# Single-Family LMI Retrofit Programs Evaluation: Natural Gas and Electric Billing Analysis Results (Program Year 2023)

***Final Report***

Prepared for:

**New York State Energy Research and Development Authority**
Albany, NY

**Elizabeth Boulton**

Project Manager, NYSERDA Market Characterization and Evaluation

Prepared by:

**DNV**

Corporate Headquarters: Katy, TX

Ari Stern, Senior Consultant, DNV

Miriam Goldberg, Sr. Principal, DNV

Antonia Sanhueza, Engineer, DNV

Patrick Hewlett, Principal, DNV

NYSERDA Contract 151666 TWO ID No. 216912 September 2025

# Record of Revision

| Document Title |
|---|
| Single-Family LMI Retrofit Programs Evaluation: Natural Gas and Electric Billing Analysis Results<br>September 2025 |

| Revision Date | Description of Changes | Revision on Page(s) |
|---|---|---|
| September 2025 | Original Issue | Original Issue |
| | | |
| | | |
| | | |
| | | |
| | | |
| | | |
| | | |
| | | |
| | | |
| | | |
| | | |

# Notice

# Table of contents

## List of figures



## List of tables

# Glossary

- **Advanced Metering Infrastructure (AMI):** In this evaluation, refers to data from advanced meters that are capable of providing consumption data at hourly intervals.
- **Adherence rates:** A percentage showing how closely estimated savings match TRM savings. 100% means estimated savings equal TRM savings.
- **Default TRM savings:** Savings calculated by the evaluation contractor team using TRM algorithms and default values for input variables when available in the TRM.
- **Disadvantaged Communities (DACs):** Communities that meet criteria developed by New York Climate Justice Working Group for disadvantaged communities based on a range of socio-economic factors. This evaluation used criteria established in 2023, though the criteria will be updated by the end of 2025.
- **EmPCalc**: The tool used by the EmPower+ program prior to PY2023 to estimate savings.
- **EmPower+:** The program that is the focus of this evaluation. It is the combination of what was previously known as the EmPower program, which served low-income customers, and the Assisted Home Performance program, which served moderate-income customers.
- **Estimated savings:** Savings estimates calculated by program project management software using inputs provided from contractors and provided by the program for this study.
- **Evaluated savings:** Savings calculated by the evaluation contractor team from billing analysis (i.e., consumption data analysis).
- **kWh:** Kilowatt-hours
- **Low-income (LI):** A descriptor for households with income below 60% State median income (SMI).
- **Low-to-Moderate-Income (LMI):** The combination of low-income and moderate-income households.
- **MMBtu:** Million British thermal units
- **Moderate-income (MI):** A descriptor for households with income below 80% State median income (SMI).
- **New York Home Energy Portal (NYHEP):** The data portal used for program participation that started being used in July 2023.
- **Quality Assurance Score (QA Score**): A 0 to 5 score used by NYSERDA to assess the quality of contractors' work in the program.
- **Realization rates:** The percentage of estimated savings that shows up in the evaluated savings (evaluated savings/estimated savings).
- **Site-specific TRM savings:** Savings calculated by the evaluation contractor team using TRM algorithms and all available site-specific data from the program data.
- **Single-family (SF):** Refers to one-to-four-unit buildings.
- **Technical Reference Manual (TRM):** In this evaluation, refers to Version 10 of the New York Standard Approach for Estimating Energy Savings from Energy Efficiency Program.
- **Verified savings:** Savings calculated by the evaluation contractor use TRM algorithms.
- **Weatherization Assistance Program (WAP):** A program that helps low-income customers weatherize their homes.

# Executive summary

This evaluation was conducted by DNV (evaluation contractor team) for the NYSERDA EmPower+ program for the year 2023. The EmPower+ program provides comprehensive home energy assessments to low- and moderate-income single-family and two-to-four-unit households, followed by the implementation of a selection of energy efficiency and beneficial electrification measures. The evaluation had two main objectives:

1. Quantify the impact of the program on participant energy consumption.
2. Provide actionable recommendations to improve the program.

Table ES-1 shows the various activities included in the evaluation, their purpose, and the achieved sample sizes for each.

**Table ES-1. Study activities and purpose**

| Activity performed | Purpose | Sample size |
|---|---|---|
| Site-level billing analysis | Quantify participation's impacts on energy consumption and utility bills. | 1,560 gas sites and 4,182 electric sites |
| Engineering analysis | Assess the adherence of the New York Home Energy Portal (NYHEP) savings estimation methodology to the New York Technical reference manual (TRM) Version 10.[1] | All PY2023 participants in NYHEP |
| Program staff interviews | Developed understanding of program process, identified knowledge gaps for the program, and inflation reduction act impacts. | 8 |
| Participant survey | Answer process questions relating to how participants entered the program, their satisfaction with various parts of participation, and perceived impacts of participation. | 1,166 participant customers |
| Contractor survey | Answer process questions relating to contractors' experience with the program, customer education, barriers to measure uptake, and trainings | 99 participating contractors |
| Delivered fuel analysis | Quantify energy savings resulting from participation of households that use propane, oil, or wood. | 75 participants |
| Add-on participant survey | Provide additional insights into beneficial snapback resulting from program participation. | 477 participant customers |

[1] https://dps.ny.gov/technical-resource-manual-version-10-filed-december-30-2022-effective-january-1-2023

This report focuses on the natural gas and electric billing analysis results. A later report will cover all other research activities.

## Key Findings

**Finding 1: The realization rates are 20% for gas savings and 18% for electric savings (row F in Table ES-2).** Realization rates show how evaluated savings compare to estimated savings and are influenced by factors like end-uses present in the home, usage patterns, and existing efficiencies of installed measures. They are not a measure of a program's impact, but they are a measure of how well the impact was estimated. Various factors can cause differences between actual savings and estimated savings, including assumptions made during the estimation process. If both estimated and evaluated savings perfectly matched actual savings, realization rates would be 1. Rates below 1 usually result from differences between assumed inputs and what actually occurred. For a full discussion on the limitations of realization rates, see Appendix B.

**Table ES-2. Overall savings summary**

| Parameter | Gas (MMbtu) | | | Electricity (kWh) | | |
|---|---|---|---|---|---|---|
| | Overall | LI | MI | Overall | LI | MI |
| Sample count | 1,560 | 1,517 | 43 | 4,182 | 4,142 | 40 |
| A. Estimated savings (kWh or MMBtu/yr) | 40 | 40 | 17 | 707 | 712 | 223 |
| B. Average evaluated baseline usage (kWh or MMBtu/yr) | 106.6 | 106.6 | 95.0 | 8,590 | 8,596 | 7,987 |
| C. Average estimated % savings (kWh or MMBtu/yr; A/B) | 38% | 38% | 18% | 8% | 8% | 3% |
| D. Average evaluated savings (kWh or MMBtu/yr) | 8.1 | 8.1 | 10.2 | 127 | 128 | 68 |
| E. Average evaluated % savings (kWh or MMBtu/yr; D/B) | 8% ±1% | 8% ±1% | 11% ±3% | 1% ±1% | 1% ±1% | 1% ±5% |
| F. Realization rate (D/A) | 20% ±16% | 20% ±17% | 58% ±46% | 18% ±60% | 18% ±60% | 31% ±998% |

**Finding 2: Changes in the way estimated savings are calculated have resulted in lower realization rates.**

ES-3 shows the realization rates over time for gas participants. While PY2023 evaluated savings values are similar to previous studies, estimated savings are much higher, leading to lower realization rates. The engineering analysis found that the program software adheres well to the NY Technical Resource Manual (TRM) methodologies for calculating estimated savings, and so the high estimated savings are a product of the TRM itself. In 2023, the program shifted to using the TRM to calculate estimated savings rather than using a proprietary tool, EmPCalc, as was done in previous years. Unlike the TRM, the EmPCalc tool accounted for interactive effects between measures and was calibrated to previous year evaluated savings. The change from EmPCalc to the TRM coincides with the increase in estimated savings.[2]

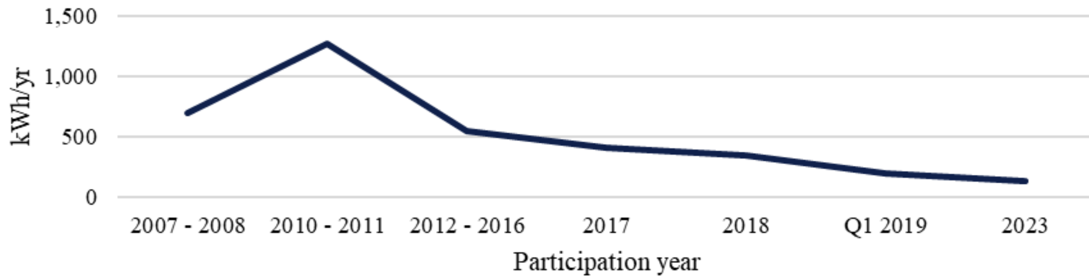
**Table ES-3. Historical LI gas evaluated savings, estimated savings, and realization rates**

| Participation year | A. Average evaluated savings (MMBtu/yr) | B. Average estimated savings (MMBtu/yr) | C. Realization rate (A/B) |
|---|---:|---:|---:|
| 2007–2008 | 10 | 14 | 71% |
| 2010–2011 | 2 | 4 | 50% |
| 2012–2016 | 12 | 27 | 44% |
| 2016–mid 2017 | 8 | 26 | 31% |
| 2017–Q1 2019 | 9 | 26 | 35% |
| **2023 (Current study)** | **8** | **40** | **20%** |

**Finding 3: Electric savings have been diminishing since 2011.** As shown in Figure ES-1, the average evaluated electric savings have been decreasing in each evaluation since the evaluation for PY2011. As LED lighting becomes more widely adopted electric baselines have gotten more efficient making it harder to achieve additional savings. The main driver of electric savings in PY2023 identified by the billing analysis, replacement refrigerators and freezers, also has diminishing savings (430 kwh yr in 2019 to 218 kwh/yr in 2023.

[2] For electric realization rates over time, see Section 2.5.2.

**Figure ES-1. LI electric evaluated and estimated savings historical trend**




**Finding 4: EmPower+ has greater impact on gas consumption than on electric consumption.** The billing analysis found evaluated gas savings representing 8% of baseline usage, while electric savings were only 1% of baseline usage.[3] The gas savings are similar to previous evaluations, while the electric savings are slightly lower than previous evaluations. Gas savings are driven by insulation and air sealing measures. For both gas and electric savings, there is no significant difference between geographic DAC status, utility, or geographic region.

**Finding 5: There is some indication of beneficial snapback which leads to positive outcomes for customers associated with an increase in energy consumption.** For example, if the program repairs broken heating, cooling, or water heating equipment, the customer can use that equipment again which would result in an increase in energy consumption. Indeed, the billing analysis found that equipment repair measures were associated with negative savings. Additionally, during the add-on participant survey, 35% of respondents who used a secondary source of heating for part of their home reported using that secondary source less often after participating.

**Recommendations**

**Recommendation 1: Consider focusing on fossil fuel savings and beneficial electrification rather than electric savings.** The evaluation found the program is more effective at achieving gas and delivered fuel savings while electric savings have been diminishing over time. New York State climate goals also call for electrification of heating and water heating systems, which increases electric consumption. As electric baselines become more efficient, achieving electric

[3] See Section 2.3.1, Overall evaluated savings.

savings will only become more difficult, and payback time on electric measures will increase. A focus on fossil fuel savings and electrification may better align the program with real-world limitations and New York state climate goals.

**Response to Recommendation**: Implemented. Beginning in March 2025, EmPower+ began to phase out incentivizing the replacement of refrigerator and freezer incentives due to limited opportunities to achieve electric savings and ability to meet cost effective payback. Further, the program expects to phase out incentives for LED lighting beginning in 2027 as a non-strategic measure identified by the NY Technical Resource Manual.

**Recommendation 2: Engage with NYS Department of Public Service and Technical Resource Manual (TRM) Management Committee to address the overestimation of program savings resulting from the TRM.** NYSERDA could explore the possibility of applying lessons learned from the EmPCalc tool to the New York Home Energy Portal (NYHEP) portal calculation methodology. This could include discussion of allowing NYHEP to be calibrated to previous estimations of evaluated savings and account for interactive effects, rather than just using the TRM methodology.

**Response to Recommendation**: Pending. The TRM management process includes use of evaluation findings and recommendations to continually improve methods. NYSERDA will consider this recommendation.

# 1 Introduction

## 1.1 Program description

The EmPower+ program is part of the Statewide Low and Moderate Income (LMI) Portfolio,[4] and is administrated by NYSERDA. Prior to 2023, EmPower+ was two separate NYSERDA-operated programs: EmPower NY and Assisted Home Performance (AHP). The program includes a comprehensive home energy assessment, followed by the implementation of a selection of energy efficiency and beneficial electrification measures, as shown in Table 1-1. NYSERDA directly pays contractors for work they have done at participating customers' homes based on the pre-approved work scopes. There are two tiers of the EmPower+ program. Tier 1 serves low-income households and Tier 3 serves moderate-income households.

Table 1-1 provides details about EmPower+ program, including income eligibility, building eligibility, homeowner incentives, measures offered, and the modeling tool used to calculate savings.

**Table 1-1. EmPower+ program details**

| | **EmPower+** |
|---|---|
| Income eligibility | Low-income: household income below 60% State median income (SMI)<br>Moderate-income: A household with income below the higher of 80% of area median income (AMI) and 80% of SMI |
| Building eligibility | Single family, multifamily (2-4 units)<br>Owner occupied and rental housing (rental housing need landlord consent for all measures except electric reduction measures) |
| Funding and constraints | • Clean Energy Fund (CEF) and New Efficiency: New York (NENY) – Condition: participants must be a customer of a Systems Benefits collecting utility<br>• Office of Temporary Disability (OTDA) – Condition: household must participate in LIHEAP; no utility restriction<br>• Regional Greenhouse Gas Initiative (RGGI) |
| Homeowner incentives | Tier 1: Up to $10,000 with special projects up to $13,000. OTDA customers receive up to $20,000 for one-unit buildings and an initial $20,000 plus $5,000 for additional units in two- to four-unit buildings |

[4] https://documents.dps.ny.gov/public/Common/ViewDoc.aspx?DocRefId=%7b7E40641A-263D-4969-9AD2-D68EE6122114%7d

| | EmPower+ |
|---|---|
| | Tier 3: 50% cost covering incentive up to $5,000 for single unit buildings, and an initial $5,000 plus $2,500 for additional units in two- to four-unit buildings |
| Offered measures | Approximately 110 measures, including:<br>• Shell measures such as insulation and air sealing<br>• HVAC such as gas furnaces, heat pumps, and water heaters<br>• Electric reduction such as refrigerators and light bulbs<br>• health and safety such as CO detectors, smoke detectors, and bath fans |
| Modeling tool | NY TRM using LMI savings calculations |
| Project Savings Estimate | The EmPower+ program uses modeling tool(s) to calculate savings by aggregating the measure level savings to create a project level estimate |

## 1.2 Study objectives and background

This evaluation of the NYSERDA EmPower+ single-family (SF) LMI Program focuses on evaluating energy and cost savings performance, by measuring first-year and lifetime energy savings and realization rates for electricity and natural gas, as described in Table 1-2.

**Table 1-2. Study objectives and methods**

| Objective | Purpose | Method |
|---|---|---|
| Evaluate energy and cost savings performance | Verify first year gross electric and gas savings for program participants | Billing analysis using Uniform Methods Project (UMP)[a] and IPMVP Option C[b] |
| | | |

a Agnew, K.; Goldberg, M. (2017). *Chapter 8: Whole-Building Retrofit with Consumption Data Analysis Evaluation Protocol, The Uniform Methods Project: Methods for Determining Energy-Efficiency Savings for Specific Measures.* National Renewable Energy Laboratory. NREL/SR-7A40-68564. http://www.nrel.gov/docs/fy17osti/68564.pdf

b *International Performance Measurement and Verification Protocol (IPMVP) Core Concepts - 2022.* Efficiency Valuation Organization. https://evo-world.org/en/products-services-mainmenu-en/protocols/ipmvp

## 1.3 Activities performed

The evaluation activities for the NYSERDA EmPower+ SF LMI Program include site-level billing analysis to quantify realization rates by specific energy efficiency measures, advanced metering infrastructure (AMI) analysis to estimate peak demand energy savings, and segmentation of sites to assess gross energy savings across various participant and program characteristics, such as income level, fuel-switching technologies (e.g., heat pumps), and contractor performance indicators like Quality Assurance (QA) scores. These methods collectively support a detailed understanding of program impacts and performance drivers.

**Table 1-3. Study activities and purpose**

| Activity performed | Purpose |
|---|---|
| Site-level billing analysis | Site-level billing analysis is used to quantify realization rate by measure. |
| Advanced Metering Infrastructure (AMI) analysis | AMI analysis is used in determining peak demand energy savings |
| Segmentation of sites from billing analysis | Quantify gross energy savings by household, program or contractor characteristics such as income classification, fuel-switching measures (e.g., heat pumps) and contractor Quality Assurance (QA) scores. |
| Engineering analysis | Engineering analysis is used to assess the adherence of the NYHEP portal savings estimation methodology to the New York Technical reference manual (TRM) Version 10.[a] |
| Program staff interviews | Identify base-level understanding of program process, knowledge gaps for the program, and inflation reduction act impacts. |
| Participant survey | Answer process questions relating to how participants entered the program, their satisfaction with various parts of participation, and perceived impacts of participation |
| Contractor survey | Answer process questions relating to contractors' experience with the program, customer education, barriers to measure uptake, and trainings |
| Delivered fuel analysis | Quantify energy savings resulting from participation of households that use propane, oil, or wood. |
| Add-on participant survey | Provide additional insights into beneficial snapback resulting from program participation |

a https://dps.ny.gov/technical-resource-manual-version-10-filed-december-30-2022-effective-january-1-2023

# 2 Results and Findings

This section provides detailed findings from the electric and natural gas billing analysis.

## 2.1 Detailed billing analysis findings

This section presents the NYSERDA EmPower+ billing analysis findings. The evaluation contractor team used site-level billing analysis as part of the UMP Chapter 8 methods[5] to calculate electric and gas savings impact. The billing analysis results focus on three key metrics:

1. Evaluated savings: the weather-normalized average annual energy savings per household.
2. Realization rates (RR): The ratio of evaluated savings to estimated savings, indicating the accuracy of the program's estimated savings.
3. Savings as a percentage of baseline usage: evaluated energy savings per participant divided by the average annual energy consumption per participant during the pre-period (i.e., before program participation).

The billing analysis estimates program impact by comparing participant energy usage to that of a matched comparison group.[6] This method assumes that the comparison group reflects how the participant group would have behaved without the program. If the control group is not a good match to participants, there might be differences in how the control and participant would have behaved in absence of the program. These differences can affect the evaluated savings results, even though they are not directly visible in the calculations.

### 2.1.1 Overall evaluated savings

**Average annual per-customer evaluated savings were 8.1 MMBtu for gas and 127 kWh for electricity, based on a sample of 1,560 gas sites and 4,182 electric sites.** Table 2-1 presents the overall gas (MMBtu) and electric (kWh) evaluated savings achieved through the program. Evaluated gas savings account for 8% of baseline consumption, while evaluated electric savings are relatively smaller, only 1% of baseline usage (see Row E).

[5] Agnew, K.; Goldberg, M. (2017). *Chapter 8: Whole-Building Retrofit with Consumption Data Analysis Evaluation Protocol, The Uniform Methods Project: Methods for Determining Energy-Efficiency Savings for Specific Measures.* National Renewable Energy Laboratory. NREL/SR-7A40-68564. http://www.nrel.gov/docs/fy17osti/68564.pdf.

[6] See Section 3.4.4, Matched comparison group selection, for a description of the matched comparison group methodology.

**Table 2-1. Overall savings summary**

| Parameter | Gas (MMbtu) | | | Electricity (kWh) | | |
|---|---|---|---|---|---|---|
| | Overall | LI | MI | Overall | LI | MI |
| Sample count | 1,560 | 1,517 | 43 | 4,182 | 4,142 | 40 |
| A. Estimated savings (kWh or MMBtu/yr) | 40 | 40 | 17 | 707 | 712 | 223 |
| B. Average evaluated baseline usage (kWh or MMBtu/yr) | 106.6 | 106.6 | 95.0 | 8,590 | 8,596 | 7,987 |
| C. Average estimated % savings (kWh or MMBtu/yr; A/B) | 38% | 38% | 18% | 8% | 8% | 3% |
| D. Average evaluated savings (kWh or MMBtu/yr) | 8.1 | 8.1 | 10.2 | 127 | 128 | 68 |
| E. Average evaluated % savings (kWh or MMBtu/yr; D/B) | 8% ±1% | 8% ±1% | 11% ±3% | 1% ±1% | 1% ±1% | 1% ±5% |
| F. Realization rate (D/A) | 20% ±16% | 20% ±17% | 58% ±46% | 18% ±60% | 18% ±60% | 31% ±998% |

**The realization rates are 20% for gas savings and 18% for electric savings (row F).** Realization rates show how evaluated savings compare to estimated savings and are influenced by factors like end-uses present in the home, usage patterns, and existing efficiencies of installed measures. They are an imperfect measure of a program's impact. Various factors can cause differences between actual savings and estimated savings, including assumptions made during the estimation process. If both estimated and evaluated savings perfectly matched actual savings, realization rates would be 1. Rates below 1 usually result from differences between assumed inputs and what actually occurred. For a full discussion on the limitations of realization rates, see Appendix B, Realization rate limitations.

**The low realization rate for gas may be attributable to overestimates of estimated savings, while the low realization rate for electricity may be due to the challenge associated with detecting the low amount of electric savings with billing analysis.** As discussed further in Section 2.1.2, a contributing factor to the low gas realization rate is the high estimated savings, which represent 38% of baseline usage, an unrealistic estimate for the EmPower+ program. The low electric realization rate may be attributed to the fact that estimated savings were only 8% of baseline electric usage (compared to 38% for gas), which is difficult to detect through monthly billing analysis due to natural consumption variability from occupancy changes or changes in usage patterns, which can mask small savings.

### 2.1.2 Savings historical trends

This section provides additional context and comparisons with prior evaluations. The focus is on Low-Income (LI) participants. Approximately 97% of projects in the program population were classified as LI, and the remaining projects were identified as serving Moderate-Income (MI) participants.[7]

**Historically high estimated gas savings are causing the low realization rates for PY2023.** Historically, LI gas evaluated savings have been between 8 and 12 MMBtu per year (Table 2-2), and this study is in line with those results. However, the estimated savings increased by nearly 54%, rising from 26 to 40 MMBtu when comparing the 2017 – Q1 2019 evaluation to the current program.

**Table 2-2. Historical LI gas evaluated savings, estimated savings and realization rates**

| Participation year | Average evaluated savings (MMBtu/yr) | Average estimated savings (MMBtu/yr) | Realization rate |
|---|---|---|---|
| 2007–2008 | 10 | 14 | 71% |
| 2010–2011 | 2 | 4 | 50% |
| 2012–2016 | 12 | 27 | 44% |
| 2016–mid 2017 | 8 | 26 | 31% |
| 2017–Q1 2019 | 9 | 26 | 35% |
| **2023 (Current study)** | **8** | **40** | **20%** |

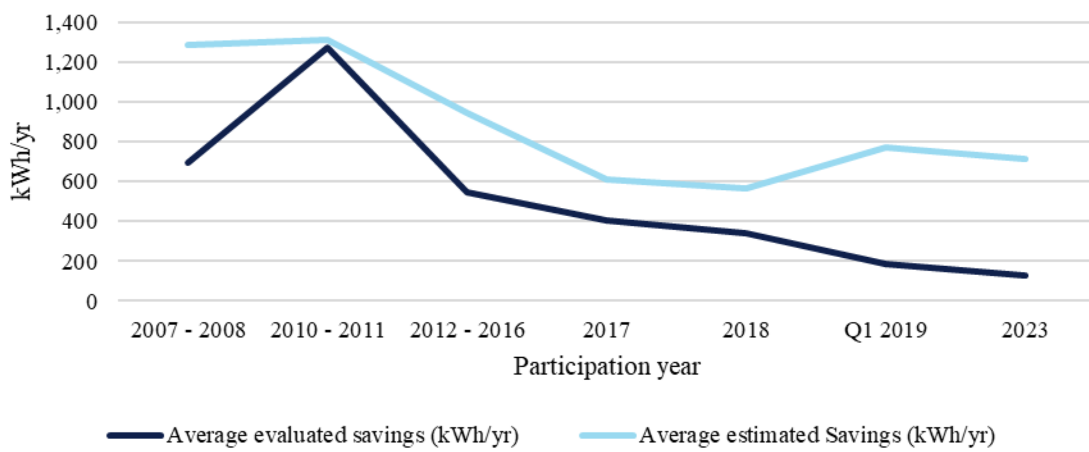
**The increase in estimated savings is driven partially by a change in the methodology used to calculate tracked savings.** In 2023, the method for calculating tracked savings was updated in 2023 to use the Technical Reference Manual (TRM), replacing the proprietary tool, EmPCalc, NYSERDA previously used. The new approach to estimating savings does not account for interactions between measures installed within the same household, which can lead to inflated

[7] After data cleaning and applying the attrition criteria, only 40 MI projects are in the electric sample and 43 in the gas sample, which represent ~1% and ~3% of the electric and gas sample respectively.

estimated savings. Additionally, EmPCalc was calibrated to evaluation results of previous evaluations. While this change may explain some of the increase, it does not fully account for the gap, suggesting that other factors may also be contributing.

**For electric savings, the evaluated and estimated savings, as well as the realization rates, are consistent with patterns seen in earlier studies, but savings are decreasing over time.** Figure 2-1 shows the historical trend for electric evaluated and estimated savings that depicts a steady decline in evaluated savings (dark blue line) since 2012, while the estimated savings have remained relatively stable since 2017.

**Figure 2-1. LI electric evaluated and estimated savings historical trend**




**While evaluated electric savings have decreased over time, the rate of the decrease has slowed.** Table 2-3 shows the percent change per year in average evaluated electric savings. For this evaluation, the average percent decrease per year has decreased to single digits, compared to double digit rate of reductions found in the previous studies.

**Table 2-3. LI electric evaluated and estimated savings, RR, and percent change per year**

| Participation year | Average evaluated savings (kWh/yr) | Average estimated Savings (kWh/yr) | Realization rate (%) | Years between measurements | Percent change per year |
|---|---|---|---|---|---|
| 2007–2008 | 694 | 1,285 | 54% | NA | NA |
| 2010–2011 | 1,275 | 1,314 | 97% | ~2 | 42% |
| 2012–2016 | 547 | 943 | 58% | ~3 | -19% |
| 2017 | 406 | 614 | 66% | 1 | -26% |
| 2018 | 339 | 564 | 60% | 1 | -17% |
| Q1 2019 | 187 | 769 | 24% | 1 | -45% |
| 2023 | 128 | 712 | 18% | 4 | -8% |

### 2.1.3 Measure-level evaluated savings

This section presents results by measure group. The evaluation contractor team identified 31 measures in the tracking data, but it is not possible to isolate savings with confidence for many of them due to limitations further discussed in Section 3.1.6, Measure-level savings. To address this, measures were grouped into categories where separate savings estimates could be developed with reliable confidence.

Some measures included in this section do not have statistically significant savings estimates. These were retained because they represent a large portion of the program activity and excluding them would risk blending distinct effects into a single, unclear result. Additionally, measures such as equipment repair, which may show negative effects, were kept separate to avoid distorting the overall findings.[8]

**The analysis reveals that the largest share of evaluated gas savings is attributed to Air Sealing and Insulation and Domestic Hot Water (DHW) replacement measures.** Table 2-4 shows the gas evaluated savings by measure group.[9] On average, these measures yielded significant savings per home of 6 MMBtu and 4.8 MMBtu, respectively.

---

[8] For example, if an equipment repair measure involved fixing a broken furnace, the home would consume more energy after participation because the furnace went from broken to operational.

[9] Please refer to Appendix A for a comprehensive list of measure-to-measure group mappings.

No other measures yielded statistically significant results, and therefore, their evaluated savings should be interpreted with caution. Equipment repair measures showed negative savings, although the result was not statistically significant. This outcome may reflect beneficial snapback effects, where repaired equipment, such as a previously non-functioning furnace, returns to normal energy consumption levels, thereby increasing energy usage post-repair.

**Table 2-4. Gas measure-level evaluated savings summary**

| Measure group | Number of sites with measure (N = 32,049) | Average evaluated savings per home that has the measure (MMBtu/yr) | Total evaluated savings (MMBtu/yr) | Average estimated savings per home (MMBtu/yr) | Realization rate (evaluated savings/estimated savings) |
|---|---|---|---|---|---|
| Air sealing and/or insulation | 22,882 | 6.0[b] | 137,755 | 31 | 19% [b] |
| DHW replacement (gas)[c] | 2,406 | 4.8 [b] | 11,545 | 11 | 44% [b] |
| Thermostat | 6,876 | 1.9 | 13,065 | 2.6 | 74% |
| Heating replacement[1] | 4,346 | 1.9 | 8,381 | 19.1 | 10% |
| Other (Pipe insulation, aerators, lighting) | 7,745 | 0.2 | 1,564 | 1.4 | 14% |
| Equipment repair | 12,145 | -0.3 | -3,453 | 1.1 | -26% |

[a] Excludes projects with heat pump installations.

[b] Indicates the result is significant at the 90% confidence level

[c] Measure group includes only sites where gas water heating is the identified fuel type.

**Electric evaluated savings were mostly driven by Refrigerator/Freezer Replacement and DHW Replacement measures.** Table 2-5 summarizes the evaluated electric savings by measure group. On average, these measures yielded significant savings per home of 218 kWh and 904 kWh, respectively. No other measure type had statistically significant results, and Equipment repair showed negative but non-significant savings, which may suggest snapback effects. Both patterns were consistent with the gas measure-level findings. Finally, Thermostats had negative electric evaluated savings, which is in line with evaluations of other programs, which have found smart thermostats can lead to an increase in electric consumption.

**Table 2-5. Electric measure-level savings summary**

| Measure group | Number of sites with measure (N = 33,099) | Average evaluated savings per home that has the measure (kWh/yr) | Total evaluated savings (kWh/yr) | Average tracked savings per home | Realization rate |
|---|---|---|---|---|---|
| Refrigerator/Freezer Replacement | 12,012 | 218 | 2,620,386 | 465 | 47%[a] |
| DHW Replacement (electric)[b] | 727 | 904 | 657,203 | 2,107 | 43%[a] |
| Other (Pipe insulation, aerators, lighting) | 19,803 | 23 | 457,509 | 305 | 8% |
| Air sealing and/or insulation | 22,882 | 11 | 248,878 | 357 | 3% |
| Equipment repair | 12,145 | -3 | -31,235 | 14 | -19% |
| Thermostat | 6,876 | -24 | -83,220 | 25 | -48% |

a Indicates the result is significant at the 90% confidence level

b Measure group includes only sites where electric water heating is the identified fuel type.

### 2.1.4 Evaluated savings by utility

This section examines how evaluated electric and gas savings vary across participating utilities. Results are for informational purposes only.

**Gas evaluated savings generally ranged between 6% and 8% of baseline usage.** New York State Electric and Gas (NYSEG) achieved the highest savings ratio, evaluated gas savings relative to baseline usage, of approximately 12% (Table 2-6). In contrast, Consolidated Edison reported the lowest savings ratio at less than 1%, National Grid, which accounted for over half of the sample (823 sites), showed a savings ratio of 8% and contributed significantly to the overall average. The current study shows a slight decline in average performance from the prior study. which reported savings ratios of 9.0% for low-income (LI) and 11.7% for moderate-income (MI) participants.

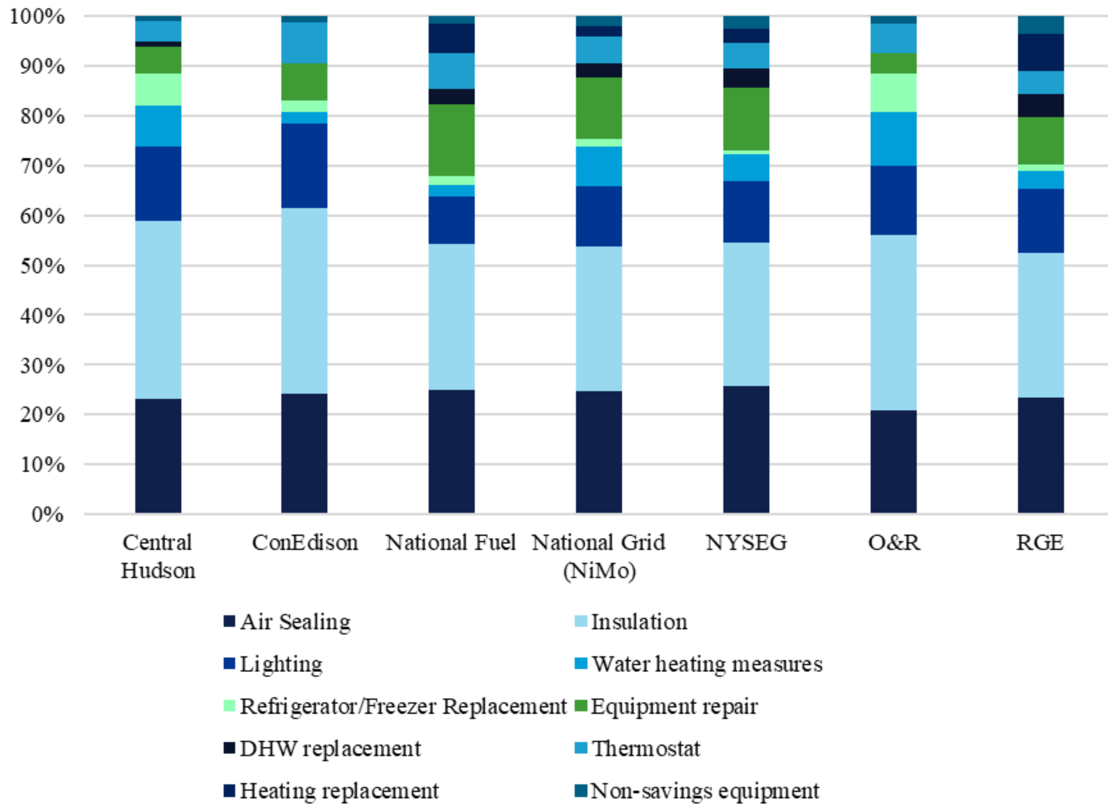
**On average, the evaluation contractor estimates customers saved $134 on their annual gas bills.** Although National Grid showed lower evaluated gas savings than NYSEG, its customers experienced higher average cost savings due to NYSEG’s lower gas rates. Consolidated Edison had the highest cost per MMBtu, which resulted in relatively large cost savings despite modest evaluated savings.

**Table 2-6. Average gas evaluated savings by utility**

| Utility | Number of sites | Average evaluated savings MMBtu/yr | Average baseline usage MMBtu/yr | Annual cost savings ($) | Savings ratio = evaluated savings / baseline usage |
|---|---|---|---|---|---|
| Central Hudson | 20 | 3.6 | 97.7 | 60 | 4% ± 8% |
| Consolidated Edison | 18 | 0.9 | 96.8 | 24 | 1% ± 16% |
| National Fuel Gas Distribution | 253 | 6.7 | 120 | 117 | 6% ± 3% |
| National Grid | 823 | 8.8 | 106.2 | 153 | 8% ± 2% |
| New York State Electric and Gas | 90 | 10.2 | 85.3 | 142 | 12% ± 6% |
| Orange and Rockland | 104 | 9 | 130.4 | 166 | 6% ± 5% |
| Rochester Gas and Electric | 252 | 7 | 87.9 | 87 | 8% ± 4% |
| **Grand Total** | **1560** | **8.1** | **105.7** | **134** | **8%** ± 1% |
| Prior Study (2017-2019) | 2543 LI | 9.3 LI | 103.7 LI | | 8.98% LI |
| | 1789 MI | 11.4 MI | 97.1 MI | | 11.73% MI |

**The distribution of the number of measures across utilities does not suggest any meaningful differences in composition that could explain the variation in evaluated savings (Figure** 2-2**).** Notably, approximately 60% of all measures correspond to air sealing or insulation, and this proportion is consistent across all utilities.

**Figure 2-2. Distribution of measures by utility – Gas sites**




**Electric evaluated savings range from 2 % to 7% of baseline usage, excluding Orange & Rockland (O&R).** Rochester Gas and Electric and Central Hudson reported the highest evaluated savings ratios, both at approximately 7%. O&R was the only utility to report negative savings, indicating the program led to an increase in consumption. As shown in Figure 2-2, O&R also has the highest proportion of Refrigerator/Freezer Replacement measures among the utilities. It is possible that these new appliances were not replacing functioning units, but rather replacing broken units, thereby increasing overall energy consumption, which would explain the utility's negative average savings.

**When compared to the prior study (2017–2019), the current results show a decline in electric savings performance.** The prior study reported savings ratios of 4.66% for low-income participants and 2.94% for moderate-income participants, both higher than the current overall average. However, as shown in Figure 2-1, evaluated electric savings have had a negative trend over the years.

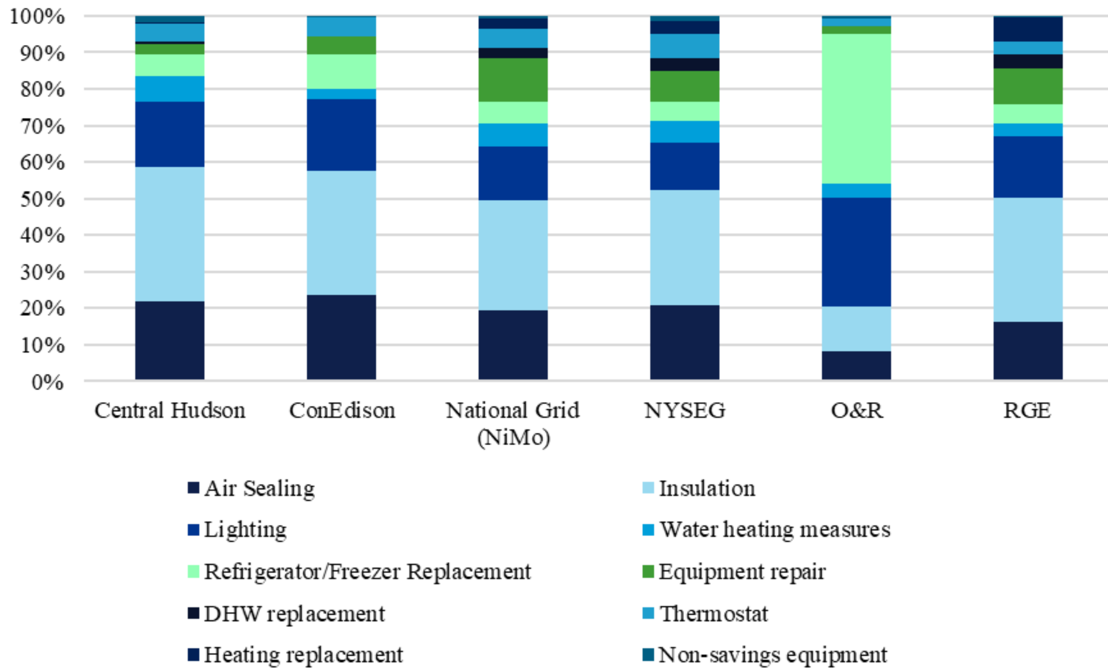
**On average, customers saved an estimated $25 on their annual electric bills as shown in Table** 2-7**.** While NYSEG showed higher evaluated savings than Consolidated Edison, the latter's higher cost per kWh led to larger cost savings. This mirrors the behavior for gas savings, where higher unit costs translated into larger financial impacts despite lower energy savings. In fact, NYSEG's customers achieved nearly half the cost savings of those served by Consolidated Edison.

**Table 2-7. Average electric evaluated savings by utility**

| Utility | Number of sites | Average evaluated savings kWh/yr | Average baseline usage kWh/yr | Cost savings ($/yr) | Savings ratio = Evaluated savings / baseline usage |
|---|---|---|---|---|---|
| Central Hudson | 111 | 601 | 8,348 | 151 | 7% ± 6% |
| Consolidated Edison | 514 | 201 | 6,148 | 63 | 3% ± 3% |
| National Grid | 2,827 | 146 | 8,378 | 25 | 2% ± 1% |
| New York State Electric and Gas | 63 | 230 | 9,430 | 36 | 2% ± 7% |
| Orange and Rockland | 559 | -176 | 11,701 | -41 | -2% ± 2% |
| Rochester Gas and Electric | 68 | 593 | 8,094 | 96 | 7% ± 8% |
| **Grand Total** | **4,182** | **127** | **8,590** | **25** | **1%** ±1% |
| Prior Study (2017-2019) | 5400 LI | 357 LI | 7669 LI | | 4.66% LI |
| | 1407 MI | 238 MI | 8079 MI | | 2.94% MI |

**O&R had a disproportionate number of refrigerator and freezer replacements. The** distribution of measures in Figure 2-3 shows that around 50% of measures across utilities correspond to Air Sealing, Insulation, or Lighting. The only exception is O&R, where New Refrigerator/Freezer Replacement is the largest portion of the measures.

**Figure 2-3. Distribution of measures by utility – Electric sites**




### 2.1.5 Fuel-switching results: Heat pump savings

This section presents findings from the 2023 participant sample who installed heat pumps. The number of sites with heat pump measures that also had enough billing data for analysis was relatively small: 26 electric sites and 19 gas sites.

**Heat pump installations led to a 40% increase in electricity usage compared to baseline consumption, alongside a 52% reduction in gas usage relative to gas baseline levels (Table** 2-8**).** These shifts reflect the expected transition in energy sources associated with heat pump adoption. Looking only at the electric analysis, heat pumps installations lead to an increase in utility bills of $635/yr, while looking only at gas analysis leads to a decrease in utility bills of $910/yr.[10]

[10] For the methodology to calculate utility bill savings, see Section 3.2.1.4

**Table 2-8. Evaluated savings for sites with heat pump installations**

| Commodity | Number of Sites | Average baseline usage | Average evaluated savings | Evaluated savings / baseline usage | Utility bill savings ($/yr) |
|---|---|---|---|---|---|
| Electric | 26 | 30.5 MMBtu/yr<br>8,935 kWh/yr | -12.1 MMBtu/yr<br>-3,555 kWh/yr | -40% | ($635) |
| Gas | 19 | 99.8 MMbtu/yr | 52.3 MMBTU/yr | 52% | $910 |

**For heat pump sites that had enough electric and gas billing data to look at utility bill impacts as a whole, the average net savings for participation was $999/year.** Only six sites in the sample had both gas and electric data available, allowing for a combined energy analysis of heat pump measures. Table 2-9 presents the aggregated energy savings for these sites, with electricity and gas consumption converted to a common unit (MMBtu). The results indicate total annual savings of 77.8 MMBtu, representing a 47% reduction relative to baseline energy usage. Overall, the homes savings $999/yr. This highlights the substantial impact of heat pump installations on overall energy consumption when both fuel types are considered, though these results may not be representative of broader trends due to the small sample size.

**Table 2-9. Heat pump evaluated savings for sites with both electric and gas data**

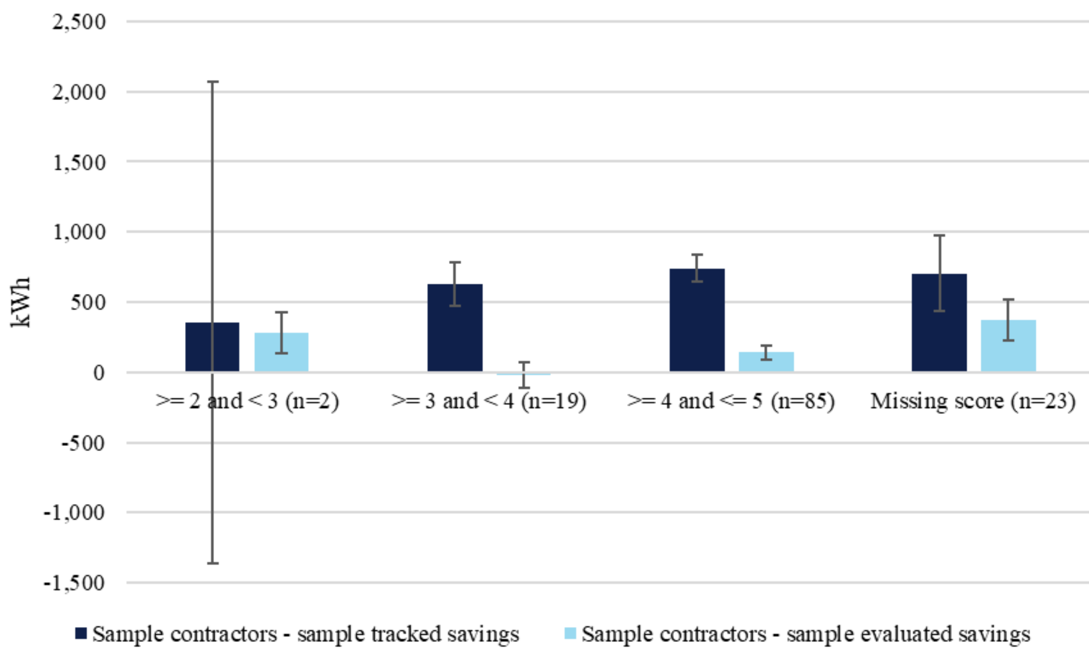
| Commodity | Number of Sites | Average baseline usage (MMBtu/yr) | Average evaluated savings (MMBtu/yr) | Evaluated savings / baseline usage | Utility bill savings ($/yr) |
|---|---|---|---|---|---|
| Electric | 6 | 35.1 MMBtu/yr<br>10,286 kWh/yr | -12.3 MMBtu/yr<br>-3,615 kWh/yr | -35% | ($701) |
| Gas | 6 | 129.3 MMBtu/yr | 90.2 MmBtu/yr | 70% | $1701 |
| **Electric and Gas** | **6** | **164.4 MMBtu/yr** | **77.8 MMBtu/yr** | **47%** | **$999** |

## 2.1.6 Evaluated savings by contractor QA Score

**Since NYSERDA typically removes contractors from the program who receive poor quality scores, there was no correlation between savings outcomes and contractor quality.** Figure 2-4 and Figure 2-5 show the analysis of gas and electric savings by contractor Quality Assurance

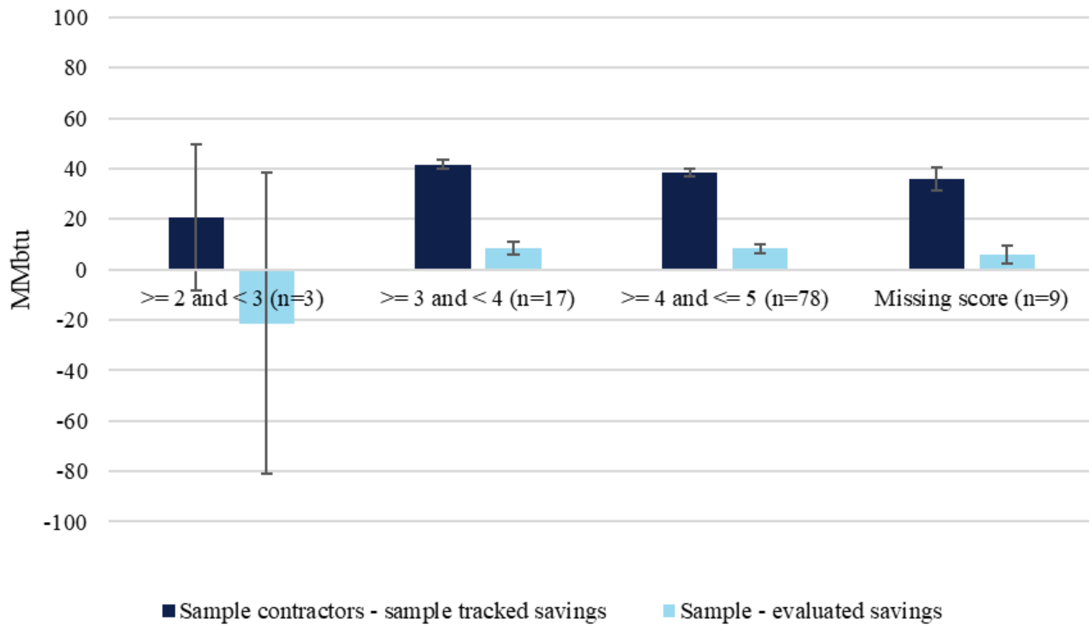
(QA) score.[11] The results show no clear correlation between savings outcomes and contractor performance ratings. Since NYSERDA typically removes contractors with scores below 2 from the program due to performance concerns, the bins are from 2 upward. The majority of contractors score 3 or higher. In fact, 65% of contractors receive scores between 4 and 5, indicating generally high performance across the evaluated group.

**Figure 2-4. Electric savings by contractor QA Score (average kWh savings by contractor score bin with 90% CI)**




[11] QA scores range from 0 to 5 with 0 representing poor quality and 5 representing high quality. Contractor's QA scores are determined based on quality assurance visits and review by NYSERDA.

**Figure 2-5. Gas savings by contractor QA score (average MMBtu savings by contractor score bin with 90% CI)**




**There was no significant difference in savings between contractors.** Additionally, the distribution of projects was relatively scattered, with no single contractor representing a disproportionately large share of the sample.

### 2.1.7 Evaluated savings by DAC geographic status

**There was no statistically significant difference between evaluated savings for DAC and non-DAC participant.** Using NYSERDA's current criteria for identifying Disadvantaged Communities (DAC), the evaluation contractor team conducted a comparative analysis of average savings by DAC geographic status. This included comparisons across income classifications (LI, MI) and regions (Downstate, Long Island, Upstate), as shown in Table 2-10. Approximately 97% of gas projects and 99% of electric projects are classified as Low-Income, and roughly 88% of all projects are in Upstate New York. The observed evaluated savings per household yielded no statistically significant differences between DAC and non-DAC participants, either overall or within the defined subgroups.

**Table 2-10. Average Evaluated savings by DAC geographic status**

| | Variable | Category | Non-geographic DAC | | | Geographic DAC | | | Statistical difference test | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | Count | Avg evaluated savings | SE | Count | Avg evaluated savings | SE | T-statistic | P value (*) |
| **Electric savings (kWh)** | **Overall** | **Overall** | **2,774** | **126.84** | **58.49** | **1408** | **128.25** | **75.97** | 0.014 | 0.989 |
| | Income Classification | LI | 2,747 | 123.54 | 58.87 | 1394 | 137.16 | 76.45 | 0.138 | 0.891 |
| | | MI | 26 | 513.98 | 518.37 | 14 | -759.30 | 633.13 | -1.505 | 0.141 |
| | Region | Downstate | 24 | 133.48 | 555.35 | 52 | 1122.55 | 377.94 | 1.471 | 0.145 |
| | | Long Island | 311 | 109.86 | 140.14 | 131 | 27.35 | 153.73 | -0.347 | 0.729 |
| | | Upstate | 2439 | 128.94 | 63.87 | 1225 | 96.83 | 84.07 | -0.297 | 0.766 |
| **Gas savings (MMbtu)** | **Overall** | **Overall** | **895** | **7.55** | **1.11** | **665** | **8.92** | **1.22** | **0.825** | **0.409** |
| | Income Classification | LI | 862 | 7.52 | 1.14 | 655 | 8.80 | 1.24 | 0.750 | 0.454 |
| | | MI | 33 | 8.14 | 3.28 | 10 | 16.85 | 5.75 | 1.289 | 0.205 |
| | Region | Downstate | 8 | -12.82 | 12.10 | 12 | 8.01 | 12.32 | 1.152 | 0.264 |
| | | Long Island | 128 | 6.01 | 2.61 | 32 | 14.05 | 5.35 | 1.370 | 0.173 |
| | | Upstate | 759 | 8.02 | 1.22 | 621 | 8.67 | 1.26 | 0.368 | 0.713 |

(*) P-values less than 0.10 are statistically significant at the 90% confidence

### 2.1.8 Cost of energy saved

The study included an evaluation of the cost of energy saved of the different measure groups by calculating the dollars spent per unit of evaluated energy savings, measured in $/MMBtu for gas and $/kWh for electric, across both sample and population data, shown in Table 2-11. The overall $ per kWh and $ per MMBtu values include all measures with associated costs, even if those measures did not have tracked savings. In contrast, the measure group savings only account for measures that have tracked savings.[12]

Cost per energy saved is presented for all evaluated measures. Statistically significant savings estimates are highlighted, while non-significant values are included for transparency but should be interpreted with caution, as they may reflect uncertainty in the savings estimate rather than true cost-effectiveness.

**For gas measures, on average, the program spent $67.1 per MMBtu saved in the sample and $55.1 per MMBtu saved in the population.** Among the gas measure groups, Air Sealing + Insulation and DHW replacement had similar cost per energy saved, with a cost of $47 and $46 per MMBtu saved in the sample, respectively. However, the population values differed, with the cost per MMBtu saved for Air Sealing + Insulation being $14 higher than that for DHW replacement. As mentioned in Section 2.1.3, Measure-level evaluated savings, equipment repair measures tend to increase energy consumption, resulting in negative cost of saved energy.

**For electric measures, the overall cost-effectiveness was $3.6 per kWh in the sample and $3.5 per kWh in the population.** DHW replacement was the most cost-effective electric measure group in the population, with $0.2 per kWh saved, followed by Refrigerator/Freezer replacement, with $0.8 per kWh saved in the population. Thermostats and Equipment Repair have negative dollars per kWh due to increases in electricity consumption from those measures.

[12] Costs reflect only those directly attributed to individual measures in the program tracking data. Overhead, administrative, or other indirect costs are not included.

**Table 2-11. Cost per unit saved by measure group**

| GAS | | |
|---|---|---|
| **Measure group** | **$/ MMBtu** | |
| | **Sample** | **Population** |
| Air Sealing + Insulation | $ 47.0[a] | $ 54.2[a] |
| DHW replacement | $ 46.0[a] | $ 39.8[a] |
| Equipment repair | $ (209.6) | $ (210.4) |
| Heating replacement | $ 110.5 | $ 102.3 |
| Thermostat | $ 8.0 | $ 8.0 |
| WH + Other | $ 79.5 | $ 160.9 |
| **Overall (*)** | **$ 67.1** | **$ 55.1** |
| **ELECTRIC** | | |
| **Measure group** | **$/ kWh** | |
| | **Sample** | **Population** |
| Air Sealing + Insulation | $ 33.5 | $ 29.6 |
| DHW replacement | $ 0.8[a] | $ 0.2[a] |
| Equipment repair | $ (24.0) | $ (23.0) |
| Refrigerator/Freezer replacement | $ 0.4[a] | $ 0.5[a] |
| Thermostat | $ (1.0) | $ (1.3) |
| WH + Lighting | $ 0.5 | $ 0.5 |
| **Overall (*)** | **$ 3.6** | **$ 3.5** |

a Indicates the result is significant at the 90% confidence level.

(*) The overall $ per kWh and $ per MMBtu values include all measures with associated costs, even if those measures did not have tracked savings, such as assessment fees. In contrast, the measure group savings only account for measures that have tracked savings.

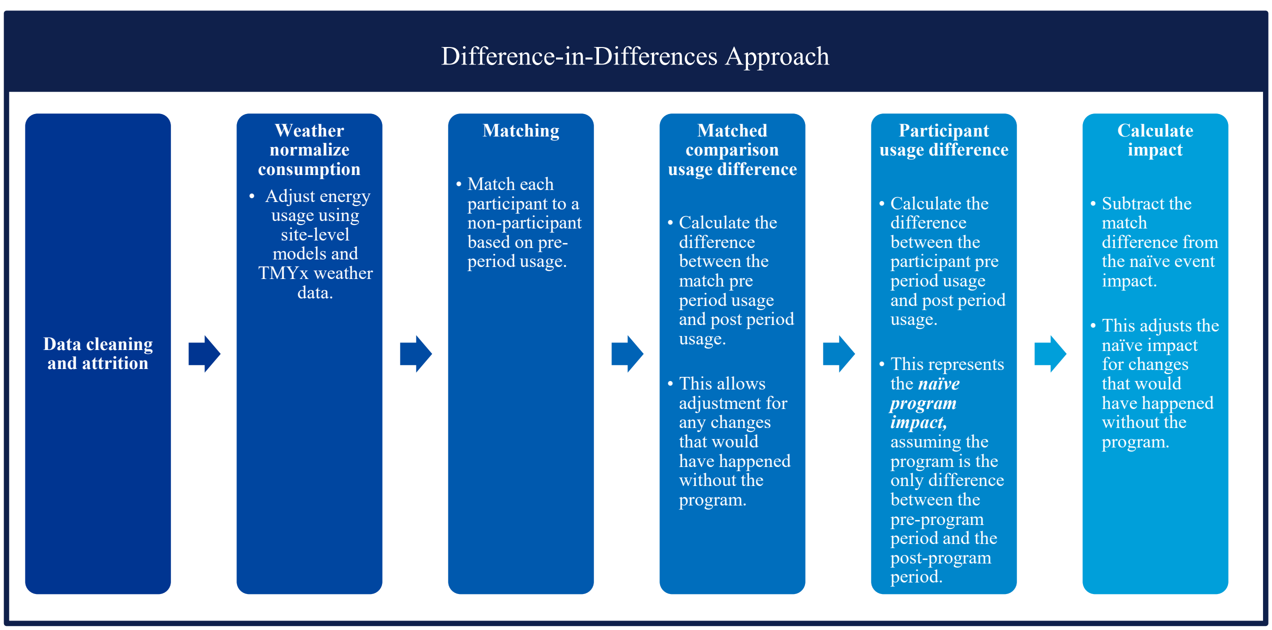
# 3 Methods

## 3.1 Detailed billing analysis methods

The impact evaluation used a difference-in-differences methodology to estimate annual program impacts. Difference-in-differences methods calculate the difference in annual usage before versus after participation, and subtracts the corresponding change for a comparison group of nonparticipants from the participants' change. The comparison group represents the change the participants would have had absent the program, so that the difference of differences is the program effect.

The team begins with data cleaning, followed by weather normalization, then matches participants to control sites, and finally calculates savings using a difference-in-differences approach. Figure 3-1 presents a conceptual overview of what each step does and how they are combined to create the most accurate impact estimates.

**Figure 3-1. Difference-in-differences approach**




### 3.1.1 Data sources

The evaluation contractor team used a comprehensive set of data sources that enabled robust analysis of energy savings, participant experience, and program performance. These included program tracking data provided by NYSERDA, monthly utility billing data, weather data from NOAA, energy cost data based on utility tariffs, and Advanced Metering Infrastructure (AMI) data from Con Edison. Together, these datasets allowed for detailed site-level billing analysis, weather normalization, peak demand estimation, and segmentation of households by key characteristics such as income level, fuel type, and contractor performance.

#### *3.1.1.1 Tracking data*

The evaluation contractor team received tracking data from NYSERDA on the EmPower+ program, which contained building and project measure details for 50,833 projects. This analysis used the following key parameter fields: utility providers, address, zip code, project installation end date, and main heating fuel, main water heating fuel, measure category, kWh tracked savings and MMBtu tracked savings. NYSERDA shared complementary data containing contractors' quality scores.

Although EmPower+ primarily serves single-family homes, some small multifamily properties are also included. During the review of project billing accounts, the evaluation contractor team identified instances where electric and heating accounts overlapped across multiple projects. This occurred for two main reasons: either a single-family home participated in the program more than once, or a multifamily building was involved. The latter presents more complexity, since analyzing multifamily properties ideally requires building-level data, and identifying which units belong to the same building can be challenging.

To accurately capture the overall change in energy consumption attributable to the program, the evaluation contractor team consolidated overlapping accounts into a single site-level record. This site-level aggregation forms the basis of the analysis. In the majority of cases, each site is associated with one electric account and one gas account.

The evaluation contractor team initiated the data request process by manually retrieving program applications for each project from NYSERDA's Salesforce system. The team verified if each program applicant had provided consent to share their billing data by providing a signature on their application. For those who provided the billing data consent, the team redacted the applications to exclude all sensitive information not related to the billing data consent and gave the applications back to NYSERDA to request data from the utilities.

#### 3.1.1.2 Monthly billing data

The evaluation contractor team received monthly gas and electricity usage data, referred to as billing data, from six utility providers.[13] The bills spanned from December 2021 through August 2024. The evaluation contractor team obtained monthly billing data from two sources: Electronic Data Interchange (EDI) and utility archive data requests. The EDI data, accessed through NYSERDA's internal connection with the utilities, includes records going back two years. Because the analysis focused on projects completed in 2023, the team supplemented the EDI data with archive requests submitted to the utilities. These requests covered accounts with signed consent forms and EmPower+ projects finalized in 2023. Upon review, the team identified overlapping records between the two sources. A comparison confirmed that the datasets were consistent, with no discrepancies observed.

#### 3.1.1.3 Weather data

The evaluation contractor team mapped each site by ZIP code to the closest National Oceanic and Atmospheric Administration (NOAA) primary station. The team then collected hourly weather data from these NOAA stations, covering the period from December 2021 through July 2024. The team also collected TMYx[14] data of vintage 2004 to 2018 for the same weather station for use in determining weather-normalized annual energy savings.

#### 3.1.1.4 Energy cost data

The evaluation contractor team conducted an assessment of residential energy bill savings in New York State by actively collecting energy cost data for both natural gas and electricity from publicly available sources. For natural gas, the team reviewed utility-specific tariff information, which includes supply, delivery, and customer charges. The team calculated a cost metric in dollars per MMBtu assuming 100 therms of usage per home to enable consistent comparisons across utilities. For electricity, the team used data from the U.S. Energy Information Administration's Form EIA-861, focusing on residential sales and revenue reported by New York State utilities. By dividing total residential revenue by total residential kilowatt-hour (kWh) sales,

---

[13] Con Edison, Central Hudson, National Fuel, National Grid, NYSEG, Orange and Rockland, and Rochester Gas and Electric.

[14] TMYx - https://climate.onebuilding.org/WMO_Region_4_North_and_Central_America/USA_United_States_of_America/index.html#IDNY_New_York-

the team derived a bundled cost in dollars per kWh, which reflects both supply and delivery charges.

#### 3.1.1.5 AMI data

The evaluation contractor team received electric AMI data for only Con Edison. Only data for participants in the matching sample, was used for peak demand savings calculations.

### 3.1.2 Data cleaning and attrition

The evaluation contractor team consolidated the six utilities' billing datasets into one standardized dataset for analysis. The team categorized the different read type codes from each utility into actual and estimated readings. To maintain continuity and accurately reflect usage over time, estimated meter reads were combined with subsequent actual reads. This is important because actual reads often correct cumulative usage across both the estimated and actual billing periods, resulting in more accurate consumption data.

While the data received was of high quality, there was some attrition. Attrition occurs when projects are removed from the analysis during data cleaning due to inadequate program tracking or utility billing data. The evaluation contractor team required a minimum of six actual bills before and after installation for gas and electric meters, with at least one actual reading in each season—winter, summer, and shoulder—within the 14-month pre- and post-installation period. These requirements ensure sufficient data coverage capturing key seasonal variations—summer, winter, and shoulder months—allowing the team to reflect typical usage patterns. While UMP Chapter 8 methods suggest longer periods when possible, six bills provide a reliable balance between statistical robustness and data availability, especially in cases where a full year of pre- and post-installation billing data is not available.

The evaluation contractor team identified multiple sites that contained zero or null estimated energy savings (kWh or MMBtu). These sites were excluded from the electric and gas savings analyses, respectively. Additionally, sites with heat pump installations were removed from the primary results and analyzed separately, as detailed in Section 2.1.5, Fuel-switching results: Heat pump savings.

When there was a variable indicating solar production, the evaluation contractor team removed those accounts. This information was not available for all utilities. A total of 4,182 sites were included in the electric billing analysis, while 1,560 sites were included in the gas billing analysis (Table 3-1).

**Table 3-1. Monthly billing analysis attrition for electricity and gas**

| Step | Data requirement | Number of sites | |
|---|---|---|---|
| | | Electric | Gas |
| A | Population | 34,361 | 33,311 |
| B | Projects with at least one good electric or gas account number | 12,001 | |
| C | Projects with confirmed consent forms | 9,364 | |
| Initial | Received billing data for 2023 Participants | 5,970 | 4,233 |
| 1 | Sites with at least 6 actual or combined bills in pre period | 5,620 | 3,313 |
| 2 | Sites with at least 1 bill each in summer, winter, and shoulder season in pre period | 5,444 | 3,255 |
| 3 | Sites with at least 6 actual or combined bills in post period | 5,131 | 2,773 |
| 4 | Sites with least 1 bill each in summer, winter, and shoulder season in post period | 5,102 | 2,758 |
| 5 | Sites with a match to a control | 4,607 | 2,272 |
| 6 | Sites where Tracked Savings were not 0 | 4,207 | 1,579 |
| 7 | Sites that didn't install a Heat Pump | 4,182 | 1,560 |
| | Final Count | 4,182 | 1,560 |

### 3.1.3 Weather-normalized energy consumption:

To accurately isolate program impacts and control for weather-related fluctuations in energy usage, the evaluation contractor team developed site-level models to estimate weather-normalized consumption before and after program implementation. These models are based on the Princeton Scorekeeping Method (PRISM®[15]) methodology, which estimates each site's energy response to outdoor temperatures and identifies the temperature thresholds (base or balance points) that trigger heating and cooling. The site level model is given by Equation 1 below.

[15] Fels, M. F. (1986). PRISM: An introduction. *Energy and Buildings, 9*(1–2), 5–18. http://www.princeton.edu/~marean/publications/prism_intro.pdf

Equation 1. Energy consumption regression model

$$E_m = \mu + \beta_H H_m(\tau_H) + \beta_c C_m(\tau_C) + \epsilon_m$$

where:

- $E_m$ is averaged daily usage during period $m$
- $\mu$ is the base load usage (intercept)
- $\tau_H$ and $\tau_C$ are base temperatures that indicate the temperature that triggers heating or cooling usage, respectively. The base temperatures are determined from a grid search.[16]
- $H_{im}(\tau_H)$ is the heating degree days (HDD) at base temperature $(\tau_H)$
- $C_{im}(\tau_C)$ is the cooling degree days (CDD) at base temperature $(\tau_C)$
- $\beta_H$ represents the increased energy use per increase in heating degree days
- $\beta_C$ represents the increased energy use per increase in cooling degree
- $\epsilon_m$ is the regression residual

The evaluation contractor team estimated the base load consumption $\hat{\mu}$, the coefficients for heating degree days $\hat{\beta}_H$ , and cooling degree days $\hat{\beta}_C$ for each site, both before and after the installation. The team then calculated normalized annual consumption (NAC) for the pre- and post-installation periods for each site and analysis time frame by combining the estimated coefficients $\widehat{\beta_H}$ and $\widehat{\beta_C}$ with the TMYx degree days $H_0$ and $C_0$ calculated at the site-specific degree-day base(s), $\hat{\tau}_H$ and $\hat{\tau}_C$. Normalized annual consumption is given by Equation 2:

Equation 2. Normalized annual consumption

$$NAC_i = (365\ x\ \hat{\mu}_i) + \hat{\beta}_H H_0 + \hat{\beta}_c C_0$$

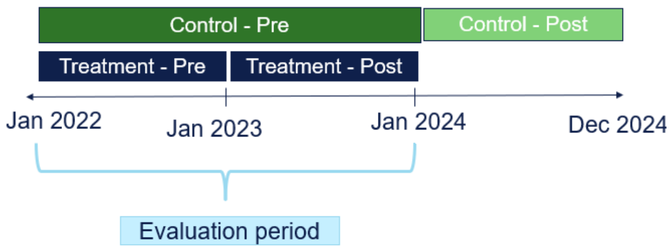
### 3.1.4 Matched comparison group selection

The second step of the impact evaluation process is selecting the matched comparison group. In this step, each participant in the program is matched with a similar comparison customer. Matched groups were created on a rolling-participation basis to account for customers enrolling

[16] Grid search involves systematically testing all possible combinations of heating and cooling balance points. By default, this process examines heating balance points ranging from 55°F to 65°F and cooling balance points from 65°F to 75°F. The optimal model is then selected based on the smallest Coefficient of Variation of the Root Mean Square Error (CVRMSE)

in the program at different times. Figure 3-2 shows an example diagram of how the comparison group (i.e., "Control") was built for January 2023 participants (i.e., "Treatment").

**Figure 3-2. Example diagram of control group selection for January 2023 participants**




All matched groups were created by finding nonparticipants within subgroups by region (Downstate, Long Island, Upstate) and main heating fuel (electricity, natural gas or other) with the most similar weather-normalized pre-period annual consumption, summer-shoulder, and winter-shoulder ratios. Using these three characteristics, the evaluation contractor team found the nonparticipants using the minimum Euclidean distance from participants. The team then selected the one nearest neighbor for each participant, with replacement. This means that each nonparticipating match could be used multiple times for different participants. The team tested different combinations of subgroups and characteristics, concluding that region and main heating fuel achieved good similarity without over-repeating controls.

Figure 3-3 and Figure 3-4 show the average monthly electric and gas consumption by treatment group before program participation. The green line is the treatment, and the blue line is the control group. In the background, the bars represent the number of sites with data for each month. Overall, the consumption patterns between the two groups are closely aligned, particularly in months with higher site counts, indicating that the treatment and control participants had similar energy usage prior to program participation. This supports the validity of the matched comparison approach used in the analysis.

**Figure 3-3. Average monthly pre-period gas consumption by treatment and control groups**

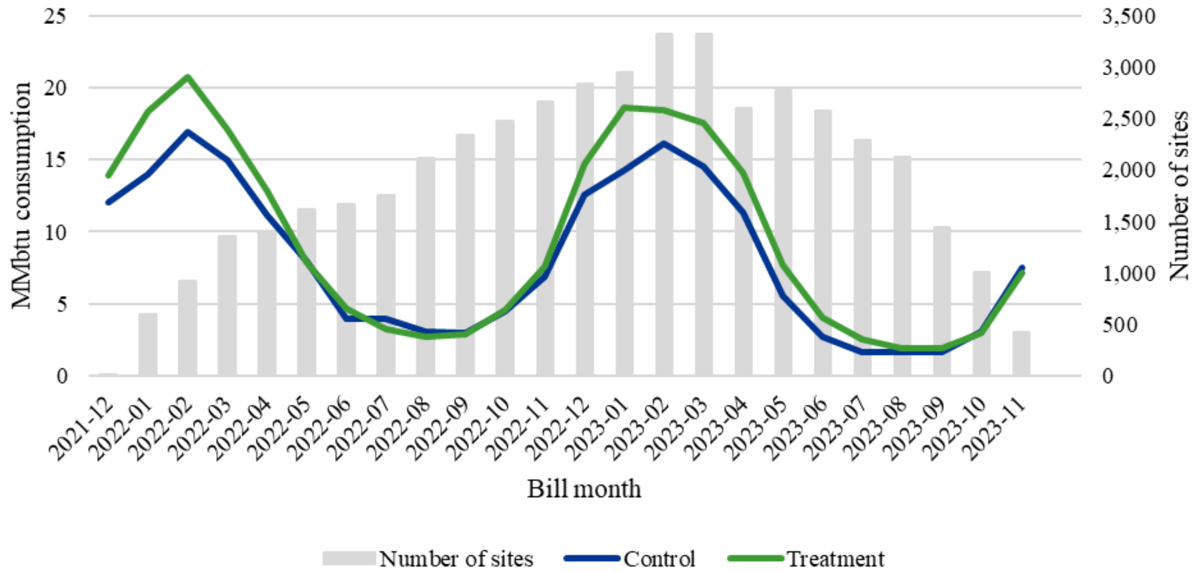
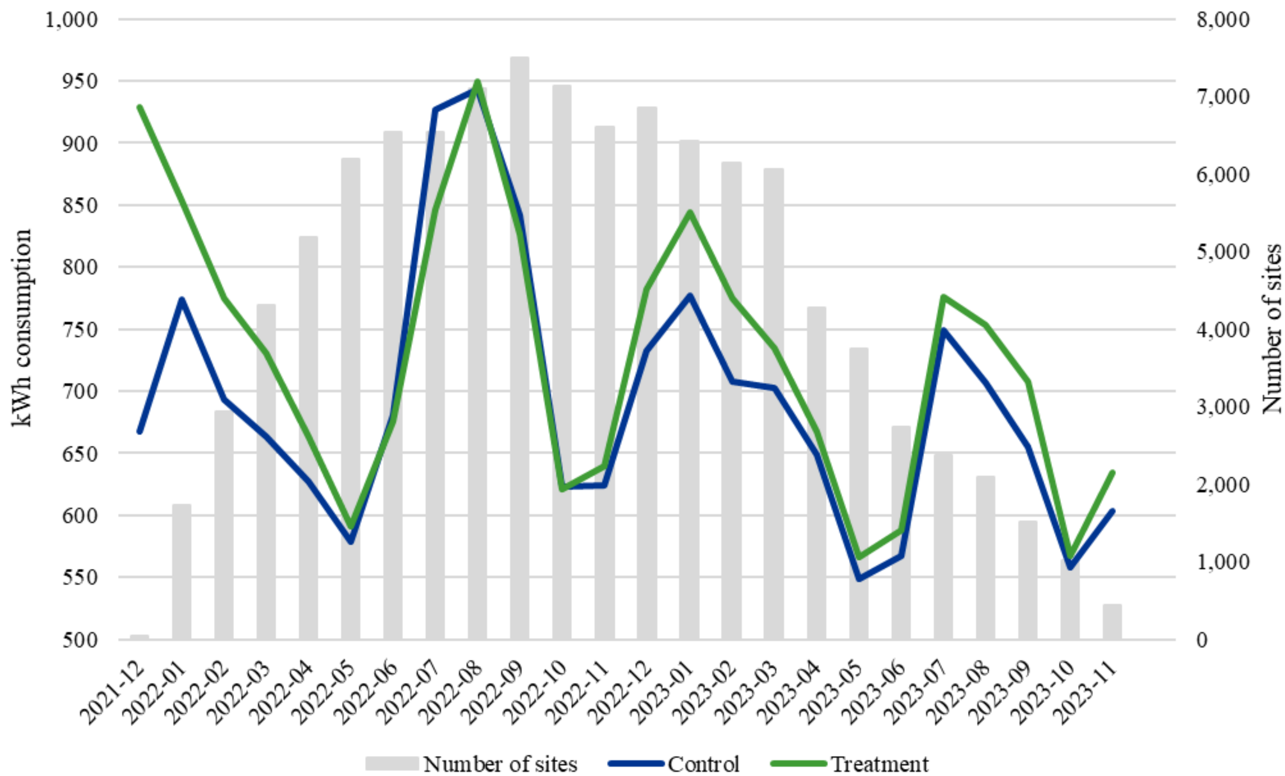



**Figure 3-4. Average monthly pre-period electric consumption by treatment and control groups**



### 3.1.5 Difference-in-differences load impact estimation

In the last step of the impact evaluation analysis, the difference-in-difference calculation, the matches found for every program participant are used to correct for differences between the pre-period and the post period. To make this adjustment, the evaluation contractor team takes the pre-vs-post usage difference for program participants and subtracts the pre-vs-post difference observed in the match group. Energy savings are therefore calculated for each participant as:

$$savings_t = (pre_t - post_t) - (pre_c - post_c)$$

Where $savings_t$ represents the weather-normalized annual energy savings for participant $t$, $pre$ and $post$ represent the weather normalized annual energy consumption in the pre and post period for participant $t$ or non-participant $c$.

Savings estimates were calculated at the annual level. We aggregated savings results across several different dimensions to understand program performance and variations. These results are shown in Section 2.1.

### 3.1.6 Measure-level savings

To better understand how individual measures contribute to overall program savings, the evaluation contractor team conducted a regression analysis using program data. The objective was to estimate the portion of evaluated savings attributable to each measure type.

Several model specifications were tested to identify the one that best explained the variation in evaluated savings. In all cases, the dependent variable was the evaluated savings derived from the billing analysis, while independent variables were drawn from program tracking data. These included indicators for whether a measure was present in a home and the tracked savings associated with each measure.

The final model selected used estimated savings per measure group as independent variables, with the following considerations to improve model performance and interpretability:

- Air Sealing and Insulation measures were grouped due to high correlation between them.

- Equipment Repair was kept separate to isolate the impact of measures that could result in negative savings (e.g., beneficial snapback, such as a repaired furnace transitioning from zero to active energy consumption).
- Smaller measures were grouped into an "Other" category to limit the number of independent variables and reduce noise.

The regression produced coefficients representing the proportion of estimated savings that were realized, effectively serving as measure-level realization rates. By multiplying the estimated savings by these coefficients, the evaluation contractor team derived the evaluated savings per measure. The regression follows the specification below:

$$evaluated\ savings_i = \sum_{mg} \beta_{mg}\ estimated\ savings_{mg,i} + \epsilon_i$$

Where:

- $evaluated\ savings_i$ corresponds to the evaluated savings of site i
- $estimated\ savings_{mg,i}$ corresponds to the estimated savings for measure group mg of site i
- $\beta_{mg}$ corresponds to the portion of estimated savings that are reflected in the evaluated savings
- $\epsilon_i$ is the error term

Finally, these results were scaled to the full program population using measure counts from the tracking data, allowing for a comprehensive view of measure-level performance across the program.

### 3.1.7 Electric peak demand energy savings

The goal of the AMI data analysis was to move beyond traditional monthly billing methods and explore the program's potential for peak demand reduction. Specifically, the analysis aimed to assess what insights AMI data could provide about the EmPower+ program's impact on summer coincident peak demand.

For this study, summer coincident peak demand was defined according to NYSERDA's criteria: the hottest non-holiday weekday between June and August, focusing on the hour ending at 5 p.m.

To isolate the effects of the EmPower+ program on peak demand, the evaluation contractor team leveraged the matched comparison group developed for the monthly billing analysis. Table 3-2 outlines the attrition steps. From the original sample of 4,182 treatment sites, only 556 sites were located within Con Edison's service territory. The final sample used for peak demand analysis includes 300 sites, where both the treatment and matched control groups had adequate data coverage to support hourly, weather-normalized comparisons.

**Table 3-2. Electric peak demand analysis attrition**

| Attrition | Number of sites |
|---|---|
| Original electric sample | 4,182 |
| ConEd in electric sample | 556 |
| Treatment with AMI data | 553 |
| Treatment and matched control with AMI data | 303 |
| **Treatment and matched control with enough pre and post AMI data** | **300** |

The evaluation contractor team estimated hourly, weather-normalized energy usage for both pre and post program periods using actual summer weekday non-holiday AMI consumption data, similar to the model used on the monthly billing data. To identify peak conditions, TMYx data was used to select the five hottest non-holiday weekdays during the summer season. Based on this selection, the team calculated the average hourly savings across the full 24-hour period.

### 3.1.8 Cost of energy saved

The cost per unit of evaluated savings was calculated using the direct measure-level costs reported in the program tracking data. These costs were annualized based on the expected measure life. Specifically, we summed the total costs for both the sample and population, divided by the average measure life and the number of sites, to estimate the annualized cost per site. This value was then divided by the evaluated annual savings to derive the cost per kWh or MMBtu saved. For measure group-level estimates, only the costs and measure lives of the measures within each group were included, along with the evaluated savings presented in Section 2.1.3, Measure-level evaluated savings.

# Appendix A Measure category mapping

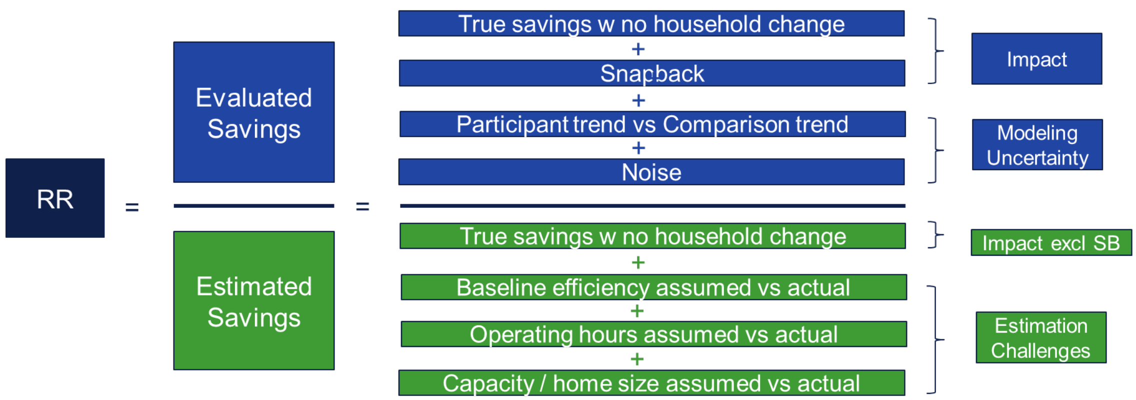
**Table A-1. Measure group classification**

| Measure group | Measure |
|---|---|
| Air Sealing and/or Insulation | Air sealing |
| | Air infiltration |
| | Attic insulation |
| | Wall insulation |
| | Other insulation |
| | Building shell |
| DHW replacement | DHW – replacement |
| Thermostat | Thermostats |
| Heating replacement | Heating replacement |
| | Primary heating and cooling |
| Equipment repair | DHW – repair |
| | Heating repair |
| Refrigerator/Freezer Replacement | Freezer replacement |
| | Refrigerator replacement |
| Other | Pipe wrapping |
| | Shower heads |
| | Water heater |
| | Other |
| | CFL/LED |
| | Appliances & lighting |

# Appendix B Realization rate limitations

Note, realization rates are not a perfect measure of program impact because they involve estimation challenges and modeling uncertainty. Figure B-1 shows the inputs to a realization rate ("RR"). Realization rates reflect only the extent to which evaluated savings, resulting from the billing analysis, match the estimated savings, from program software. An RR of 100% indicates that the evaluated savings match the estimated savings, and that the program therefore did a good job estimating savings.

**Figure B-1. Realization rate composition**




Four components make up evaluated savings. Two of those components are related to the actual impact of the program and include the following:

- **True savings with no household change** represents actual changes in energy consumption resulting from program participation without accounting for any changes in the household such as changes in the number of occupants or behavioral changes.
- **Snapback** represents changes resulting from program participation that result in increased energy consumption. For example, if the program repairs a broken furnace, gas consumption would increase. Additionally, if the program leads the participant to reduce their reliance on delivered fuels in favor of using gas, it would lead to an increase in gas consumption. Snapback reduces evaluated savings and therefore reduces the realization rate.

The remaining two components of evaluated savings are sources of modeling uncertainty:

- **Participant trend vs comparison trend** represents the attempt to account for all factors impacting energy consumption other than program participation. The approach matches participants to a comparison group of a similar population. The analysis looks at how the comparison group energy consumption changed over the same period of time examined for participants and removes that change in energy consumption from the participant observed change thus isolating only the change resulting from program participation. The uncertainty arises from the assumption that the comparison group is a good match to the participant group and had similar behaviors.

- **Nosie** refers to simple uncertainty related to the statistical analysis of the billing data.

The estimated savings is also comprised of four components. The first is the estimated impact:

- True savings with no household change for estimated savings is calculated using TRM algorithms rather than billing data. Estimated savings does not include any snapback (SB)

The remaining three components are estimation challenges in which assumptions from the TRM may not match the actual values for a site:

- Baseline efficiency assumptions may not match actual baseline efficiency values.
- Assumed operating hours may not match actual operating hours.
- Assumed capacity and home size may not match the actual values.