reasons besides energy savings. As discussed in the NTG section above, such energy audits might also give the LEUP program opportunities for increase program attribution due to the ability to get involved with projects earlier and to impact project scope.

Evidence for the potential benefits of such energy audit subsidization can be found in a program right across the state border: NYSERDA's FlexTech program, which offers subsidies of 50% of the cost of an EE technical assistance study. DNV's 2019 evaluation of this program found some promising program outcomes including:

- By year 5, 55% of the recommended energy savings had been adopted
- By year 7, 66% of the recommended energy savings had been adopted
- The majority of participants reported installing recommended measures without additional funding from EE programs

However, the interview with the TRC staff who manage the LEUP program revealed that they do not view such detailed audits as part of the current program design.

Decarbonization. Program Implementers received interest from customers in implementing decarb measures but noted there was no mechanism in the current program for valuing or incentivizing these measures. This population, which includes heavy industry, has the potential for large GHG reductions.

Market expansion. Another opportunity for program improvement lies in expanding marketing and outreach efforts. In the seven-year history of the LEUP program it has only had 16 unique participating customers. To estimate the remaining

commercial market potential for the program, DNV did a quick analysis of utility commercial customer data.[7] This analysis found approximately 80 unique commercial customers who consumed more than the LEUP program's minimum threshold of 25 million kWh annually. Table 4-13 shows a sampling of the largest of these qualifying customers with multiple examples of data centers, universities, transit facilities, and offices.

Table 4-13. A sample of LEUP program qualifying commercial customers

Building type	Number of buildings	Aggregate customer kWh
Data center*	5	373,768,752
University	107	265, 254, 236
Banking/Data center	1	226,322,394
Transit*	2	142,752,452
Data center	2	128,104,918
Transit	31	119,299,751
Data center	1	112,903,302
Property management	1	109,709,690
Banking	8	100,664,547
University	41	96,594,971
Data center	1	95,447,454
Data center	6	91,509,302
Data center	1	91,420,911
Telecom	131	88,490,136
Offices	4	87,650,446
Offices	1	85,455,084
Data center	2	84,832,120
		Sample of remainder
Telecom*	267	73,515,514
Big Box Retail	19	51,241,966
Convenience	151	34,951,076

[7] This data set does not include medium or heavy industrial and manufacturing enterprises.

Building type	Number of buildings	Aggregate customer kWh
Hospital	4	29,684,238

5 PROGRAM COMPARISONS

This section reviews process and impact evaluation findings from a sampling of other large custom programs with features comparable to LEUP. Table 5-1 outlines process and impact evaluation findings from a sampling of other programs similar to LEUP for comparison.

The Atlantic City Electric (ACE) Commercial and Industrial (C&I) Prescriptive and Custom Program serves C&I customers, of all sizes, while LEUP serves only the largest C&I customers. LEUP is a custom-only program offering incentives for electric and natural gas savings, whereas ACE C&I is an electric-only program offering both custom and prescriptive projects. Approximately 99% of the total ACE C&I PY1 savings were claimed through lighting measures, while the LEUP first year savings measure mix was 57% lighting, 25% HVAC and 18% process measures. The purpose of the ACE evaluation, as stated in the report, was to support "future evaluations, check completeness of the tracking data, check the evaluability of the data, and verify savings calculation methods". The ACE C&I PY1 evaluation was based on desk reviews supplemented with an online survey to verify measures impact and collect the NTGR battery. Five surveys were completed representing a total of 91 participants. The ACE evaluation study is included in the comparison because it is the only other New Jersey C&I program with published evaluation results at the time of this writing, however since the customer mix and the objective and rigor of the study are different, the results are not comparable.

The New York State Energy Research and Development Authority (NYSERDA) previously managed an Industrial and Process Efficiency (IPE) Program. The IPE evaluation is included because its program design is similar to LEUP's. Through this program, incentives for electric and natural gas savings were offered to eligible industrial customers and data centers. Lighting projects constituted approximately 13% of total electric projects completed 2014-2017. Process improvement projects accounted for approximately 24% of total electric projects and approximately 73% of total natural gas projects completed 2014-2017. Lighting projects constituted approximately 13% of total electric projects completed 2014-2017. IPE required rigorous development of ex ante estimates like LEUP, thus the sample design referenced the IPE error ratio in developing the LEUP sampling. Also, the IPE gross evaluation study design was like the LEUP design, thus the gross results are comparable, although IPE had been evaluated at least once previously. The LEUP GRRs are lower than IPE, primarily due to the undercounting of interactivity effects in LEUP.

The Massachusetts (MA) Custom Electric Evaluation and MA Custom Gas Evaluations are parallel impact evaluations of Custom Program Installations for the Massachusetts PAs and Energy Efficiency Advisory Council (EEAC) Consultants. Unlike LEUP, the MA program serve a range of C&I customers and while the application process is rigorous, it does not require ex ante M&V. This evaluation investigated custom electric non-lighting projects for commercial and industrial customers, while a parallel evaluation assessed natural gas savings. The MA study design entails on-site M&V, which includes a desk review, customer interview, and on-site verification and metering. The MA protocols have required application of dual baseline methods since 2019. The MA Custom Evaluations includes a focus on the classification and use of dual baseline for applicable measures and has established practices for continuous ISP research and collaboration through the "Baseline Advisory Group". A default outyear factor (OYF) of 90% is applied for non-lighting dual baseline measures. Currently, the evaluation results see little discrepancies in savings due to dual baseline factors. The MA programs are evaluated every year using a rolling sample targeting 90/10 precision for each three-year rolling window, thus the programs and evaluations are stable. The LEUP GRRs are lower than the MA GRRs, primarily due to the undercounting of interactivity effects in LEUP.

The last program included in the comparison is the California Public Utility Commission (CPUC) evaluation of the statewide electric and gas utilities' portfolio of custom programs. California is included because it has been a leader in program design and evaluation methods. California fields dozens of custom programs targeting specific niche markets (i.e., commercial property, agricultural, education) sponsored by five program administrators, but evaluated under one custom study. The study design entails on-site M&V, which includes a desk review, customer interview, and on-site verification and metering. The CPUC evaluation framework is harsh; for example, site savings are evaluated as zero, if a customer refused an M&V site visit or if the verified measure installation date falls in December of the previous year (where most jurisdictions use an easier to track invoice date). California also requires a high burden of proof to claim early replacement savings; for example, the evaluator reclassified 38% of the projects from early replacement to time of sale in the reference study. Thus, the GRRs tend to be lower in California than what they would be under other state evaluation regimes.

Table 5-1. Large custom and similar programs from other states, process findings

Program name	Program type	State	Evaluation year	Electric energy GRR	Electric demand GRR	Natural gas GRR	NTGR	Lifetime GRR
New Jersey Large End User Program	Industrial, custom projects	New Jersey	2024	78%	N/A	62%	60% electric energy; 56% natural gas	73% electric energy; 59% natural gas
Atlantic City Electric Commercial and Industrial Prescriptive and Custom[8]	C&I, prescriptive & custom projects	New Jersey	2023	102%	100%	N/A	102%	N/A
California Statewide Utilities' Portfolio of Custom Programs[9]	Commercial, industrial, and agricultural custom	California	2024	37%	50%	45%	61% electric energy; 63% electric demand; 76% natural gas	38% electric; 19% natural gas

[8] Guidehouse. "Commercial and Industrial (C&I) Prescriptive and Custom Program Evaluation Report." njcleanenergy.com, 11/18/2022. https://njcleanenergy.com/files/file/Utility%20Info/Evaluation%20Studies/ACE_PY1_CI_Pres_Custom_Evaluation_Report_Final.pdf.

[9] DNV. "Custom Industrial, Agricultural, and Commercial (CIAC) 2022 Impact Evaluation." pda.energy.dataweb.com, 5/14/24. https://pda.energydataweb.com/api/view/3971/CIAC%202022%20Evaluation%20Final%20Report.pdf.

Program name	Program type	State	Evaluation year	Electric energy GRR	Electric demand GRR	Natural gas GRR	NTGR	Lifetime GRR
Massachusetts Custom Electric[10] and Custom Gas[11]	Large C&I, custom projects	Massachusetts	2023	83.4%	80.3% summer; 77.5% winter	80%	N/A	82.8%
NYSERDA Industrial and Process Efficiency[12]	Industrial, custom projects	New York	2018	86%	N/A	91%	N/A	N/A

GRR – Gross Realization Rate
NTGR – Net-to-Gross Ratio

[10] DNV. "Massachusetts Impact Evaluation of PY2020/2021/2022 Custom Electric Installations." ma-eeac.org, 7/26/2023. https://ma-eeac.org/wp-content/uploads/MA22C02-E-CUSTELEC-PY2020-21-22-Custom-Electric-Report-w-Site-Reports-FINAL.pdf.

[11] DNV. "Massachusetts Impact Evaluation of PY2019 Custom Gas Installations." ma-eeac.org, 3/31/2022. https://ma-eeac.org/wp-content/uploads/MA20C13-G-CUSTGAS_Stage5_Report_Custom_Gas_Impact_Evaluation_PY2019_Final.pdf.

[12] ERS. "2014-2017 Industrial and Process Efficiency Program Impact Evaluation." nyserda.ny.gov, September 2018. https://www.nyserda.ny.gov/-/media/Project/Nyserda/Files/Publications/PPSER/Program-Evaluation/2018-IPE-Evaluation-FinalReport.pdf.

6 CONCLUSIONS AND RECOMMENDATIONS

Based on the key findings and results of this evaluation, our conclusions and recommendations for the LEUP program are detailed in this section.

6.1 Retrospective gross and net results

Table 6-1 and Table 6-2 present the summaries of LEUP first year retrospective savings and retrospective lifetime savings for projects Initiated in PY2016-PY2020 and completed in CY2021-CY2022.

Table 6-1. Summary of LEUP first year retrospective evaluation savings

Fuel	First year							
	Tracking savings	Evaluated savings	GRR	Relative precision	Confidence interval	Lower limit	Upper limit	Average tracking savings per location
Electric energy (MWh)	28,576	22,283	78%	23%	90%	17,064	27,502	370,253
Electric demand (kW)	3,881	2,938	76%	25%	80%	2,201	3,675	50
Natural gas (x1,000 therms)	3,828	2,361	62%	9%	90%	2,145	2,577	43,207

GRR – Gross realization rate

Table 6-2. Summary of LEUP retrospective lifetime evaluation savings

Fuel	Tracking savings	Evaluated savings	GRR
Electric energy (MWh)	489,224	355,281	73%
Natural gas (x1,000 therms)	75,615	44,351	59%

Gross realization rate

These findings are based on confirmation of site conditions through an in-depth interview (IDI) and a review of a statically selected sample of project sites. In some cases, a reanalysis of program-provided monitoring and verification (MV) was completed. As the program implementers, TRC reviewed all M&V workbooks for the evaluated sites. DNV reviewed and resolved TRC's review comments. While the electric relative precision falls short of the target of ±10%, the results are representative of the program sites. The NTGR research was conducted following the New Jersey recommended NTG methods and the sites that made up the NTG sample accounted for 91% of the LEUP program electric savings.

Recommendations

- Apply the retrospective evaluation metrics when reporting evaluated savings following the process defined in the "New Jersey Energy Efficiency Triennium 2 Evaluation Framework".

6.2 Accounting for measure interactivity

The New Jersey Technical Reference Manual (TRM) and protocols in effect at the time of the initiation of the LEUP projects directed LEUP participants to follow the Pay-for-Performance Existing Building Protocols.[13] These protocols permit an existing conditions baseline, but also require that savings should "ensure interactive savings are taken into account." However, interactive savings were rarely captured in the savings estimates of the sites that were evaluated. For example:

- Lighting measures did not account for the cooling bonus or the heating penalty from the reduced heating load of the more efficient lighting systems.
- A site with a combined heat and power (CHP) plant, operating in a thermal following mode, did not account for the additional electricity purchases from the grid due to the reduction in thermal load from the installed measures. This error reduced the program GRR by 16%.
- A fuel switching measure at one site involved converting a gas turbine drive to an electric motor. As an exception, the site claimed gas savings and accounted for the increase in electrical usage from the electric motor.

The tracking savings largely did not account for interactivity, which led to an overestimation of the effect on the grid and to a smaller extent, the gas distribution. This is the foundation of determining the benefits of the installed measures.

Recommendations

- Follow TRM instruction for this program to account for measure interactivity including any impacts on grid purchases from differential operation of CHP output because of the measure implementation, the impact of lighting on HVAC systems, as well as other measure interactivity.

6.3 Quality of the program developed M&V

The program design requires savings estimates to be supported by metered pre and post high-quality information. We found the overall quality of the M&V to be good. All projects included substantial documentation of the measures, the mechanisms of savings, and very importantly, evidence of meter-backed supporting data.

However, we did observe that some of the M&V Plans were not well-conceived or executed. Examples include the following:

- For several projects, small samples of lighting circuits were selected for pre/post metering. These samples were not an accurate representation of the full project scope and resulted in unreasonable unit savings. For one site, the savings were negative. In some projects, the metered results were consequently rejected as they should have been during project closeout. In other cases, metered data was used without appropriate adjustment.
- For one project, the analysis included only the metered data from the summer months for a cooling tower project and extrapolated the results to a full year. However, the full dataset clearly showed a large difference in winter operation with substantially less savings. By using only the summer data, the savings were overstated.

In some cases, the results produced grossly underestimated savings, and in other cases, the savings were grossly overestimated. There were several cases where the savings were reasonably estimated. Cumulatively, the results produced slightly overestimated savings.

[13] BPU – New Jersey's Clean Energy Program. "Pay for Performance (P4P) Program Guide – Commercial & Industrial (C&I) New & Existing Buildings Fiscal Year 2023." njcleanenergy.com, 2023. https://njcleanenergy.com/files/file/Pay%20for%20Performance/FY23/P4P%20FY23%20Program%20Guide.pdf.

Recommendations

- Review the applicant metering sample and analysis plan to ensure the meter data is representative of the population of installed equipment and representative of operation across the year.
- Include metered data and savings calculations referenced in the M&V report in their stand-alone native files in the project folder.

6.4 Application of dual baseline calculation methods

New Jersey program administrators have been directed by the New Jersey 2023 Triennial TRM[14] to use dual baseline methods for estimating lifetime savings of early replacement measures. Dual baseline methods adjust the lifetime savings to account for the fact that, for early replacement projects, the old equipment would have failed within a few years and would have been replaced with more efficient standard practice equipment.

The New Jersey 2023 Triennial TRM does not fully specify the methods or requirements for implementing dual baseline savings estimates. Dual baseline lifetime savings were estimated in this study using two methods:

- **Adjusted Measure Life (AML).** AMLs provide a simplified approach to calculating dual baseline lifetime savings

$$Evaluated\ lifetime\ savings = AML \times \text{First year evaluated savings}$$

This study used the AMLs defined in the NJ 2023 Triennial TRM for lighting. AMLs were not defined for other measures.

- **Dual baseline equation.** Dual baseline calculations discount the savings of the latter two-thirds of the measure life using an Out-Year-Factor (OYF) to account for what the customer would have done when the original equipment failed and had to be replaced:

$$Lifetime\ savings = EUL \times \left(\frac{1}{3} + \frac{2}{3} \times OYF\right) \times First\ year\ evaluated\ savings$$

DNV engaged subject matter experts to estimate the OYF non-lighting measure of ISP baseline practice when the original equipment would have failed. As a point of reference, the default OYF defined in the Massachusetts Baseline Framework is 90%.15 A default OYF of 90% produces a default AML equal to the product of 93% and the EUL.

Implementers may require further direction in the dual baseline methods to guide implementation in PA tracking systems. BPU will need to consider the trade-offs between incremental accuracy of methods versus the level of effort required by PAs to adapt their tracking systems to the methods.

The PAs should consider how dual baseline savings methods will impact program cost-effectiveness since dual baseline will reduce lifetime savings compared to a single baseline. The PAs should be more selective with program funds to systematically target and prioritize measures that will not significantly reduce lifetime savings. Lighting measures are becoming increasingly less cost-effective as LED technology becomes standard practice which will be embodied in poorer NTGR and reduced lifetime savings.

[14] BPU – New Jersey's Clean Energy Program. "New Jersey 2023 Triennial Technical Reference Manual for 2024 Filings." nj.gov, 4/3/2023. https://nj.gov/bpu/pdf/publicnotice/4.%20EE%20T2%20Technical%20Reference%20Manual%202023.pdf.

[15] DNV GL, Energy & Resource Solutions. "Massachusetts Commercial/Industrial Baseline Framework." ma-eeac.org, 4/26/2017, https://ma-eeac.org/wp-content/uploads/MA-Commercial-and-Industrial-Baseline-Framework-1.pdf.

Recommendations

- Reconsider the program measure mix that will provide the best value under a dual baseline framework.
- Lay the groundwork for the application of dual baseline savings and identify a pathway for modifying current systems for the new calculations and establish an internal process for defining second period baselines.

6.5 Prospective evaluation results

The retrospective evaluation metrics (i.e., GRR) reflect the sampled population's performance in following TRM guidance that was in effect at the time of the measure installation and in how well the tracking savings reflected the M&V data provided in the files. Prospective evaluation metrics are intended to inform planning and could potentially be applied to future program tracking savings to report evaluated savings.

Application of the prospective evaluation metrics to LEUP tracking estimates after the PY2021-PY2023 without adjustments will depend upon to what extent the LEUP program has adopted the recommendations of this study. When assessing this application, stakeholders should consider:

- The LEUP program serves large complex sites with long development cycles, thus many of the applications have been developed under older protocols and TRM guidance. These older initiated projects will likely exist in future evaluation sample populations and may not be subject to the newer requirements which are embedded in the prospective evaluation metrics.
- Lighting accounts for almost two-thirds of the first-year tracking savings, but about one-third of the expected lifetime savings due to its relatively shorter EUL compared to process and HVAC EULs. Lighting faired relatively well with an uptick in first year evaluated savings but experienced a reduction in prospective lifetime savings. HVAC measures accounted for one-quarter of tracking first year savings, but 50% of lifetime savings due to the longer EULs. Section 4.1.4 provides more information on savings by end-use.
- If the LEUP applicant savings factors in interactivity in savings estimates as recommended, the prospective GRR may overly penalize program savings.
- If the LEUP program implementers adopt dual baseline savings methods in part (for example, adopting the lighting AML only) or fully as recommended, the prospective GRR will not apply.

If stakeholders agree to immediately evaluate subsequent program years (starting with PY2024), the evaluation will resolve the degree to which the program has adopted recommendations. However, if instead, the results of this evaluation study are applied to future program years, it would be prudent to review the population of participants to determine the rates of adoption to develop fair adjustments to the GRRs.

Recommendations

- Either revise ex-ante savings estimation method to reflect the findings of this study (for example by better accounting for interactive effects) or apply the prospective realization rates. Alternatively, apply an appropriate combination of both recommendations.

6.6 Program delivery improvements

LEUP program satisfaction was very high with participants giving satisfaction ratings of 4.5 or higher (on a 5-point scale) for the program overall as well as for most program aspects. However, there was room for improvement in the program NTGRs and so many of the study's recommendation focus on ways to increase these ratios.

Asking screening questions to filter out projects with free ridership risk. One way for the LEUP program to improve its NTGRs is to ask screening questions of end users who are proposing energy efficiency projects for LEUP program participation. One important screening question would try to determine whether the proposed project has already been approved by company/organization management. Another possible screening question would ask how important financial criteria such as payback periods or returns-on-investment (ROI) were as project drivers. If a customer says that payback/ROI considerations are not important for a given energy efficiency project, this indicates that something else is driving the project besides energy savings considerations (e.g., corporate sustainability mandates or manufacturing productivity gains) and that program incentives are unlikely to be very influential. Section 4.2.2 discusses evidence from recent California custom program evaluations that supports the use of such NTG screening questions.

Reducing the volume of lighting projects. Recent cross-jurisdictional studies have found that custom lighting projects have much lower NTGRs on average than other custom electric measures such as HVAC. Likely reasons for these lower NTGRs include LEDs becoming standard technologies and payback periods for lighting projects typically being short. Therefore, reducing the volume of lighting projects should increase the overall LEUP program NTGRs. While the LEUP program recently introduced a 50% cap on lighting projects, further reductions in this cap are warranted. Besides the net savings benefits, reducing the volume of lighting projects would also have gross savings benefits because these projects are subject to adjusted measure live (AML) calculations which significantly reduce the second period gross energy savings in the dual baseline scenario.

Getting involved earlier in projects. Earlier program involvement in energy efficiency project development and approval processes should also improve NTGRs. The earlier the LEUP program can be involved in energy efficiency projects, the greater the chance the program can influence project scoping. And even in cases where the LEUP program does not significantly impact project scope, the earlier involvement of the program in the discussion of the project financing should also increase program attribution. Early program involvement in projects can also reduce the incidence of the program providing incentives to projects that have already been greenlighted.

Clarifying and enhancing program information. One potential area for modest program improvement relates to the level of detailed information that the LEUP program provides concerning the requirements for the Draft Energy Efficiency Plan (DEEP) in particular, as well as for the program overall. Participant satisfaction levels were lower for both program information (4.1-4.3) and the DEEP stage (4.1-4.2) than for other program aspects. Participant verbatims revealed that dissatisfaction stemmed from either program information being unclear or the required information from the participants appearing onerous or unnecessary (see Sections 6.3 and 10.2 for more detail on these participant concerns). DNV's review of the program documents and website also identified opportunities to clarify and enhance program information. For example, the program's website landing page had only very high-level information about program eligibility requirements and incentive levels in a single compressed page. If prospective participants want to learn more the program, they are given a link to a program guide over 20 pages long. It would be useful if the program could provide a program description of intermediate length.

Energy audit subsidies. Another potential area for improvement would be for the LEUP program to offer subsidies for detailed energy audits, especially for the commercial sites. As discussed in the detailed participant interviews (Section 4.3), the commercial real estate industry finds it difficult to get energy efficiency projects approved by management without the supporting evidence of a detailed audit that estimates potential energy savings. Yet the costs of these detailed energy audits can be prohibitive without program support. As discussed in Section 4.3, New York's FlexTech C&I audit program has had promising results with over half of the recommended energy efficient improvements installed within five years.

The interviews with TRC staff who manage the LEUP program revealed that they do not view such detailed audits as part of the current program design which is intended to be self-directed. Therefore, if the program were to add this audit

subsidization component, it would require a revision to the program design and the implementer's scope of work. We note that self-directed programs are sometimes constrained by the services offered and audits may not fit within that framework.

Decarbonization. Program Implementers received interest from customers in implementing decarb measures but noted there was no mechanism in the current program for valuing or incentivizing these measures. This population, which includes heavy industry, has the potential for large GHG reductions.

Recommendations

- Improve the NTG ratio by engaging with customers earlier in the project discovery and development phase and screen out projects that are at or near standard practice, particularly lighting measures.
- Simplify or clarify the application process.
- Offer audit subsidies for less technical commercial properties (versus institutional or industrial customers) to spur activity and to mitigate free-ridership risk.
- Use the decarbonization framework and CET to formulate appropriate decarb incentives and an initial set of cost-effective decarb measures.

APPENDIX A. SAMPLE DESIGN

This section presents the tasks evaluators completed to generate the sample for this study.

Participant population

DNV received tracking data from the Large Energy User Program for all projects completed in 2021 and 2022, although projects were initiated in PY2016-PY2020, and thus are subject to different versions of the TRM. Table A-1 shows the distribution of participants by customer type, number of unique locations, and electric and gas savings. The building types are diverse ranging from office to manufacturing. A single customer with facilities at 68 locations accounts for half of the electricity savings, while another single customer at one site accounts for 90% of the natural gas savings. In addition, three sites report negative electric savings which are summed into the reported savings.

Table A-1. LEUP participant distribution for completed projects Calendar Year 2021–2022

Telecom facility	Number of locations	Electricity savings (kWh)	Natural gas savings (therms)
Telecom facility	68	10,393,031	-
Transit facility	3	6,673,947	-
Office	5	6,375,364	84,576
Manufacturing facility	4	4,803,285	3,573,959
Data center	3	2,115,274	-
Pharmaceutical company	1	1,055,194	-
Research and manufacturing facility	1	320,312	40,611
Wastewater mgt facility	1	105,350	16,631
Grand total	**86**	**31,841,758**	**3,715,777**

The evaluation team also reviewed available project files and summarized the mix of measures in Table A-2. Half the electric savings are from lighting projects from two customers (one of which is the multi-location (68) Telecom) and the balance is from a diverse mixture of HVAC and process measures. Note, some locations have two or more measures.

Table A-2. LEUP measure distribution for completed projects Calendar Year 2021–2022

Measure type	Number of measures	Electricity savings (kWh)	Natural gas savings (therms)
Lighting	74	17,767,063	-
Process	13	8,086,120	3,658,535
HVAC	4	4,123,665	-
Unknown	3	1,864,909	16,631
Insulation	1	-	40,611
Grand total	**94**	**31,841,758**	**3,715,777**

Sample design

DNV employed a stratified ratio estimation (SRE) sample design targeting 90/10 precision for each fuel. The SRE method is an efficient sampling strategy for LC&I programs because typically there are (a) a few sites with a disproportionally large savings and (b) the tracking savings are produced using site-specific parameters. At the highest level, the population is segmented by fuel (Electric and Natural Gas) and then further segmented by sites with single fuel and mixed fuel savings which optimizes the selection of dual savings sites to support electric and natural gas savings. The electric population has a

dedicated segment for the Telcom customer due to its numerous locations with a single measure installation (lighting). The populations are further stratified by the magnitude of savings at the location.

The sampling unit for this design is all fuel savings at each unique location whether that location consists of a single building or is a campus. This is a straightforward approach to managing the diverse locations of some customers (like the Telecom customer) and to capturing the diversity of the customer building types and measures. Three sites with negative electric savings were included in the populations using the absolute value of the savings, thus, the negative savings sites are subject to sample selection and evaluation.

This program has not been evaluated before, so there is no pre-existing error ratio. The 2004 CA Evaluation Framework notes typical error ratios in the range of 0.4 to 1.0. We expect the project estimates of parameters such as number of units installed, nameplate efficiency and hours of operation to be quite good due to the implementer M&V and the reliance on desk reviews as the primary mode of site evaluation. However, large capital-intensive programs can be vulnerable to revisions of measure baselines which can dramatically impact the savings of a site. While DNV has observed error ratios as low as 0.34 for a similar program, 0.4 is a more prudent selection given that LEUP has no prior evaluation history and there is a possibility of baseline errors creating large discrepancies.

Table A- 3 summarizes the key sampling parameters.

Table A- 3. Summary of sample design

Parameter	Description
Population	Tracking data for projects completed in the 2021 and 2022 calendar years.
Explicit sampling strata	By fuel (electric only, gas only, and both) By Telecom customer/non-Telecom customer By size
Gross sample allocation	Targeted 90/10 by fuel
NTGR sample allocation	Same as the gross sample
Sample design approach	Stratified ratio estimation
Target parameters	GRR
Analysis domains	By fuel
Error ratios	DNV observed error ratio for similar programs of 0.4
Projected precision at 90% confidence (80% confidence for electric demand)	±10% relative precision at the 90% confidence interval for electric and natural gas energy ±10% relative precision at the 80% confidence interval for electric demand
Savings size stratification	Up to 3 levels based on savings

NTGR – Net-to-gross ratio
GRR – Gross Realization Rate

Table A-4 shows the segmentation and stratification discussed, the number of participating sites in the population, the proposed sample, and the expected relative precisions.

Table A-4. Sample design strata and sample target

High level segment	Sub-segment	Strata	Forecasted gross savings (MWH)	Forecasted gross savings (therms)	Popu-lation	Error ratio	Sample
Natural gas	Gas	L	0	3,573,959	2	0.4	2
		S	0	215,430	3	0.4	1
Electricity	TeleCom customer	L	3,324	0	5	0.4	2
		M	3,527	0	17	0.4	3
		S	3,541	0	46	0.4	2
	Site w/o gas savings	L	4,729	0	1	0.4	1
		M	5,706	0	3	0.4	3
		S	5,538	0	9	0.4	2
	Gas sites w/ electric savings	Census	5,475	0	0	0.4	*
			31,841	3,715,777	86		16

*The three gas sites are included in the electricity sample with a weight of 1 since they are selected with certainty.

The sample design above approximates the targeted precision of ±10% relative precision at the 90% confidence level by fuel while efficiently representing both electric and natural gas savings.

Sample disposition

Table A-5 summarizes the final disposition of the sample for both gross and net data collection. In three cases, the gross savings interview was completed, but in two cases, the decision maker could not be recruited for the NTG interview. DNV had contact with DNV12 staff but could not recruit the person most knowledgeable of the project. Since this site accounted for 90% of the gas savings and the MV provided in the M&V project file was substantial, the gross savings were evaluated using a desk review. One primary sample site was replaced because the site was nonresponsive.

Table A-5. Sample disposition*

High level segment	Sub-segment	Strata	Sample target	Refused, no response	Replaced sites	Gross com-pletes	Net com-pletes
Natural gas	Gas	L	2	1*	-	2	1
		S	1	-	-	1	0
Electricity	TeleCom customer	L	2	-	-	2	2
		M	3	-	-	3	3
		S	2	1	1	2	2
	Site w/o gas savings	L	1	-	-	1	1
		M	3	-	-	3	2
		S	2	-	-	2	1
	Gas sites w/ electric savings	Census	*	-	-		*
			16			16	12

* The three gas sites are included in the electricity sample with a weight of 1 since they are selected with certainty.

Potential outlier analysis

This section presents the analysis evaluators conducted to identify anomalous electric energy site results that could be potential outliers and candidates for down-weighting in the analysis. To screen for potential outliers, the analysis first mathematically screens for high influence and then consider whether the site and/or measure circumstances are different enough from the rest of the population to consider alternate weighting. This analysis follows protocols developed for identifying outliers in impact evaluations in Massachusetts.

Influence is the extent to which the estimate is changed if a particular value is omitted or changed. A point can have high leverage but not be influential if the y value is in line with the pattern in the rest of the data. To look at influence we can simply look at whether the result changes a lot if we leave the point out. High influence results from the combination of high leverage with high deviation (a large residual y – Rx).

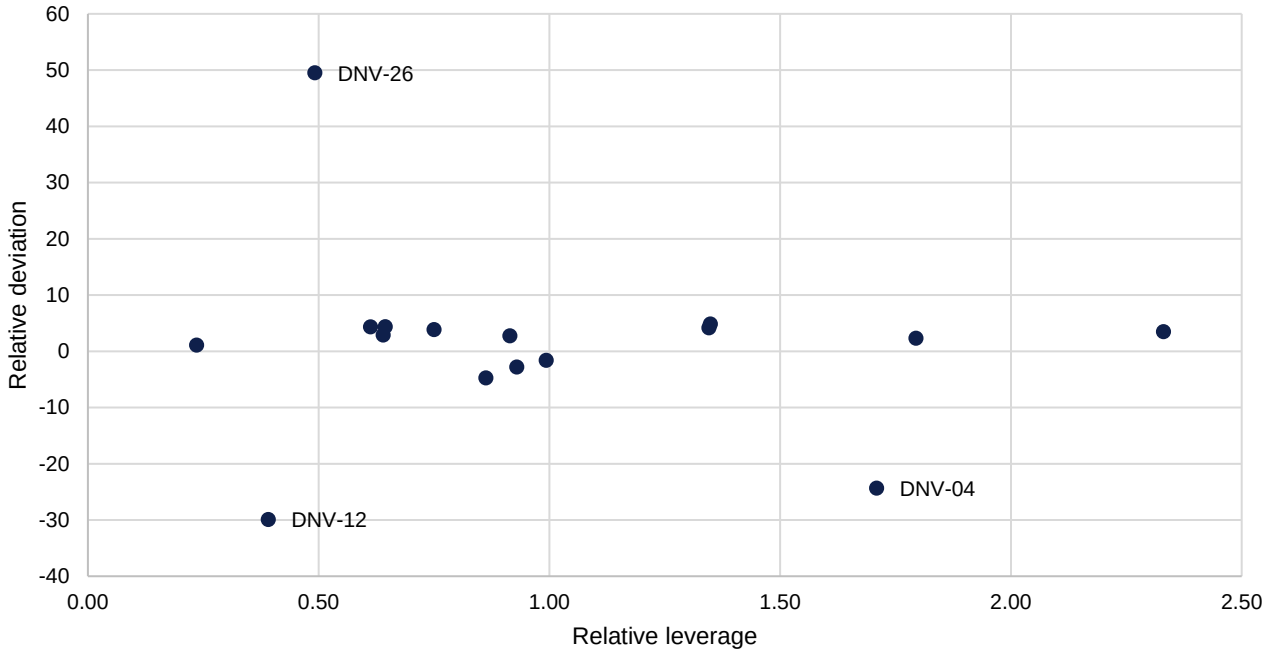
If we identify a potential outlier, we don't want to exclude or reduce its weight just because it has leverage or influence. We do our best to create close to uniform leverage but we can't create a perfectly optimal sample. Moreover, we don't want to automatically down-weight very large sites, especially on a retrospective basis. Those sites represent a large fraction of the savings, which is why they are sampled at high rates.

We would want to down-weight a high-leverage case if:

1. It's highly influential
2. We have reasons to believe its realization rate is anomalous and somehow much more extreme than is typical of its stratum or the population overall.

Figure A-1 presents the relative leverage and the influence of each of the sixteen evaluated sites. All sixteen sites have a relative leverage of less than 2.5 (which is less than the threshold value of 4 above which a site has a high relative leverage). The relative deviation of site DNV-26 tested above the relative deviation threshold of 50 and indicates an anomalous deviation, but the relative leverage is relatively low.

Figure A-1. Potential outlier analysis results



Although none of the 16 evaluated sites featured a combination of relative leverage and relative deviation that would have qualified them as potential outliers, evaluators reviewed three evaluated sites that have greater relative deviations than the other thirteen. Site DNV-26 that tested for an anomalous relative deviation installed a lighting project that erroneously claimed electric energy penalties. Sites DNV-04 and DNV-12 featured relative deviations that were not anomalous but different from the ones of the other thirteen sites, and installed projects that resulted in electric energy penalties and natural gas savings.

The evaluators determined an error in the electric energy saving reported for the lighting project installed at site DNV-26 due to an ill-conceived metering plan that could have occurred at other sites and determined the site does not feature any anomalies when compared to the rest of the evaluated sites.

The evaluators determined that the projects installed at sites DNV-04 and DNV-12 claimed valid electric energy penalties and natural gas savings and did not feature any anomalies when compared to the rest of the evaluated sites (project scope and documentation, savings calculations, and evaluation approach were similar to the ones used for the other evaluated projects).

Based on the details provided in Figure A-1 and on the findings of the additional review evaluators conducted on sites DNV-26, DNV-04, and DNV-12, the evaluation team determined the results of all sixteen evaluated sites can be used to calculate the program evaluation results.

Outlier screening methodology

Evaluators conducted an in-depth analysis to identify anomalous electric energy site results that could be potential outliers and candidates for down-weighting in the analysis.

Relative leverage

The relative leverage is a characteristic of the sample design, that is the x values (tracking results) and their weights w, independent of the value of a particular observed y (evaluated results). A point has high leverage if the parameter being estimated [such as the gross realization rate (GRR)] is very sensitive to deviations between the observed y and the underlying line at that value of x.

The leverage concept is most familiar in the context of a linear regression. For a linear regression (unweighted) a point is said to have high leverage if its x value is far from the mean of the other x values. When that happens, the regression line can be swung a lot according to whether the y observation at that x is high or low relative to the true regression line. How far a particular x value is from the mean of the x values is measured relative to the variance of x values, and the overall sample size.

For this study, we calculated a GRR with unequal sampling across strata, leading to unequal weighting. We further have a design assumption that the standard deviation of the error—that is, the dispersion around the ratio line--is roughly proportional to the x_j value. Specifically, we assume the standard deviation of the deviations from the GRR line is proportional to x_j^{γ} where the power γ=0.8.

For this weighted ratio situation, the leverage – that is, the sensitivity of the GRR to unusually large deviations from the ratio line – is proportional to the product of x_j^{γ} and the case weight w_j. We therefore screen for high leverage by looking for large values of "relative leverage", defined as:

$$L_i = \frac{x_j^{\gamma} \times w_j}{\frac{1}{16}\sum_{j=1}^{16}\left(x_j^{\gamma} \times w_j\right)}$$

Where:

L_i	= relative leverage of site j (j = 1 to 16)
x_j^{γ}	= tracking electric energy savings of site j raised to the power of γ (γ = 0.8)
w_j	= site j weight

The selection probability pj for an individual site j is set to be proportional to x_j^{γ} and its weight w_j is proportional to $\frac{1}{p_j}$. As a result, the product $x_j^{\gamma} \times w_j$ is fairly uniform. The sample design thus mitigates most extreme leverage situations. This protection from high leverage can break down in a few situations:

- An x value accounts for such a large fraction of the total of x that its allocation would be much more than 1. In that case its selection probability is lower than the general proportional rate, so that $x_j^{\gamma} \times w_j$ is larger than average.
- The completed sample doesn't meet all its targets, resulting in higher weights than planned in the design.
- The value of γ assumed for the design is very different from what is observed in the collected data.

As a rule of thumb, we suggest that high leverage is indicated if the leverage metric is greater than 4. This is based on general experience that in typical samples for the reasons indicated above, variations in this range are not uncommon.

Relative deviation (residual outliers)

To look at how extreme or anomalous a value is, we compare the residual with the "typical" residual. Now, we know that the magnitude of the residual in general increases with x_j. An outlier residual is one that is large given the magnitude of x_j. That is, we're looking not for large absolute residuals, but for large "relative residuals". We identify a site is an outlier based on a simple t test, using the following formulae:

$$Residual_j = \frac{y_j - x_j \times GRR}{x_j^{\gamma}}$$

Where:

$Residual_j$	= relative residual of site j (j = 1 to 16)
y_j	= evaluated electric energy savings of site j ((j = 1 to 16)
x_j	= tracking electric energy savings of site j ((j = 1 to 16)
GRR	= program gross realization rate
x_j^{γ}	= tracking electric energy savings of site j raised to the power of γ (γ = 0.8)

If:

$$\left[Residual_j - \frac{1}{16}\sum_{j=1}^{16}\left(Residual_j\right)\right] > \left[t \times STDV\left(Residual_{j=1}^{16}\right)\right]$$, then site j is a residual outlier

Where:

t	= test value (set to 2)
$STDV\left(Residual_{j=1}^{16}\right)$	= standard deviation of all 16 relative residuals

Residual outliers

To look at how extreme or anomalous a value is, we compare the residual with the "typical" residual. Now, we know that the magnitude of the residual in general increases with x_j. An outlier residual is one that is large given the magnitude of x_j. That is, we're looking not for large absolute residuals, but for large "relative residuals". We identify a site is an outlier based on a simple t test, using the following formulae:

$$Residual_j = \frac{y_j - x_j \times GRR}{x_j^{\gamma}}$$

Where:

$Residual_j$	= relative residual of site j (j = 1 to 16)
y_j	= evaluated electric energy savings of site j ((j = 1 to 16)
x_j	= tracking electric energy savings of site j ((j = 1 to 16)
GRR	= program gross realization rate
x_j^{γ}	= tracking electric energy savings of site j raised to the power of γ (γ = 0.8)

If:

$$\left[Residual_j - \frac{1}{16}\sum_{j=1}^{16}\left(Residual_j\right)\right] > \left[t \times STDV\left(Residual_{j=1}^{16}\right)\right]$$, then site j is a residual outlier

Where:

t	= test value (set to 2)
$STDV\left(Residual_{j=1}^{16}\right)$	= standard deviation of all 16 relative residuals

APPENDIX B. DETAILED NTGR METHODOLOGY

This appendix discusses the methodology used in the impact evaluation of the Large Energy User Program (LEUP). The methodology follows the followed New Jersey's Statewide Evaluator EM&V Guidelines Net-to-Gross (NTG) Guidance for Downstream Rebate Programs. Methodological issues were observed.

However, due to changes in the TRM to require the application of dual baseline methods and likely more scrutiny of baselines going forward, a baseline battery of questions was added to both the net survey instrument and to the gross evaluation protocol.

This appendix addresses the overall NTG method and the added panel of questions.

NTGR methods

The Net-to-Gross (NTG) methodology for the impact evaluation of the LEUP primarily relies on responses from in-depth interviews with LEUP participants.

The relevant NTG questions in the participant interview guide fall into three categories.

1. *Baseline/market event questions:* These are a series of questions asking the interviewee about the baseline conditions and market event scenarios for the projects. In earlier discussions with SWE, the study team proposed that the gross savings data collection instrument would primarily be used to collect information about baselines/market events. However, it was deemed prudent to also collect some baseline/market event information in the net interview guide because it was possible that the interviewees for the gross and net interviews might be two different people.
2. *NTG setup and framing questions:* These setup/framing questions are useful for helping the interviewees recall the projects and the decision-making that went into them that later NTG scoring questions will be asking about. This is important because some time has passed since the implementation of the evaluated projects. In addition, many LEUP participants have implemented multiple energy efficiency projects in recent years and therefore it is important for the evaluators to help the participant distinguish the evaluated project from other recent energy efficiency projects. Finally, these questions can also help the evaluators know whether there are other key decision-makers they should be talking to estimate net savings.
3. *NTG scoring questions:* These questions closely follow the recommended survey questions for nonresidential programs that appear in the New Jersey NTG guidelines and are the primary source of the NTG scoring.

Besides these three batteries of NTG-relevant questions the interview guide also contains some process evaluation questions. The following subsections describe the three NTG question categories in more detail and explain how they will be used to estimate NTG ratios.

Baseline/market event questions

As noted, these questions are designed to supplement similar, more detailed baseline/market event questions being asked in the gross energy savings data collection instrument. The baseline/market event questions in the NTG interview guide include:[16]

- Q8: “Which of the following scenarios best describes the situation concerning the installation of your [MEASURE GROUP]?” This Q8 offers seven different market event scenarios to choose from ranging from the common

[16] For the sake of brevity, the questions in this methodology do not include the response options.

replacement of existing equipment in existing spaces to the addition of controls, retro-commissioning, and major plant expansions.

- Q9: [IF Q8=REPLACEMENT OF EXISTING EQUIPMENT IN EXISTING SPACE] “Was this equipment replacement part of a major remodeling, gut rehab, or a tenant fit-out?”
- Q10: “Was there a recent or planned change to the use of the building system that your [MEASURE GROUP] was part of? These might include changes to system output due to changed lighting needs, greater heating or cooling demand in the same space, or expansion of a production line.”
- Q11: “[If Q10 = YES] You mentioned that you were making or planning changes to the building system that your [MEASURE GROUP] was part of to address new needs in the building. How well could these new needs have been met by the old system, assuming the old system was still functioning well?”
- Q12: IF Q10=YES] Apart from any adjustments to meet the new system needs, how well was the old [MEASURE GROUP] system working at the time the decision was made to replace it?
- Q13: [IF 10=YES AND THEY SAID, IN RESPONSE TO Q12, THAT THE EQUIPMENT WAS WORKING BUT THEY COULD NOT FULLY REPAIR IT] You said that the equipment you replaced was working but you could not fully repair it. For approximately how long was your old equipment in a state where it could not be fully repaired?
- Q14: [IF 10=YES AND THEY SAID, IN RESPONSE TO Q12, THAT THE EQUIPMENT WAS WORKING BUT THEY COULD NOT FULLY REPAIR IT] How serious a concern was it that your old equipment wasn't working? Would you say it was … (interviewees were then given various categories of “seriousness” to choose from}
- Q15: Which of the following best describes the recent repair history for the old equipment? (interviewees were then given various options for characterizing the repair history to choose from}
- Q16: [ASK ONLY IN DELAMPING SITUATIONS] Was the previous equipment decommissioned?
- Q17: [ASK ONLY IN DELAMPING SITUATIONS] Was the use of other equipment (other than the new system} increased to compensate for the retirement of the old equipment?

The net evaluation team then compared the responses of the NTG interviewees to these questions with the responses of the gross savings interviewees to a battery of similar baseline/market event questions. If these comparisons revealed inconsistencies in the market event or baseline characterizations, the evaluation team made follow-up queries with the participants to better understand these inconsistencies and to determine whether prior interview responses needed to be changed or clarified.

NTG setup and framing questions:

As noted, these framing questions are useful for helping the interviewees recall the evaluated projects and the decision-making that went into them while also helping the evaluators know whether there are other key decision-makers they should be interviewing to measure net savings. The NTG setup and framing questions in the interview guide included:

- Q4: What was your specific role in the project?
- Q5: Were others involved with the project decision-making, particularly the go-no go decision?
- Q6: Please tell us a little more about this PROJECT [at %ADDRESS]
    - What key motivations, considerations or factors were driving the implementation of this project?
    - [IF PROJECT WAS PART OF A LARGER EFFORT WITH MULTIPLE RELATED PARTS] How does this project tie in with those other project(s)?
- Q7: [FOR PARTICIPANTS WITH MULTIPLE EE MEASURE GROUPS] Was the decision-making process for the installation of this equipment a singular event, or did some measures have separate decision-making process than others?

- Q18: How did the idea for installing this project originate?
    - a. [IF NOT ALREADY MENTIONED] Was a vendor or consultant involved in the decision?
        - [IF YES] Was this a vendor referred to you by the program's implementation contractor TRC? Or was the vendor someone your company selected on its own?
            - [IF PROGRAM VENDOR] Did the vendor play a role in convincing your company to do these projects?
    - b. [IF NOT ALREADY MENTIONED] Was it suggested in an energy audit?
        - [IF YES] Who conducted this audit?

- Q19. About when was this project first put forward for your company's consideration?
- Q20. About when was the final decision to go ahead with the project?
- Q21. About when did your organization first begin discussions with TRC regarding funding/incentives and/or measurement and verification for this project?

NTG scoring questions

These NTG questions were recommended in the NJ SWE NTG guidelines and are used for NTG scoring.

- Q23. Before you heard about Large Energy Users Program and its financial incentives, had you already planned to purchase and install the [MEASURE X]?
    - A. [IF Q23=YES OR DK] Prior to hearing about the program incentive, was the purchase of [MEASURE X] included in your organization's capital budget?
        - i. [IF Q23A = YES] Had your organization ALREADY ordered or purchased the [MEASURE] BEFORE you heard about the program?
            - a. [IF Q23Ai = YES] Just to be clear, is it correct that you installed ordered or purchased the [MEASURE] before you heard anything about the Large Energy Users Program?
- Q24. Without the incentives and information from the Large Energy Users Program, would you most likely still have purchased the same [MEASURE X]?
    a. [IF Q24 = NO] So, without the incentives and information or education from the Large Energy Users Program, you would not have installed purchased [MEASURE X] at all. Is that correct?
- Q25. Without the Large Energy Users Program and its financial incentives and information, what efficiency level of [MEASURE X] would you most likely have purchased? (The interviewee was then provided with a list of efficiency options]
- Q26. Thinking about timing, without the Large Energy Users Program and its financial incentives and information, when would you have most likely purchased the [MEASURE X]? (The interviewee was then provided with a list of timing options]
- Q27. Did the incentive from the Large Energy Users Program help [MEASURE X] receive implementation approval from your organization?
- Q28. Without the Large Energy Users Program and its financial incentives and information, how many [MEASURE Xs] would you most likely have purchased?

- Q29. Please rate how important the Large Energy Users Program and its financial incentives were in your decision to purchase and install the [MEASURE X]. Use a scale from 1 to 5, with 1 meaning "not at all important" and 5 meaning "extremely important."
- Q30. Please rate how important a recommendation from a staff member of TRC, the company which administers the Large Energy Users Program, was in your decision purchase and install the [MEASURE X]. Use a scale from 1 to 5, with 1 meaning "not at all important" and 5 meaning "extremely important."
- Q31. Please rate how important energy efficiency information that the Large Energy Users program may have provided you through websites or program materials was in your decision to purchase and install the [MEASURE X]. Use a scale from 1 to 5, with 1 meaning "not at all important" and 5 meaning "extremely important."
- Q32. Please rate how important information from a contractor was in your decision to purchase and install the [MEASURE X]. Use a scale from 1 to 5, with 1 meaning "not at all important" and 5 meaning "extremely important."
- Q33. Please rate how important previous participation in a New Jersey energy efficiency program was in your decision to purchase and install the [MEASURE X]. Use a scale from 1 to 5, with 1 meaning "not at all important" and 5 meaning "extremely important."
- Q34. In your own words, can you please describe how important the Large Energy User Program and its financial incentives was in your decision to purchase and install the [MEASURE X]?

NTGR calculations

The evaluation team used the methodology prescribed in the NJ SWE NTG protocols to calculate the NTGR ratio. This methodology bases the free-ridership ratio on the average of the Intention Score and the Influence Score as shown in the following formula:

$$Final\ Freeridership\ Ratio = \frac{Intention\ Score\ +\ Influence\ Score}{2}$$

Each participant's intention FR score starts at 100%, then decreases based on responses to the survey questions Q23 through Q28 as shown in the following matrix:

Table B-1. Matrix for intention free-ridership scoring

Question #	Scoring instructions
Q23.	If response = Yes then 0% is subtracted If response = No then 50% is subtracted
Q23a.	If response = Yes then 0% is subtracted If response = No then 50% is subtracted
Q23a1.	If response = Yes then 100% Intention FR Score assigned
Q24.	If response = Yes then 0% is subtracted If response = No then 50% is subtracted

Question #	Scoring instructions
Q24a.	If response = Yes/correct, would not have installed without the program incentive/information then 100% is subtracted
Q25.	If response is same efficiency installed or higher then 0% is subtracted If response is lower efficiency then 50% is subtracted If response = lowest efficiency or lowest cost option available then 0% is subtracted
Q26.	If response is would have installed in same year then 0% is subtracted If response is within 1-2 years, then 50% is subtracted If response is longer than 2 years or never, then 100% is subtracted
Q27.	If response = Yes then 50% is subtracted If response = No then 100% is subtracted
Q.28	OPEN ENDED (Final intention free-ridership score = Initial intention free-ridership score multiplied by (Q28 response ÷ installed quantity).

The first step in calculating the Influence FR score is examining the importance scores for questions Q29-Q33 as shown Table B-2. The highest of these importance scores is the one used for the Influence FR score calculations. The NJ SWE NTG guidelines imply that any contractor or vendor information can be included in the Influence FR score calculations. However, our evaluation team believed that the vendor/contractor influence should only be credited to the program if the program introduced that contractor/vendor to the program participant. In cases where the participant was using a vendor/contractor for the project that they had used before program involvement or had found on their own, we gave the project no FR influence score for vendor/contractor information.

Table B-2. General FR influence component question

Q29.The program incentives	1 (Not at all important)	2 (Slightly important)	3(Moderately important)	4 (Very important)	5 (Extremely important)	Not applicable
Q30. Recommendation from program staff or implementer	1 (Not at all important)	2 (Slightly important)	3 (Moderately important)	4 (Very important)	5 (Extremely important)	Not applicable
Q31.EE information that the program provided	1 (Not at all important)	2 (Slightly important)	3 (Moderately important)	4 (Very important)	5 (Extremely important)	Not applicable
Q32.Contractor or vendor information	1 (Not at all important)	2 (Slightly important)	3 (Moderately important)	4 (Very important)	5 (Extremely important)	Not applicable
Q33.Previous EE program participation	1 (Not at all important)	2 (Slightly important)	3 (Moderately important)	4 (Very important)	5 (Extremely important)	Not applicable

EE – Energy Efficiency

Since high program influence and FR have an inverse relationship, the team then took the highest importance score from Table B-2 and converted it to an FR score using the scale shown Table B-3 (as provided in the NJ NTG guidelines). Finally, we used Q34 – where the program participant describes the program's influence on the project in their own words – as a consistency check for the NTG scoring.

Table B-3. Influence free-ridership implied by response to influence items

Influence rating	Influence FR score
1 (Not at all important)	100%
2 (Slightly important	88%
3 (Moderately important)	58%
4 (Very important)	37%
5 (Extremely important	0%
Not applicable	Remove from analysis

FR – Free ridership

NTGR methodological shortcomings

Table B-4 summarizes the NTGR findings for each of the 13 sites with a completed NTGR interview. As can be seen in the tables, there is a bifurcation between 0.5 and 1.0. Sites DNV-18 through DNV-80 were geographically dispersed sites for the same customer and the interview was completed with a single decision-maker with responsibility for the projects across these sites.

Table B-4. NTGR results by site

Site ID	Measure description	NTG
DNV-04	Heat of compression dryer, chiller upgrade	0.56
DNV-07	LED Lighting, steam blankets, gas compressor	No IDI
DNV-08	Cooling tower retrofit	1.0
DNV-10	Lighting retrofit & controls	1.0
DNV-12	Gas compressor replacement, pump	No IDI
DNV-14	Lighting retrofit & controls	1.0
DNV-16	Cooling tower replacement	1.0
DNV-17	AC to DC IMM motor replacement	1.0
DNV-18	Lighting retrofit, chiller VFD	No IDI
DNV-19	Lighting retrofit & controls	0.50
DNV-26	Lighting retrofit & controls	0.50
DNV-32	Lighting retrofit & controls	0.50
DNV-35	Lighting retrofit & controls	0.50
DNV-56	Lighting retrofit & controls	0.50
DNV-60	Lighting retrofit & controls	0.50
DNV-80	Lighting retrofit & controls	0.50

This bifurcation of the NTGRs is driven partially by the NTG protocol itself. As the equation below shows, the final free ridership ratio is prescribed to be the average of two components: the intention score and the influence score. The intention score starts at 100% free ridership and then this percentage is potentially decremented through a series of questions about the quantity, efficiency, and timing of the project in counterfactual scenarios (e.g., without the program interventions) and also questions about how far along the project decision-making had progressed before the program became involved. All but two of the LEUP interviewees finished this intention battery with 0% free ridership. For example, in response to the question: “Did the incentive from the Large Energy Users Program help the [MEASURE X] project receive implementation approval from your organization?,” all the participants said “Yes” which led to all the intention free ridership percentage being decremented by 50% just in response to this one question.

Final free-ridership ratio components

$$Final\ Freeridership\ Ratio = \frac{Intention\ Score\ +\ Influence\ Score}{2}$$

The influence component of the final free ridership ratio is based on responses to a series of questions about potential program influences including incentives, recommendations from the program implementers, program information, information from a contractor, and previous participation in a New Jersey energy efficiency program. For each of these potential program influences, the questions asked the participants to assign an importance score where five equaled "extremely important" and one equaled "not at all important." The state NTG protocol prescribes that the max value of these program influence importance scores be used to calculate the influence free ridership values. If the max importance value was five, the influence free ridership ratio would be 0%. All but one of the LEUP participants had at least one program influence importance score of five and so these all had influence free ridership ratios of 0%. This result meant there was a de facto "floor" on the overall NTGR because it was rare when the influence ridership was not 0%.

Due to this great potential for both the intention score and the influence score to be reduced to zero, the end result was that all but two of the LEUP interviews scored NTGRs of 100%. The final program NTGRs did end up in the 56%-58% range (depending on fuel type) but this was solely due to the fact that the participants with partial free ridership had some of the highest savings weights in the sample. If this had not been the case, the program NTGR would have been much higher.

After implementing the state NTG protocol for the LEUP program, DNV had some concerns about both the method and its outcomes. First, unlike NTG methods for custom program in other jurisdictions, there are no recommended questions about non-program project drivers such as a corporate carbon reduction goals, manufacturing production improvements, or regulatory compliance. Anecdotal information from the LEUP interviews, and evidence from custom program NTG evaluations in other jurisdictions, indicate that these non-program project drivers can be very important. However, for the state NTG protocol, none of the prescribed NTG scoring questions accounted for these potential non-program project drivers.

Second, one of the prescribed program influence questions is: "Please rate how important information from a contractor was in your decision to purchase and install the [MEASURE X]." In other jurisdictions, the influence of a contractor on the project NTG is contingent on that contractor being program-supplied/referred. If the participating end user used a contractor that they had selected without the program's assistance, the program gets no attribution for the contractor influence.

However, the New Jersey NTG protocol has no such restrictions and so even customer-sourced contractors can contribute to lowering the influence free ridership score. This is concerning because evidence from both the participant interviews and the program implementer interviews indicated that the LEUP participants were more likely to choose their own contractors.

A final concern is the inclusion of previous New Jersey energy efficiency participation as one of the program influence factors that can increase the overall NTGRs. We object to this factor on three grounds. First, if previous program participation has helped end users better understand and value the benefits of energy efficiency projects, then this attitudinal transformation would imply that these end users are less likely to need incentives on the next projects because they have already learned the benefits from the previous ones. But by using this past participation factor to potentially increase the NTGRs, this is sending the opposite signal: that prior participants need more incentives.

A second objection to the inclusion of previous New Jersey energy efficiency participation as one of the program influence factors is that one objective of calculating NTGRs is to assess whether the energy efficiency program in evaluated year X spent its resources wisely. So, in the purest sense this should involve looking at the program activities and expenditures in year X and then comparing these with attributable (net) energy savings that these year X activities/expenditures were used

to achieve. Once you add past program influences (and implicitly past program expenditures) into the mix, it "muddies the waters" as to whether one is truly measuring the effectiveness of the program expenditures in year X or something else.

Finally, we believe that a frequent argument for using such a past program participation factor – that it is designed to compensate the energy efficiency program for past unmeasured spillover – is unsupported. First, New Jersey NTG protocols do require the measurement of spillover and so this argument will be less credible in future years. Second, other jurisdictions that have measured participant spillover for C&I programs have historically found levels below 10% of total energy savings. So even if New Jersey was not measuring spillover, it would be incongruous that 100% of current projects should get credit for past unmeasured spillover that occurred, on average, in less than 10% of past projects. Finally, it is also worth noting that in other jurisdictions such as California, previous program participation is treated as a factor that can potentially lower NTGRs, not increase them.

M&V assessment

As part of the site-specific impact assessment, the engineer reviews project files, interviews a knowledgeable person about the operation of equipment, and potentially gathers additional information through on-site inspections. The baseline condition assessment protocol will leverage these gross impact activities to determine the state of the existing equipment and to determine if there were other external mandates for the measure's installation. The next section describes the protocol for gathering evidence for the baseline condition.

Protocol description

The engineer will gather evidence from these sources and select an initial market event type that best matches the evidence.

1. Review of the project file to determine the full scope of the project, including project context, financial information (like costs, payback or BCRs), other concurrent work or measures, alternate quotes or configurations that were identified.
    a. Using the market event interview guide and information from the file review, the engineer will identify the key probing questions. See the Interview Guide
    b. Interview the operational manager. This is typically a first interview and will also have the objective of confirming the installation and operation of the measure.
    c. The engineer will summarize the evidence in a narrative form and make an initial recommendation for a market event type for each of the measure groups/systems.
        i. The engineer will complete a check-list of key elements that define the selected event type, and indicate which portion of the narrative supports these.
2. The NTGR surveys will be completed.
    a. If the operational manager who completed the M&V interview also completes the NTGR survey, the baseline condition battery be skipped.
    b. If another decision maker completes the NTGR survey and that decision maker survey results in a different baseline type DNV will resolve any discrepancy based on a preponderance of evidence, including conducting follow-up interviews with both parties as needed to reconcile the differences.
3. The engineer will update the narrative and the check-list with any survey findings, and recommend a market event type for each of the measures.
4. The evidence and conclusions will be reviewed by senior DNV team, and also made available to the SWE and BPU staff.

Conducting the interview

There are three outcome scenarios:

- Scenario 1: The event type is clearly correct. After the customers summary of the existing equipment, the engineer will summarize and have the customer confirm. The details will be recorded.
- Scenario 2: The event type is clearly wrong. After the customer summary of the existing equipment, the engineer will summarize what the customer said – and pose a counter-factual. For example, if the customer states the equipment had failed (although the event type was listed as an early retirement), the engineer will probe by asking about whether the equipment was at all operational or whether easy repairs would have restored its function. The details will be recorded.
- Scenario 3: There is some ambiguity as to whether the event type is correct or not: The difference between an early replacement and a time-of-sale (end-of-life) requires extra care. The engineer must probe for red flags, maintenance practices, and how well the existing equipment was serving the site.

M&V plan template and check-list

The M&V Assessment protocol is embedded in a M&V Template spreadsheet which is used by the engineer to plan the site M&V, including identifying specific customer questions developed through the protocol. The M&V Plan template is also designed to capture the site specific findings and includes the check-list noted in Section 2.1.

Comparison of NTGR and engineering assessments of baseline conditions

DNV added a battery of questions for identifying the baseline conditions of a given project to both the gross savings and net savings data collection instruments. DNV had developed these questions in cooperation with the SWE and had produced a memo describing the protocol and scoring. The following is a summary of the baseline questions in the NTG interview guide:

- *Questions 8-9:* These questions asked participants to choose from a list of baseline market conditions to identify which best characterized their project. If the participant selected a scenario where the energy-efficient measure was replacing existing equipment in an existing building, the survey asked them whether the equipment replacement was part of a major remodeling, gut rehab, or a tenant fit-out.
- *Questions 10-14:* These questions asked participants about possible changes to the use of the building systems such as changes in lighting/HVAC needs or expansion of a production line. If such changes are identified, the questions ask about the performance level of the old systems and whether they could have accommodated the new building changes.
- *Question 15:* This question asks for a description of the recent repair history for the old equipment.
- *Questions 16-17:* These questions, asked only in delamping situations, query whether the previous equipment was decommissioned and whether the use of other equipment (other than the new system} increased to compensate for the retirement of the old equipment.

The NTG interview guide with the complete question language appears in APPENDIX D.

After data collection was complete, DNV's gross savings and net savings teams independently assigned baseline conditions to each project and then met to compare notes. The meetings revealed that the gross and net teams had reached the same conclusions as to the baseline conditions for all but two sites. After discussion and re-examination of the evidence, the teams settled on one of the proposed baseline conditions for each of the two sites.

Reflecting on this baseline condition analysis process, the team DNV identified the following advantages and disadvantages of the process:

- Advantages:
    - *The baseline questions in the NTG guide help remind program participants which projects the evaluation is focusing on:* These memory triggers are important because LEUP participants are large C&I customers who typically have recently implemented multiple EE projects and therefore it is important to ensure they are focusing on the evaluated project.
    - *The baseline questions in the NTG guide help remind program participants of non-program project drivers:* As discussed elsewhere in this report, one limitation of the current NJ NTG methodology is that it does not remind participants of non-program project drivers such as corporate carbon reduction goals or manufacturing process improvements. The baseline questions in the NTG guide can help remind the participants of these non-program drivers.
    - *The process provides an additional perspective if the interviewees for gross and net are different people:* Sometimes, for a given participating site, the onsite contact whom the DNV team interviews for the gross savings calculations is a different person than the project decision-maker whom the team interviews for the net savings calculations. In such cases, it can be useful to have multiple participant perspectives on what the baseline conditions should be.
    - *The formalized structure of baseline questions in net interview guide serves as useful crosscheck for less structured baseline questions from the gross savings team.* As noted, the engineers who ask the baseline condition questions from the gross savings data collection instrument have some latitude in how they ask these questions. In contrast, the net interviews ask the baseline condition questions more formally and therefore may collect some information about the baseline conditions that the more informal gross survey did not collect.
    - *Occasional differences between the gross and net savings teams as to the identification of the baseline condition for a given site may force a deeper examination of these assumptions and the evidence that supports them.*
- Disadvantages
    - *Lengthening an already long NTG interview guide:* As noted, these interview guides already have separate batteries of framing/setup and NTG scoring questions and so adding the baseline condition questions runs the risk of tiring out some interviewees.
    - *If the gross savings and net savings interviewees are the same person, such an interviewee may view the repetition of the baseline questions in two different data collection instruments to be onerous.*
    - *If there is a long period between when the gross savings and net savings teams complete their participant surveys, and the two teams reach different conclusions about the baseline conditions, it can be onerous for the gross savings team to redo their gross savings analysis due to a change in the baseline condition assumptions.*

NTGR methods findings, conclusions and recommendations

After implementing the state NTG protocol for the LEUP program, DNV had some concerns about both the method and its outcomes.

- Unlike NTG methods for custom program in other jurisdictions, there are no recommended questions about non-program project drivers such as a corporate carbon reduction goals, manufacturing production improvements, or

regulatory compliance. Evaluations of custom programs in other jurisdictions indicate that these non-program project drivers can be very important.

- The New Jersey NTG protocol allows the customer-sourced contractors to contribute to lowering free ridership score. This is concerning because the NTGR should reflect program influence, not a non-program influence.
- The New Jersey NTG protocol includes previous energy efficiency participation as one of the program influence factors that can increase the overall NTGRs. We object for several reasons.
  - If previous program participation has helped end users better understand and value the benefits of energy efficiency projects, then these end users should be less likely to need incentives to encourage them to do future energy efficiency projects. Yet this prior program credit, by artificially increasing NTGRs, sends the opposite signal that the program should maintain past incentive levels because this spending is purportedly justified by these higher NTGRs.
  - This prior program credit confounds the use of NTG as a tool for assessing whether program expenditures for the evaluated years were prudent, by giving the program additional credit for program activities (and implicitly program expenditures) that occurred before the evaluation period.
  - A frequent justification for these past program credits is that energy efficiency programs need to be paid back for possible uncredited past spillover. However, since New Jersey requires the measurement of spillover, these arguments are less compelling. In addition, jurisdictions who have measured spillover for custom programs have found its incidence to be low (less than 10% of total savings). Therefore, even if spillover was not being measured, it would not be justified to give 100% of all projects a past program credit for uncredited spillover which occurred, on average, less than 10% of the time.
  - Other jurisdictions (e.g., California, Illinois) consider past program participation as a factor that *reduces* NTGRs in custom programs.
- Finally, the NTGR calculations produce a bifurcated pattern of NTGRs that were either 100% or in the 50%-56% with no NTGRs below or between these levels. These NTGRs are very different than the NTGRs coming out of custom programs in other jurisdictions which have a broader range of values.

This evaluation tested the value of adding a battery of questions to the NTG survey instrument and to the engineering interview protocol for determining the baseline condition for individual measures.

The findings, conclusions, and recommendations regarding methods and guidance are summarized in Table B-5.

Table B-5. Summary of methods and guidance findings, conclusions and recommendations

Findings and conclusions	Recommendations
DNV noted that the current NTGR individual site results cluster around either 0.50 or 1.00. DNV had some concerns about the NTGR method and outcomes. The methods do not align with other jurisdiction best practices in accounting for program and non-program influences.	DNV recommends that a stakeholder meeting be convened to discuss possible adjustments and enhancements to this NTG method. Possible changes to discuss might include: • Adding questions to the NTG protocol which ask directly about possible non-program project drivers and revise the program influence factor to better reflect both program and non-program project drivers (as is done in other jurisdictions such as California and Illinois)

Findings and conclusions	Recommendations
	• Limiting the contractor influence factor to only program provided/recommended contractors as appropriate for measuring program influence • Removing the prior program participation influence factor.
DNV tested a battery of questions which were incorporated into the NTG survey instrument and added to the site engineer in-depth interview protocols. The results were mixed. • The battery added significantly to length of time required to administer the survey which increased the risk that survey respondents might pay less attention to other survey questions • However, the baseline protocol added to the engineering instrument improved baseline documentation and consistency between projects. • In the NTG survey instrument, the baseline battery had some benefits in improving participant recall of the projects being evaluated and reminding them of non-program project drivers. • The process of reconciling the engineering interview with the NTGR interview was helpful, but the timelines of the two efforts are difficult to synchronize.	DNV recommends adding the baseline protocols to the engineering data collection protocols and see little downside in doing so. DNV recommends further testing of the baseline battery in the NTG survey to reduce its length without sacrificing the collection of useful baseline information.

Future evaluation

Table B-6 reports the error ratios derived from the results of this evaluation. The error ratio characterizes the site-to-site variance in the ratio of the site-specific evaluated savings and the tracking savings. Higher error ratios indicate higher levels of variance and consequently larger sample sizes to meet precision targets. This evaluation was planned to use an error ratio of 0.40, based on DNV's experience with similar programs. The unexpectedly high variance was created by large swings in savings, especially those that switched signs from negative tracking savings, like DNV26 that ultimately showed positive savings. If the program adopts the recommendations of this study, the error ratios in subsequent evaluations are likely to be lower.

Table B-6. Evaluation error ratio findings

Measure type	Sample planning error ratio	Study determined error ratio findings	Current study GRR RP
Electric measures	0.40	0.70	23%

Measure type	Sample planning error ratio	Study determined error ratio findings	Current study GRR RP
Natural gas measures	0.40	0.27	25%

GRR – Gross realization rate
RP – Relative precision

Another small data challenge was getting participants to identify the years when projects were approved as these “go, no go” project decisions often happened many years ago. DNV recommends asking questions about project timing via email before the interview, so the interview flows are not interrupted.

Recommendations

- Future LEUP impact evaluation sample designs should adopt the error ratios determined in this study with some potential modifications depending on program adoption of recommendations.
- Participant interviewees should be emailed ahead of the interview with a list of questions that may require a modest amount of preparation.

APPENDIX C. DETAILED PROCESS EVALUATION FINDINGS

This section summarizes key findings from in-depth interviews with TRC representatives who were involved in the LEUP program design and delivery.

LEUP Program Staff Interviews

Changes in program design

The implementer interviews revealed that the program had gone through several iterations since it first began as a pilot program back in 2011, mostly in response to participant feedback. For example, an earlier iteration of the program had limited project enrolments to a single 45-day period each year. However, due to participant feedback the program eventually removed this requirement in favor of a more open-ended enrolment process. One TRC representative explained this change:

> *We found that the [45-day enrolment requirement], after a number of years, was a challenge because some people didn't know what the project was going to be … So, we thought that … flexibility in the process was important. So we took that 45-day enrolment requirement away and went to a full-fledged program, not a pilot. And it's now a rolling enrollment. They can enroll anytime they wish to.*

Besides more flexibility in the timing of program enrolment, the program also added flexibility as to project scope. One TRC representative explained:

> *We then had it restricted to mostly one energy plan being submitted, which could be one project, 20 projects across one building, several buildings. So we made a change to that to say: "Well, maybe they have $1 million worth of eligibility, but they only have a project of that's going to consume half of that. And then to not have the other half go unused, you can submit another project within that same current year, and that could be considered … so again, flexibility*

A third change to the program design was to allow participants to "bank" their incentive dollars. A TRC representative elaborated on this:

> *We also did what's called banking between program years. So in a very simple example, let's say a customer is eligible for, $1 million a year… and they didn't have a project this year to use that money, but next year, they have an even bigger project coming, which they would love to do … but the $1 million isn't quite enough to get it over the finish line from their decision branch, … So, [the program rules were changed to] say: … "If you're eligible for $1 million last year, and you didn't use it, but you want to use one this year, you can now do $2 million." So that helped move that bigger project along because of the larger amount of money.*

Over time the program also reduced the minimum energy consumption thresholds for program eligibility. "We changed the threshold for the eligibility of these large energy customers because over the years, there were some customers that were large that were over the threshold, but because they did a good job [implementing energy efficiency projects], they now fell under it," said a TRC representative.

The TRC staff indicated that it was a delicate balancing act to both accommodate the demands of these large energy users and maintain the rigor and integrity of the program. "We've made changes to the program to try to be as accommodating as possible while maintaining the intent of the program to not essentially drop off a suitcase in cash and say: 'Go ahead,'" said one TRC representative.

Program marketing and outreach

The TRC staff described the program marketing and outreach efforts as including:

- *Targeting of large energy users by TRC outreach team:* The TRC representatives said that this outreach team helps reach out to large energy users on behalf of the LEUP program and also tracks potential program participant leads using Customer Relationship Management (CRM) software. The representatives noted, however, that this outreach team is not a dedicated marketing resource for the LEUP program, since the team also markets other programs in TRC's portfolio.
- *Maintaining and cultivating relationships with large energy users who have previously participated in the LEUP program:* "Given the history of this program, a lot of the applicants that we work with come back every year or every other year, … so generally speaking, we've had the same points of contact for quite some time," one TRC representative observed. However, this representative also observed that there is turnover in some of these participating companies and therefore the LEUP program sometimes must identify new contacts at the companies and then educate them about the program benefits and participation requirements.
- *Program-knowledgeable contractors/consultants:* The TRC representatives said that there is a small group of contractors and consultants who are familiar with the LEUP program and who can help guide large energy users through the program requirements. "They check out the website, they know what's happening, … they tend to try to be as knowledgeable as they possibly can," said one TRC representative. "They're staying in tune to the set of opportunities … they obviously want to go after these larger entities … to sell that bigger project."
- *Working with organizations that represent C&I customers in New Jersey.* The TRC representatives identified both the Large Energy Users Coalition and the New Jersey Business Action Center as organizations that they work with. The New Jersey Business Action Center aims to both attract new companies to New Jersey and keep existing companies from leaving the state and it often advertises energy efficiency programs like LEUP as part of a larger package of benefits that the state can offer.

Recent changes in program focus

The TRC representatives identified some recent changes in the focus of the TRC program including:

- *Greater focus on non-energy-impacts (NEIs) such as greenhouse gas emission reduction:* The TRC representatives said that they have been trying to convince the BPU representatives who must approve the larger projects (e.g., those using more than $500,000 in incentives) to look beyond traditional measurements of project value – such as avoided costs or simply payback calculations – and consider NEIs such as greenhouse gas emission reduction and reduced maintenance costs. The TRC representatives noted that they have proposed a decarbonization pilot program where the program incentives would be based on carbon reduction rather than electric or gas savings. Representative verbatims from the TRC staff on these topics included:
    - *"Simple payback has been a method that we've used in the past, and we've been working with [BPU] staff because simple payback doesn't really tell the whole story. Because greenhouse gas emissions savings seems to a bigger and bigger driver of why customers are doing some of these projects, it's requiring more of an explanation [to the BPU staff] of … what's happening. … What are the avoided [equipment] repair, or replacement costs that are being saved that wouldn't be reflected in the simple payback?"*
    - *"We have to provide [the BPU reviewers] with some additional context. This [project] is going to help the comfort of the building … or it's for greenhouse gas emission reductions, and things of that nature … It's not an all energy-based [focus], but it is a focus where [energy savings] is one piece of the bigger puzzle."*

- *Greater EE measure diversity:* The TRC representatives said that in response to New Jersey energy policies encouraging more comprehensive energy efficiency projects, the LEUP program had recently taken steps to encourage a more diverse EE measure mix. “Over the last couple of years, we’ve implemented some strategies to try to promote comprehensiveness, we vetted a 50% cap on savings coming from lighting,” said one TRC representative. The representative also said that the LEUP program has insisted that participants only pursue lighting projects if all other energy efficiency options have been exhausted. “In any instance where a building has excessive lighting, we’re asking the customer applying to demonstrate that there’s nothing else in the building that can be done, that’s not lighting, just showing us that they’ve us that they’ve done the due diligence and that truly there’s nothing else to be done, no old equipment to replace that they’re putting off,” she said.
- *Competition with other New Jersey C&I programs:* DNV asked the TRC representatives whether they were concerned about competition to the LEUP program from other C&I energy efficiency programs that the New Jersey IOUs had recently introduced. They acknowledged that this was the case. “For lack of a better word, there’s a bit of competition,” said one TRC representative. “The LEUP program is the only program we continue to run that’s for existing buildings and so now, in some cases, the utilities have programs that these same customers are eligible for.” One TRC representative claimed that this broadening of program choices had the potential for customer confusion.

  “This is one of the rare instances where these customers kind of have to shop around and it’s not abundantly clear that you go here if you have this kind of a project. They could use some of the utility offered programs, or they could use the LEUP. So, there is a little bit more confusion I think in the marketplace than there was before for this segment of customers.”

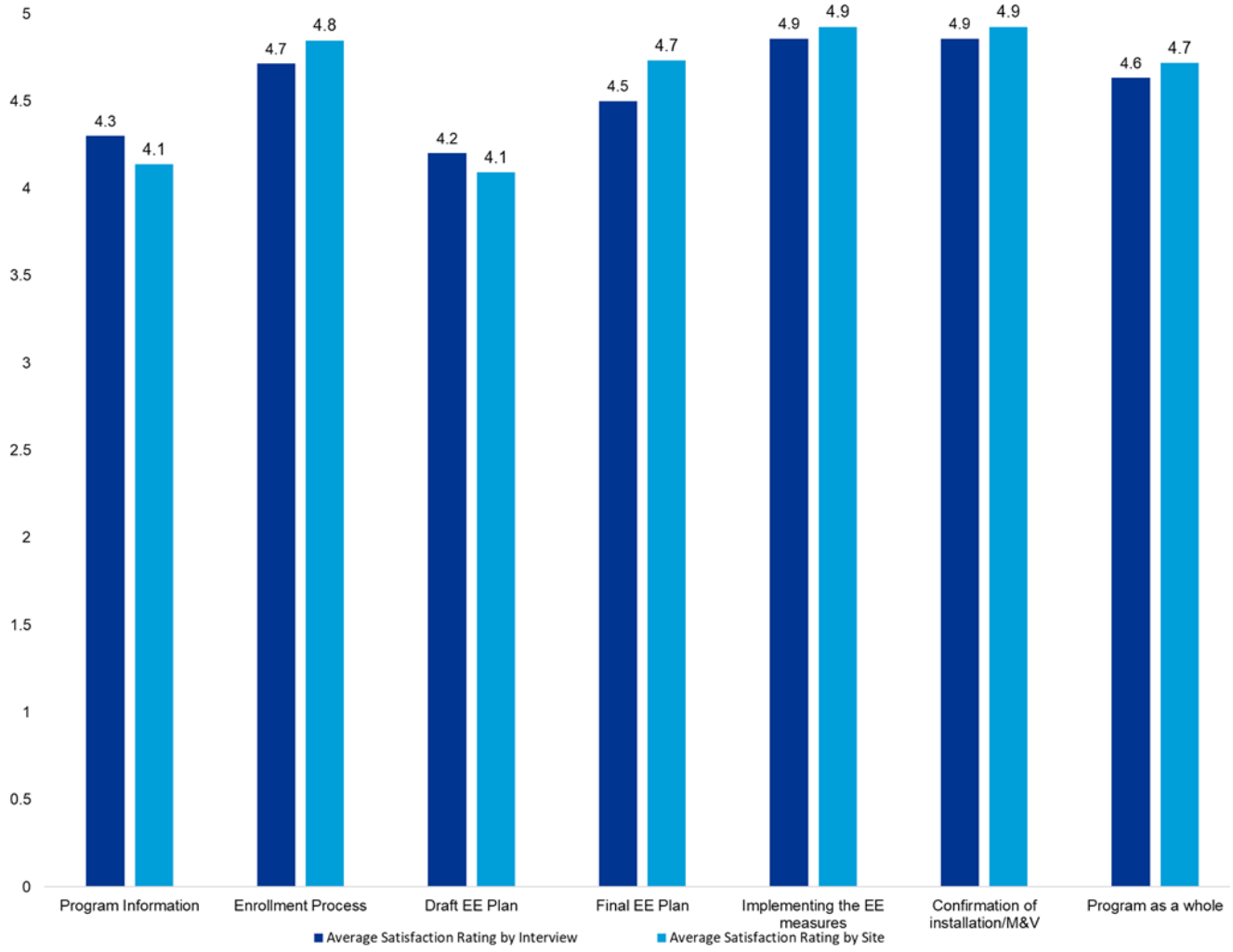
  Financing was one utility program offering that the TRC representatives thought might cause them to lose some customers. “PSE&G’s [programs] offer financing, which we don’t,” said one TRC representative. “So certain projects may fit better over on PSE&G than here. … Financing at 0% can be very compelling versus whatever incentive could be offered elsewhere.”

  Despite this competition, the TRC representatives claimed that they were open with large energy users about their range of options. “Our angle, and what our outreach team does is, to want these customers to understand everything that’s available to them, not just what we offer, but what the utilities offer, and help them make the best decision for them.”

Participant interviews

Besides these in-depth interviews with the LEUP program implementers, DNV also completed in-depth interviews with program participants. These interviews used the same guide described above in the NTGR section. For the process evaluation component of these interviews, the main focus was to collect information about participant satisfaction with program information, requirements, and processes. The satisfaction questions used a five-point Likert scale where 5 equaled “very satisfied” and 1 equaled “very dissatisfied.” DNV calculated the average satisfaction levels using both the number of interviews and the number of sites (Figure C-1). The chart shows that overall program satisfaction was very high (4.6-4.7 average satisfaction levels) and satisfaction with most of the program attributes was also very high with two exceptions. Satisfaction with the program information and the requirements for the draft energy efficiency plan were both lower (4.1-4.3 average satisfaction levels).

Figure C-1. Participant satisfaction with LEUP requirements and processes



Onerous program requirements. Some participants elaborated on their lower satisfaction ratings for the program information and draft energy efficiency plan. The following verbatims illustrate the viewpoints of some participants that the program information requirements were either unclear clear or onerous to comply with:

- *Program information requirements unclear:*
    - "When it comes to submitting the nitty-gritty details of the paperwork, some of the aspects could have been clearer."
    - "Sometimes [the program information] is a little ambiguous … which results in a follow-up conversation."
    - "We've had back and forth [with the program implementer] …. with the exact calculations and what needs to be included in coming up with the final calculations, like the labor and how to calculate that and materials and how to calculate that. There was perhaps some more back and forth that isn't necessarily as clear as day in the program documentation."

- *Onerous program requirements:*
    - "I feel like there's a lot of [requested] information that's not applicable. … We're a site with 40-some-odd buildings. They wanted to know the total square footage of the site. We were doing a project in this one building. So, who cares what the size of the building over there is? It has nothing to do with my project.
    - "[It is difficult] just trying to keep track of all the materials and the quotes and making sure that everything is organized to submit everything in one package."

Financing. Interviews with the LEUP implementation staff revealed their interest in knowing whether LEUP participants viewed the program's absence of financing as a barrier to participation. So DNV asked the LEUP participants whether this lack of financing was a barrier to their participation in the program. None of the participants said that it was a barrier. The following are some representative quotations from these participants:

- "I would say [the lack of financing was] not a barrier because, in some ways, it was almost simpler to just do it, then get the rebate, even though it took a lot of paperwork and a long time."
- "With the corporate office, when the project is as good as it was, as we described it, I don't think [the lack of financing] would have even mattered."
- [The lack of financing] is not a barrier. I mean, [our company] is a $2 billion company. So once projects get approved, you very rarely have any issue in securing the funds."
- "[The lack of financing was] not a barrier, no. We prefer not to finance. We budget for these expenses. So we're not looking to finance."

Competition with Utilities. The SWE team was interested in learning why participants had selected the LEUP program for their energy efficiency projects instead of C&I energy efficiency programs that the NJ utilities offer. The participants cited a wide variety of reasons for choosing the LEUP program including unawareness of alternative programs, the LEUP program being the only one offering adequate incentives, recommendations from their energy managers, successful previous experience with the program, the fact that their projects were more custom than prescriptive in nature, and the fact that, for a period of time, the LEUP program was the only relevant program available. The following verbatims illustrate some of these reasons:

- *Unawareness of alternative programs:* "I don't believe we were aware of any other programs at the time."
- *LEUP was only one offering enough incentives:* "I don't know of any other rebate program that pays you to this extent. … I don't think there's anything competing with that … The LEUP, to me, is the best way to go … we have five factories in New Jersey."
- *Energy managers recommended it:* "We have a department within our corporate office which handles the purchasing of all utilities, electrical, natural gas, oil … and they're the ones who recommended that we use [the LEUP program]."
- *Successful previous experience with the program:* "We've used this [LEUP] program successfully for other projects prior to these."
- *LEUP is a custom program:* "[Our EE measures] they're all custom. And you have to look at the building holistically … I used to do building modeling for a living, and in the energy sphere for a long time. When you do these projects, you have to understand how the whole building works."

- *LEUP was the only option available:* I think at the time, this was really the only one available. I know that now a lot of [programs] are offered by the utilities directly."

Besides these process evaluation findings that were in response to direct survey questions, the participants offered some other suggestions for improving the LEUP program on a more *ad hoc* basis. One participant, with extensive experience promoting energy efficiency projects in the commercial real estate sector, recommended that the LEUP program subsidize the costs of energy audits. "The real solution to [increasing LEUP EE project volume] would be to frontend the audit process [costs]," he said, "You'd see projects happening all over the place."

This participant explained that these audits were very expensive and that commercial real estate energy managers usually can only get energy efficiency projects approved if they have an analysis demonstrating future energy savings. He elaborated on this predicament:

> *"It's a constant battle. I've got to convince the operations VP to spend the money to do the audit. Then you do the audit. … The results, hopefully, are interesting enough to where the investment is compelling. Then you've got to fight for the money to do the analytics to be able to do the FEEP [Final Energy Efficiency Plan]. And then when you do the FEEP, you've got to fight to get the project approved with the builder."*

This participant also said that this barrier – the high cost of energy audits – was likely more significant for the commercial real estate sector than other sectors such as manufacturing:

> *"In some businesses, maybe it would be easy. … If you're running a manufacturing plant, where they're used to looking at procedures, protocol, … and constantly figuring out how to improve processes, they're used to these kinds of decisions. The commercial real estate world is not used to these kinds of decisions. They're building buildings. They're renting space. They're doing tenant fit up. It's all about generating rent dollars, not about generating savings. … So. the easier the [LEUP] program can make it to get to that end, so much the better."*

The interview with the TRC staff who manage the LEUP program revealed that they do not view such detailed audits as part of the current program design. One TRC representative said:

> *Our program doesn't provide things like that [a detailed audit]. … this was meant to be a self-directed program. So the large energy users themselves are developing their scope of work. They're handling that more detailed audit, things like that. Our team comes in, and we'll do a pre, post inspection of the site to confirm the information that they're providing in their reports are accurate.*

Therefore, if the program was to add this audit subsidization component, it would require a shift in the implementer's perspective that the program was meant to be "self-directed."

APPENDIX D. CUSTOMER INTERVIEW GUIDE

Process and NTG Interview Guide

For LEUP participants

NTG, Process and Baseline Interview Summary	
Respondent:	Measure decision-maker
Mode:	In-depth interview conducted by an energy professional, by phone or in-person if the decision-maker is also the facility manager and the site is subject to on-site gross impact evaluation
Target Duration: 30 minutes	
Key Objectives:	Free ridership
	Spillover
	Baseline event type (participant perception)
	Participant satisfaction by program stage
	Recommended program improvements
Algorithms:	Based on NJ EM&V Guidelines Net-to-Gross (NTG) Guidance for Downstream Rebate Programs

Interview Information

Interviewer		Survey Length (min)	
Completion Date			

Contact Information

Phone	
Email	

Call Tracking

Date/Time	Notes

Introduction

[NOTE: THE QUESTIONS IN THIS INTERVIEW GUIDE WILL NOT NECESSARILY BE READ VERBATIM BUT MAY BE MODIFIED TO SUIT THE INTERVIEW]

1. [IF END USER CONTACT OTHER THAN LEAD END USER CONTACT IS ANSWERING THE PHONE] Hi, my name is X OF DNV. We are calling on behalf of the New Jersey Board of Public Utilities (BPU) and New Jersey Clean Energy Program (NJCEP). According to our records, you recently installed [EE MEASURES] at [ADDRESS] which received financial incentives through the Large Energy User's Program, currently being administered by TRC. Can we please speak to [LEAD END USER CONTACT] about this project?

[IF THEY ASK WHY WE WANT TO TALK TO LEAD END USER CONTACT] We are interviewing customers that participated in the Large Energy User's Program to gain a better understanding of how and why they decided to install energy efficiency measures through this program.

2. [IF LEAD END USER CONTACT IS ANSWERING THE PHONE] Hi, my name is X OF DNV. We are calling on behalf of the New Jersey Board of Public Utilities (BPU) and New Jersey Clean Energy Program (NJCEP). According to our records, you recently installed [EE MEASURES] at [ADDRESS] which received financial incentives through the Large Energy User's Program, currently being administered by TRC. Are you okay answering some questions about this project? Any information that you provide will remain strictly confidential. We will not identify or attribute any of your comments or organization information.

    a. [IF YES, SKIP TO Q3]

    b. [IF YES, BUT THEY CAN'T DO THE INTERVIEW AT THAT TIME, SCHEDULE ANOTHER TIME]

    c. [IF NO, AND IF CONTACT NAME WAS OBTAINED BY TRC] I was told by TRC that you were the most knowledgeable and the most involved with the decision to implement the project I just mentioned. Is that correct?

        i. [IF YES] So, we will need to interview you. Can we go ahead with this interview?

            1. [IF YES BUT IT'S NOT A CONVENIENT TIME, RESCHEDULE THE INTERVIEW AND THEN BEGIN WITH Q3]

        ii. [IF NO, OBTAIN ALTERNATE CONTACT INFO AND SCHEDULE INTERVIEW WITH NEW PERSON]

d. [IF THEY ASK HOW LONG THE INTERVIEW WILL TAKE] The interview will take approximately 30 minutes

e. [IF THEY ASK TO VERIFY WITH THE THAT THE RESEARCH IS LEGITIMATE, GIVE THEM ONE OF THE FOLLOWNG CONTACT NAMES]

Philip Chao, BPU, (609)322-9618 Philip.chao@bpu.nj.gov

Leigh Cignavitch, TRC, (732) 603-1054 LCignavitch@trccompanies.com

CONFIRMATION OF CORRECT RESPONDENT, PROJECT BACKGROUND

3. According to our records, your company implemented a project involving <%MEASURE> at <%ADDRESS> on approximately <%INSTALL_DATE>, is all this information correct?

    a. [IF YES, SKIP TO Q4]

    [IF NO, MARK ANY CORRECTED INFORMATION IN THE FOLLOWING TABLE]

Project Information	Information from Tracking Data (pre-entered)	Corrected information (if relevant)
Measure		
Address		
Install Date		

4. What was your specific role in the project?
5. Were others involved with the project decision-making, particularly the go-no go decision?
    a. [IF YES, OBTAIN THEIR NAMES AND CONTACT INFO]
6. Please tell us a little more about this PROJECT [at %ADDRESS]
    a. What key motivations, considerations or factors were driving the implementation of this project?
    b. [IF PROJECT WAS PART OF A LARGER EFFORT WITH MULTIPLE RELATED PARTS] How does this project tie in with those other project(s)?
7. [IF RELEVANT] Our records show that your organization installed more than one MEASURE GROUP through the LARGE ENERGY USER'S PROGRAM. They are … <%QTY_1> <%MEASURE GROUP>, <%QTY_2> <%MEASURE GROUP>, <%QTY_3> <%MEASURE GROUP>. Was the decision-making process for the installation of this equipment a singular event, or did some measures have separate decision-making process than others?
    a. [IF RESPONDENT SAID SOME MEASURES HAD SEPARATE DECISION-MAKING PROCESSES THAN OTHERS] Please explain which measures had separate decision-making process than others.
8. [ASK THESE QUESTIONS ONLY IF THE INTERVIEWEE DID NOT ALREADY ANSWER THESE MARKET EVENT TYPE QUESTIONS AS PART OF THE MV TEMPLATE] Which of the following scenarios best describes the situation concerning the installation of your [MEASURE GROUP]?
    a. Replacement of existing equipment in existing space
    b. Adding an improvement to existing equipment, including controls or control points
    c. Retro-commissioning
    d. A new facility
    e. A major addition of space to a facility
    f. A gut rehab of existing space or tenant fit-out

g. An expansion of the plant or the addition of a new production line

h. Don't know

i. Refused

9. [IF Q8=A] Was this equipment replacement part of a major remodelling, gut rehab, or a tenant fit-out?

   a. Yes

   b. No

   c. Don't know

   d. Refused

10. Was there a recent or planned change to the use of the building system that your [MEASURE GROUP] was part of? These might include changes to system output due to changed lighting needs, greater heating or cooling demand in the same space, or expansion of a production line.

    a. Yes

    b. No

    c. Don't know

    d. Refused

11. [If Q10 = YES] You mentioned that you were making or planning changes to the building system that your [MEASURE GROUP] was part of to address new needs in the building. How well could these new needs have been met by the old system, assuming the old system was still functioning well?

    a. Fully with no adjustments

    b. Fully after some adjustments

    c. Fully only with some major adjustments

    d. Would not have fully met the new need even with adjustments

    e. Don't know

    f. Refused

12. [IF Q10=YES and Q11≠E OR F] Apart from any adjustments to meet the new system needs, how well was the old [MEASURE GROUP] system working at the time the decision was made to replace it?

    a. Working with no need of repair

    b. Working with need of minor repair

    c. Working but can't fully repair

    d. Working with need of major repair

    e. No longer working

f. Don't know

g. Refused

13. [IF 10=YES and Q12=C] You said that the equipment you replaced was working but you could not fully repair it. For approximately how long was your old equipment in a state where it could not be fully repaired?

    a. _____# of years

    b. Don't know

    c. Refused

14. [IF Q12=C AND Q13≠ B OR C] How serious a concern was it that your old equipment wasn't working? Would you say it was …

    a. A minor nuisance, mostly not noticed

    b. Noticeable but not a major disruption or inconvenience

    c. A major drag on operations or major source of complaints

    d. Other situation ____ [PLEASE SPECIFY]

    e. Don't know

    f. Refused

15. Which of the following best describes the recent repair history for the old equipment?

    a. Typical maintenance for its age, no major issues

    b. High/increasingly high frequency or cost of repairs

    c. No longer able to be repaired to full effective operation

    d. Major breakdown, unable to be repaired

    e. Could be repaired but not worth the money

    f. Other situation ____ [PLEASE SPECIFY]

    g. Don't know

    h. Refused

16. [ASK ONLY IN DELAMPING SITUATIONS] Was the previous equipment decommissioned?

    a. Yes

    b. No

    c. Refused

    d. Don't know

17. ASK ONLY IN DELAMPING SITUATIONS] Was the use of other equipment (other than the new [system] increased to compensate for the retirement of the old equipment?

a. Yes

b. No

c. Refused

d. Don't know

18. How did the idea for installing this project originate?

    a. [IF NOT ALREADY MENTIONED] Was a vendor or consultant involved in the decision?

        i. [IF YES] Was this a vendor referred to you by the program's implementation contractor TRC? Or was the vendor someone your company selected on its own?

            1. [IF PROGRAM VENDOR] Did the vendor play a role in convincing your company to do these projects?

    b. [IF NOT ALREADY MENTIONED] Was it suggested in an energy audit?

        i. [IF YES] Who conducted this audit?

19. About when was this project first put forward for your company's consideration?

20. About when was the final decision to go ahead with the project?

21. About when did your organization first begin discussions with TRC regarding funding/incentives and/or measurement and verification for this project?

22. You received financial incentives for this energy efficiency project through the Large Energy Users Program. Had your company considered implementing this project through other New Jersey energy efficiency programs?

    a. [IF YES] Which other programs?

    b. [IF YES] Why did your company decide to implement this project through the Large Energy Users Program instead of possible alternative New Jersey energy efficiency programs?

    c. [IF YES] Did the availability of multiple New Jersey energy efficiency programs for which your project was eligible make it difficult for you to decide which program to select?

Free Ridership

Now I'm going to ask you a few questions about possible factors that may have influenced your decision to go ahead with this project.

23. Before you heard about Large Energy Users Program and its financial incentives, had you already planned to purchase and install the [MEASURE X]?

    a. [IF Q23=YES OR DK] Prior to hearing about the program incentive, was the purchase of [MEASURE X] included in your organization's capital budget?

        i. [IF Q23A = YES] Had your organization ALREADY ordered or purchased the [MEASURE] BEFORE you heard about the program?

1. [IF Q17A = YES]

24. Without the incentives and information from the Large Energy Users Program, would you most likely still have purchased the same [MEASURE X]?

    a. [IF Q18 = NO] So, without the incentive and information or education from the Large Energy Users Program, you would not have installed purchased [MEASURE X] at all. Is that correct?

25. Without the Large Energy Users Program and its financial incentives and information, what efficiency level of [MEASURE X] would you most likely have purchased?

    a. Same efficiency or higher

    b. Lower efficiency

    c. Lowest efficiency/lowest cost [MEASURE X]

    d. Don't know

    e. Refused

26. Thinking about timing, without the Large Energy Users Program and its financial incentives and information, when would you have most likely purchased the [MEASURE X]?

    a. In the same year

    b. 1 - 2 years later

    c. Within 3-5 years

    d. In more than 5 years

    e. Never

    f. Don't know

    g. Refused

27. Did the incentive from the Large Energy Users Program help [MEASURE X] project receive implementation approval from your organization?

    a. Yes

    b. No

28. Without the Large Energy Users Program and its financial incentives and information, how many [MEASURE Xs] would you most likely have purchased?

29. Please rate how important the Large Energy Users Program and its financial incentives were in your decision to purchase and install the [MEASURE X]. Use a scale from 1 to 5, with 1 meaning "not at all important" and 5 meaning "extremely important."

    a. 1 - not at all important

    b. 2 - slightly important

c. 3 - moderately important

d. 4 - very important

e. 5 - extremely important

f. Don't know

g. Refused

30. Please rate how important a recommendation from a staff member of TRC, the company which administers the Large Energy Users Program, was in your decision purchase and install the [MEASURE X]. Use a scale from 1 to 5, with 1 meaning "not at all important" and 5 meaning "extremely important."

    a. Not applicable, did not receive recommendation from TRC staff member

    b. 1 - not at all important

    c. 2 - slightly important

    d. 3 - moderately important

    e. 4 - very important

    f. 5 - extremely important

    g. Don't know

    h. Refused

31. Please rate how important energy efficiency information that the Large Energy Users program may have provided you through websites or program materials was in your decision to purchase and install the [MEASURE X]. Use a scale from 1 to 5, with 1 meaning "not at all important" and 5 meaning "extremely important."

    a. Not applicable, did not receive information from the Large Energy Users program

    b. 1 - not at all important

    c. 2 - slightly important

    d. 3 - moderately important

    e. 4 - very important

    f. 5 - extremely important

    g. Don't know

    h. Refused

32. Please rate how important information from a contractor was in your decision to purchase and install the [MEASURE X]. Use a scale from 1 to 5, with 1 meaning "not at all important" and 5 meaning "extremely important."

    a. Not applicable, did not receive information from a contractor

    b. 1 - not at all important

c. 2 - slightly important

d. 3 - moderately important

e. 4 - very important

f. 5 - extremely important

g. Don't know

h. Refused

33. Please rate how important previous participation in a New Jersey energy efficiency program was in your decision to purchase and install the [MEASURE X]. Use a scale from 1 to 5, with 1 meaning "not at all important" and 5 meaning "extremely important."

    a. Not applicable, did not previously participate in a New Jersey energy efficiency program.

    b. 1 - not at all important

    c. 2 - slightly important

    d. 3 - moderately important

    e. 4 - very important

    f. 5 - extremely important

    g. Don't know

    h. Refused

34. In your own words, can you please describe how important the Large Energy User Program and its financial incentives was in your decision to purchase and install the [MEASURE X]?

    a. Record answer if provided: _____

    b. Don't know

    c. Refused

    [REPEAT QUESTIONS 17-27 IF PARTICIPANT RECEIVED MULTIPLE EQUIPMENT TYPES AND THEY SAID IN RESPONSE TO QUESTION 7 THAT THERE A SEPARATE DECISION-MAKING PROCESS FOR EACH MEASURE]

35. If you were to summarize, in your words, what were the biggest impacts of the Large Energy Users Program and its implementation contractor TRC on your installation decision, what would this be?

Spillover

36. Since participating in Large Energy Users Program, have you made any other energy efficient improvements in New Jersey that you did NOT receive rebates or financing from a New Jersey energy efficiency program?

    a. Yes

b. No [SKIP TO Q33]

c. Don't know [SKIP TO Q33]

d. Refused [SKIP TO Q33]

37. What energy efficient improvements have you made since you participated in the Large Energy Users Program?

    a. Please Describe) _____

    b. Don't know

    c. Refused

38. On a scale from 1 to 5, with 1 meaning "not at all important" and 5 meaning "extremely important", please rate how important your experience with the Large Energy Users Program and its funded measure(s) were in your decision to install the [ANY MEASURE MENTIONED IN RESPONSE TO Q29]? [REPEAT QUESTION FOR EACH MEASURE MENTIONED IN RESPONSE TO Q29].

    a. 1 - not at all important

    b. 2 - slightly important

    c. 3 - moderately important

    d. 4 - very important

    e. 5 - extremely important

    f. Don't know

    g. Refused

[IF Q29=A (YES INSTALLED UNINCENTED MEASURE) AND Q31=b to f (SOME PROGRAM INFLUENCE). REPEAT QUESTION FOR EACH Q29 MEASURE.]

39a. Can you estimate the size of the project, either in terms of its cost, annual savings, or savings relative to the measures we already have been talking about?

    a. Yes – Measure cost $_____

    b. Yes – Measure annual savings:

        i. $_____

        ii. _____ kWh

        iii. _____ MMBtu

        iv. _____ Other units such as production rate, square feet, etc. _____

    c. Yes – Relative size

        i. _____% bigger

        ii. _____% smaller

d. No but name and contact information for person who can

_____

e. No

f. Don't know

g. Refused

39b. Why did you NOT apply for and receive a rebate or financing for the [ANY MEASURES MENTIONED IN RESPONSE TO Q29]?

a. Unaware if rebate/financing is available

b. Product did not qualify

c. Other (please describe) _____

d. Don't know

e. Refused

40. In what year did you purchase and install the [MEASURE MENTIONED IN RESPONSE TO Q29]?

a. 2022

b. 2023

c. Don't know

d. Refused

41a. Was this measure installed at the same site as the incented site or another?

a. Same site

b. Another site

c. Don't know

d. Refused

Process Evaluation, Program Satisfaction

42.How did you first hear about this Large Energy Users program?

43. What motivated your company to join this program?

Now I am going to ask you about your level of satisfaction with various aspects of the Large Energy Users program.

Program Information

44.When you were considering joining the program, you likely received some information about how the program worked. This might have been from one of TRC's sales representatives or from print materials or from information on the program website. Do you recall this program information?

45. [IF Q44 = YES ELSE SKIP TO Q47] Using a five-point satisfaction scale where 5 equals “very satisfied” and 1 equals “very dissatisfied”, how satisfied were you with this information about the program?

46. [IF SATISFACTION LEVEL IN Q45 ≤ 3] Why were you less than satisfied with the program information?

Enrollment Process

47. The first stage in the Large Energy Users Program is the enrollment stage. In this stage, program applicants submit a Pre-Qualification application as well as provide the program manager with their utility usage summary. Using a five-point satisfaction scale where 5 equals “very satisfied” and 1 equals “very dissatisfied”, how satisfied were you with this enrollment process?

48. [IF SATISFACTION LEVEL IN Q47 ≤ 3] Why were you less than satisfied with this enrolment stage of the program?

Draft Energy Efficiency Plan

49. The second stage in the Large Energy Users Program is the Draft Energy Efficiency Plan (DEEP) stage. In this stage, program applicants submit a draft report with high-level project details to reserve incentives. Using that same five-point satisfaction scale [REPEAT SCALE IF NECESSARY], how satisfied were you with this Draft Energy Efficiency Plan stage of the program?

50. [IF SATISFACTION LEVEL IN Q49 ≤ 3] Why were you less than satisfied with this Draft Energy Efficiency Plan stage of the program?

Final Energy Efficiency Plan

51. The third stage in the Large Energy Users Program is the Final Energy Efficiency Plan (FEEP) stage. In this stage, program applicants submit a complete report with project details fully supported by appendices including savings calculations, specification sheets, cost quotes, M&V plans and other project documents in order to get a formal incentives commitment from the program. Using that same five-point satisfaction scale [REPEAT SCALE IF NECESSARY], how satisfied were you with this Final Energy Efficiency Plan stage of the program?

52. [IF SATISFACTION LEVEL IN Q51 ≤ 3] Why were you less than satisfied with this Final Energy Efficiency Plan stage of the program?

53. When you were developing either your draft energy efficiency plan or your final energy efficiency plan, did the staff with the Large Energy Users Program encourage you to install a wider range of energy efficiency measures than you had previously planned to?

54. [IF Q53 = YES] Please describe how the program encouraged you to adopt a wider range of energy efficiency measures?

Project Implementation

55. The next stage in program participation is the implementation of the energy efficiency project. Using that same five-point satisfaction scale [REPEAT SCALE IF NECESSARY], how satisfied were you with this project implementation stage?

56. [IF SATISFACTION LEVEL IN Q55 ≤ 3] Why were you less than satisfied with this project implementation stage?

Confirmation of Installation

57. The last stage of the program is the Confirmation of installation stage. In this stage, the program applicant submits to the Program Manager a closeout package which includes the as-built scope of work, project invoices, M&V report and other necessary for review and payment authorization. Using that same five-point satisfaction scale [REPEAT SCALE IF NECESSARY], how satisfied were you with this confirmation of installation stage?

58. [IF SATISFACTION LEVEL IN Q57 ≤ 3] Why were you less than satisfied with this confirmation of installation stage?

59. One important component of the Confirmation of Installation Stage of the program is required measurement and verification (M&V) of the installed equipment. Using that same five-point satisfaction scale [REPEAT SCALE IF NECESSARY], how satisfied were you with these M&V requirements?

60. [IF SATISFACTION LEVEL IN Q59 ≤ 3] Why were you less than satisfied with these M&V requirements?

Overall Program

61. We have covered your levels of satisfaction with various stages of the program, but now I would like to know your level of satisfaction with the Large Energy Users Program as a whole. Using that same five-point satisfaction scale, how satisfied are you with the program as a whole?

62. [IF SATISFACTION LEVEL IN Q61 ≤ 3] Why were you less than satisfied with the program as a whole?

63. This Large Energy Users program did not offer financing options. How significant a barrier was that to your participation in this program. Would you say it was a …

    a. Extreme barrier
    b. Moderate barrier
    c. Somewhat of a barrier
    d. Not a barrier
    e. Don't know
    f. Refused

Firmographics

And finally, I have a few questions about the characteristics of your business.

64. Roughly, about what percentage of your operating costs are spent on energy?

65. Approximately how many square feet of heated or cooled floor area is the building/facility where the energy efficiency project was implemented?

66. What is the main business activity at this facility?

67. Approximately how many people are currently working at the building/facility where the measure energy efficiency project was implemented, including both full and part time?

68. Does your business own, lease or manage this building/facility?

a. Own

b. Lease/Rent

c. Manage

69. How many locations in New Jersey does your company/organization have?

70. Besides the energy efficiency project you implemented through the program, have you made any changes to the building/facility where the project was implemented that might have significantly impacted energy usage?

    a. [IF YES] What changes did you make?

That's all the questions I had. Thank you so much for your time.

APPENDIX E. SITE-SPECIFIC RESULT TABLES

Table E-1. Site-specific result tables

DNV measure ID	Measure description	Reported electric savings	Electric GRR	Reported natural gas savings	Natural gas GRR	NTGR
DNV-04.1	Heat of compression dryer	-	-	215,430	99%	56%
DNV-04.2	Upgrade existing chillers	4,450,575	68%	-	-	56%
DNV-07.1	LED Lighting	136,373	55%	-	-	NI
DNV-07.2	Removable steam blankets	-	-	84,576	125%	NI
DNV-07.3	Gas compressor upgrade	1,167,561	100%	-	-	NI
DNV-08.1	Cooling tower retrofit for a condenser water utility.	2,078,739	63%	-	-	100%
DNV-10.1	Lighting retrofit & controls	1,590,686	100%	-	-	100%
DNV-12.1	Gas compressor replacement	-	-	586,660	71%	NI
DNV-12.2	Cold crude pump upgrade	(705,000)	280%	2,771,869	52%	NI
DNV-14.1	Lighting retrofit & controls	4,729,372	89%	-	-	100%
DNV-16.1	Cooling tower replacement for a process system	57,190	90%	-	-	100%
DNV-17.1	Motor replacement - Ph 1	585,950	97%	-	-	100%
DNV-17.2	Motor replacement - Ph 2	414,570	104%	-	-	100%
DNV-18.1	Lighting retrofit	1,631,967	109%	-	-	NI
DNV-18.2	Chiller VFDs	338,943	16%	-	-	NI
DNV-18.3	Resheave CRAH fans	65,718	100%	-	-	NI
DNV-19.1	Lighting retrofit & controls	1,050,033	104%	-	-	50%
DNV-26.1	Lighting retrofit & controls	(18,638)	(615%)	-	-	50%
DNV-32.1	Lighting retrofit & controls	65,737	131%	-	-	50%
DNV-35.1	Lighting retrofit & controls	141,247	118%	-	-	50%
DNV-56.1	LED lighting	150,466	118%	-	-	50%
DNV-60.1	Lighting retrofit & controls	258,316	65%	-	-	50%
DNV-80.1	Lighting retrofit & controls	602,351	45%	-	-	50%

GRR – Gross realization rate
NTGR – Net-to-gross ratio

About DNV

DNV is an independent assurance and risk management provider, operating in more than 100 countries, with the purpose of safeguarding life, property, and the environment. Whether assessing a new ship design, qualifying technology for a floating wind farm, analyzing sensor data from a gas pipeline, or certifying a food company's supply chain, DNV enables its customers and their stakeholders to manage technological and regulatory complexity with confidence. As a trusted voice for many of the world's most successful organizations, we use our broad experience and deep expertise to advance safety and sustainable performance, set industry standards, and inspire and invent solutions.