

EEPS Commercial & Multifamily Close-Out Impact Evaluation, including National Fuel Gas Distribution Corporation's Non-Residential Rebate Program

Final Report

Prepared for:

New York State Energy Research and Development Authority
Albany, NY

National Fuel Gas Distribution Corporation
Williamsville, NY

Judeen Byrne,
Project Manager, NYSERDA

Evan Crahen
Manager, Rates and Regulatory Affairs
National Fuel Gas Distribution Corporation

Prepared by:

ERS

Corporate Headquarters: North Andover, MA
Jon Maxwell, Vice President of Engineering and Evaluation
Lucy Neiman, Existing Facilities Lead
Patrick Hewlett, Multifamily Lead
Levon Whyte, Commercial New Construction Lead

With DNV GL, Tom Ledyard PM and Cx Associates, Daniel Tuhus-Dubrow PM

Record of Revision

EEPS Close-Out Impact Evaluation Final Report
EEPS Close-Out Impact Evaluation Final Report January 2020

Revision Date	Description of Changes	Revision on Page(s)
12/16/2019	Original Issue	Original Issue
1/13/2020	Requested edits from NYSERDA and responses to comments	Throughout report

Notice

This report was prepared by ERS in the course of performing work contracted for and sponsored by the New York State Energy Research and Development Authority and National Fuel Gas Distribution Corporation (hereafter the "Sponsors"). The opinions expressed in this report do not necessarily reflect those of the Sponsors or the State of New York, and reference to any specific product, service, process, or method does not constitute an implied or expressed recommendation or endorsement of it. Further, the Sponsors, the State of New York, and the contractor make no warranties or representations, expressed or implied, as to the fitness for particular purpose or merchantability of any product, apparatus, or service, or the usefulness, completeness, or accuracy of any processes, methods, or other information contained, described, disclosed, or referred to in this report. The Sponsors, the State of New York, and the contractor make no representation that the use of any product, apparatus, process, method, or other information will not infringe privately owned rights and will assume no liability for any loss, injury, or damage resulting from, or occurring in connection with, the use of information contained, described, disclosed, or referred to in this report.

NYSERDA makes every effort to provide accurate information about copyright owners and related matters in the reports we publish. Contractors are responsible for determining and satisfying copyright or other use restrictions regarding the content of the reports that they write, in compliance with NYSERDA's policies and federal law. If you are the copyright owner and believe a NYSERDA report has not properly attributed your work to you or has used it without permission, please email print@nyscrda.ny.gov.

Information contained in this document, such as web page addresses, are current at the time of publication.

Table of Contents

List of Tables	v
List of Figures	v
1 Introduction	1
1.1 Objectives.....	1
1.2 Evaluation History.....	2
2 Existing Facilities	3
2.1 Introduction	3
2.2 Program Background – Existing Facilities	3
2.3 Methodology – Existing Facilities.....	4
2.3.1 Sample Design	5
2.3.2 Data Collection and Site-Specific Analysis	6
2.3.3 Aggregation	7
2.4 NYSERDA Existing Facilities Results	7
2.4.1 EFP Electric Energy and Demand Management Program Savings	7
2.4.1.1 Discrepancies between Program-Reported and Evaluated Electric Energy Savings	9
2.5 NYSERDA Existing Facilities Conclusions and Recommendations	10
2.6 NFGDC Non-Residential Rebate Program Results	10
2.6.1 Discrepancies between Program-Reported and Evaluated Natural Gas Savings.....	11
2.7 NFGDC Conclusions and Recommendations	12
2.8 Comparison with Past Evaluation Results	13
3 Multifamily Performance Program	14
3.1 Introduction	14
3.2 Program Background – Multifamily	14
3.3 Methodology – Multifamily	15
3.3.1 Sample Design	15
3.3.2 Data Collection and Site-Specific Analysis	16
3.3.3 Savings Persistence	17
3.3.4 Aggregate Analysis	17
3.4 Multifamily Results	18
3.4.1 Discrepancy Analysis	19
3.4.2 Fuel-Specific Results	21
3.4.3 Savings Persistence	21
3.5 Multifamily Conclusions and Recommendations.....	22

3.5.1 Conclusions 22

3.5.2 Recommendations 24

4 New Construction 26

4.1 Introduction 26

4.2 Program Background – CNCP 26

4.3 Methodology – CNCP 26

4.3.1 Sample Design 26

4.3.2 Data Collection and Site-Specific Analysis 27

4.3.3 Aggregation 27

4.4 New Construction Results 27

4.4.1 Discrepancy Analysis 30

4.5 New Construction Findings and Conclusions 32

4.5.1 Findings and Conclusions 32

List of Tables

Table 1-1. Project Completion Periods Studied in Impact Evaluation	2
Table 1-2. Program Evaluation History	2
Table 2-1. EFP and NFGDC Summary Statistics of Final Sample Frame by Company, Fuel Type, and Subset	5
Table 2-2. Existing Facilities and NFGDC Final Sample Design Summary	6
Table 2-3. Existing Facilities Overall Electric Results.....	7
Table 2-4. EFP Demand Management Program Results	9
Table 2-5. EFP Overall Electric Discrepancy Breakdown.....	9
Table 2-6. NFGDC Overall Natural Gas Results	10
Table 2-7. NFGDC 2014–2015 Discrepancy Breakdown.....	11
Table 2-8. NFGDC 2016–Q2 2018 Discrepancy Breakdown	12
Table 2-9. EFP and NFGDC Prior Impact Evaluation Realization Rates.....	13
Table 3-1. MPP Final Sample Design by Market Type and Payment Stage.....	16
Table 3-2. MPP Impact Evaluation Results – All Fossil Fuels.....	18
Table 3-3. MPP Key Contributors to Fossil Fuel Realization Rate	20
Table 3-4. MPP Impact Evaluation Results by Fuel Type	21
Table 4-1. CNCP Electric Sample Design Summary	26
Table 4-2. CNCP Natural Gas Sample Design Summary	27
Table 4-3. CNCP Overall Electric Results	28
Table 4-4. CNCP Overall Natural Gas Results	28
Table 4-5. CNCP Current and Prior Impact Evaluation Realization Rates.....	28
Table 4-6. Investigation of Key Contributors to CNCP Electric Realization Rate	30
Table 4-7. Investigation of Key Contributors to CNCP Natural Gas Realization Rate	31

List of Figures

Figure 2-1. NYSERDA EFP Program-Reported and Evaluated Custom Electric Energy Savings	8
Figure 2-2. NFGDC Reported and Evaluated Natural Gas Energy Savings	11
Figure 3-1. MPP Impact Evaluation Results: Overall (Left) and Close-Up (Right).....	19
Figure 3-2. MPP Year-by-Year Savings Persistence	22
Figure 4-1. Comparison of CNCP Reported and Evaluated Electric Energy Savings: Overall (Left) and Close-Up (Right)	29
Figure 4-2. Comparison of CNCP Reported and Evaluated Natural Gas Savings	29

1 Introduction

This impact evaluation studies the gross impact of three NYSERDA Energy Efficiency Portfolio Standard (EEPS)–funded¹ legacy programs and one National Fuel Gas Distribution Corporation (NFGDC) program. The projects included in the evaluation were initiated through NYSERDA’s Energy Efficiency Portfolio Standard (EEPS-2).² NFGDC’s program from 2016 on is funded by the Energy Efficiency Transition Implementation Plan and System Energy Efficiency Plan (ETIP/SEEP).³ The four programs included in this evaluation are the following:

- NYSERDA Existing Facilities Program (EFP)
- NYSERDA Multifamily Performance Program (MPP)
- NYSERDA Commercial New Construction Program (CNCP)
- NFGDC Non-Residential Rebate Program, administered by NYSERDA and delivered with EFP during the early part of the evaluation period, run by NFGDC in the latter part of the period.

NYSERDA’s EFP has ended. CNCP and MPP are continuing under a new funding source but in substantially altered form. The NFGDC Non-Residential Rebate program is robust and continuing.⁴

ERS served as the lead and prime contractor for this effort with subcontractors DNV GL/KEMA and Cx Associates providing sample design and site review for conflict of interest and selected additional sites.

1.1 Objectives

The primary objective was to determine verified gross savings (VGS)⁵ for electric energy, electric demand, and natural gas energy and calculate corresponding realization rates for each of the four programs, with no worse than 10% relative precision at 90% confidence.

¹ In May 2007, the New York Public Service Commission (PSC) issued an order instituting a proceeding to develop an Energy Efficiency Portfolio Standard (EEPS1). In October 2011, the PSC issued a further order reauthorizing EEPS programs through December 31, 2015. In December 2015, the PSC issued an order extending the Energy Efficiency Portfolio Standard (EEPS2) through Feb 29, 2016. Collectively, EEPS1 and EEPS2 activities and funding sources are referred to as EEPS.

² Department of Public Service, Filing #4779 Case No. 07-M-0458, 2016

³ <https://www.nationalfuelforthought.com/wp-content/uploads/2017/07/December-1-2016-Energy-Efficiency-Transition-Implementation-PlanETIP-for-the-2016-2018-Period.pdf>

⁴ The program structure remained relatively consistent from 2014 through 2018Q2. There are two major areas of change: In 2014 and 2015, the program was restricted to customers using less than 12,000 Mcf per year (i.e., small commercial customers only). From 1/1/2016 through 6/30/2018, the program was “opened up” to all commercial and industrial customers, regardless of their size. Few new measures were added to the program over the multi-year period.

⁵ Called “evaluated gross savings” in prior NYSERDA evaluation reports.

With one of the programs no longer being offered and two substantially different from the version offered during the evaluation period, recommending program improvements was a limited secondary objective. ERS has included recommendations for NFGDC’s continuing program and where there are recommendations that apply generally to NYSERDA’s other programs.

The evaluated period of performance covers projects completed in 2014 through 2018Q2 for EFP natural gas (both NYSERDA and NFGDC) and MPP, and 2016 through 2018Q2 for EFP electricity (NYSERDA only) and CNCP, as illustrated in Table 1-1.

Table 1-1. Project Completion Periods Studied in Impact Evaluation

Program	Evaluation Period				
	2014	2015	2016	2017	2018
Existing Facilities – Electricity					
Existing Facilities and NFGDC NRCIP – Natural Gas					
Multifamily Performance Program					
Commercial New Construction Program					

1.2 Evaluation History

The programs have a substantial evaluation history. Table 1-2 summarizes research completed since 2012.⁶ Historic impact results are presented for context with this evaluation’s results.

Table 1-2. Program Evaluation History

Program	Evaluation Final Report Year of Completion				
	Impact, Most Recent (years evaluated)	Impact, Earlier (years evaluated)	Process	Market Characterization	Nonparticipant Spillover & Market Effects
EFP	2015* (2012-Sept.2013)	2012 (2006-2009)	2012	2012	2013
MPP	2015 (2009-2011)		2013	2013	
CNCP	2016 (2009-2013)	2012 (2007-2008)	2013	2013	

*NYSERDA and NFGDC’s NRCIP

⁶ All reports are available at <https://www.nyscrda.ny.gov/About/Publications/Program-Planning-Status-and-Evaluation-Reports/Evaluation-Contractor-Reports>.

2 Existing Facilities

2.1 Introduction

This impact evaluation studies the gross impact of NYSERDA Energy Efficiency Portfolio Standard (EEPS)–funded⁷ legacy Existing Facilities Program (EFP) and the National Fuel Gas Distribution Corporation (NFGDC) Non-Residential Rebate Program, which was administered by NYSERDA and delivered with EFP and now is overseen by NFGDC. The projects included in the evaluation were initiated through NYSERDA’s EEPS-2⁸. The evaluated period of performance covers natural gas projects completed in 2014 through 2018Q2 and electric projects completed in 2016 through 2018Q2. NYSERDA’s EFP has since ended and no longer exists. NFGDC’s Non-Residential Rebate Program is funded via ETIP/SEEP,⁹ remains operational under NFGDC’s administration, and has implemented programmatic enhancements during the evaluation period.¹⁰

2.2 Program Background – Existing Facilities

The EFP encouraged the adoption of electric and natural gas energy efficiency measures¹¹ across a range of sectors, including commercial and industrial businesses, health care facilities, universities and colleges, state and local governments, and mission-critical facilities such as data centers and communications facilities. Though EFP offered both pre-qualified and performance-based incentives, this study only examined custom measures that were incentivized through the performance-based track. During the timeframe covered by this study, some sites that installed electric measures receiving EEPS-2 incentives (through EFP for their kWh savings) also received incentives for peak demand reductions through the Demand Management Program¹² (DMP), a joint effort

⁷ In May 2007, the New York Public Service Commission (PSC) issued an order instituting a proceeding to develop an Energy Efficiency Portfolio Standard (EEPS1). In October 2011, the PSC issued a further order reauthorizing EEPS programs through December 31, 2015. In December 2015, the PSC issued an order extending the Energy Efficiency Portfolio Standard (EEPS2) through Feb 29, 2016. Collectively, EEPS1 and EEPS2 activities and funding sources are referred to as EEPS.

⁸ Department of Public Service, Filing #4779 Case No. 07-M-0458, 2016

⁹ <https://www.nationalfuelthought.com/wp-content/uploads/2017/07/December-1-2016-Energy-Efficiency-Transition-Implementation-PlanETIP-for-the-2016-2018-Period.pdf>

¹⁰ The NFGDC program structure remained relatively consistent from 2014 through 2018Q2. There are two major areas of change: in 2014 and 2015, the program was restricted to customers using less than 12,000 Mcf per year (i.e., small commercial customers only). From 1/1/2016 through 6/30/2018, the program was “opened up” to all commercial and industrial customers, regardless of their size. Also a few new measures were added to the program over the multi-year period.

¹¹ EFP’s promotion of demand management ended in 2011, when demand management measures were transitioned to the Technology and Market Development Program (T&MD) portfolio of programs.

¹² MW savings will only be assessed for measures supported through the Demand Management Program for sites that fall within EFP’s EEPS2 sample.

between NYSEDA and Con Edison.¹³ The evaluation team assessed peak demand savings for sites that fell naturally within the EEPS-based sample and received incentives through both EFP and DMP.

NFGDC's Non-Residential Rebate Program is an equipment replacement program that offers rebates for both prequalified and custom projects. NFGDC's custom incentives are based on estimated gas energy savings, which are further confirmed through a technical engineering analysis. This program was administered by NYSEDA's EFP staff for program years 2014 and 2015, and by NFGDC's own implementer for the 2016–2018 program years. Both time periods included both pre-qualified and custom projects.

The evaluation of these programs was a joint effort by the administrators, with NYSEDA and NFGDC staff sharing responsibilities for tasks as appropriate, including but not limited to decision making, document review, weekly call participation, and meeting attendance.

2.3 Methodology – Existing Facilities

A critical component of the evaluation was the development of estimates of the project RRs for program-reported natural gas savings, and electric custom savings for EFP, which includes verifying the installation of efficiency measures and the generation of an independent savings analysis. The evaluation team conducted desk reviews of project documentation for all sites in a representative sample of participating custom electric and natural gas projects; the team also followed up with site personnel through phone conversations or emails to verify equipment installation and operational parameters. Per NYSEDA's instruction, no site visits or metering were conducted for EFP or NFGDC sites. Billing analysis was conducted for those gas sites where the evaluation team was able to obtain billing data. The project level savings were then weighted to calculate the program-level RR.

The program-level RR is applied to the program-reported savings, resulting in the evaluated gross savings estimates. RRs represent an adjustment to the program-reported savings, upward or downward, to account for differences between the evaluated gross savings and program-reported savings.

$$RR = \frac{Savings_{Evaluated}}{Savings_{Reported}}$$

where,

RR = Realization Rate

$Savings_{Evaluated}$ = Savings as per measurement and verification (M&V) evaluation

$Savings_{Reported}$ = Savings as reported by the Program

¹³ See the Indian Point Energy Center Energy Efficiency, Demand Reduction, and Combined Heat and Power Implementation Plan for details.

The team evaluated all projects in the sample at a verification level of rigor with desk reviews and interviews with participants. Billing analysis was attempted for all NFGDC natural gas projects where billing data was available and was used as the most appropriate means of analysis for a small percentage of the projects.

The following subsections describe the EFP/NFGDC evaluation sample design and analysis techniques.

2.3.1 Sample Design

The goal of the EFP/NFGDC sample design is to statistically select a sample of EEPS and EEPS-2 only projects (electric and gas) for EFP and Non-Residential Rebate Program gas measures for NFGDC. The electric projects in the sample frame were installed from January 1, 2016, through June 30, 2018, and the gas projects from January 1, 2014, through June 30, 2018. The original design targeted $\pm 10\%$ precision at the 90% confidence interval around first-year energy savings at the program-level design. Changes in the participant population due to project movement into other programs caused an adjustment to the expected precision from $\pm 10\%$ to $\pm 12\%$. The design is broken out by electric and gas projects as an upper stratification level. Other subsets of interest include company (NYSERDA and NFGDC), upstate and downstate regions for NYSEDA activity, and program period for NFGDC.

The sample unit was defined as the smallest combination of project and site. Projects that contained both electric and gas measures were broken out into a gas project and an electric project. Due to concerns about representativeness, the bottom 3% of program savings were removed from the population as part of developing the final sample frame. More information on the conditions used to establish the final sample frame are provided in Appendix A. Table 2-1 summarizes the sample frame by fuel type, company, and subset of interest. In total, there are 720 projects totaling 836,579 MMBtu savings in the final sample frame.

Table 2-1. EFP and NFGDC Summary Statistics of Final Sample Frame by Company, Fuel Type, and Subset

Fuel Type	Company	Subset	Project Count	MMBtu Savings	Mean Savings	Minimum Savings	Maximum Savings
Electric	NYSERDA	Upstate	87	71,237	819	78	7,597
Electric	NYSERDA	Downstate	390	373,206	957	73	16,144
Gas	NYSERDA	Upstate	59	216,158	36,664	746	27,296
Gas	NYSERDA	Downstate	13	48,023	3,694	854	15,903
Gas	NFGDC	2014 to 2015	100	77,057	771	34	8,101
Gas	NFGDC	2016 to Q2 2018	71	50,897	717	35	6,565
Overall Program			720	836,579	1,162	34	27,296

The sample design uses a stratified ratio framework to optimize the selection of sample points to target the desired precision around first-year MMBtu savings at the program level. The final sample design used a stratified ratio framework and an assumed error ratio of 0.5 for the electric sample and 0.90 for gas. Table 2-2 shows the final

sample design achieved, which included 81 sites. The electric projects have a total of 25 sample points selected and the gas projects have a total of 56 projects selected. Overall, the program is expected to obtain a relative precision of $\pm 12\%$ with the gas accounting for about 15.5% precision and the electric 18.3% precision. The fully stratified final sample design is provided in Appendix A.

Table 2-2. Existing Facilities and NFGDC Final Sample Design Summary

Fuel Type	Company	Subset	Project Count	Total MMBtu Savings	Error Ratio	Sample Size	Expected Relative Precision
Electric	NYSERDA	Upstate	87	71,237	0.50	8	29.42%
Electric	NYSERDA	Downstate	390	373,206	0.50	17	21.02%
Electric	NYSERDA		477	444,443	0.50	25	18.27%
Electric			480	444,443	0.50	25	18.27%
Gas	NYSERDA	Upstate	59	216,158	0.90	21	21.62%
Gas	NYSERDA	Downstate	13	48,023	0.90	5	48.43%
Gas	NYSERDA		72	264,182	0.90	26	19.76%
Gas	NFGDC	2014 to 2015	100	77,057	0.90	15	35.05%
Gas	NFGDC	2016 to Q2 2018	71	50,897	0.90	15	29.11%
Gas	NFGDC		171	127,955	0.90	30	24.08%
Gas			243	392,136	0.90	56	15.46%
Overall Program			720	836,579	0.69	81	12.11%

Since many projects had multiple measures installed, a final step of the sample design was to limit the number of measures being asked about in the phone survey. This was done to minimize survey fatigue to ensure that the evaluation team systematically received information on all measures of interest. Survey responders were not selected in the final sample on more than four projects, and their total number of measures represented did not exceed 20.

2.3.2 Data Collection and Site-Specific Analysis

The evaluators requested tracking data for the participant population and project documentation for all sites sampled for electric and gas savings evaluation from NYSEDA and NFGDC. After the engineer reviewed project file information to find customer contact information, verify tracking savings, and ensure that the file was sufficiently complete for evaluation, he or she attempted to conduct a telephone interview to collect critical site-specific data that impacts the analysis approach and results.

For EFP, the analysts used three site-specific methods. In order of priority they were:

1. Custom spreadsheet review
2. Deemed savings analysis
3. Billing analysis when data was available

All options made use of telephone interview data and included baseline assessment. The applicant analysis files, custom or deemed, were updated by the evaluators with revised inputs per project file review or phone interview responses where necessary, to estimate the evaluated savings. A deemed savings approach was used for a few projects where no project documentation was available. For the projects where data was available, billing analysis was employed.

The data collection and site-specific analysis procedures used to calculate the verified gross savings and RRs were generally common to all three program evaluations. Appendix B describes them in detail.

2.3.3 Aggregation

After all project-level results were calculated, the evaluators calculated program-level evaluation results for the 81 analyzed projects through statistical expansion analysis. RRs were calculated for all sampled projects as the ratio of project-level evaluated savings to reported savings. Aggregate methods are described in Appendix B. The aggregate analysis also included investigation of site-specific discrepancy analyses, as presented in Section 2.4.1.1.

2.4 NYSERDA Existing Facilities Results

This section presents the results and findings from the NYSERDA EFP evaluation. This includes electric energy, electric demand, and NYSERDA natural gas savings.

2.4.1 EFP Electric Energy and Demand Management Program Savings

The total RR of 0.98 for the electric custom projects installed in period of January 1, 2016, through June 30, 2018, is calculated by dividing the evaluated gross savings by the program-reported savings. Table 2-3 provides a comparison of the upstate, downstate, and total program-reported and evaluated electric energy savings for custom projects. There was no clear pattern to the RR with respect to geographic region.

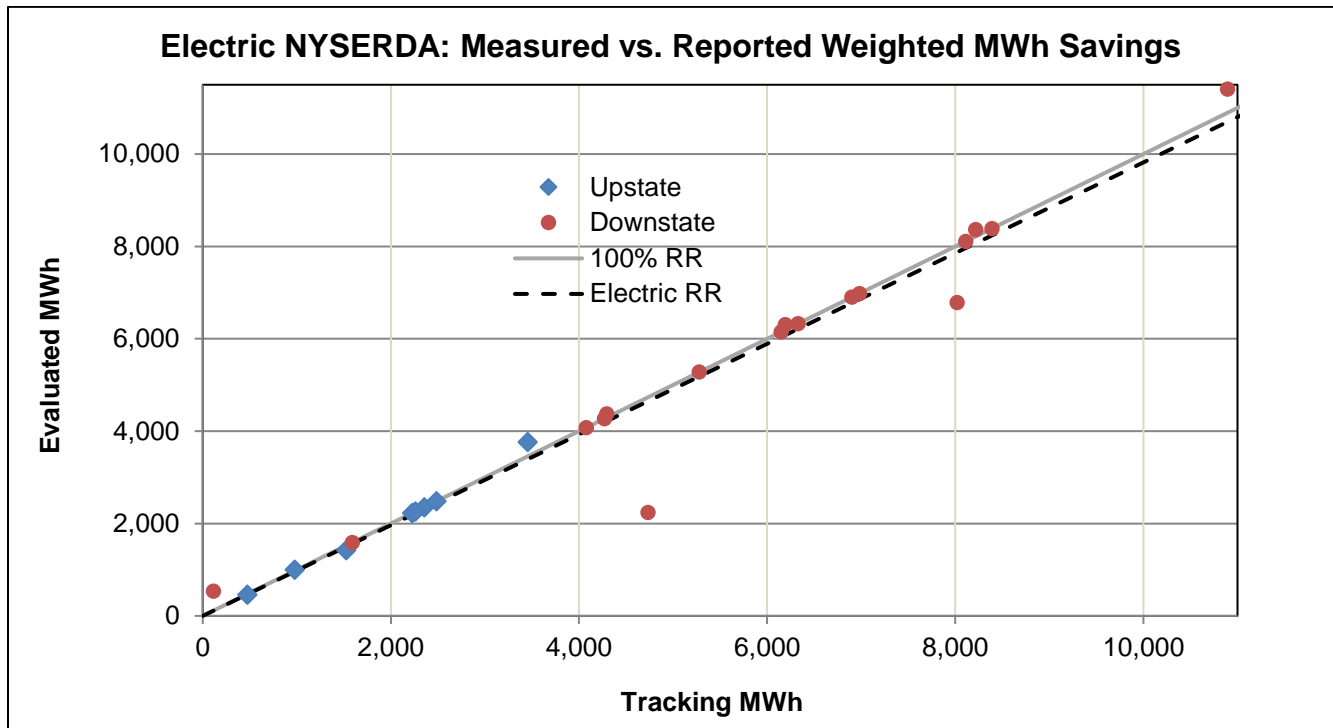
Table 2-3. Existing Facilities Overall Electric Results

Subset	Achieved Sample	Total Reported Savings (MWh)	Total Evaluated Savings (MWh)	Evaluated Gross RR	Relative Precision
Upstate	8	20,878	21,173	1.01	3.0%
Downstate	17	109,380	106,724	0.98	2.4%
Total	25	130,259	127,897	0.98	2.0%

In Figure 2-1, below, the ideal result for which the evaluated gross savings would always match the program-reported savings, i.e., a RR of 1, is shown as a solid gray line. Actual findings are plotted as points on the graph.

A pattern of points below the ideal line illustrates an RR of less than 1; points above the line illustrate an RR greater than 1.

Figure 2-1. NYSERDA EFP Program-Reported and Evaluated Custom Electric Energy Savings



The error ratio is the key measure of the variability in the relationship between the evaluated estimate of gross savings and the program-reported savings. The larger the error ratio, the larger the sample size required to meet the targeted level of confidence and precision. Conversely, the smaller the error ratio, the smaller the sample size required to meet the targeted level of confidence and precision. The electric energy error ratio is 0.16. This is a low value, and in particular is lower than assumed in the sample design, which led to evaluation results with better than planned precision level of confidence and precision.

Table 2-4 provides the RR for the projects that also participated in NYSERDA’s Demand Management Program (DMP). There was no DMP stratification built into the sample design; however, the team evaluated these demand savings when they occurred in projects selected in the sample. As these results are not statistically representative of the DMP population, no relative precision or error ratio was calculated.

Table 2-4. EFP Demand Management Program Results

Subset	Achieved Sample	Total Reported Savings (MW)	Total Evaluated Savings (MW)	Evaluated Gross RR	Relative Precision
DMP	12	3.20	3.26	1.02	N/A
MW savings were only assessed for measures supported through the Demand Management Program for sites that fell within EFP's EEPS2 sample. Although a RR was calculated; it is not representative of the DMP population.					

2.4.1.1 Discrepancies between Program-Reported and Evaluated Electric Energy Savings

For each project with an evaluated RR other than 1, a discrepancy analysis was performed to identify the major driver or drivers to the evaluated RR. The results from the discrepancy analysis are divided into categories and aggregated to identify systematic differences.

The results of the discrepancy analysis are presented in Table 2-5. This table shows the positive and negative savings impact. Differences that increased RR are shown as the green bars, and differences that reduced RR are shown as the red bars.

Table 2-5. EFP Overall Electric Discrepancy Breakdown

Discrepancy Category	Negative Impact		Positive Impact		Overall Impact	
	Frequency	RR Impact	RR Impact	Frequency	Frequency	RR Impact
Tracking/Clerical	2	-2.2%	0.3%	1	3	-1.9%
Algorithm/Adherence to TRM	0	0.0%	0.3%	5	5	0.3%
Quantity	0	0.0%	0.4%	1	1	0.4%
Wattage/Size	1	0.0%	0.0%	0	1	0.0%
Hours/Load	0	0.0%	0.2%	2	2	0.2%
HVAC Interactivity	0	0.0%	0.2%	2	2	0.2%
Other	1	-1.1%	0.0%	0	1	-1.1%
Total	4	-3.3%	1.5%	11	15	-1.8%

Table 2-5, above, demonstrates key contributors to the 0.98 RR for electric energy savings. The tracking/clerical category had the largest effect on the realization rate. This category generally includes values that were estimated in the project reporting correctly, but the value entered into the tracking database varied. This occurred in two separate instances during the evaluation.

These differences generally represent stand-alone events. The analysis found no systematic differences when it was compared with the evaluation.

2.5 NYSERDA Existing Facilities Conclusions and Recommendations

Overall, the RRs in the EFP evaluation were very close to 1. The impact evaluation team found that the program estimated savings well for both electric and natural gas saving measures. The evaluators did, however, encounter challenges with the tracking data and project documentation in several instances. Some projects had missing files or the files contained information that did not match the tracking database. Additionally, the tracking data itself did not contain all the information needed to evaluate the programs effectively. As this program is not continuing, these issues do not justify a recommendation but the need for accurate tracking data correlated with project documentation can be relevant to other ongoing NYSERDA programs.

2.6 NFGDC Non-Residential Rebate Program Results

The overall RR for NFGDC’s Non-Residential Rebate Program is 0.86; the calculation is the evaluated gross savings divided by the program-reported savings. As shown below, the RR is significantly different between the 2014–2016 and 2016–2018 time periods. The evaluators have provided these results separately as they represent a shift in program administration and program funding. The Program should use the 2016–2018 result as the prospective RR moving forward, as it is expected to be more representative of the current Program; however, evaluations of the program should continue in the future.

Table 2-6 provides a comparison of the NFGDC’s program-reported and evaluated natural gas energy savings.

Table 2-6. NFGDC Overall Natural Gas Results

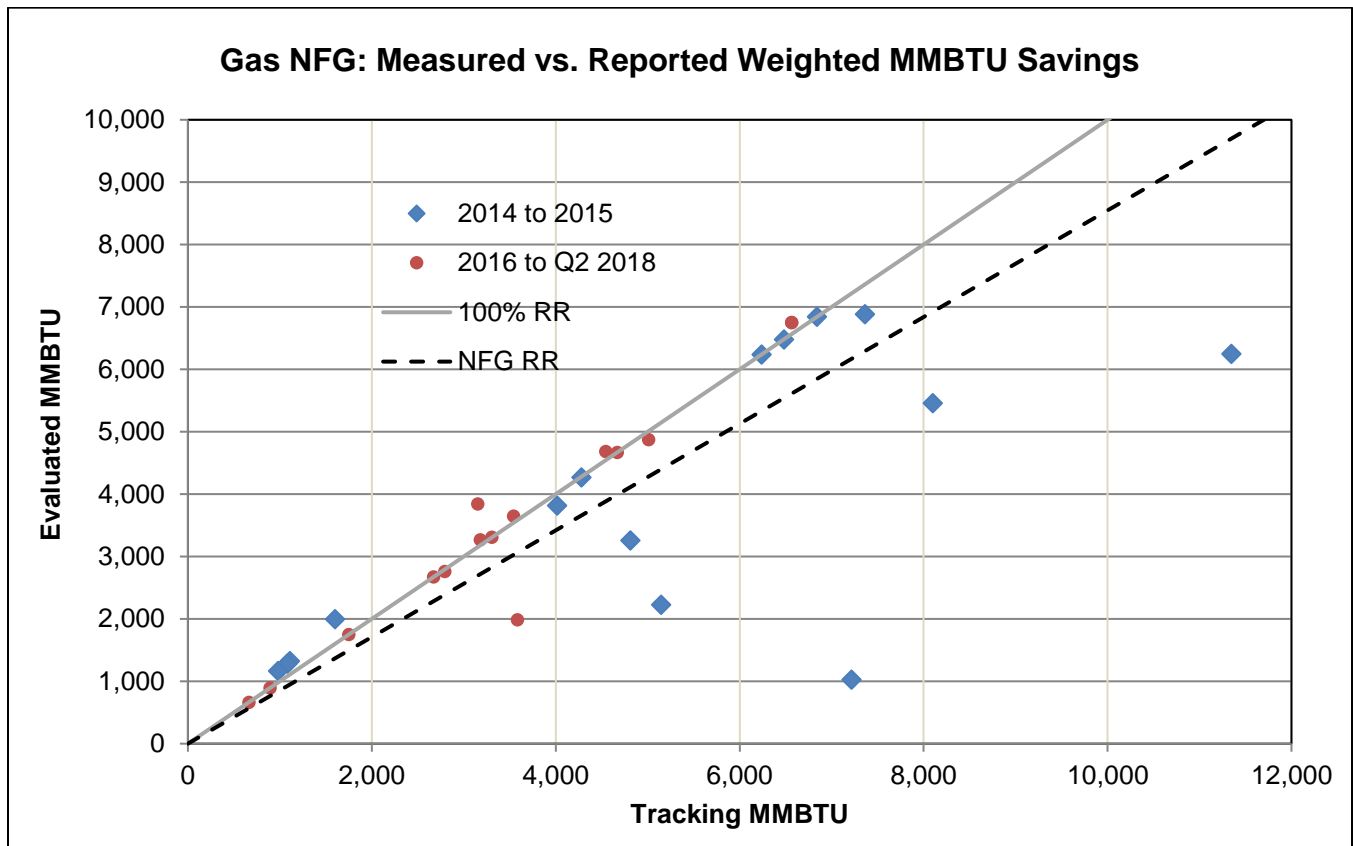
Subset	Achieved Sample	Total Reported Savings (MMBtu)	Total Evaluated Savings (MMBtu)	Evaluated Gross RR	Relative Precision
2014 to 2015 ¹	15	77,057	58,837	0.76	9.5%
2016 to Q2 2018 ²	15	50,897	50,530	0.99	4.9%
Total	30	127,955	109,367	0.86	5.6%

¹ EEPS funded and administered by NYSERDA through EFP.

² 2016 – Q2 2018 results for NFGDC are reported separately because this portion of the program was ETIP/SEEP funded and administered by NFGDC.

Figure 2-2 illustrates the evaluated annual natural gas energy savings compared with those reported by NFGDC. Ideally, the evaluated gross savings would always match the program-reported savings. This ideal is shown as a solid gray line on the chart. The actual findings are plotted as points on the graphs.

Figure 2-2. NFGDC Reported and Evaluated Natural Gas Energy Savings



2.6.1 Discrepancies between Program-Reported and Evaluated Natural Gas Savings

For each project with an evaluated RR other than 1, a discrepancy analysis was performed to identify the major driver or drivers to the evaluated RR. The results of this analysis are divided into categories and aggregated to identify systematic differences to better inform Program staff.

The results of the discrepancy analysis are presented in Tables 2-7 and 2-8. These tables show the positive and negative savings impact for the various discrepancies. The discrepancies have been separated into two tables because of the different performance between the two time periods.

Table 2-7. NFGDC 2014–2015 Discrepancy Breakdown

Discrepancy Category	Negative Impact		Positive Impact		Overall Impact	
	Frequency	RR Impact	RR Impact	Frequency	Frequency	RR Impact
Tracking/Clerical	3	-8.4%	0.5%	1	4	-7.8%
Algorithm/Adherence to TRM	4	-8.1%	0.8%	3	7	-7.3%
Efficiency	2	-4.7%	0.0%	0	2	-4.7%
Hours/Load	1	-3.8%	0.0%	0	1	-3.8%
Total	10	-24.9%	1.3%	4	14	-23.6%

Table 2-8. NFGDC 2016–Q2 2018 Discrepancy Breakdown

Discrepancy Category	Negative Impact		Positive Impact		Overall Impact	
	Frequency	RR Impact	RR Impact	Frequency	Frequency	RR Impact
Tracking/Clerical	2	-0.3%	1.3%	6	8	1.0%
Algorithm/Adherence to TRM	0	0.0%	1.3%	1	1	1.3%
Efficiency	0	0.0%	0.0%	0	0	0.0%
Hours/Load	1	-3.0%	0.0%	0	1	-3.0%
Total	3	-3.3%	2.6%	7	10	-0.7%

The tables above show that most of the discrepancies occurred in the 2014–2015 time period. Some of the major drivers to RR are discussed below:

- The tracking/clerical category had the largest impact on the RR during the 2014–2015 time period. This includes values incorrectly reported in the tracking database. One project in particular had a large savings value reported that was not supported by the project documents or the customer interview, and therefore received a low RR.
- The second-largest category was algorithm/adherence to the TRM. This includes differences in the algorithms used to calculate savings. In some instances, the TRM algorithms were applied incorrectly, and in other instances the values used in the algorithms needed to be modified. The baseline efficiency was the most common variable updated in this category.
- The largest category in the 2016 – Q2 2018 time period was the hours/load category. The project that was the primary contributor to this discrepancy received a billing analysis that revealed that the equipment was operating at lower load conditions than anticipated in the applicant’s analysis.

2.7 NFGDC Conclusions and Recommendations

The impact evaluation team offers NFGDC two recommendations based on the impact evaluation research.

1. The program is estimating natural gas savings well.

The RR’s for projects in the latter time period of this evaluation were very close to 1. The program has instituted practices that are effective in estimating natural gas savings. These include post-installation inspections as well as some project M&V performed by their implementation contractor.

- **Recommendation:** Continue the existing practice of performing M&V for custom projects. This process seems to be accurately estimating natural gas savings. Additionally, we recommend doing site-level M&V during the next evaluation cycle if practical from a budget perspective, to continue to review savings estimates being calculated by the program.

2. Project documentation is critical for evaluation and verification of savings estimates.

There were some missing project documents from a time period in the evaluation that was administered by a different implementation contractor than the Program is currently using. This lack of documentation caused some minor logistical challenges with this evaluation. This is one of the reasons why the former implementation contractor was relieved from their duties associated with NRCIP. These issues have been resolved with the newer implementation contractor.

- **Recommendation:** Ensure that the program continues the existing practice of collecting project documents at a regular interval from the implementation contractor to ensure consistent and complete tracking of projects and the associated documentation.

2.8 Comparison with Past Evaluation Results

Compared with past evaluation results shown in Table 2-9, this evaluation shows similar results for the electric program and improvements over previous program performance.

Table 2-9. EFP and NFGDC Prior Impact Evaluation Realization Rates

Program	Electric Energy	Electric Demand	Natural Gas
EFP – NYSERDA – 2015	1.01	0.98	0.64
EFP – NYSERDA – 2012	1.03	0.81	N/A
EFP – NFGDC – 2015	N/A	N/A	0.66

3 Multifamily Performance Program

3.1 Introduction

This impact evaluation studies the gross impact of the NYSERDA Energy Efficiency Portfolio Standard (EEPS)–funded¹⁴ Multifamily Performance Program (MPP). The evaluated period of performance covers projects completed in 2014 through Q2 2018. The program has been substantially redesigned since 2018.

3.2 Program Background – Multifamily

The MPP was developed in 2006 by consolidating the multifamily components of various NYSERDA programs to better serve the market by providing building owners and developers with a single portal of energy efficiency solutions. The MPP offers staged incentives to reduce overall energy use in multifamily buildings by a minimum of 15%. During the EEPS funding period, the Program promoted energy efficiency to owners of existing multifamily buildings as well as developers of new construction for both market rate and affordable housing. The MPP offered incentives for the installation of energy efficiency measures that comprehensively reduce energy use among a variety of multifamily systems. During the evaluation time frame of projects completed between January 2014 and June 2018, the MPP awarded more than \$70 million in incentives among 387 completed projects, reporting 73,184 MWh and 1,075,669 MMBtu of electric and fossil fuel savings, respectively.

To achieve aggressive energy reduction targets during both EEPS-funded and current iterations of the program, the MPP funds a comprehensive analysis of a suite of energy-saving measures at participating facilities. The cornerstone of this process is the energy reduction plan (ERP) analysis, which is typically completed by a program-approved performance partner, who recommends a package of efficiency improvements for the facility with associated incentive offerings from the Program. The ERP must demonstrate, through building modeling and simulation, that the suite of proposed measures will lead to a 15% or greater reduction in source energy use at the participating multifamily building. Once the MPP approves the ERP, an initial incentive is paid (payment #1), and the project is implemented with oversight by the Program, partner, and associated contractors. The MPP awards two additional incentives (payment #2 and #3) at the 50% and 100% installation milestones, respectively. Participating facilities are eligible for a bonus incentive payment (the “performance payment” or payment #4) if the target energy reduction is evident when comparing pre- and post-project utility consumption data.

¹⁴ In May 2007, the New York Public Service Commission (PSC) issued an order instituting a proceeding to develop an Energy Efficiency Portfolio Standard (EEPS1). In October 2011, the PSC issued a further order reauthorizing EEPS programs through December 31, 2015. In December 2015, the PSC issued an order (Department of Public Service, Filing #4779 Case No. 07-M-0458, 2016) extending the Energy Efficiency Portfolio Standard (EEPS2) through Feb 29, 2016. Collectively, EEPS1 and EEPS2 activities and funding sources are referred to as EEPS.

As the MPP has changed with NYSERDA's initiation of the Clean Energy Fund in 2016,¹⁵ the program now focuses exclusively on energy efficiency in the affordable housing sector at existing multifamily buildings only.. Program savings targets have become more aggressive, requiring ERPs to demonstrate 20% or greater energy reduction to qualify.

3.3 Methodology – Multifamily

The following subsections describe the MPP evaluation sample design and analysis techniques.

3.3.1 Sample Design

The goal of the MPP evaluation sample design was to select a statistically representative sample of projects designated with a payment date between January 1, 2014, through June 30, 2018. Upon receiving project- and measure-level tracking data through request from NYSERDA, the evaluators designed the sample with upper-level variables that are anticipated to potentially demonstrate different results: market type (affordable and market rate) and stage completed¹⁶ (100% complete and performance payment). This stratification provides a more efficient sample design by ensuring representation of each group, mitigating bias in the overall results and allowing more granular analysis. The sample was designed to provide first-year fossil fuel MMBtu savings results at the program level within $\pm 10\%$ precision at the 90% confidence interval.¹⁷

The sample unit was defined as the smallest combination of project and site. Other conditions, that are further detailed in Appendix A, were applied to the population data to establish the final sample frame. One of these conditions was to exclude the smallest projects in the population that cumulatively contribute less than 3% of total claimed savings, reducing the project count to 256 in the sample frame. Evaluators next applied a *stratified ratio estimation* technique with an assumed error ratio of 1.0 for all strata.¹⁸ Table 3-1 shows the final sample design. The total of 111 sampled projects was expected to provide a program-level fossil fuel savings result that is better than $\pm 10\%$ precision at the 90% confidence interval. Note that this design was also expected to produce precisions of at or better than $\pm 15\%$ for market-level results and better than $\pm 18\%$ for results by payment stage.

¹⁵ New York Department of Public Service, Case No. 14-M-0094, <http://documents.dps.ny.gov/public/Common/ViewDoc.aspx?DocRefId={B23BE6D8-412E-4C82-BC58-9888D496D216}>

¹⁶ The previous evaluation of the MPP showed statistically different fossil fuel realization rates for projects that received a performance payment as compared with those that did not.

¹⁷ While the evaluation was originally intended to focus on natural gas impacts only, evaluators observed inaccuracies in fuel-specific reporting in the program's tracking database. To ensure that no projects were unintentionally excluded from the sample frame, evaluators expanded the sample design to assess fossil fuel impacts overall at the desired confidence and precision.

¹⁸ The prior MPP evaluation's results featured a fossil fuel error ratio of 1.09.

Table 3-1. MPP Final Sample Design by Market Type and Payment Stage

Market Type	Payment Stage	Projects in Sample Frame	Savings (MMBtu)	Error Ratio	Sample Size	Expected Relative Precision
Affordable	100% Complete	156	532,313	1.0	55	±14.3%
Affordable	Performance Payment	34	40,913	1.0	20	±17.6%
Market Rate	100% Complete	56	288,733	1.0	30	±16.0%
Market Rate	Performance Payment	10	59,405	1.0	6	±40.8%
Overall		256	1,021,364	1.0	111	±9.4%

Stratified sampling offers a way to optimize the sample to produce higher levels of precision than might otherwise be attained through simple random sampling. Overall, this sample design includes 111 of the 256 MPP participants in the sample frame (43%) yet accounts for nearly 67% of the total program reported savings.

3.3.2 Data Collection and Site-Specific Analysis

After the sample was designed, evaluators next requested all available, program-archived files and utility consumption data for the 111 sampled projects, in addition to a number of replacement projects. The MPP maintains a commendable, comprehensive archive of project files, including the 12+ months of pre-project data required for submission during the ERP phase; therefore, sufficient pre-project data was available for all sampled projects. Evaluation engineers successfully collected sufficient post-project data for an estimated 90% of sampled projects. After the engineer reviewed project file information to find customer contact information, verify tracking savings, and ensure that the files were sufficiently complete for evaluation, he or she conducted a telephone interview with a facility representative to collect critical site-specific data which impacts the analysis approach and results.

For MPP, the analysts used two site-specific methods as the primary savings estimation approach. In order of priority they were:

1. Billing analysis – Successful for 85% of sampled projects.
2. Engineering desk reviews – Completed for remaining 15% of projects.

All options made use of telephone interview data and included baseline assessment. The data collection and site-specific analysis procedures were generally common to all three program evaluations. Appendix B describes them in detail.

For MPP, the evaluators gave special attention to central plant equipment baselines. Considering the MPP design and the NY TRM Appendix N,¹⁹ the evaluators interviewed contacts and examined documentation to determine the age, operating condition, and maintenance/repair history of replaced or retrofitted equipment. Preexisting conditions were confirmed as baseline when pre-project documentation and/or interviews with facility staff indicated that the existing equipment were expected to operate for at least another year within the facility's typical maintenance practices. If evaluators could not collect sufficient evidence to definitively establish a code baseline, an early replacement baseline was assumed by default. While the baseline typically was the pre-existing condition, for 8 of 106 evaluated projects, evaluators determined that pre-project equipment had failed or was otherwise unsuitable to meet the facility's heating or hot water load. In such cases, the baseline reflected the New York State Energy Conservation Code (NYSECC) at the time of project application.

3.3.3 Savings Persistence

The MPP staff expressed interest in examining the year-by-year persistence of savings from evaluated projects. Evaluators designed the site-specific billing analysis model to delineate individual years of post-project utility consumption data to assess year-by-year shares of savings as compared with pre-project consumption. For 49 projects in the evaluation sample, evaluation engineers successfully collected at least two years of post-project utility consumption data required to analyze second-year persistence. Analyzing persistence beyond the second year was challenging due to most utilities' archive limits; nonetheless, evaluators also analyzed third-year persistence for 10 sampled projects. Evaluators paired this study's persistence results with prior, MPP-sponsored analysis that included 20 projects.

3.3.4 Aggregate Analysis

Evaluators aggregated results from 106 analyzed projects²⁰ to support program-level analysis of realization rates and discrepancies. Statistical expansion analysis was developed to quantify the program-level RR among the upper-level sampling strata discussed in Section 3.3.1. Post-hoc analysis of realization rates was conducted among other segments of interest, including: region (upstate/downstate), completion year, performance partner, and measure. Aggregate results are examined in detail in Appendix C. The aggregate analysis also included investigation of contributors to the program-level RR using site-specific discrepancy analyses, as presented in Section 3.4.1

¹⁹ New York Technical Reference Manual Version 6, page 612.

[http://www3.dps.ny.gov/W/PSCWeb.nsf/96f0fec0b45a3c6485257688006a701a/72c23defff52920a85257f1100671bdd/\\$FILE/TRM%20Version%206%20-%20January%202019.pdf](http://www3.dps.ny.gov/W/PSCWeb.nsf/96f0fec0b45a3c6485257688006a701a/72c23defff52920a85257f1100671bdd/$FILE/TRM%20Version%206%20-%20January%202019.pdf)

²⁰ Among the original sample of 111 projects, evaluators could not assess five unreplaceable projects in census or medium-saver strata. For these projects, facility staff were either unaware of project details after a change in ownership or management or they were unwilling to participate in the study.

3.4 Multifamily Results

Table 3-2 illustrates the results of the MPP impact evaluation overall and by upper-level sampling strata for all fossil fuels combined.

Table 3-2. MPP Impact Evaluation Results – All Fossil Fuels

Upper-Level Stratum		Count Projects 2014 – Q2 2018	Reported Savings (All Fossil Fuels, MMBtu/year)	Evaluated Savings (All Fossil Fuels, MMBtu/year)	RR	Relative Precision at 90% Confidence
Affordable	100% Complete	220	536,823	392,528	0.73	±5.6%
Affordable	Performance Payment	50	127,709	150,106	1.17	±5.2%
Market Rate	100% Complete	77	345,142	263,951	0.76	±10.3%
Market Rate	Performance Payment	12	59,872	57,404	0.95	±4.4%
Total^a		359	1,069,545	867,301	0.81	±4.0%

^a Totals may not sum due to rounding.

The MPP realized 81% of reported fossil fuel MMBtu savings over the studied program years. As observed in prior evaluation cycles and hypothesized for this study, the projects that received the performance payment performed better (111% RR overall) than those that did not (74% RR). On the other hand, results between affordable and market rate strata are not significantly different. The evaluated savings correspond to a 23% reduction in site-level, pre-project fossil fuel consumption on average, as compared with the program’s targeted 15% energy reduction (electricity and fossil fuels).

Comparing Table 3-2 with the original sample design in Table 3-1, evaluators observe that the results exceeded the precision targets for each of the upper-level strata considered in the design, leading to overall relative precision of ±4% (90% confidence interval) as compared to the target ±10%. The favorable achieved precision reflects the lower variability in results than predicted—evaluators assumed an error ratio of 1.0 in the design, while the results yield an error ratio of 0.6.

Variability in results is best illustrated using a scatter plot. Figure 3-1 presents the project-by-project results as points graphed by reported (x-axis) and evaluated (y-axis) first-year fossil fuel MMBtu savings. The ideal evaluation result, defined as a RR of 100%, is illustrated as a solid black line. Since many small- and medium-saver projects are clustered in the bottom left of the figure, the rightmost figure presents a “close up” of the leftmost figure’s shaded area.

Figure 3-1. MPP Impact Evaluation Results: Overall (Left) and Close-Up (Right)

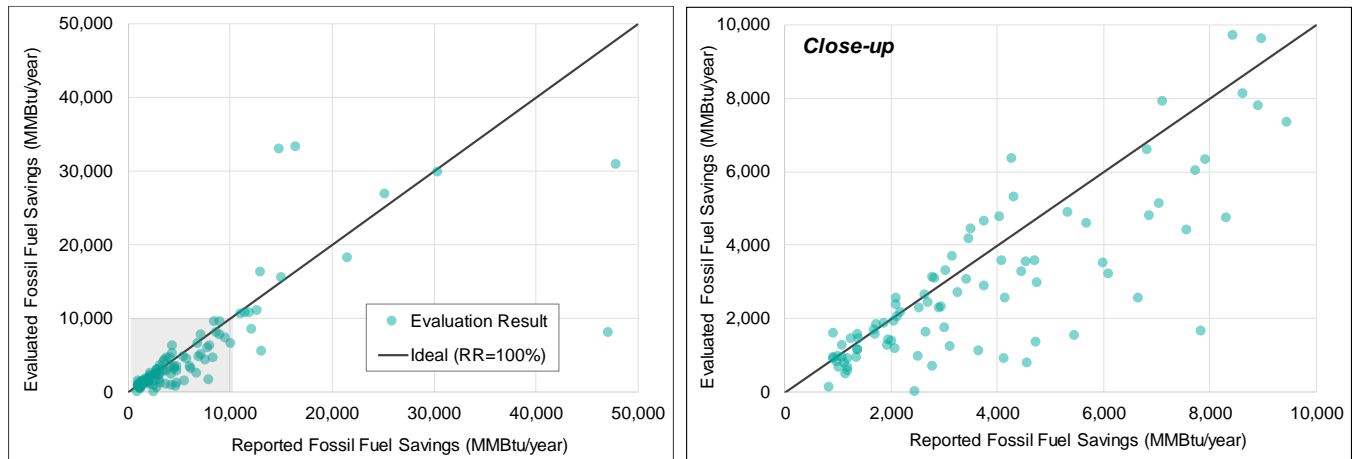


Figure 3-1 illustrates variability in results for some large-saver projects, but the majority of medium-saver projects cluster around the ideal line. Overall, most points fall below the ideal line, corroborating the 81% RR. The rightmost close-up shows that points that exceed the ideal 100% RR often do not exceed by much, while those that fall short do so significantly in some cases. This study’s results cluster near ideal more closely than in the program’s prior impact evaluation, leading to lower error ratio and better precision than predicted. Please note that Figure 3-1 excludes the lone evaluated project with a negative RR.

The prior impact evaluation focused on electricity, as that was the program’s focus during the studied 2009-2011 period. The electric energy savings realization rate was 0.79. While the program did not emphasize natural gas savings as much as electric, MPP projects typically reported gas savings as well. Evaluators determined a “performance factor,” analogous to realization rate, of 0.60 for all fossil fuels. This program’s 0.81 fossil fuel realization rate represents a marked improvement.

3.4.1 Discrepancy Analysis

Evaluators next investigated the key contributors to a realization rate 19% lower than ideal, as shown in Table 3-3.

Table 3-3. MPP Key Contributors to Fossil Fuel Realization Rate

Discrepancy Category	Negative Impact		Positive Impact		Overall	
	Frequency	RR Impact	RR Impact	Frequency	Frequency	RR Impact
Inaccurate pre-project characterization	6	-2%	1%	6	12	-1%
Incorrect baseline reference	8	-2%	0%	0	8	-2%
Incorrect reference to billing data	2	0%	0%	0	2	0%
Unknown applicant algorithm or assumptions	6	-2%	0%	0	6	-2%
Insufficient assessment of measure interactivity	1	0%	0%	0	1	0%
Differences in modeling software	0	0%	0%	1	1	0%
Inaccurate normalization to typical weather	16	-5%	1%	4	20	-4%
Inaccurate occupancy estimate	3	0%	4%	16	19	4%
Inoperable installed equipment	5	-2%	0%	0	5	-2%
Difference in quantity installed	15	-1%	1%	13	28	0%
Difference in equipment efficiency	69	-8%	5%	30	99	-4%
Difference in installed equipment size	4	0%	0%	3	7	0%
Difference in installed control strategy	17	-5%	1%	4	21	-4%
Difference in equipment load profile	17	-5%	1%	9	26	-3%
Difference in equipment hours of operation	3	0%	0%	1	4	0%
Administrative error	2	0%	0%	2	4	0%
Total	174	-33%	14%	89	263	-19%

The table presents positive and negative frequencies and RR impacts for 16 categories spanning various phases of an MPP project: the application process, ERP analysis, pre- and post-project inspections, and real-world performance. Three notable discrepancies are described below.

- Evaluators found that differences in **equipment efficiency** were the most frequently observed discrepancy. These savings differences are primarily caused by installed condensing boilers failing to operate at claimed efficiencies of 92% or greater. In practice, condensing boilers often do not achieve peak efficiency due to higher-than-expected return water temperatures that prevent condensing capability. Evaluators examined post-project inspection documents and frequently found return water temperatures in excess of 140°F, which exceeds the 90°F – 110°F condensing range.
- Impacts due to **normalization to typical weather** affected all evaluated projects in varying degrees. As described in the Billing Analysis subsection of Appendix B, the evaluators defined typical weather conditions as degree day averages over the past 11 years as opposed to TMY3. The 11-year-average approach resulted in 4% fewer total heating degree days overall, reducing savings accordingly as described in Appendix E.
- Savings differences due to **occupancy** changes occurred in 19 instances when occupancy fluctuations were identified through telephone interviews. Typically, the pre-installation period occupancy was lower than the

post-installation occupancy. As a result, the pre-project utility data was normalized to reflect the higher post-installation occupancy, which led to increased savings.

3.4.2 Fuel-Specific Results

The above results reflect overall fossil fuel impacts, encompassing all affected fuel types: natural gas, fuel oils, and district steam. Evaluators further examined results by fuel type as presented in Table 3-4.

Table 3-4. MPP Impact Evaluation Results by Fuel Type

Metric	All Fossil Fuels	Natural Gas	Fuel Oils	Steam
A – Projects in population (2014 – Q2 2018)	359	348	11	1
B – Projects in sample	106	103	26	5
C – Population reported savings (MMBtu/year)	1,069,545	1,000,581	68,320	644
D – Population evaluated savings (MMBtu/year)	867,301 ^a	-108,460 ^a	612,257 ^b	57,709 ^b
E – Realization rate	0.81	-0.11 ^c	N/A	N/A

^a For all fossil fuels and natural gas, the evaluated savings reflect the product of population reported savings (row C) and realization rate (row E).

^b For fuel oils and district steam, the evaluated savings reflect the total savings from the 106 evaluated projects, not the population as a whole, as RRs could not be developed for these fuels due to erroneously low reported savings. Therefore, the fuel-specific evaluated savings (row D) do not sum to the fossil fuel total in the second column.

^c Evaluators acknowledge that, over the studied years 2014 through 2018, many of the program's funding sources targeted gas savings only. This limitation provides important context around the fuel accounting issues and subsequently low gas RR.

Table 3-4 indicates varying results by fuel type. The program typically claimed impacts as natural gas, even for projects involving fuel conversions. On the other hand, evaluators analyzed all impacted fuels separately, including the removed fuel's savings and introduced fuel's penalty from fuel conversion projects. Table 3-4 illustrates that, because of the prominence of fuel conversion projects sponsored by MPP, the program *increased* natural gas consumption over the studied program years while removing "dirtier" fuel oils significantly more than reported. Appendix C includes an analysis of the MPP's tracked and achieved carbon emissions reduction.

3.4.3 Savings Persistence

The evaluators examined the year-by-year persistence of savings for 49 evaluated projects with multiple years of post-installation utility data. The MPP had previously sponsored persistence analysis that spanned an additional 20 projects within the evaluation timeframe. To solidify persistence results as much as possible, evaluators paired the prior analysis with this study's results to quantify two-year persistence for 69 projects and three-year persistence for 10 projects.²¹ Figure 3-2 illustrates the persistence analysis results aggregated by MMBtu impact and presented as reduction in pre-project annual fossil fuel consumption.

²¹ Due to typical utility archive limits of two years, evaluators had difficulty obtaining three or more years of post-project utility consumption data for more extended analysis.

Figure 3-2. MPP Year-by-Year Savings Persistence

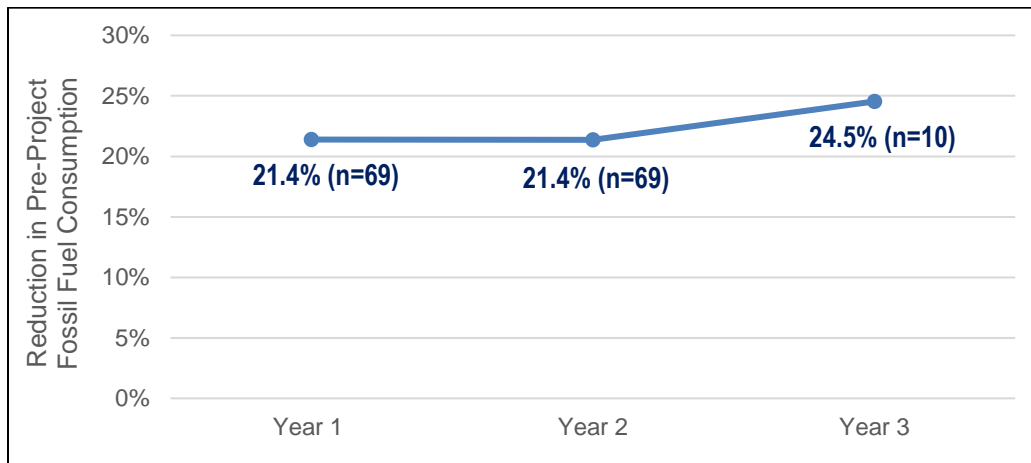


Figure 3-2 shows effectively no change in savings between the first and second years after project completion. Third-year persistence results actually show a savings increase, though the limited number of third-year data points likely introduces noise in the result.

3.5 Multifamily Conclusions and Recommendations

The following subsections present key findings from this evaluation study, followed by opportunities for program improvement.

3.5.1 Conclusions

1. The MPP achieves aggressive reductions in fossil fuel consumption at multifamily buildings. While this evaluation addressed fossil fuels only, the evaluators found that MPP projects result in a 23% reduction in pre-project fossil fuel consumption on average, compared with the program’s target of 15% energy reduction (electricity plus fossil fuels) at each participating facility.
2. The program’s MMBtu savings claims have become more accurate. The prior impact evaluation²² of the MPP (2013-14) determined a fossil fuel performance factor of 0.60, whereas this evaluation found a fossil fuel realization rate of 0.81.
3. Projects receiving the performance payment²³ performed significantly better (RR = 111%) than those that did not (RR = 74%). The evaluators found no significant differences in performance between affordable

²² Evaluators note that projects over the program years previously evaluated, 2009-11, were primarily SBC-funded. Therefore, fossil fuel savings were not the focus of such projects, but the program nonetheless reported those impacts and evaluators assessed them in the prior study.

²³ As detailed in Section 3.2, projects receiving the performance payment had achieved the targeted energy savings as demonstrated through independent analysis of pre- and post-project utility consumption data.

and market rate projects. Upstate projects performed better (RR = 89%) than downstate (RR = 75%), possibly due to lower likelihood of a fuel conversion measure.

4. The evaluators commend the MPP for its comprehensive data tracking and archiving practices. For all sampled and backup projects, the evaluators requested the program's archived project files. In nearly all cases, program staff returned sufficient project files for evaluation, including the energy reduction plan, inspection documents, photographs, and pre-project consumption data.
5. The MPP does not, however, adequately track impacts by specific fuel type. The program frequently claimed all fossil fuel impacts as natural gas, even when fuel conversions occurred (e.g., #2 fuel oil to natural gas). Such inaccuracies underestimate the program's carbon emissions reduction impacts, as summarized in Appendix C. The evaluators acknowledge that many of the program's gas-focused funding sources over the 2014-18 evaluation timeframe limited the program's ability to claim impacts for delivered fuels and district steam.²⁴
6. Projects with an "anchor" measure performed better than those without. While measure-specific impacts were difficult to glean from whole-building billing analysis, the evaluators nonetheless examined aggregate results among projects including a specific measure or not. Most strikingly, projects with a boiler measure (n = 71) achieved 87% of reported MMBtu savings, while projects without a boiler measure (n = 35) achieved 70%. This finding suggests that high-impact, predictable measures like central plant upgrades will typically result in more accurate savings claims compared with less predictable measures like controls, weatherization, and re-commissioning.
7. The program's savings models generally differed from real-world performance in three cases:
 - a. Condensing boilers typically did not achieve the modeled performance efficiencies due to higher-than-expected return water temperatures (RWTs). Most sampled projects' post-installation inspection reports showed RWTs of 140°F or higher, limiting condensing capability and reducing achieved efficiency. While conclusion #6 describes how projects with boilers performed better than those without, this finding indicates that condensing boilers could be performing even better with RWTs closer to design conditions.
 - b. Controls and weatherization measures are difficult to characterize using modeling software. In some cases, these measures had significant savings claims (i.e., 10% of pre-project consumption or higher) but often did not achieve the expected savings.

²⁴ Table 3-4 illustrates the significant evaluated savings for fuel oils compared to reported savings.

- c. The program’s simulation software uses TMY3 weather files to represent typical weather conditions, whereas evaluators used 11-year (2008–2018) NOAA weather averages to define typical conditions. This difference resulted in slightly lower HDDs and evaluated savings.
8. The evaluators found that MPP savings persist from the first year after project completion to the second year. In fact, savings are constant from the first to second year, while third-year savings increase slightly, though the number of third-year data points is limited.

3.5.2 Recommendations

1. The evaluators recommend that the MPP enhance its savings tracking protocols to allow more accurate reporting of fuel-specific impacts. While prior program versions and funding sources did not necessitate fuel-specific tracking or quantification of carbon impacts, NYSERDA’s Clean Energy Fund currently emphasizes fuel-specific MMBtu and carbon emissions reduction as a path to achieving New York’s broader greenhouse gas emission goals. After discussions with program staff, the evaluators are encouraged to hear that such improvements are underway.

The evaluation team suggests two different techniques, in order of preference, for enhancing the MPP’s tracking systems to report fuel-specific impacts more accurately:

- a. The current tracking database contains a single field for MMBtu reported savings, and another field in which the “most affected fuel” is identified. The evaluators recommend enhancing the tracking system with fuel-specific reported MMBtu fields: natural gas, #2 fuel oil, #4 fuel oil, #6 fuel oil, and district steam. These added fields would allow a single measure line item, such as a boiler fuel conversion, to claim impacts for multiple fuels.
 - b. If adding fields to the tracking database is not feasible, the evaluators alternatively recommend adding more line items per measure to characterize the different fuels affected. In the case of the fuel conversion example, one measure line item should reflect the removed fuel’s MMBtu savings (i.e., positive fuel oil impacts) and another line item should reflect the added fuel’s MMBtu penalty (i.e., negative natural gas impacts). For this alternative technique, the “most affected fuel” field should be expanded to include different grades of fuel oil as well as any other delivered fuels impacted by the program.
2. The evaluators commend the program for continually improving its Simulation Guidelines for performance partners and recommend three specific refinements as a result of conclusion #7.
 - a. Modeled condensing boiler measures should reflect a realistic efficiency value based on the installed boiler’s performance curve and the RWTs identified in the post-installation inspection

- report. If it is not feasible to revise project claims at that point, the evaluators suggest that the program update its Simulation Guidelines to require performance partners to model the boiler at an efficiency no greater than five percentage points less than peak design efficiency, unless site-specific evidence of more favorable real-world RWTs is provided.
- b. The program should require supplemental supporting evidence, such as measurement-based justification for model inputs, for any proposed controls, weatherization, or re-commissioning measures that claim savings of more than 10% of pre-project whole-building consumption.
 - c. The MPP (and NYSERDA) should establish a uniform definition of typical weather. NYSERDA's evaluation management instructed the evaluation team to consider 2008–2018 NOAA degree-day averages as typical, while the program's performance partners use modeling software that incorporate weather files based on Typical Meteorological Year data. The evaluator's analysis shows that the 11-year averages are slightly milder than TMY3.

4 New Construction

4.1 Introduction

This impact evaluation studied the gross impact of the NYSERDA Energy Efficiency Portfolio Standard (EEPS)–funded²⁵ Commercial New Construction Program (CNCP).²⁶ The evaluated period of performance covers projects completed in 2016 through 2018 Q2.

4.2 Program Background – CNCP

The EEPS-funded CNCP provided technical assistance and financial incentive to business customers building new facilities or undertaking extensive renovations of existing buildings. Projects fall under three broad categories based on the analysis approach: Whole Building (WB), Custom, and Prequalified (PQ). More information on the program background and history can be found in Appendix D.

4.3 Methodology – CNCP

The following subsections describe the CNCP evaluation sample design, data collection methods, and site-level and aggregate analysis approach.

4.3.1 Sample Design

The goal of the CNCP sample design was to statistically select a sample of EEPS-2 (electric and gas) only projects performed under PON 1601. These projects were installed from January 1, 2016, through June 30, 2018, with an application received between January 4, 2012, and March 23, 2015. This design was developed to target $\pm 10\%$ precision at the 90% confidence interval around first-year energy savings at the program level. The design is broken out by electric and gas projects as an upper stratification level. Tables 4-1 and 4-2, below, shows the final sample design achieved for electric and natural gas respectively.

Table 4-1. CNCP Electric Sample Design Summary

Measure Type	Projects	Savings (MWh)	Error Ratio	Sample Size	Expected Relative Precision
Electric	251	85,004	0.70	50	12.7%

²⁵ In May 2007, the New York Public Service Commission (PSC) issued an order instituting a proceeding to develop an Energy Efficiency Portfolio Standard (EEPS1). In October 2011, the PSC issued a further order reauthorizing EEPS programs through December 31, 2015. In December 2015, the PSC issued an order extending the Energy Efficiency Portfolio Standard (EEPS2) through Feb 29, 2016. Collectively, EEPS1 and EEPS2 activities and funding sources are referred to as EEPS.

²⁶ Department of Public Service, Filing #4779 Case No. 07-M-0458, 2016

Table 4-2. CNCP Natural Gas Sample Design Summary

Measure Type	Projects	Savings (MMBtu)	Error Ratio	Sample Size	Expected Relative Precision
Gas	27	141,869	0.90	17	14.6%

Additional details on the sampling approach can be found in Appendix A.

4.3.2 Data Collection and Site-Specific Analysis

After the sample was designed, the evaluators requested all available program-archived files for the 67 sampled projects, in addition to several replacement projects; whole-facility billing data for CNCP was not provided.

After the engineer reviewed project file information to find customer contact information, verify tracking savings, and ensure that the file was sufficiently complete for evaluation, he or she attempted to conduct a telephone interview to collect critical site-specific data which impacts the analysis approach and results. CNCP's preparatory desk reviews were intensive and involved extra steps not necessary for EFP and MPP. They included reviews of:

1. CNCP project application
2. Final TA study report
3. TA analysis files (energy simulation models, custom/prequalified spreadsheets etc.)
4. Commissioning reports, post-installation inspection documentation and photographs
5. CNCP incentive calculator and whole building calculator templates

The evaluators attempted phone interviews of managers at all CNCP sampled facilities or, if they preferred, email. Interviews were completed for 66% of sampled projects. For CNCP, the evaluators revised the original TA calculations with information gathered from the project files and customer interviews. Appendix D describes CNCP-specific data collection and analysis in more detail.

4.3.3 Aggregation

After all project-level results were calculated, the evaluators calculated program-level evaluation results through statistical expansion analysis. This is described in Appendix B and is common to MPP and EFP evaluations as well.

4.4 New Construction Results

Tables 4-3 and 4-4 provide the overall impact evaluation results for the CNCP projects claiming electric and natural gas savings respectively, under EEPS-2 installed from January 1, 2016, through June 30, 2018.

Table 4-3. CNCP Overall Electric Results

Measure Type	Count Projects – 2016 – Q2 2018 (EEPS 2)	Sampled Projects	Reported Savings (MWh/year)	Evaluated Savings (MWh/year)	Realization Rate	Relative Precision at 90% Confidence	Error Ratio
Electric	251	50	85,004	288,477	0.99	5.52%	0.32

Table 4-4. CNCP Overall Natural Gas Results

Measure Type	Count Projects – 2016 – Q2 2018 (EEPS 2)	Sampled Projects	Reported Savings (MMBtu/year)	Evaluated Savings (MMBtu/year)	Realization Rate	Relative Precision at 90% Confidence	Error Ratio
Gas	27	17	141,869	101,191	0.71	6.81%	0.43

Table 4-5 compares this cycle’s RRs to those of prior evaluation evaluations.

Table 4-5. CNCP Current and Prior Impact Evaluation Realization Rates

Program	Electric Energy	Electric Demand	Natural Gas
CNCP – 2018	0.99	N/A	0.71
CNCP – 2016	0.89	0.70	N/A
CNCP – 2012	0.71	0.52	N/A

The evaluators determined an overall annual electric energy savings gross RR of 99%, at a relative precision of $\pm 5.52\%$ at the 90% confidence interval. Factors contributing to the slightly lower than 100% kWh RR are discussed in Section 4.4.1. The evaluated results for kWh achieved the $\pm 10\%$ relative precision bound targeted for the CNCP program as a whole.

Below, Figure 4-1 compares program-reported and evaluated annual kWh savings for the sample of CNCP projects studied. Ideally, the evaluated savings would always match the reported savings; this ideal is shown as a solid blue line on the charts, with the rightmost figure a close-up of the gray box in the left.

Figure 4-1. Comparison of CNCP Reported and Evaluated Electric Energy Savings: Overall (Left) and Close-Up (Right)

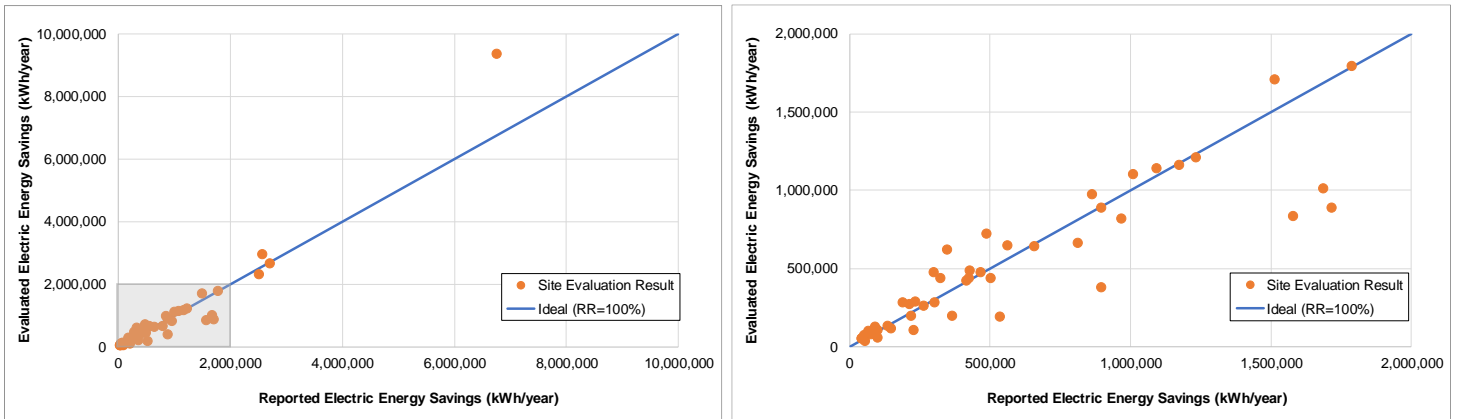
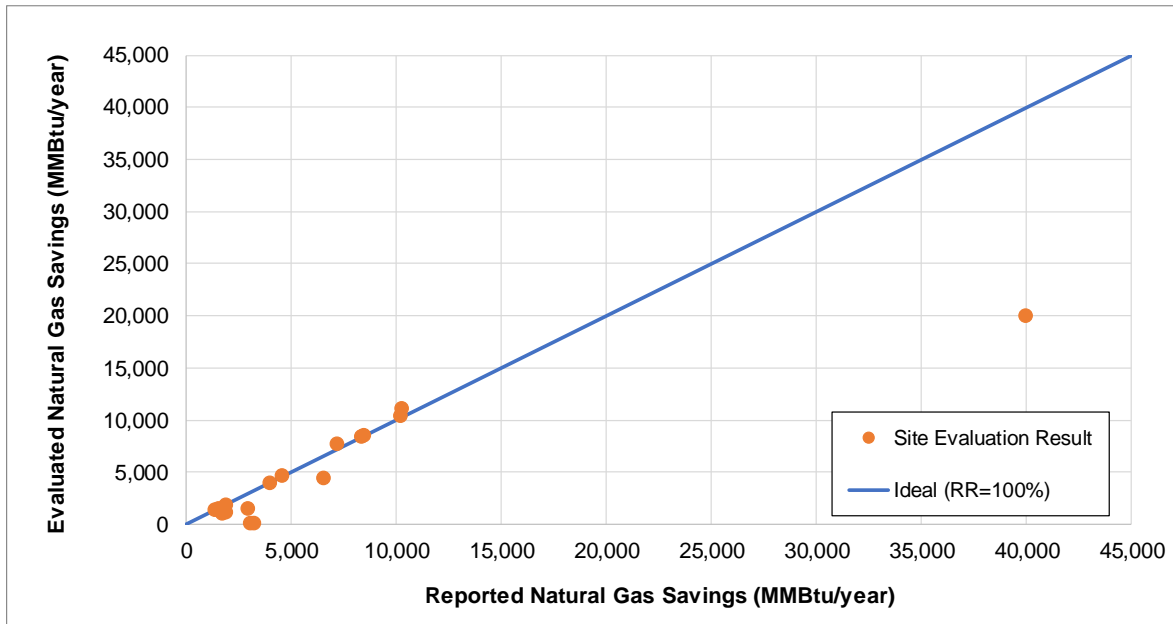


Figure 4-1 shows most of the smaller CNCP projects, in the range of 0 to 2,000,000 reported annual kWh savings, cluster around the ideal line. The majority of points fall below the ideal line, however, the largest kWh-saving project illustrated in Figure 4-1 had a significantly high evaluated savings explaining the near 100% RR for the kWh projects overall.

The evaluators determined a gross RR of 71%, at a relative precision of $\pm 6.81\%$ at the 90% confidence interval, for annual natural gas savings. Contributing factors for the 29% less natural gas savings than anticipated are explained in Section 4.4.1. Figure 4-2 compares the program-reported and evaluated annual natural gas savings for the sample of CNCP projects studied.

Figure 4-2. Comparison of CNCP Reported and Evaluated Natural Gas Savings



As illustrated in Figure 4-2, the majority of the projects resulted in near-ideal RR for natural gas impacts. Additionally, Figure 4-2 shows two of the smaller natural gas projects with no evaluated savings and the largest natural gas saving project in the sample led to significantly lower evaluated savings. These three projects were the primary contributors to the 29% lower natural gas savings than reported.

4.4.1 Discrepancy Analysis

The primary objectives of this study include identifying why the evaluated gross savings estimates differ from the program-reported savings estimates. The evaluators completed separate discrepancy analyses for the electric and natural gas energy savings.

Although the electric projects resulted in a realization rate of 99%, the evaluators investigated the key differences between reported and evaluated savings and identified several site-specific discrepancies as shown in Table 4-6.

Table 4-6. Investigation of Key Contributors to CNCP Electric Realization Rate

Discrepancy Category	Negative Impact		Positive Impact		Overall	
	Frequency	RR Impact	RR Impact	Frequency	Frequency	RR Impact
Algorithm/Adherence to TRM	7	-2%	3%	7	14	1%
Baseline	24	-7%	3%	10	34	-4%
Efficiency	7	-1%	0%	7	14	0%
Inaccurate normalization to typical weather	5	0%	2%	11	16	2%
Load	4	0%	1%	11	15	1%
Measure Not Installed	12	-2%	0%	0	12	-2%
Operations	20	-2%	4%	20	40	1%
Other	9	-2%	1%	8	17	-1%
Quantity	12	-2%	3%	12	24	1%
Size	4	0%	0%	3	7	0%
Tracking/Clerical	9	-1%	1%	9	18	0%
Total	113	-19%	18%	98	211	-1%

The table above presents positive and negative RR impacts and the frequency of these occurrences for eleven categories spanning various phases of a CNCP project. Selected discrepancy categories are examined in more detail below:

- The evaluators found that differences in **baseline characterization** between the evaluators and the applicant most significantly impacted the electric kWh RR. In 34 instances, the evaluators identified that the measure baselines needed to be updated based on the applicable ASHRAE 90.1 2007 code, and for certain whole building projects, the baselines were adjusted by the evaluators to match baseline modeling requirements. In one project example, evaluators identified that variable frequency drives (VFD) were

required by code on the applicable HVAC fan motors, whereas the applicant modeled the baseline HVAC fan motors without VFD.

- Impacts due to **normalization to typical weather** is a discrepancy that affected most of the evaluated whole building projects. The evaluators utilized TMY3-based weather files to run baseline and as-built simulation models, whereas the TA utilized TMY2 based weather files for the applicable location. Overall, this discrepancy led to a 2% increase to the electric kWh RR.
- Based on conversations with the site contact and detailed review of the post-installation documentation provided by NYSERDA, the evaluators identified that some of the **measures were not installed** as described in the TA report. This discrepancy occurred in 12 instances and led to a 2% reduction to the electric kWh RR overall.
- Differences in **equipment operation** estimated by the applicant and the evaluators were the most frequently observed discrepancy. The evaluators adjusted lighting operational hours and measure-specific HVAC set points in 40 instances based on conversations with the site contact, resulting in a 1% increase to the electric kWh RR overall.

For natural gas projects, with an RR of 29% lower than ideal, the evaluators investigated the key differences between reported and evaluated savings in Table 4-7.

Table 4-7. Investigation of Key Contributors to CNCP Natural Gas Realization Rate

Discrepancy Category	Negative Impact		Positive Impact		Overall	
	Frequency	RR Impact	RR Impact	Frequency	Frequency	RR Impact
Algorithm/Adherence to TRM	2	0%	0%	0	2	0%
Baseline	6	-4%	1%	2	8	-3%
Efficiency	3	0%	0%	0	3	0%
Inaccurate normalization to typical weather	7	-1%	3%	4	11	2%
Measure Not Installed	3	-7%	0%	3	6	-7%
Operations	3	-11%	0%	2	5	-11%
Other	1	0%	0%	0	1	0%
Quantity	1	0%	0%	1	2	0%
Size	1	0%	0%	1	2	0%
Tracking/Clerical	8	-10%	1%	2	10	-10%
Grand Total	35	-33%	5%	15	50	-29%

The table above presents positive and negative frequencies and RR impacts for ten categories spanning various phases of a CNCP project. Selected discrepancy categories are examined in more detail below:

- Differences in **equipment operation** estimated by the applicant and the evaluators most significantly impacted the natural gas RR. For one example project, evaluators verified through detailed conversations

with the site staff and supplemental billing analysis that the space heating load at the facility was less than half of what was estimated by the TA, resulting in significantly lower natural gas savings than anticipated. This discrepancy occurred in 5 instances and led to a 11% reduction to the natural gas RR overall.

- Evaluators observed **tracking/clerical** errors in 10 instances leading to a 10% reduction to the natural gas RR. On two of these instances, the program inaccurately tracked the fossil fuel savings resulting from the projects entirely to natural gas. The fossil fuel savings occurring at the facilities were not attributable to natural gas, as the facilities neither had an active gas connection nor did not have any plans to utilize natural gas in the future.
- Based on conversations with the site contact and detailed review of the post-installation documentation provided by NYSERDA, the evaluators identified that some of the **measures were not installed** as described in the TA report. This discrepancy occurred in 6 instances and led to a 7% reduction to the natural gas RR overall.

4.5 New Construction Findings and Conclusions

This section outlines the key findings resulting from the CNCP evaluation efforts. NYSERDA staff should review the current CEF program to see if the program is at risk for the process and methodological errors found in this evaluation.

4.5.1 Findings and Conclusions

1. Baseline-related discrepancies had the largest impact (-4%) on program electric RR and a notable impact (-3%) impact on gas RR. Overall, the evaluators verified that there were 42 occurrences where the applicant baseline did not match the value stipulated by ASHRAE 90.1-2007. Most baseline discrepancies were due to the baseline energy models not being developed correctly or the wrong value being used in the model to define a baseline parameter (for example, using a baseline of U-0.306 for wall insulation instead of the ASHRAE 90.1-2007 baseline of U-0.064). The evaluators anticipated that most of these errors would have been identified during the model review stage.
2. There were two occurrences where energy savings from other fuels were reported as natural gas savings. For these two projects, the evaluators assigned a realization rate of 0 as no natural gas savings were obtained even though the project may have saved on the non-incented fuel. Reporting zero gas savings for these two projects resulted in a -10% impact on the gas RR. As more emphasis is placed on carbon reduction, accurate fuel characterization is important.
3. The evaluators recorded 18 instances where reported measures were not installed, resulting in a -2% impact on the electric RR and a -7% impact on the gas RR. The evaluators confirmed that the measures

were not installed either through a phone interview with the site contact or via the post-inspection and commissioning reports present in the project files. Because post-inspection and commissioning reports indicated that measures were not installed yet tracking savings did not reflect this, the evaluators believe the tracking savings are not consistently updated to reflect post-installation and commissioning report findings.

4. There were at least 14 occurrences where the evaluators observed that the TA calculations had discrepancies or referenced the wrong TRM values (hours, energy savings factor, etc.) for the application. While the total number of occurrences account for about 5% of recorded discrepancies and the impact on the gas and electric RR was low (0% and 1%, respectively), it is a reminder to the program administrators to maintain a strong technical review process so as to keep such occurrences low.
5. For several project files, neither the as-built mechanical drawings or the complete modeling files were available, preventing the evaluators from replicating the TA evaluator modeling approach. Referencing these documents would have improved the efficiency and accuracy of the evaluation.