State of Rhode Island:
Energy Efficiency Programs Evaluation Study

Task 3 Report: Analysis of Utility Bills and Customer Experience Evaluation

March 31, 2021

Prepared by:

BrightLine Group - Lynn Roy, Nicole Wobus, Mary-Hall Johnson, McKenna Patterson

Demand Side Analytics – Jesse Smith, Steve Morris

Contents

1 Introduction

The overarching goal of the Energy Efficiency Programs Evaluation Study conducted by the BrightLine team is to understand whether there are improvements that can be made to the current evaluation measurement and verification (EM&V) process for National Grid's energy efficiency programs. The study is categorized into three tasks each with its own key objective:

- Task 1 Key Objective: Assess "Does the current Evaluation, Measurement, and Verification (EM&V) process in Rhode Island comply with national industry best practices for programs of its size and scope?"
- Task 2 Key Objective: Understand "Quantitatively, to what extent are National Grid's claimed energy savings accurate?"
- Task 3 Key Objective: Assess "Are there savings estimation and program implementation improvements that can be identified to help customers that have or are likely to experience a substantial difference in estimated gross energy savings versus installed gross energy savings and visible bill savings?"

This report outlines the Task 3 activities and findings. Activities included conducting billing analysis, customer interviews, and project documentation review to assess whether National Grid's reported energy savings align with customer utility bill savings and if not, whether there is a reasonable explanation for the variance.

There are two key sub-component objectives within Task 3. The first is to develop weather-normalized estimates of annual energy savings via pre/post billing analysis for C&I customers that participated in National Grid's gas and/or electric retrofit programs. The annual energy savings estimates drawn from the billing analysis were then compared to the gross savings estimates stored in National Grid's tracking data.[1] This comparison informed the second key objective, which was to conduct follow-up customer interviews aimed at understanding any large discrepancies between the gross savings estimates stored in the tracking data and the savings estimates derived from the billing analyses.

1.1 Summary of Findings

Overall, the BrightLine team found that the gross reported savings values for the C&I customers in the project sample are reasonable. As there are two sub-component objectives to this task, the overall findings and recommendations for each are summarized as follows:

- Objective 1: Develop Weather-Normalized Estimates of Annual Energy Savings:

[1] "Gross savings" refers to "the change in energy consumption that results directly from program-related actions taken by participants in an efficiency program, regardless of why they participated." Gross Savings and Net Savings: Principles and Guidance. Northeast Energy Efficiency Partnerships. April 2016. https://neep.org/sites/default/files/FINAL%20GS%20and%20NS%20Principles%20and%20Guidance%20Document_2016May17.pdf

- Electric analysis: Savings produced by the electric billing analysis were positively correlated with the savings estimates stored in the NGRID tracking data. However, the estimates derived via the billing analysis were generally lower, with BrightLine estimates about 80% of NGRID estimates on average. This result is not wholly unexpected, as National Grid will adjust gross savings using realization rates and in-service rates – adjustment factors that typically reduce the savings estimate based on findings from prior evaluations. (The unadjusted savings values were the point of comparison in this analysis.)
- Gas analysis: Savings estimates produced by the billing analysis were largely uncorrelated with the savings estimated stored in the National Grid tracking data. However, the sample size was also significantly smaller than the sample size for the electric billing analysis (n=34 and 298, respectively), which makes it harder to draw conclusions.
- The differences in the estimates for both the electric and gas analysis could be explained by a number of factors: business expansion resulting in greater energy use, operational changes leading to changes in hours of operation, random year-to-year variations in energy use, tracking data entry errors, faulty or outdated TRM assumptions, etc.

- Objective 2: Follow-up Customer Interviews and Project Documentation Review:
  - Customer interviews: One key theme that emerged is that facility and operational changes likely account for a significant amount of the misalignment in energy savings. Of the 34 interviewed, over two-thirds of respondents reported that changes in their facility or operations have occurred since project completion, or that they may be experiencing equipment failure or installation errors. Of those respondents, 21 described circumstances that may account for the findings of the billing analysis.
  - Project documentation review: Follow-up project documentation review was conducted for five (5) projects and we identified reasons for the variance in the billing analysis savings and National Grid's tracking data for four out of the five projects reviewed. No further project reviews or customer interviews were recommended.

BrightLine's overall assessment is that National Grid's reported electric energy savings are reasonable and that reasons for variances between National Grid's reported value and the savings calculated as part of the billing analysis are explainable. The remaining sections of this report outline the methods, activities and detailed findings for each objective of Task 3.

2 Objective 1: Develop Weather-Normalized Estimates of Annual Energy Savings

To complete the first objective for Task 3, our team requested billing data for each C&I premise that installed a retrofit measure within the past five years with the intent to perform a billing analysis for each premise. We readily acknowledge that a billing analysis is not the best option to measure savings for each and every C&I retrofit measure – some projects will be better candidates for a billing analysis than others. Predictable load patterns will facilitate the precise measurement of savings, as will larger effect sizes (relative to annual consumption at the premise). Because a billing analysis is not always the best tool for estimating savings, a number of sites were filtered out of the analysis. Table 4 summarizes the filtering process.

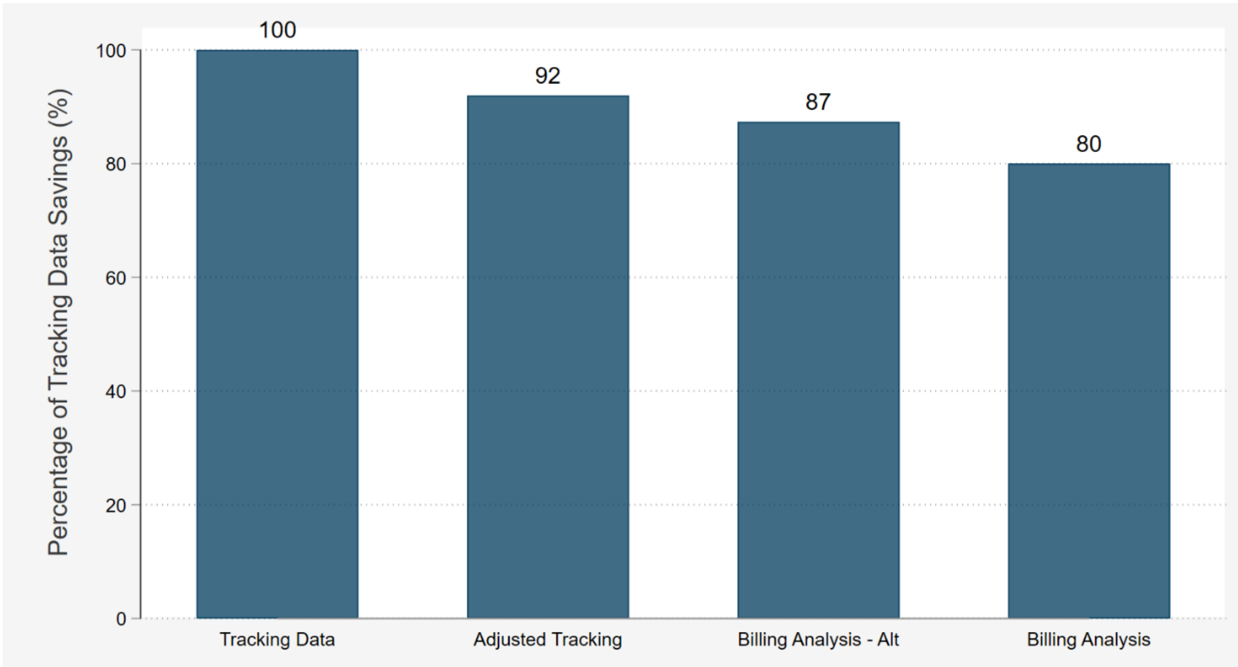
Subsequent sections describe steps taken towards completing the first objective of Task 3. Findings from the billing analysis are presented and discussed as well.

At a high level, our team reached different conclusions in the electric billing analysis and the gas billing analysis. On the electric side, the weather-normalized savings estimates derived from the billing analysis were correlated with the energy savings tracked by National Grid (Figure 13), though our estimates averaged approximately 80% of the National Grid tracking data estimates (Figure 14).[2] This result is not wholly unexpected and is generally in line with recent National Grid C&I impact evaluations. Gross savings values are commonly adjusted downwards with factors such as realization rates and in-service rates – adjustment factors that typically reduce the savings estimate based on findings from prior evaluations. (For example, a prior evaluation might find that businesses tend to overstate annual hours of use for lights in their facility. Overstating hours of use would result in an overestimate of savings. A realization rate – which is simply a ratio between realized savings and estimated savings – corrects for this.) The tracking data savings estimates used in this analysis have not been adjusted by realization rates or in-service rates. Thus, the billing analysis savings estimates should be expected to be slightly lower than the tracking data savings estimates.

Figure 1 shows how the different savings values described in the previous paragraph compare to the unadjusted savings values stored in the National Grid tracking data, which represent the benchmark used in this billing analysis. On average, adjusted tracking data savings estimates (which reflect realization rates and in-service rates) are about 92% of the unadjusted savings estimates (bar 2). As noted, the resulting billing analysis savings estimates averaged approximately 80% of the National Grid tracking data estimates (bar 4). In an alternative version of the billing analysis which used more stringent customer screens, savings estimates averaged approximately 87% of the National Grid tracking data estimates (bar 3).

[2] Under an alternative set of customer screens, this 80% figure was calculated to be 87%. Results for the alternative scenario are presented in an appendix.

Figure 1: Comparing Savings



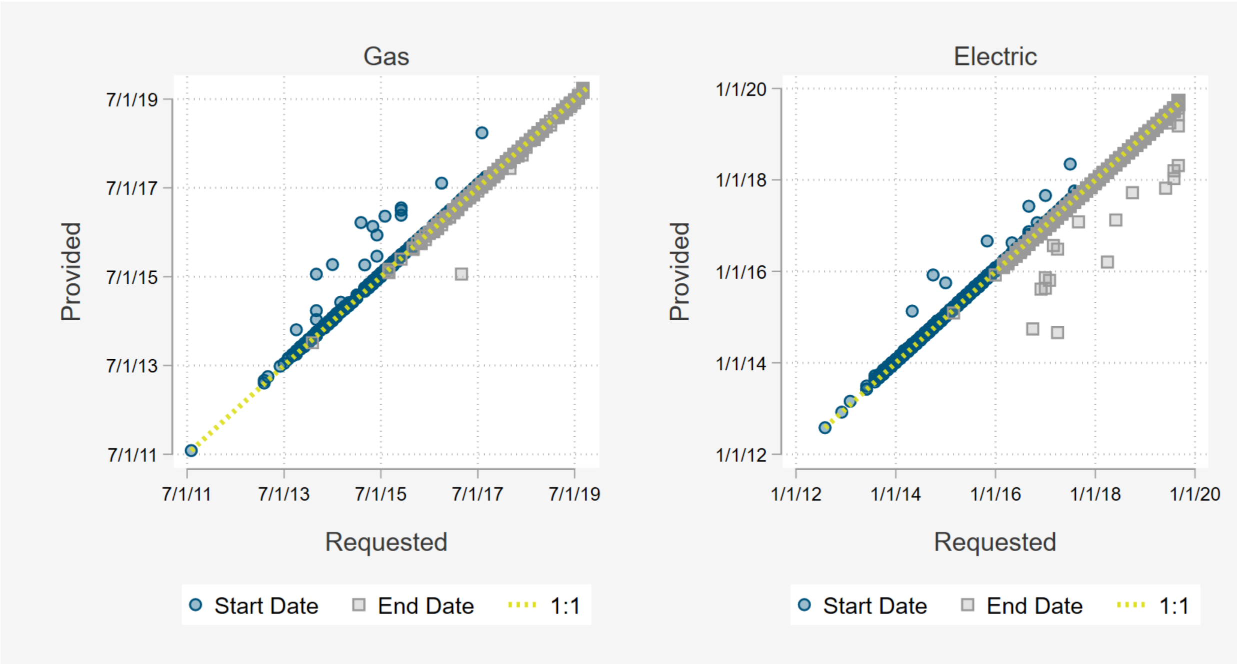
On the gas side, the weather-normalized savings estimates were not well correlated with the energy savings tracked by National Grid (Figure 17). In many cases, we found evidence of increased consumption in the post-installation period. It is important to note that the sample size for the gas billing analysis was much smaller than the electric analysis. Large sample sizes allow the results to converge around the true value because individual sites where savings are over or under-estimated can cancel each other out.

Potential problems associated with billing analyses are discussed in Section 2.3.2.4. It is possible that some of these problems – non-routine events, economic factors – are influencing the results of the billing analysis. That said, the directional effect of these factors cannot be determined without knowing when they occurred or how they altered energy consumption. Findings from the customer interviews and project documentation review sought to supplement the billing analysis findings, and did illuminate some of the reason for changes in energy consumption that were not accounted for in the billing analysis.

2.1 Review of Billing Data

Shortly after receiving customer billing data from National Grid, our team reviewed the billing data for completeness – did it match our billing request? The answer to this question was largely yes. Figure 2 compares requested billing ranges and provided billing ranges for both gas billing data (left pane) and electric billing data (right pane). Blue dots above the 1:1 line indicate premises for which we received fewer pre-installation months than expected, and grey boxes below the 1:1 line indicate premises for which we received fewer post-installation months than expected. Most points fall on the 1:1 line, indicating the data we received largely matches what we requested. Not shown in the figure is that a handful of premises from our initial data request were not found by the National Grid team (24 on the electric side and 1 on the gas side). In total, we received electric billing data for 1,551 premises and gas billing data for 420 premises.

Figure 2: Comparison of Dates



We also ran a number of checks to familiarize ourselves with the data and to highlight any potential issues with the data. Our team sent National Grid a memo ("Billing Data Questions") with the findings and a couple of questions. A summary of their responses is shown below.

- Finding 1: The relationship between account number and premise number in the billing data is not one-to-one. In cases where there are multiple account numbers per premise number, there is commonly a drastic change in the magnitude of consumption that coincides with the account number change. How should the Brightline team interpret/handle such cases?
  - National Grid response: This generally indicates a change in tenant/ownership.
  - Brightline action: Remove any premises for which the relationship between premise and account is not one-to-one, as a change in tenant/ownership likely results in a change in energy consumption behaviors. This reduces the billing analysis's ability to estimate savings. (Note: Table 4 shows how many premises were removed from our analysis file by each of the various filters our team applied.)
- Finding 2: There were a handful of premises that showed zero consumption in the billing data. Additionally, there were a handful of premises that showed zero consumption in either the pre-installation or post-installation period and some others that showed implausibly little consumption. Some instances of zero consumption could indicate shutdowns or other non-

routine events. Are there other reasons we should be seeing premises with no consumption or implausibly little consumption?

  - National Grid response: These could be shutdowns or other non-routine events. We are looking further into the premise numbers provided.
  - Brightline action: Premises with zero consumption (overall or just during the pre-installation or post-installation periods) were dropped, as zero consumption is indicative of non-routine events occurring at the premise. Such events reduce the reliability of a billing analysis. Premises with implausibly little consumption fell out of the analysis when a filter related to percent savings was applied. This filter is discussed below (Finding 4).

- Finding 3: There are a few Strategic Energy Management (SEM) programs in the tracking data. In some jurisdictions, tracked savings for such measures are non-additive (e.g., year 3 savings include year 2 savings). Is it safe to assume that all savings values shown in the tracking data sets are additive? Or will the savings values for some measures need to be adjusted?

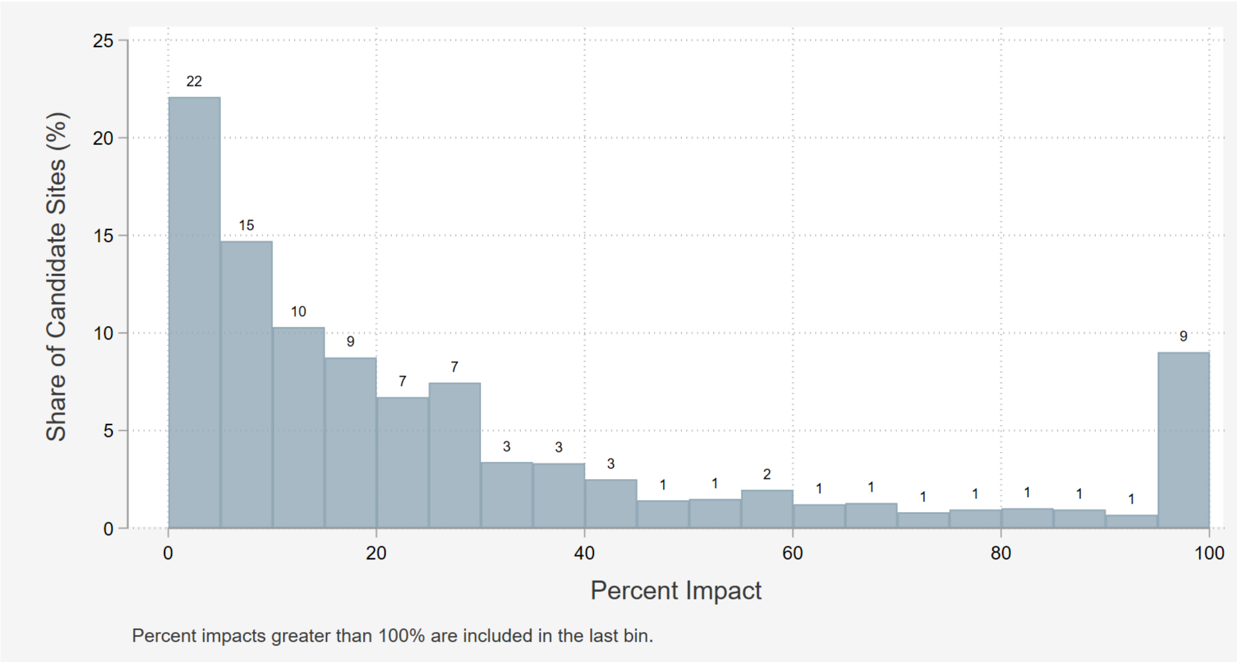
  - National Grid response: SEMP savings are additive. SEMPs typically indicate that the customer is among our largest. The savings included in one SEMP application generally cover all, or a large portion, of the energy projects that a customer will do across their buildings over the course of the year, but SEMP customers usually participate each year for a number of years. The savings likely apply to more than just the single account to which the application was attributed; they could apply to multiple accounts at a premise, or to multiple premises operated by that customer.
  - Brightline action: Since the savings likely apply to multiple premises over multiple years, any premises with SEMs were removed from the analysis as there is likely a disconnect between where the savings occur and the premise for which we requested billing data. The billing analysis would otherwise just estimate savings for one premise, likely missing some of the SEM savings.

- Finding 4: Figure 3 shows the distribution of percent savings for electric accounts and Figure 4 shows the distribution of percent savings for gas accounts.[3] For these figures, note that percent savings were capped at 100% (i.e., any percent savings greater than 100% was replaced with 100%). By and large, these distributions look reasonable – most percent savings fall under 20% and larger percent savings bins show fewer premises. That said, approximately 9% of electric accounts and 7% of gas accounts showed percent savings exceeding 100%. In some of these cases, it was obvious that the savings were not truly associated with the

[3] Note the percent savings used in these figures were based on our initial review of the billing data and were not weather-normalized.

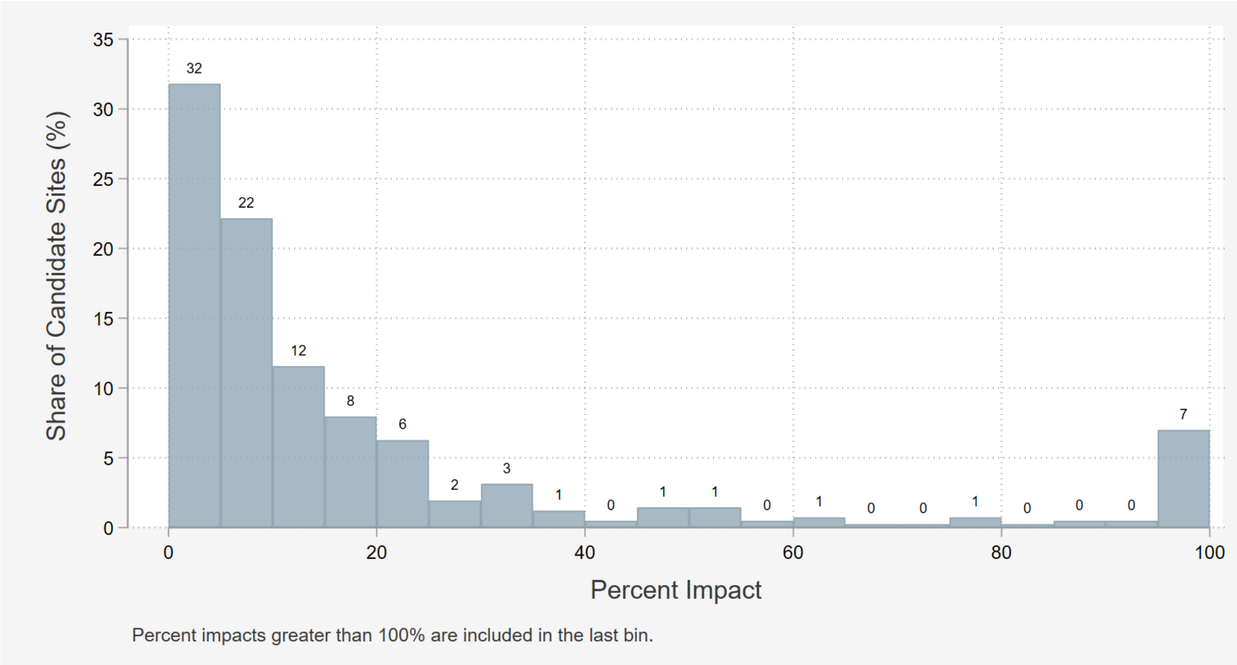
relevant premise number in the tracking data.[4] Are there other issues National Grid is aware of that might explain some of the large percent savings? Is it possible that the billing data is missing a meter multiplier for some premises, or an incorrect meter multiplier was applied to the underlying read?

Figure 3: Percent Savings – Electric



[4] In one case, annual energy savings in the tracking data for the premise add up to 3,037,985 kWh but annualized pre-retrofit consumption at the premise is only 26,052 kWh. Upon digging, we discovered that this premise is an electrical wholesaler. Here, it seems certain that the measures are being installed at a location other than the location associated with the premise number in the tracking data.

Figure 4: Percent Savings – Gas



- **National Grid response:** The examples you provided are probably common occurrences within the data. Streetlighting is billed not based on metering, but based on the installed fixtures and the operational schedule. Another likely occurrence in both the greater than 100% data and some of the other high percentage bins is that the application is associated with one account, but that the measures impact buildings covered by multiple accounts. It's also possible some of these sites have solar or other distributed generation that impacts their net consumption.
- **Brightline action:** Remove any premises with a percent savings greater than 100% from the analysis (where the numerator in the percent savings is total annual savings at the premise per the NGRID tracking data, and the denominator is weather-normalized pre-installation annual consumption), as the tracking data savings likely span multiple premises. The billing analysis would only capture the portion of the savings that occur at the premise for which we requested billing data.

Based on the feedback discussed above, our team dropped 45 gas premises (11% of total) and 257 electric premises (17% of total) from the billing analysis. For various other reasons, a number of other premises were dropped from the analysis as well – these are discussed in the next sections. Table 4 shows how many premises were removed by each filter.

2.2 Data Preparation

There were four data sources for this analysis – C&I gas/electric tracking data provided by National Grid, gas/electric billing data provided by National Grid, historical weather data downloaded from the National Oceanic and Atmospheric Administration (NOAA), and typical meteorological year (TMY) weather data downloaded from the National Renewable Energy Laboratory. These sources needed to be weaved together in order to perform the billing analysis. Each data source, its relevance to the analysis, and any steps taken to prepare the data are discussed in the sections below. Our initial review of the billing data was discussed in Section 2.1.

2.2.1 C&I Tracking Data

The C&I tracking data shows measure-level details for each retrofit measure. Measure types other than retrofits were included in the tracking data, but this investigation focused just on the retrofit measures since pre-existing condition is the appropriate baseline for retrofits (as opposed to code minimum, which would be problematic for a pre/post billing analysis). Three relevant tracking data fields for this research effort are the installation completion date, the account number, and the premise number. We used this information to develop a list of electric and gas premises that installed one or more retrofit measures. This list was the basis for our billing data request. Installation completion dates in the tracking data informed the date ranges used in the billing data request (e.g., if one premise had two measures with completion dates of 1/5/2017 and 5/5/2017, then we would need data from 1/5/2016 through 5/5/2018 to have at least twelve months of pre-installation billing data and twelve months of post-installation billing data). These also helped us to identify which periods in the billing data were pre-installation and which periods were post-installation. (For a site with multiple retrofit measures, there would potentially be a third period between pre-installation and post-installation. This third period did not factor into the analysis.) In total, we requested electric billing data for 1,575 premises and gas billing data for 421 premises.

The other critical component of the tracking data is an estimate of gross savings (gross kWh savings in the electric tracking data and gross therms savings in the gas tracking data). These values are essentially the benchmarks for this analysis – they represent what we compare the billing analysis results to. It is important to note that these gross kWh or therm values are not necessarily what is ultimately claimed and reported by National Grid. Additional factors such as in-service rates or realization rates from prior evaluation studies are used to adjust the data for reporting purposes.

The tracking data also included information on the building type (though this was sparsely populated in the gas tracking data) and the measure type.[5] This information was used to filter analysis results (e.g., examining results for building type or by measure type).

2.2.2 Billing Data

There were five main fields in the billing data we received:

- Account number
- Premise number
- Cycle start date

[5] In the C&I custom gas tracking data, 87% of retrofit measures have a missing building type. In the C&I prescriptive gas tracking data, 63% of retrofit measures had either a missing or "other" building type. In the C&I custom electric tracking data, 15% of retrofit measures have an "other" building type. In the C&I prescriptive electric tracking, 27% of retrofit measures have an "other" building type.

- Cycle end date
- Total consumption (either kWh or therms)

Our initial review of the billing data is discussed in Section 2.1. This review also looked for duplicate records and gaps/overlaps in the billing data. Duplicate records were rare, as were gaps between billing cycles. Overlaps were also rare other than common cycle start and end dates for consecutive cycles (e.g., one cycle goes from 10/15/2015 to 11/15/2015, then the next cycle goes from 11/15/2015 to 12/15/2015).

To check for erroneous readings from the billing data, our team calculated standardized consumption values for each billing record. This allowed us to identify and tag unusually large or small consumption values for any given customer (i.e., outliers).[6] Approximately 0.33% of electric billing records were tagged as outliers and 0.28% of gas billing records were tagged as outliers. Such records were not included in the billing analysis.

Premises with less than 12 months of pre-installation or post-installation billing records were also dropped from the analysis. With this filter, 233 premises were dropped from the electric billing analysis and 77 premises were dropped from the gas billing analysis. For the majority of these instances, it was the case that the measure was installed too recently for 12 months of post-installation billing data to be available.

2.2.3 Historical Weather Data

As noted, weather data was downloaded from NOAA. In our initial Work Plan, we proposed mapping each premise to the nearest weather station (using zip codes) and downloading weather data for all of the relevant weather stations. We had to tweak this approach, as not all weather stations provide TMY data and TMY data is needed to produce weather-normalized savings estimates. The new approach was to map each premise to one of two TMY weather stations – T. F. Green State Airport (Providence) or North Central State Airport (Pawtucket) based on proximity.

After downloading the historical weather data, some general data preparation was needed. The two main steps were removing erroneous temperature values (e.g., -9999) and interpolating missing values. As an example of the latter point, suppose temperature data for North Central State Airport is missing for a three-week period in June 2017. Here, we used the relationship between temperature readings at the two stations to predict what the weather was like at North Central State for the missing period.

After interpolating any missing data, the weather data had to be assembled in such a way that it could merge with the billing data – records in the billing data represent cycles rather than days. Thus, transforming the daily records into cycles was necessary (although there are other approaches one

[6] Any consumption value more than three standard deviations away from the average consumption value at the premise was tagged as an outlier. As an example, suppose the average daily consumption value at premise 123 is 1,000 kWh and the standard deviation is 200 kWh. Any billing record with an average daily consumption value greater than 1,600 kWh or less than 400 kWh would be tagged as an outlier.

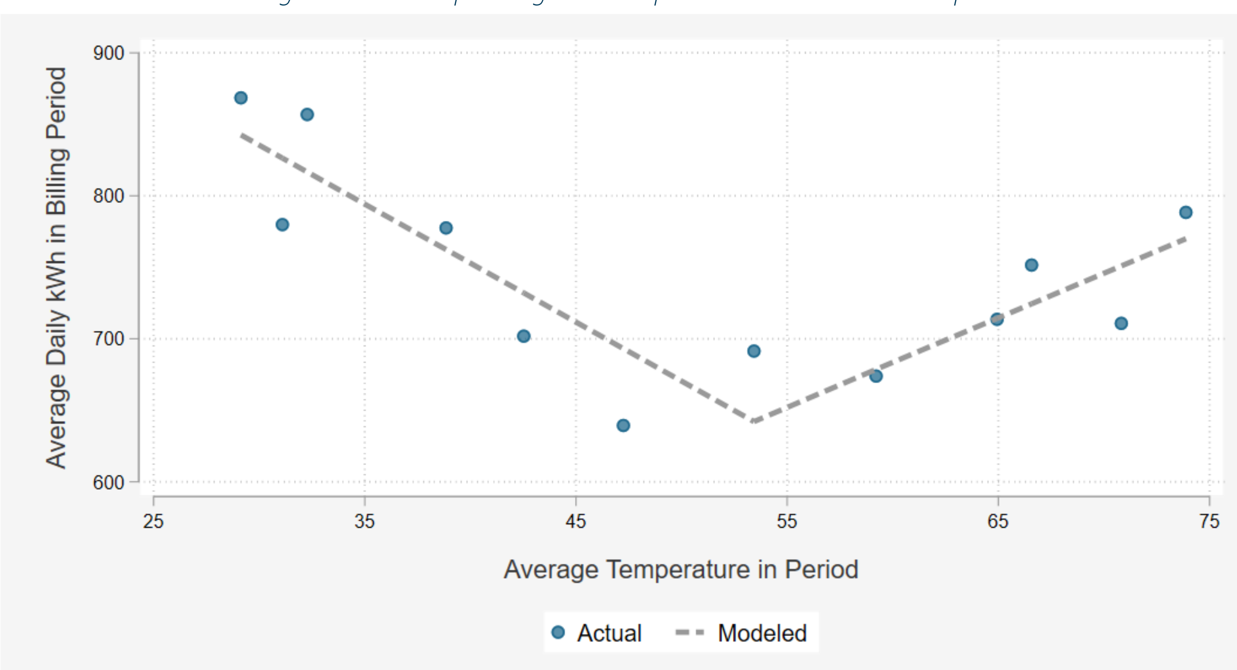
could take here). As an illustrative example, suppose one billing cycle runs from 4/5/2017 to 5/6/2017. For the same period, we would calculate (1) the average daily temperature, (2) the sum of cooling degree days (CDD), and (3) the sum of heating degree days (HDD). These values could be merged with the billing record chosen for this illustrative example.

Part of this analysis entails estimating the best degree day base for each premise. This means that the merge described in the example above was slightly more complex. Components (2) and (3) were really arrays rather than single values. That is, we would merge in multiple CDD and multiple HDD values, each using a different degree day base. Multiple CDD and HDD terms were needed so that we could determine the best degree day base to use at each site. The degree day base selection is discussed in the next section.

2.2.4 Degree Day Base Selection

The regression models we use to estimate weather-normalized consumption rely on two independent variables: average daily CDD in billing cycle and average daily HDD in the billing cycle. CDD and HDD themselves are functions of average daily temperature and the degree day base being used. For example, consider a day where the average temperature is 80° (F). Using a degree day base of 55, CDD would be 25. Using a degree day base of 60, CDD would be 20. For some sites, using a base of 55 may be appropriate, but a base of 60 may be better for others. It depends entirely on the relationship between load and weather at the premise. Figure 5 shows an example. For this hypothetical customer, a degree day base around 54 or 55 would be the most appropriate, as that is the point where the relationship between temperature and consumption changes. (When temperatures are between 30° and 54°, there is a negative correlation between consumption and temperature. Above 55°, the correlation becomes positive.)

Figure 5: Consumption Against Temperature – Illustrative Example



The example visualized in Figure 5 is useful, but exploring similar graphs for each premise would be an inefficient use of time (and would create the possibility of error through guesswork). Our team used a more empirical approach to determine what degree day base should be used for each customer. Steps taken were as follows:

- Trim the data to just the pre-installation period.
- For each premise, fit a series of regression models. The regression models will use the same general specification, but the degree day base used will change from model to model. An example model is shown below. Note that "CDD55" means CDD with a degree day base of 55 degrees (F). Our team tested out bases from 40 to 75.

$$Avg.Daily\ Consumption = \beta_0 + \beta_1 \times (Avg.Daily\ CDD55) + \beta_2 \times (Avg.Daily\ HDD55)$$

- For each model, store the R^2 statistic. This statistic measures how much of the variation in consumption can be explained by the regression model. In general, higher R^2 values are better.
- For each premise, determine which degree day base provides the highest R^2 value. Use that degree day base when estimating weather-normalized consumption for the pre and post periods.

Across all customers, Table 1 shows the median degree day base for each data set.

Table 1: Degree Day Bases

Billing Set	Median DD Base	Average R^2
Gas	57	89.9%
Electric	58	68.9%

2.2.5 Analysis Data Set

Once the pieces were merged together, there was one final step – consumption records (and also CDD and HDD records) needed to be standardized to allow for an "apples to apples" comparison. The idea here is that some billing cycles are longer than others, and the longer cycles may show higher total consumption just by virtue of being longer. The fix is to divide total consumption for the billing cycle by the number of days in the billing cycle. For electric consumption, as an example, this produces a kWh/day value for each billing cycle. These standardized values were the basis of our analysis.

A few example rows from the electric analysis data set are shown in Table 2. "Pre" is an indicator variable that equals 1 for any record that predates the retrofit installation, 0 otherwise. "Post" is an indicator variable that equals 1 for any record that follows the retrofit installation, 0 otherwise. Some billing records will fall between the pre-installation and post-installation periods. For such records, both "Pre" and "Post" will equal 0. Also recall that the degree day base used for the CDD and HDD terms varied from one premise to another.

Table 2: Example of Analysis Data

Premise	Start Date	End Date	Pre	Post	Avg. Daily kWh	Avg. Daily CDD	Avg. Daily HDD
12345	4/4/16	5/2/16	1	0	731.0	0.5	6.9
12345	5/3/16	6/1/16	1	0	633.3	7.3	1.0
12345	6/2/16	7/4/16	0	0	709.1	15.0	0.0
12345	7/5/16	8/2/16	0	1	865.5	25.6	0.0

2.2.6 Typical Meteorological Year (TMY) Data

Per the National Renewable Energy Laboratory's (NREL) user manual, a TMY data set is an "annual data set that holds hourly meteorological values that typify conditions at a specific location over a longer period of time, such as 30 years".[7] These "typical" weather values can be used to predict weather-normalized energy consumption – an estimate of what energy consumption *would have been* if observed weather conditions during the relevant period had been typical (as opposed to whatever actual conditions were observed, which could be unusually warm/cold).

TMY data sets can be found online and do not require cleaning like historical weather data sets do.[8] Their use in this analysis is described in greater detail in Section 2.3.1. For this research, two TMY stations were used: T. F. Green State Airport and North Central State Airport (Figure 6). Each premise in the analysis was mapped to one of these two stations based on geographic proximity.

[7] Users Manual for TMY3 Data Sets, S. Wilcox, W. Marion. Revised May 2008. Available at https://www.nrel.gov/docs/fy08osti/43156.pdf

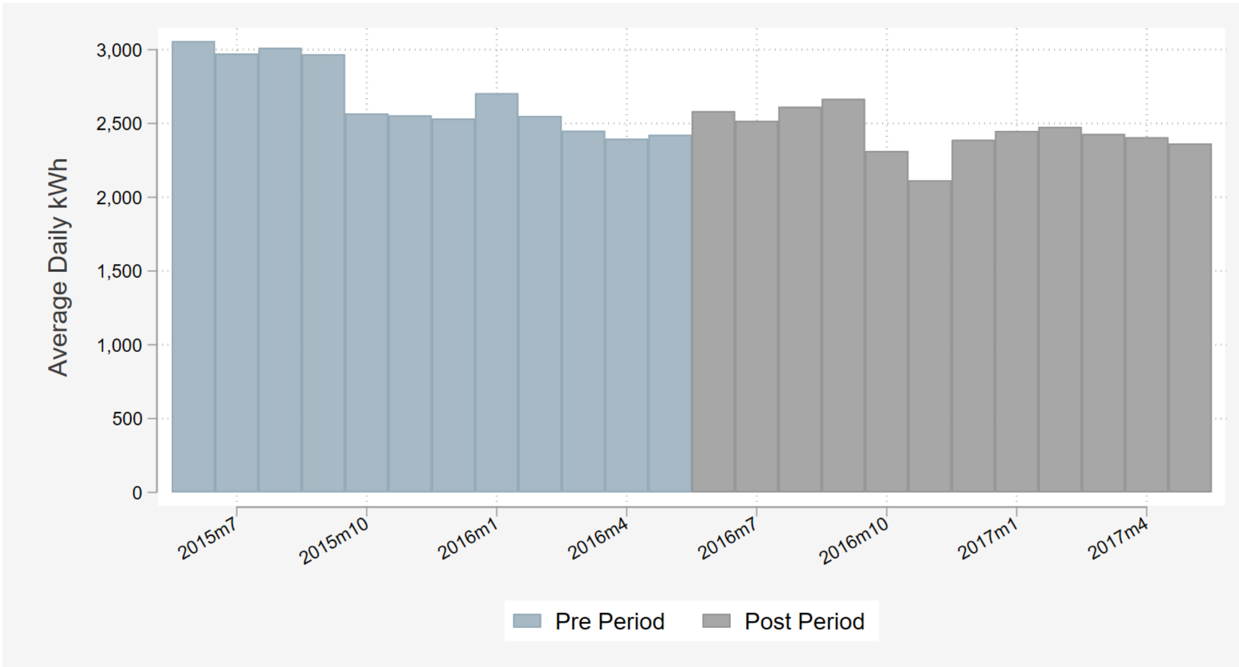
[8] https://rredc.nrel.gov/solar/old_data/nsrdb/1991-2005/tmy3/by_state_and_city.html

Figure 6: TMY Station Map

2.3 Billing Analysis

A billing analysis seeks to estimate the effect an energy conservation measure (ECM) has on energy consumption through an investigation of gas/electric bills. Figure 7 shows 24 months of electric bills for a hypothetical participant – 12 months of bills that predate the ECM and 12 months of bills that follow the installation of the ECM. Clearly, the bills in the post-installation period are a bit lower – but why? Is the difference attributable to the ECM? Could the difference be explained by differences in weather in the two periods? Is the difference just noise (i.e., random year-to-year variation that cannot be explained through observable/available variables)? Does the difference result from a combination of these factors? These are the questions that a billing analysis seeks to answer. The final output in a billing analysis is an estimate of savings that are attributable to the ECM (and a corresponding measure of uncertainty – the "margin of error" – which represents how much we think the savings estimate could be off by given the unexplained variation in the data).

Figure 7: Monthly Electric Bills – Illustrative Example



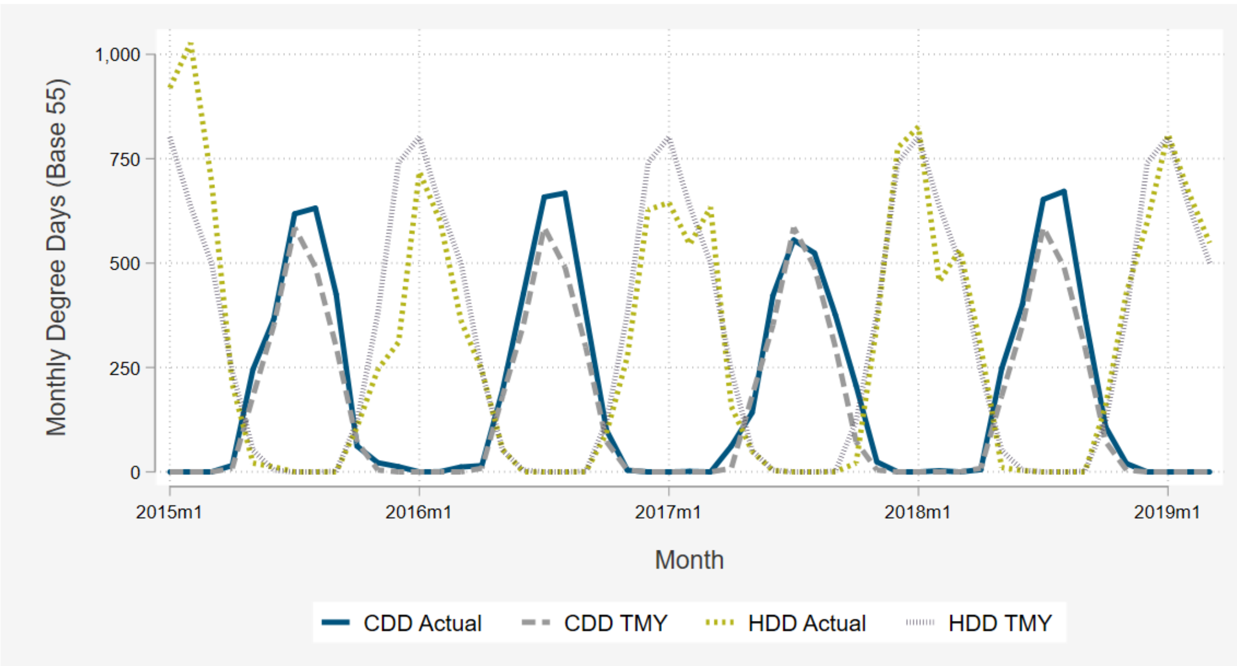
The amount of uncertainty around the savings estimate is a function of several factors, notably (1) the magnitude of the expected savings relative to annual consumption, (2) the ability to explain variation in energy consumption with external variables like weather, and (3) the amount of random variation in consumption. We expect some of the billing analyses to produce more precise savings estimate (less uncertainty) than others, as some ECMs will produce more savings (i.e., a larger signal to detect) than others and the relationship between consumption and weather will be stronger for some sites than others.

For each premise, we produced savings estimates with two different approaches. The primary approach will yield an estimate of weather-normalized savings. This approach controls for the fact that the weather observed during the pre and post periods were different and may have been atypical for the region. In other words, the primary approach answers this question: What amount of savings would we expect to observe annually given typical regional weather?

The secondary approach, which also relies on weather data as an input, yields an estimate of avoided energy use. This approach removes the notion of a "typical weather year" from the equation. Instead, the secondary approach answers this question: If the relationship between weather and consumption during the pre period carried over into the post period, how much energy use was avoided in the post-period? Avoided energy use is calculated as the difference between predicted post energy use (via a pre period regression model and post period weather conditions) and actual post energy use. Steps taken in producing an estimate of savings – be it a weather-normalized estimate or not – are discussed in subsequent sub-sections.

Our belief is that the primary approach will yield a more robust estimate of savings because it accommodates for changes in weather from year to year. As an illustration, see Figure 8. This figure shows actual Rhode Island weather against typical meteorological year (TMY) weather for the region. The 2017 summer was fairly typical and the 2018 summer was a bit warmer than usual. Supposing no ECMs were installed at a given facility, it is likely that consumption during the 2018 summer is higher than consumption during the 2017 summer (due to a greater cooling load). A chiller measure installed in spring 2018 would likely show more savings in an avoided energy use model than a normalized model because summer 2018 was warmer than usual. That said, we believe the secondary approach (avoided energy use) is useful in that it mimics the viewpoint of program participants – they see their bills, not what their bills *would have been* if the weather had been more typical.

Figure 8: Historical and Typical Weather in Providence, RI



Assumptions

The savings estimated derived via this billing analysis are based on one key assumption: the only differences between the pre period and the post period are the weather and the ECM itself. In other words, the pre period acts as a baseline against which the post period is compared. If other changes occur between or during the periods, then the regression models used to estimate savings will confound the effect of the ECM with whatever other changes occurred between or during the two periods. For example, consider a premise that reduces its operating hours near the beginning of the post period. Presumably, this reduction in operating hours would result in a reduction in electric and/or gas consumption. Unless this reduction is explicitly controlled for in the billing analysis, the analysis will attribute this decrease in consumption to the ECM. Thus, the savings estimate will be overstated. It will actually be an estimate of the effect of the ECM plus an effect related to the operating hours change. These two components cannot be untangled without including additional information in the regression models used in the billing analysis.

As the billing analysis relies on regression modelling, there are a handful of other assumptions built into the savings estimates. Readers interested in the assumptions upon which linear regression models are built should consult a Statistics textbook.[9]

When comparing the savings estimates stored in the tracking data to the weather-normalized savings estimates calculated via the billing analysis, there is one other important consideration. The billing analysis returns an estimate of savings observed at one single premise. For some premises, it is possible that the savings shown in the tracking data are realized over multiple premises/locations (i.e., not all savings are truly associated with the premise number in the tracking data). To the extent possible, our team tried to filter such instances out of this analysis, but there may be cases where the comparison of National Grid savings to the Brightline savings estimate is not an apples-to-apples comparison.

2.3.1 Weather-Normalized Savings

2.3.1.1 Approach

A weather-normalized savings estimate is simply the difference between weather-normalized pre-installation consumption and weather-normalized post-period consumption. Note that these weather-normalized consumption values are annualized, so the weather-normalized savings value is an annual metric as well.

Finding the difference between two values is easy enough. The question thus becomes: How does one estimate weather-normalized consumption? To do this, we start by creating a mathematical model that relates consumption and the relevant temperature variables (CDD and HDD):

$$Avg.Daily\ Consumption = Intercept + \beta_1 \times (Avg.Daily\ CDD) + \beta_2 \times (Avg.Daily\ HDD)$$

In this equation, β_1 represents the change in average daily consumption per each additional average daily CDD, and β_2 represents the change in average daily consumption per each additional average daily HDD. These values (as well as the "Intercept" term) are estimated separately for each premise and for each of the relevant periods (pre-installation and post-installation).

After the model has been created, the relationship is cast over TMY data to develop an estimate of consumption under typical weather conditions. Suppose our model for pre period consumption is as follows:

$$Avg.Daily\ Consumption = 617 + 6.89 * (Avg.Daily\ CDD55) + 10.08 * (Avg.Daily\ HDD55)$$

And suppose our model for post period consumption is as follows:

$$Avg.Daily\ Consumption = 496 + 9.54 * (Avg.Daily\ CDD55) + 9.82 * (Avg.Daily\ HDD55)$$

[9] One of the main assumptions is that a linear relationship exists between the response variable (consumption) and the explanatory variables (CDD and HDD). There is generally some prediction error in regression modeling. These errors are assumed to be independently distributed with a constant variance. That is, the prediction error for the billing cycle in July is assumed to be unrelated to the prediction error for the billing cycle in June. Regarding "constant variance", this simply means that prediction errors do not get larger/smaller for larger/smaller consumption values.

With the models in hand, we can estimate weather-normalized consumption. The second and third columns of Table 3 show average daily CDD55 and HDD55 for Providence, Rhode Island, based on the TMY data file. Plugging those values into the equations shown above will yield the values in the "Predicted Daily Consumption" columns. Those are the estimates of weather-normalized daily consumption. The difference between these estimates represents savings. Multiplying the daily savings estimate by the number of days in the month will yield an estimate of monthly savings. Summing the monthly savings estimates across the year yields the weather-normalized annual savings estimate (39,843 kWh in the example below). The annual savings estimate can be expressed as percent savings by dividing by average annual consumption in the pre period (and multiplying by 100%).

Table 3: Weather-Normalized Savings – Illustrative Example

Month	Average Daily TMY Weather		Predicted Daily Consumption		Average Savings	
	CDD55	HDD55	Pre	Post	Daily	Monthly
1	0.0	25.9	878	750	127.9	3,964
2	0.0	22.7	846	719	127.1	3,558
3	0.0	16.2	780	655	125.4	3,886
4	0.3	8.0	700	578	122.5	3,675
5	5.9	1.6	674	568	106.0	3,287
6	11.7	0.1	699	609	90.1	2,704
7	18.9	0.0	747	676	71.1	2,204
8	15.8	0.0	726	647	79.2	2,456
9	10.0	0.0	686	592	94.7	2,842
10	2.3	3.9	673	557	116.0	3,596
11	0.1	12.6	745	621	124.1	3,722
12	0.0	23.9	858	731	127.4	3,949
Total	---	---	---	---	---	39,843

2.3.1.2 Filters Applied

Before presenting the results, it should be noted that a number of premises were removed from the analysis either because (1) the size of the savings signal relative to annual load at the facility was expected to be too small to detect or (2) because variations in consumption were largely unrelated to the explanatory variables in the regression models (CDD and HDD). Regarding (1), a billing analysis cannot reliably estimate savings if the savings signal is small – it's simply the wrong tool for the job. To this end, a percent savings threshold of 5% was used. An upper threshold of 100% was also in place, as it is not possible for a site to save more energy than it consumes. Percent savings were calculated as:

$$\%\,Savings = \frac{Gross\ savings\ estimate\ per\ NGRID\ tracking\ data}{Weather\text{-}normalized\ pre\text{-}installation\ annual\ consumption} * 100\% \qquad (1)$$

For premises that installed multiple retrofit measures, the numerator in the percent savings calculation would be the sum of the gross savings values for the different measures.

Regarding (2), a relative precision threshold of 50% was used. Relative precision is the ratio of uncertainty to savings. In this case, uncertainty is represented by the margin of error around the savings estimate produced by the billing analysis. Relative precision at the 95% confidence level was calculated as:

$$Relative\ Precision = \frac{Margin\ of\ Error\ Around\ Savings\ Estimate}{Savings\ Estimate} * 100\% \qquad (2)$$

Readers will note that the calculation of relative precision relies on findings from the billing analysis. The 50% relative precision threshold was therefore applied after the analysis was completed. This filter served to remove sites from the reporting rather than screen sites out of the analysis itself. As such, sites that did not meet the threshold are not included in any of the tables, figures, or statistics in the following sections.

A number of premises did not meet the percent savings and relative precision thresholds noted above. Additionally, premises without at least 12 months of pre-installation billing data and 12 month of post-installation billing data were withheld from the analysis, as we felt we did not have enough information to predict weather-normalized annual consumption (in either the pre period or post period) for such premises. For a handful of premises, one or more measures had savings that expired during the analysis period (one electric, three gas) – these premises were dropped as well.[10] One additional post-analysis filter was applied: Sites where there was a considerable difference between the tracking data savings estimate and the billing analysis savings estimate were removed. Here, "considerable" is relative to the distribution of savings ratios, which are discussed later. In total, eight electric customers and two gas customers were removed by the filter. These premises were not removed from the pool of interview candidates for the second component of Task 3.

After the filtering noted here and elsewhere in this document[11], a total of 298 premises remained in the electric billing analysis and 34 premises in the gas billing analysis (down from 1,575 and 421 in the

[10] For these premises, it was the case the multiple measures were installed across a few years. For example, a measure with a useful life of three years was installed in January 2015 and a measure with a useful life of five years was installed in July 2017. The savings from the first measure would expire in the middle of the post-installation period.

[11] Other filters: premises with multiple account numbers were dropped, premises with zero consumption in either the pre-installation or post-installation period (or both) were dropped, premises with SEM projects were dropped, and premises with percent savings greater than 100% were dropped.

initial data request, respectively). Table 4 shows how the sample size decreases after different filters are applied.[12] Though the number of premises in the final sample is considerably less than the number of premises from the initial data request, the filters were in place to remove premises that are not good candidates for a billing analysis. We believe that the premises that passed through all of the filters are strong candidates for a billing analysis. We did not expect to run billing analyses for all 1,575 premises in the electric billing data request or all 421 premises in the gas billing data request.

Table 4: Number of Premises Removed by Filtering

Filter #	Filter	Electric Premises		Gas Premises	
		Removed	Remaining	Removed	Remaining
--	Original population of retrofit premises	---	1,575	---	421
1	Premise not found by National Grid	24	1,551	1	420
2	Premise/account relationship is not 1:1	130	1,421	31	389
3	Zero consumption in pre or post period	6	1,415	1	388
4	Remove premises with SEM measures	3	1,412	0	388
5	Less than 12 months of pre/post data	224	1,188	76	312
6	Percent savings < 5% or > 100%	399	789	129	183
7	Relative precision > ±50%	482	307	145	38
8	Savings expire during analysis window	1	306	2	36
9	Savings ratio outliers	8	298	2	34
	Final analysis data set	**1,275**	**298**	**2**	**34**

Regarding Filter 6, we will note that an upper bound of 100% represents a theoretical limit. For a site to save 80% or 90% would be improbable but not impossible. One possible explanation for such cases would be missing meter data. In Section 3 , results are presented with an upper bound 50% for Filter 6 rather than 100%. In general, the more stringent filter produced results that were better aligned with the tracking data savings values.

Table 5 shows an end-use breakdown for the 298 electric customers in the final analysis file. The sum of the numbers in the table exceeds 298 – this is because several customers installed multiple retrofit measures. A majority of the retrofit measures were lighting measures. End-use information in the gas

[12] If the order of the filters was rearranged, the number of premises removed at each step would change. However, the begin and ending values would be the same.

tracking data was less informative (the end-use categories are C&I Gas Custom, Large Commercial Retrofit, and Steam Initiative).

Table 5: End-Use Information for Electric Customers

End-use	Numbers of Measures	Percentage of Measures (%)	Sum of Tracking Data Savings (MWh)	Percentage of Tracking Data Savings (%)
Misc. Custom	23	1.1	2,942	5.2
HVAC	115	5.3	7,678	13.4
Lighting	1,756	80.5	39,336	68.9
Refrigeration	87	4.0	5,453	9.5
VFD	200	9.2	1,713	3.0
Total	2,181	100%	57,123	100%

2.3.1.3 Distribution of Consumption

To help contextualize the results of the analysis, a discussion on the distribution of consumption is presented in this section. Figure 9 and Figure 10 show the distribution of annual electric consumption and annual gas consumption, respectively, for the premises that remained in the final analysis file. Note that annual consumption was calculated solely based on the pre period for the figures and summaries shown in this section. The distribution of annual electricity consumption has a long right tail – this is a common shape for variables that have a lower bound (zero consumption) but no upper bound. The distribution of annual gas consumption has a similar, though less pronounced, shape.

Figure 9: Distribution of Annual Electricity Consumption (kWh)

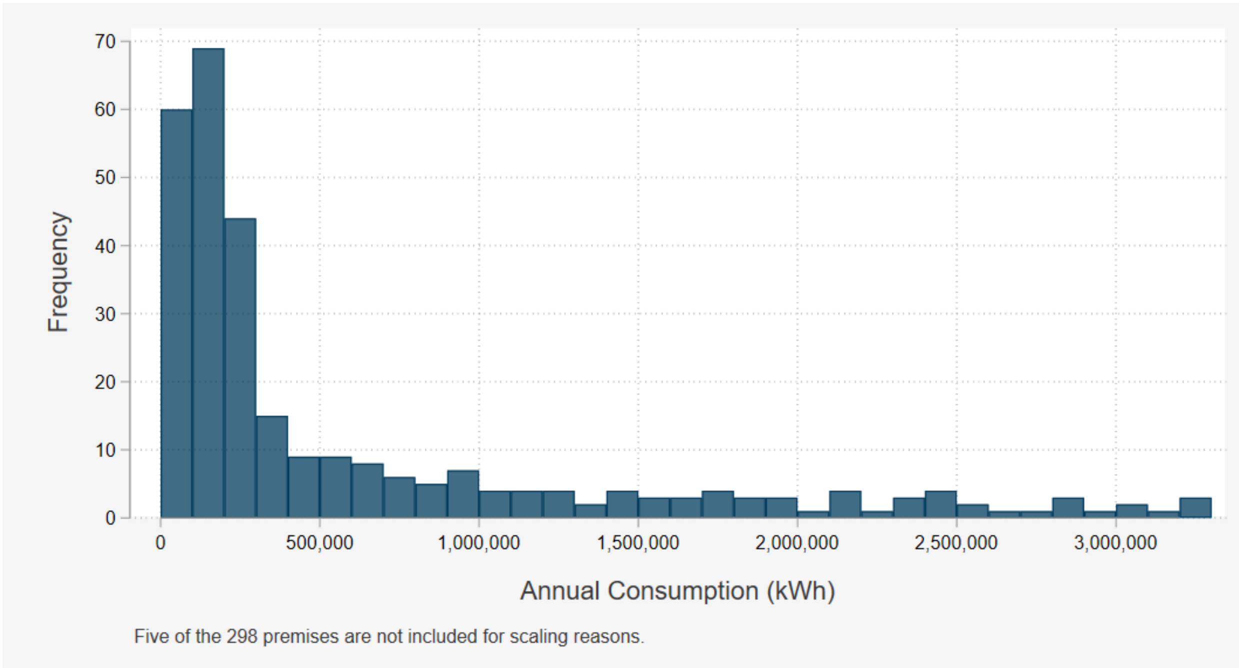


Figure 10: Distribution of Annual Gas Consumption (therms)

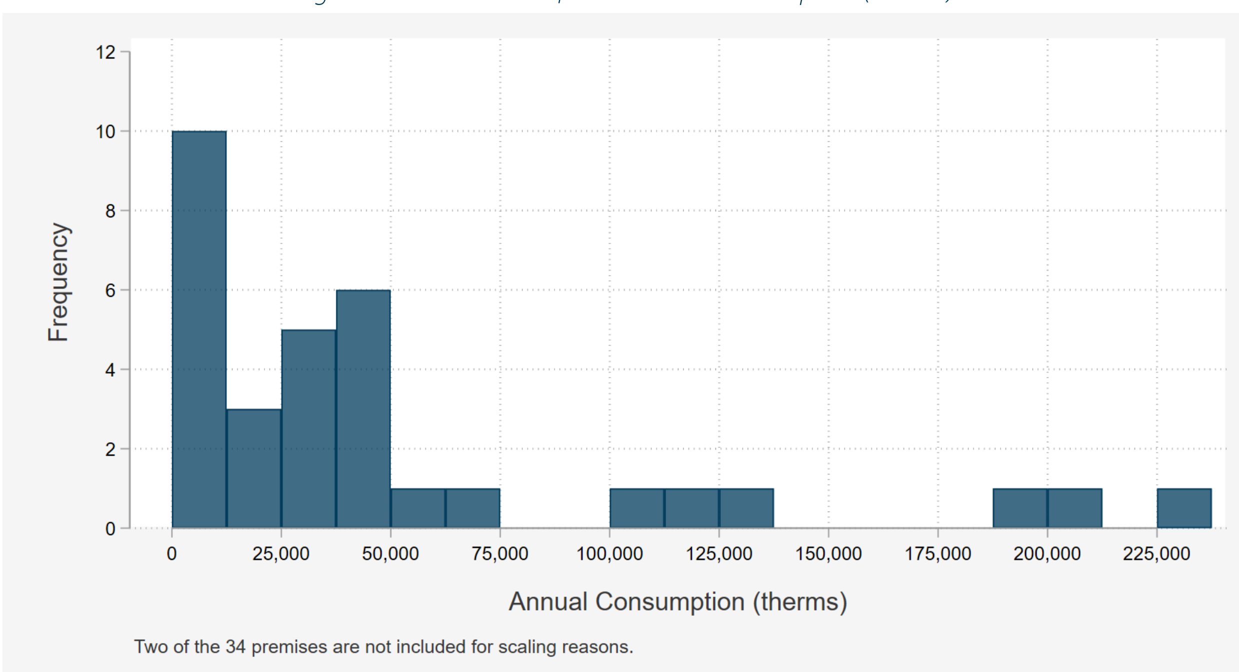


Table 6 shows some summary statistics describing average daily consumption. For both data sets, average consumption exceeds the median consumption value. This can be explained by the long right tails in the figures above. The average is highly sensitive to the premises with considerably higher consumption, while the median is not. In other words, sites with considerably higher consumption will pull the overall average in their direction but have little effect on the median. The standard deviation is also a sensitive statistic, which is why it is so large.

Table 6: Daily Consumption Summary

Fuel	Statistic	Value
Electric	Number of sites	298
	Average	2,234 kWh
	Median	632 kWh
	Standard deviation	5,748 kWh
Gas	Number of sites	34
	Average	621 therms
	Median	98 therms
	Standard deviation	2,487 therms

2.3.1.4 Results – Electric

All impact estimates produced by the Brightline billing analysis were found to be statistically significant at the 95% confidence level.[13] Table 7 shows average and median annual kWh savings for the premises that remained in our electric analysis file. Both NGRID tracking data savings and the savings estimate from the Brightline billing analysis are shown. For both the average and the median, the Brightline estimate comes in around 70% of the NGRID estimate. The median estimates are lower than the average estimates because there are a few premises with very high savings values that inflate the average. The median is unaffected by these values and is likely the better measure of central tendency in this case.

Table 7: Savings Comparison – Electric

Fuel	Measure	# of Premises	NGRID	Brightline
Electric	Sum	298	57,123 MWh	38,781 MWh
Electric	Average	298	191,687 kWh	130,139 kWh[1]
Electric	Median	298	53,328 kWh	39,733 kWh

[1] A 95% confidence interval for the average savings value was calculated to be (96,446, 163,832).

[13] This is as expected. Filter 7 in Table 4 removes premises where uncertainty exceeds the savings estimate which would render the estimate statistically insignificant.

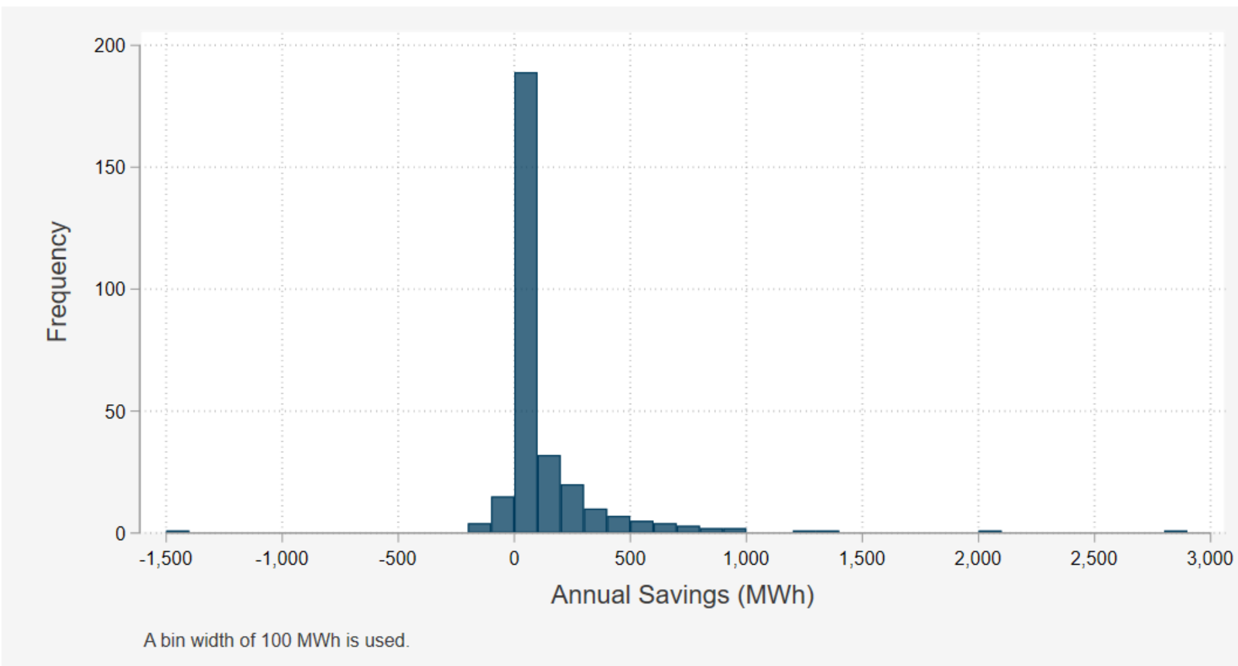
Note that the tracking data savings values shown in Table 7 are not adjusted with realization rates and in-service rates.[14] The savings values that National Grid actually claims have these adjustments factored in. Table 8 compares unadjusted savings values, adjusted savings values, and the savings values calculated via the billing analysis. The remainder of the figures, tables, and comparisons in this report use the unadjusted savings values.

Table 8: Comparison with Adjusted Savings Values

Source	# of Premises	Total Savings (MWh)	% Difference Relative to Unadjusted Savings
NGRID – Unadjusted	298	57,123	---
NGRID – Adjusted	298	50,863	-32%
Brightline Analysis	298	38,781	-11%

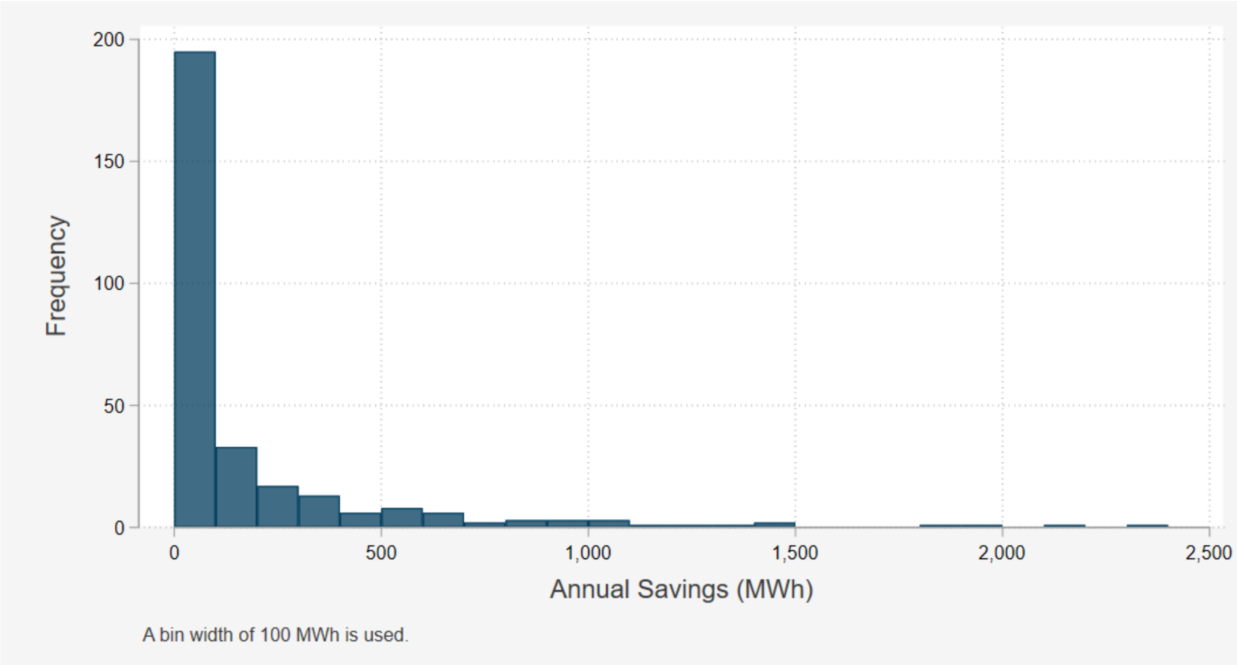
Figure 11 and Figure 12 show the distribution of savings derived via the billing analysis and in the tracking data, respectively.

Figure 11: Distribution of Electric Savings – Billing Analysis



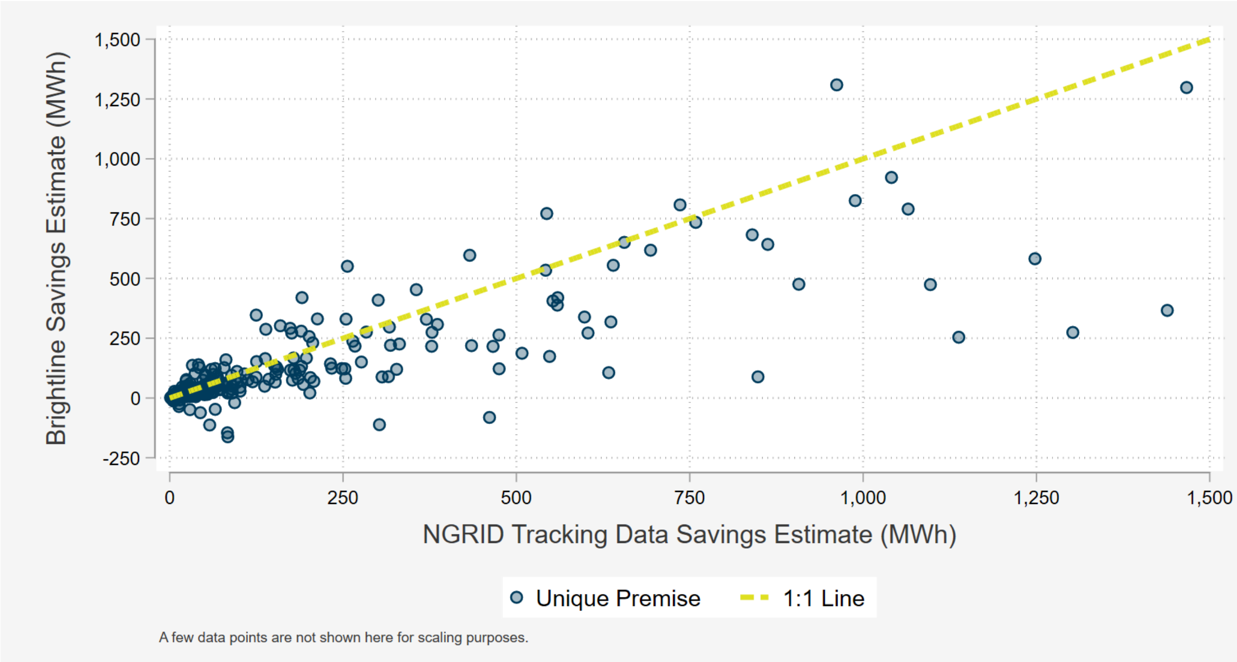
[14] A realization rate represents the percentage of expected savings that are actually realized. Consider a lighting upgrade that is expected to save 100,000 kWh per year. Measurement and verification of the project returns an annual savings estimate of 90,000 kWh per year. In this case, the realization rate of the project is 90%. The realization rates stored in NGRID's tracking system are averages based on measurement and verification results over time. An in-service rate represents the percentage of program-supported equipment that is installed.

Figure 12: Distribution of Electric Savings – Tracking Data



By premise, Figure 13 shows the Brightline electric savings estimates plotted against the NGRID tracking data savings estimates. A few data points were not included for scaling reasons. The figure certainly shows a positive linear trend, though there are a number of premises for which the Brightline billing analysis produced a negative savings estimate (i.e., increased consumption in the post period). The negative values may be indicative of premises that are undergoing expansion (e.g., producing more widgets per day) or some other change between the pre period and post period other than the ECM itself and weather. Recall that one key assumption used herein is that the only differences between the two periods are the ECM and the weather. Any other changes that effect energy consumption will get confounded with the savings estimate produced via the billing analysis.

Figure 13: Comparison of MWh Savings



Due to inherent noise in the data, there is some error in the savings estimates produced by the billing analysis. For this reason, the Brightline team also calculated a 95% savings confidence interval for each premise. Such an interval is a range of values which we expect to contain the true savings value (assuming the only differences between the pre-installation and post-installation periods are the ECM itself and weather conditions). We compared the NGRID savings values with these intervals. Results are shown in Table 9. Most commonly, the NGRID savings value fell above the confidence interval calculated by the Brightline team. Of course, there is also uncertainty associated with the National Grid tracking estimates. Although they are reported as integer values, TRM assumptions and engineering rules-of-thumb are generally based on averages from studies that also have a margin of error.

Table 9: Confidence Interval Comparisons – Electric

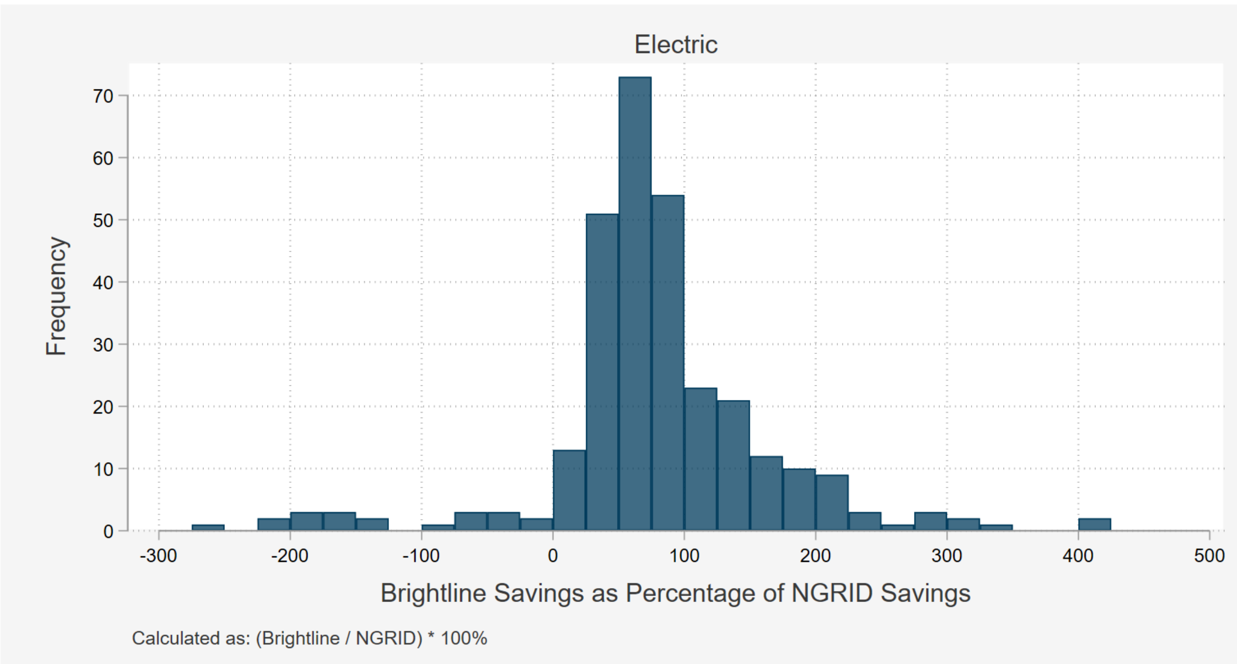
Result	Count	Average NGRID Savings (kWh)	Average Brightline Savings (kWh)
NGRID Estimate Below Savings Interval	50	135,472	222,594
NGRID Estimate Within Savings Interval	84	165,794	159,815
NGRID Estimate Above Savings Interval	164	222,087	86,751

The Brightline team also looked at savings ratios. For each premise, the Brightline team calculated a savings ratio as:

$$Savings\ Ratio = \frac{Brightline\ Savings\ Estimate}{NGRID\ Savings\ Estimate} * 100\% \quad (3)$$

In cases where the Brightline and NGRID savings estimates were virtually equal, this ratio would be near 100%. In cases where the Brightline savings estimate was much larger, the ratio would be much larger than 100%. Similarly, in cases where the Brightline savings estimate was much lower than the NGRID estimate, the ratio would be much lower than 100%. Figure 14 shows the distribution of these ratios for the electric billing analysis.[15] A negative ratio indicates that the Brightline billing analysis produced a negative savings estimate (e.g., consumption increases after the ECM is installed). The average and median of the distribution are 80% and 72%, respectively – this is approximately where the distribution in Figure 14 peaks. A 95% confidence interval for the average ratio was calculated to be (70.3%, 89.8%).

Figure 14: Distribution of Electric Savings Ratios



The Brightline team reviewed the results by percent savings bins, number of measures installed, building type, and measure type (custom or prescriptive). Results for these breakouts are discussed below.

2.3.1.4.1 Results by Percent Savings

The calculation of percent savings was discussed in 2.3.1.2. Importantly, percent savings were calculated using the NGRID savings estimate in the numerator rather than savings estimated calculated by the Brightline team. Our team divided premises into 10 bins based on percent savings –

[15] One of the post-analysis filters noted in Table 4 in Section 2.3.1.2 concerned savings ratios. For reporting, our team removed premises that fell more than three standard deviations away from the mean savings ratio value. The distribution shown in Figure 14 has already been trimmed. In total, eight electric customers were removed from the reporting component of this report based on this filter.

5%-10%, 10%-20%, 20%-30%, and so on.[16] Summary stats for each bin are shown in Table 10. The table also shows average and median savings ratios for each bin. (Equation (3) shows how savings ratios are calculated at the premise level.) In the first two bins, the estimates produced by the Brightline team exceed the estimates in the NGRID tracking data. For the larger percent savings bins, the difference between NGRID savings estimates and Brightline savings estimates grows. In the larger bins, there is an increased likelihood that there is a data mismatch between the premise number in the tracking data and the premise (or premises) where the savings occur. Or there may be some non-measure, non-program issue going on with the site that has not been identified. In such cases, we would expect the billing analysis to underestimate savings, as the billing analysis will only measure savings for one premise.

Table 10: Results by Percent Savings Bin – Electric

Percent Savings Bin	# of Premises	Statistic	NGRID (kWh)	Brightline (kWh)	Savings Ratio (%)
5% - 10%	34	Average	137,138	169,860	164
		Median	42,132	76,054	162
10% - 20%	74	Average	148,016	163,333	107
		Median	51,345	51,249	106
20% - 30%	86	Average	135,425	98,630	61
		Median	50,162	34,277	70
30% - 40%	31	Average	226,046	135,704	69
		Median	78,435	36,530	68
40% - 50%	25	Average	248,436	123,357	40
		Median	46,593	28,442	52
50% - 60%	20	Average	435,792	232,890	56
		Median	233,752	82,524	51
60% - 70%	8	Average	485,483	-112,400	34
		Median	130,676	38,398	43
70% - 80%	8	Average	187,727	67,196	37
		Median	127,091	26,512	35
80% - 90%	5	Average	51,834	15,428	34

[16] Recall that premises with percent savings less than 5% were filtered out of the analysis data set. This is why the first bin runs from 5% to 10% rather than 0% to 10%.

Percent Savings Bin	# of Premises	Statistic	NGRID (kWh)	Brightline (kWh)	Savings Ratio (%)
		Median	18,567	5,279	23
90% - 100%	7	Average	325,884	110,457	21
		Median	327,212	120,142	27

2.3.1.4.2 Results by Number of Measures Installed

A number of premises installed multiple retrofit measures.[17] For such measures, our billing analysis treated the pre-installation period as the period before any measures were installed. The post-installation period was the period after all measures were installed. Thus, the savings estimate derived from the billing analysis is an estimate of the combined savings, not the savings for any particular measure.

The majority of sites installed three measures or fewer, but there were a number of premises with a measure count in the 20s and one with a measure count above 100. We reviewed results by the number of measures installed, though we took a binary approach here. The comparison bins were "just one measure" and "more than one measure." Table 11 shows the results. The table also shows average and median savings ratios for each bin. (Equation (3) shows how savings ratios are calculated at the premise level.) For sites with just one measure installed, the Brightline savings estimates came out slightly higher than the NGRID estimates, on average. For sites with multiple measures installed, Brightline savings estimates were lower than the NGRID estimates, on average.

Table 11: Results by Number of Measures – Electric

# of Measures	# of Premises	Statistic	NGRID (kWh)	Brightline (kWh)	Savings Ratio (%)
Just one	45	Average	70,663	75,618	94
		Median	19,343	23,969	117
More than one	253	Average	213,213	139,836	78
		Median	57,900	43,852	70

2.3.1.4.3 Results by Building Type

Results by building type are shown in Table 12. Building type assignments were drawn from the tracking data. Not all premises were assigned a building type. Tables entries are sorted by savings

[17] We're defining measures as unique rows in the tracking data. It's possible that one single tracking data entry could incorporate multiple technologies, such as a lighting upgrade combined with occupancy sensors. Such an example would be counted as one measure in this analysis.

ratio (using the median value) from least to greatest. For building types at the top of the table, there was not much agreement between the tracking data savings estimate and the billing analysis findings. Towards the bottom of the table, there was more agreement between the savings estimates. Warehouses produced some of the largest discrepancies, with an average savings value of 317,970 kWh in the tracking data and -21,274 kWh according to the billing analysis.

Table 12: Results by Building Type – Electric

Building Type	# of Premises	Statistic	NGRID (kWh)	Brightline (kWh)	Savings Ratio (%)
Hospital	1	Average	83,791	-162,039	-193
		Median	83,791	-162,039	-193
Small Office	8	Average	53,299	2,240	-23
		Median	28,637	11,407	24
Refrigerated Space	2	Average	87,652	41,444	48
		Median	87,652	41,444	48
Warehouse	11	Average	317,970	-21,274	40
		Median	21,905	10,215	57
Multifamily low-rise	2	Average	133,522	66,336	60
		Median	133,522	66,336	60
Secondary School	4	Average	188,348	87,484	59
		Median	98,027	65,719	65
Other	59	Average	225,772	130,950	72
		Median	51,461	30,160	65
Primary School	16	Average	145,820	63,157	54
		Median	126,799	63,450	66
Multi Story Retail	3	Average	199,789	198,452	86
		Median	179,166	117,974	66
University	5	Average	266,529	101,287	58
		Median	174,009	116,553	67
Fast Food Restaurant	67	Average	47,327	34,727	78
		Median	46,844	32,363	70
Large Office	17	Average	149,758	151,066	76
		Median	50,868	31,489	72
Light Industrial	10	Average	362,141	408,552	76
		Median	187,858	155,124	77
Heavy Industrial	8	Average	483,116	402,047	117
		Median	259,184	283,798	89
Religious	4	Average	54,719	8,364	73
		Median	6,276	10,071	93
Grocery	24	Average	400,379	274,143	98
		Median	284,834	272,954	95
Big Box Retail	17	Average	399,264	350,766	138

Building Type	# of Premises	Statistic	NGRID (kWh)	Brightline (kWh)	Savings Ratio (%)
		Median	192,629	127,271	99
Automobile	5	Average	73,849	71,643	86
		Median	46,593	46,805	100
Multifamily high-rise	4	Average	192,579	172,322	121
		Median	221,016	186,344	108
Small Retail	24	Average	46,826	36,811	114
		Median	17,079	25,543	114
Hotel	7	Average	322,064	293,141	116
		Median	177,027	123,580	128

2.3.1.4.4 Results by Measure Type

Some premises installed only custom measures, some premises installed only prescriptive measures, and some premises installed a mixture of custom and prescriptive measures. Table 13 compares National Grid savings with Brightline savings for these three bins. The table also shows average and median savings ratios for each bin. (Equation (3) shows how savings ratios are calculated at the premise level.) Savings ratios were highest for premises that only installed prescriptive retrofit measures (i.e., no custom measures).

Table 13: Results by Measure Type – Electric

Type of Measure	# of Premises	Statistic	NGRID (kWh)	Brightline (kWh)	Savings Ratio (%)
Custom measures	82	Average	137,887	90,640	79
		Median	51,587	36,218	71
Prescriptive measures	178	Average	161,842	105,988	83
		Median	47,762	35,433	77
Mixture of measures	38	Average	447,578	328,503	68
		Median	231,219	191,552	71

2.3.1.4.5 Results by Technology

Lighting measures accounted for approximately 80% of the records in the electric tracking data and approximately 70% of the savings. To investigate whether the two savings estimates were better aligned for lighting measures, we put each premise in one of three bins: no lighting measures, some lighting measures, and only lighting measures. Premises in the first bin did not have any lighting measures in the tracking data, while premises in the other bins had at least one lighting measure in the tracking data. Premises in the "only lighting measures" did not show any non-lighting measures in the tracking data. Comparisons are shown in Table 14. In each bin, average and median estimates from the billing analysis were less than average and median estimates from the tracking data. The table also shows average and median savings ratios for each bin. (Equation (3) shows how savings ratios are calculated at the premise level.) On average, savings ratios were better in the "no lighting upgrades" bin compared to the other two bins.

Table 14: Results by the Presence of Lighting Upgrades

Lighting Bin	# of Premises	Statistic	NGRID (kWh)	Brightline (kWh)	Savings Ratio (%)
No Lighting Upgrades	36	Average	182,639	114,340	105
		Median	116,483	76,034	89
Some Lighting Upgrades	110	Average	264,201	189,693	66
		Median	57,538	42,623	70
Only Lighting Upgrades	152	Average	141,352	90,782	84
		Median	42,020	29,072	80

2.3.1.4.6 Results by Annual Consumption Bins

Table 15 shows the results by customer size (based on pre-retrofit annual consumption). For this breakout, five annual consumption bins were defined: less than 50 MWh, between 50 and 100 MWh, between 100 and 500 MWh, between 500 and 1,000 MWh, and greater than 1,000 MWh. Savings ratios were highest in the 50-100 MWh bin and lowest in the under 50 MWh bin. Overall, no patterns related to annual consumption bin are present other than savings ratios being lowest in the smallest consumption bin. The BrightLine team ran the same comparison using non-weather-normalized savings and found nearly identical results. These results are shown in Table 16. Greater discussion on non-weather-normalized savings is presented in Section 2.3.2.2.

Table 15: Results by Annual Consumption – Weather-Normalized

Annual Consumption	# of Premises	Statistic	NGRID (kWh)	Brightline (kWh)	Savings Ratio (%)
< 50 MWh	35	Average	10,583	4,713	45
		Median	9,198	4,577	60
50 - 100 MWh	27	Average	24,108	18,037	98
		Median	17,015	17,643	93
100 - 500 MWh	134	Average	59,498	37,270	77
		Median	49,344	33,877	70
500 - 1,000 MWh	36	Average	227,829	132,758	97
		Median	178,896	118,757	65
> 1,000 MWh	66	Average	604,951	429,635	88
		Median	409,367	323,718	86

Table 16: Results by Annual Consumption – Non-Weather-Normalized

Annual Consumption	# of Premises	Statistic	NGRID (kWh)	Brightline (kWh)	Savings Ratio (%)
< 50 MWh	35	Average	10,583	4,597	43
		Median	9,198	4,602	56
50 - 100 MWh	27	Average	24,108	18,119	96
		Median	17,015	15,740	93
100 - 500 MWh	134	Average	59,498	38,088	79
		Median	49,344	34,434	71
500 - 1,000 MWh	36	Average	227,829	133,422	99
		Median	178,896	114,744	63
> 1,000 MWh	66	Average	604,951	429,337	89
		Median	409,367	311,515	84

2.3.1.5 Results – Gas

All impact estimates produced by the Brightline billing analysis were found to be statistically significant at the 95% confidence level.[18] Table 17 shows average and median annual therm savings for the premises that remained in our gas analysis file. Both NGRID tracking data savings and the savings estimate from the Brightline billing analysis are shown. The average savings estimate for the Brightline billing analysis is higher than the average from the tracking data, but this is solely due to the premise with the largest savings value – the Brightline savings estimate for this premise was more than twice as large as the NGRID estimate. This premise had 154 measures installed (153 steam trap measures accounting for 452,063 annual therms saved and one custom insulation measure account for 22,206 annual therms saved). Without this premise, the average savings values would be 11,785 therms and 6,724 therms for NGRID and Brightline respectively. The median is not affected by this one premise and is a better measure of central tendency for this collection of premises. The median savings from the Brightline analysis comes in around 35% of the NGRID estimate.

Table 17: Savings Comparison – Gas

Fuel	Statistic	# of Premises	NGRID	Brightline
Gas	Sum	34	863,176 therms	1,357,691 therms
Gas	Average	34	25,388 therms	39,932 therms[1]
Gas	Median	34	6,335 therms	2,264 therms

[1] A 95% confidence interval for the average savings value was calculated to be (-29,156, 109,020).

Note that the tracking data savings values shown in Table 17 are not adjusted with realization rates and in-service rates.[19] The savings values that National Grid actually claims have these adjustments factored in. Table 18 compares unadjusted savings values, adjusted savings values, and the savings values calculated via the billing analysis. The bottom three rows remove the site with 154 measures from the totals. The remainder of the figures, tables, and comparisons in this report use the unadjusted savings values.

[18] This is as expected. Filter 7 in Table 4 removes premises where uncertainty exceeds the savings estimate which would render the estimate statistically insignificant.

[19] A realization rate represents the percentage of claimed savings that are actually realized. Consider a lighting upgrade in a building where hours of use are believed to be 5,000 hours annually. If lights in this building are actually on for 4,500 hours annually, then the actual savings value will be less than the expected savings value since the initial assumption overstated use by 500 hours. An in-service rate represents the percentage of program-supported equipment that is installed.

Table 18: Comparison with Adjusted Savings Values

Source	# of Premises	Total Savings (therms)	% Difference Relative to Unadjusted Savings
NGRID – Unadjusted	34	863,176	---
NGRID – Adjusted	34	796,028	-8%
Brightline Analysis	34	1,357,691	+57%
NGRID – Unadjusted	33	388,907	---
NGRID – Adjusted	33	324,202	-16%
Brightline Analysis	33	221,879	-42%

Figure 15 and Figure 16 show the distribution of savings derived via the billing analysis and in the tracking data, respectively.

Figure 15: Distribution of Gas Savings – Billing Analysis

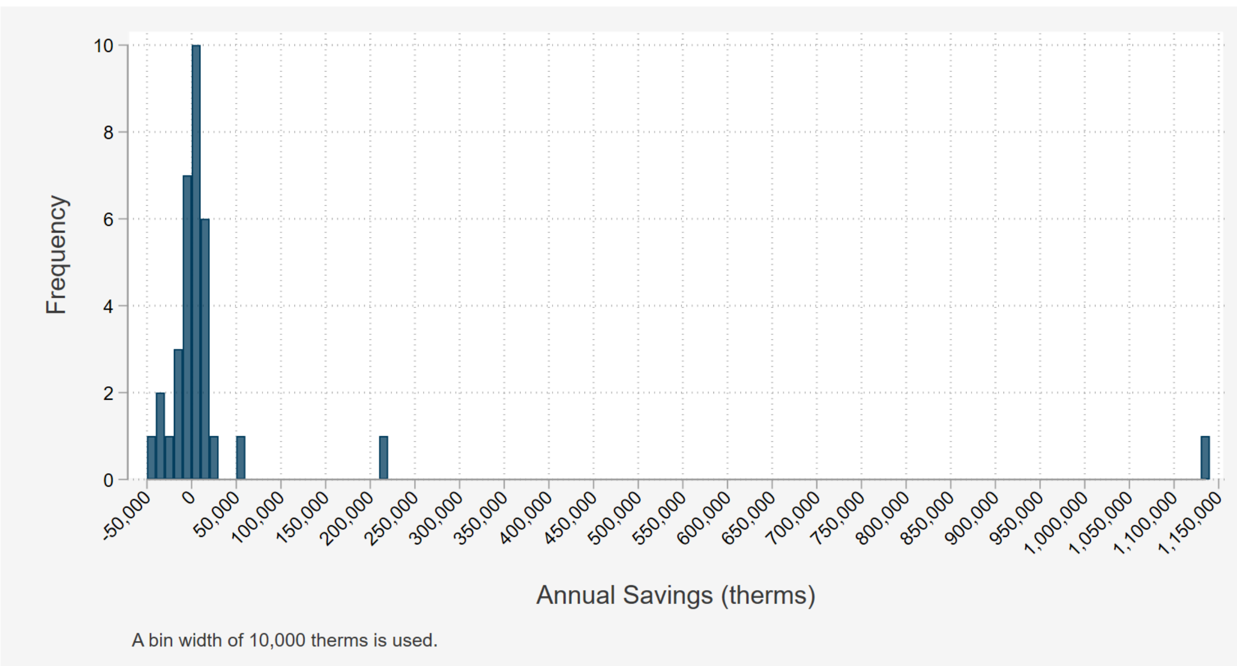
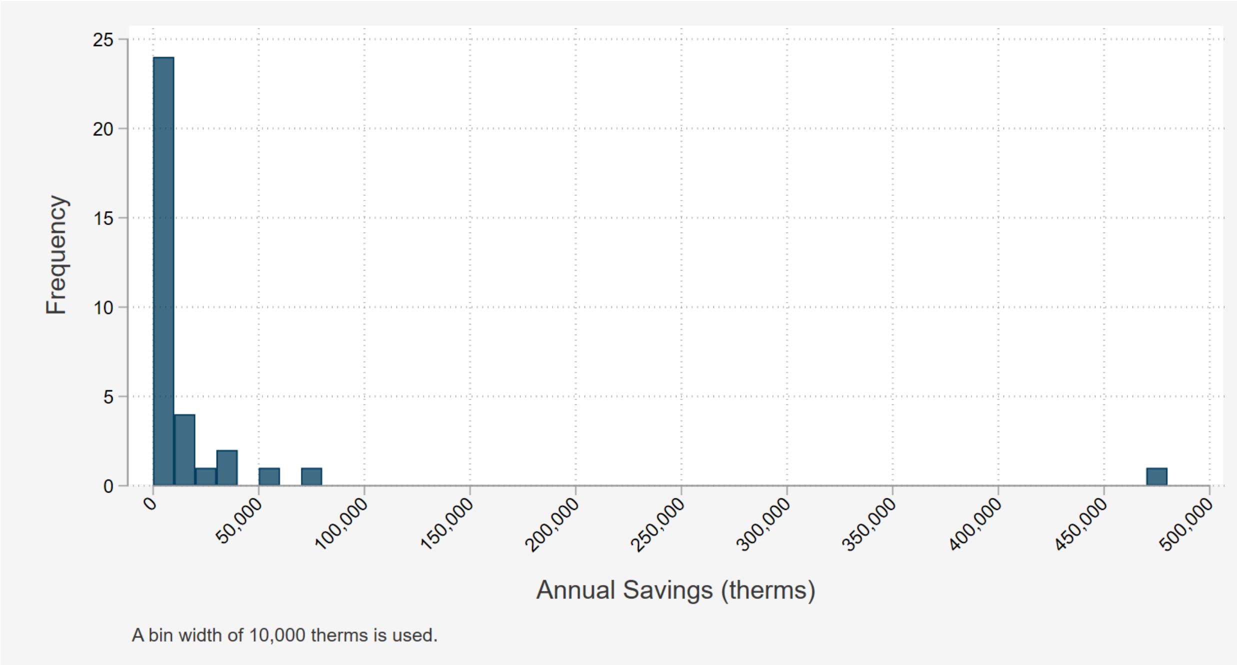
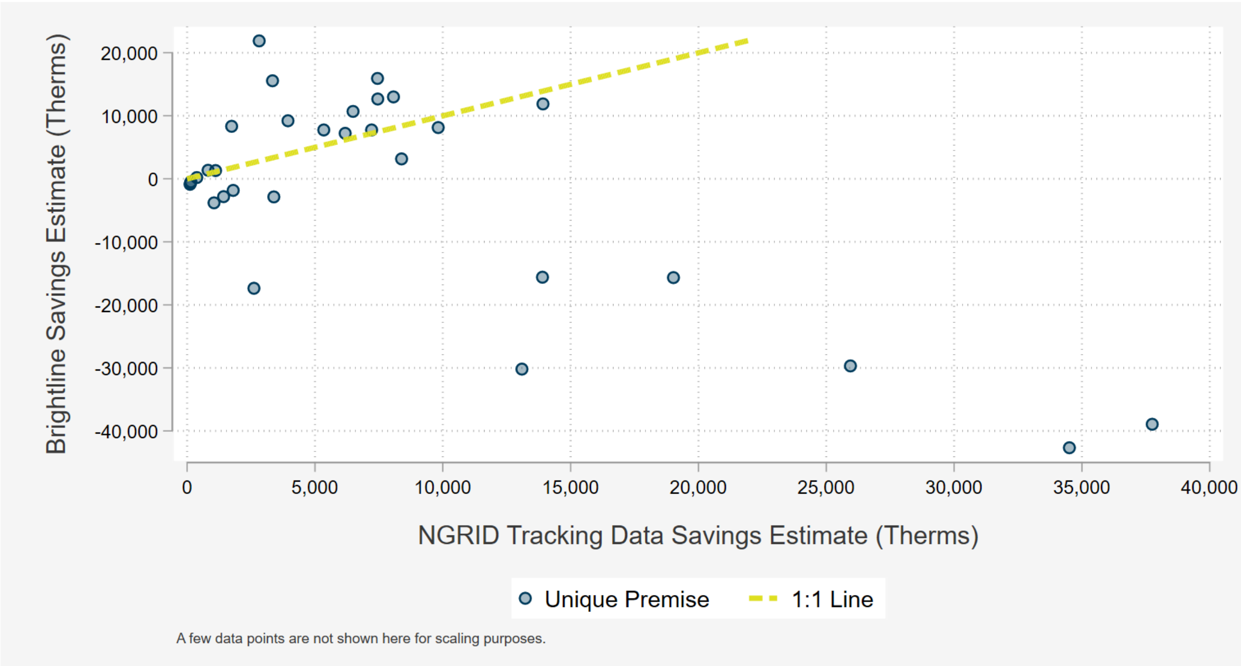


Figure 16: Distribution of Gas Savings – Tracking Data



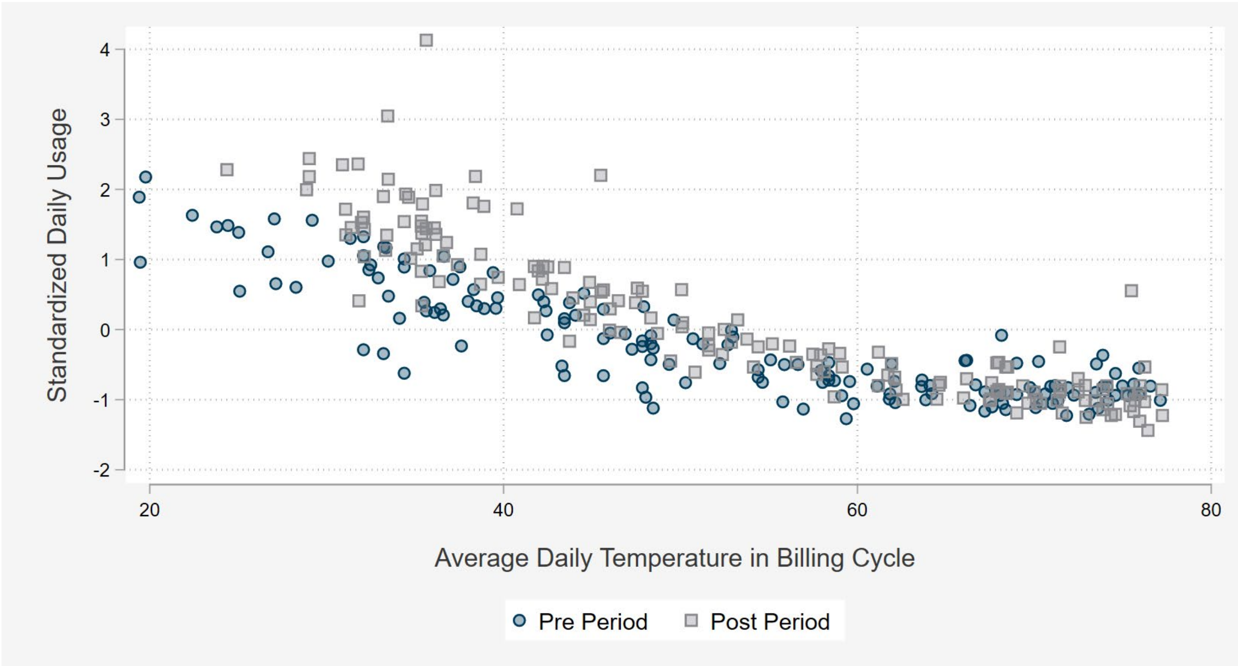
By premise, Figure 17 shows the Brightline gas savings estimates plotted against the NGRID tracking data savings estimates. A few data points were not included in the figure for scaling reasons. Unlike the trend in the figure showing kWh impacts (Figure 13), there is not much of a trend between Brightline gas savings estimates and NGRID gas savings estimates. Additionally, a higher percentage of the Brightline gas savings estimates are negative (compared to the electric savings estimates). As with the negative savings estimates for electric, it might be the case that the negative values are indicative of premises that are undergoing expansion (e.g., producing more widgets per day), other non-routine events, or meter-matching issues. The billing analysis will confound the effects of the expansion and the ECM, though the second component of Task 3 (site visits and interviews) could provide additional insights for such premises.

Figure 17: Comparison of Therms Savings



Our team reviewed the raw billing data for any premise for which the Brightline savings estimate was less than -10,000 therms per year. Indeed, the raw billing data for these premises showed an increase in average daily consumption in the post-installation period relative to the pre-installation period. Figure 18 shows the results using standardized consumption values as a function of outdoor temperature. "Standardizing" is a transformation of the data that puts all of the premises on a common consumption scale (i.e., it removes differences in magnitude and spread), which makes it easier to compare trends across premises.

Figure 18: Standardized Daily Gas Consumption for Premises with Negative Savings



Due to inherent noise in the data, there is some error in the savings estimates produced by the billing analysis. For this reason, the Brightline team also calculated a 95% savings confidence interval for each premise. Such an interval is a range of values which we expect to contain the true savings value (assuming the only differences between the pre-installation and post-installation periods are the ECM itself and weather conditions). We compared the NGRID savings values with these intervals. Results are shown in Table 19. Most commonly, the NGRID savings value fell above the confidence interval calculated by the Brightline team.

Table 19: Confidence Interval Comparisons – Gas

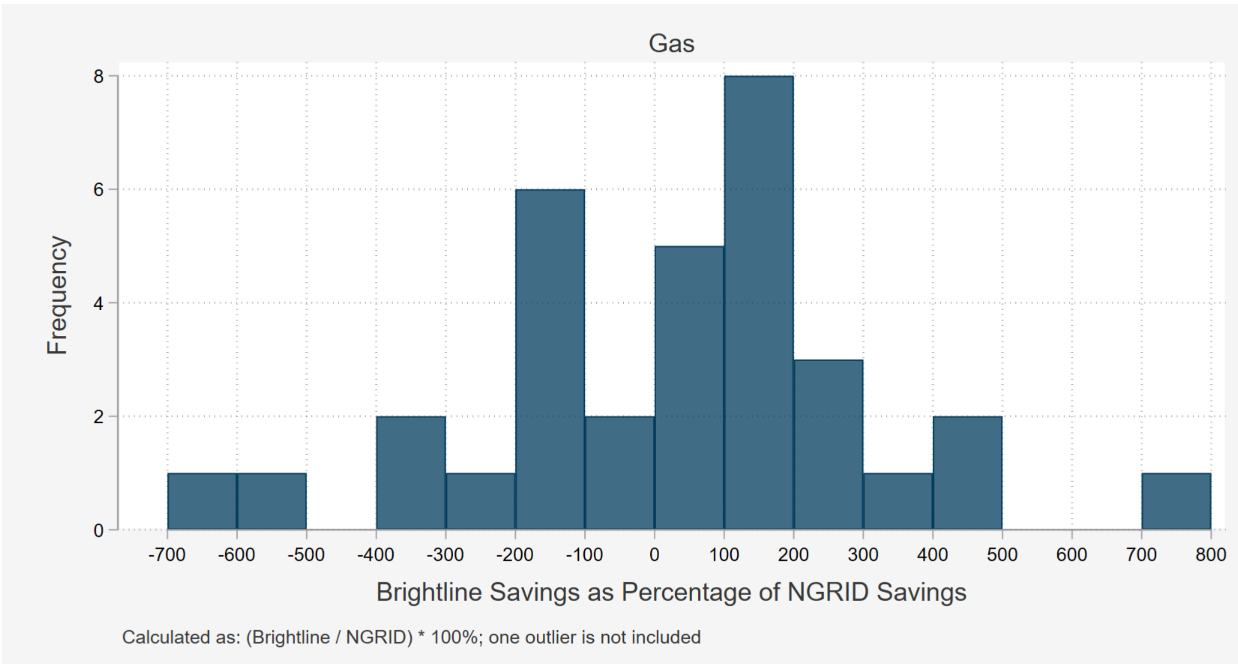
Result	Count	Average NGRID Savings (therms)	Average Brightline Savings (therms)
NGRID Estimate Below Savings Interval	9	63,045	160,872
NGRID Estimate Within Savings Interval	8	6,560	7,308
NGRID Estimate Above Savings Interval	17	14,311	-8,742

The Brightline team also looked at savings ratios. For each premise, the Brightline team calculated a savings ratio as:

$$Savings\ Ratio = \frac{Brightline\ Savings\ Estimate}{NGRID\ Savings\ Estimate} * 100\% \tag{4}$$

In cases where the Brightline and NGRID savings estimates were virtually equal, this ratio would be near 100%. In cases where the Brightline savings estimate was much larger, the ratio would be much larger than 100%. Similarly, in cases where the Brightline savings estimate was much lower than the NGRID estimate, the ratio would be much lower than 100%. Figure 19 shows the distribution of these ratios for the gas billing analysis.[20] A negative ratio indicates that the Brightline billing analysis produced a negative savings estimate (e.g., consumption increases after the ECM is installed). There is a considerable amount of spread in the figure, with ratios ranging from -700% up to 800%. This implies there is not much agreement between the tracking data savings estimate and the billing analysis savings estimate.

Figure 19: Distribution of Gas Savings Ratios



The Brightline team reviewed the results by percent savings bins and by number of measures installed. (Building types were sparsely populated for gas measures, so that breakout was not examined for gas savings.)

2.3.1.5.1 Results by Percent Savings

The calculation of percent savings was discussed in 2.3.1.2. Importantly, percent savings were calculated using the NGRID savings estimate in the numerator rather than savings estimated calculated by the Brightline team. Our team divided premises into bins based on percent savings – 5%-10%, 10%-20%, 20%-30%, and so on. Summary stats for each bin are shown in Table 20. The table also shows average and median savings ratios for each bin. (Equation (4) shows how savings ratios are calculated at the premise level.) In each bin, both median and average Brightline estimates fall below the NGRID estimates with one exception – the Brightline average for the 5% – 10% bin. The

[20] One of the post-analysis filters noted in Table 4 in Section 2.3.1.2 concerned savings ratios. For reporting, our team removed premises that fell more than three standard deviations away from the mean savings ratio value. The distribution shown in Figure 19 has already been trimmed. In total, two gas customers were removed from the reporting component of this report based on this filter.

NGRID average savings estimate is much higher in this bin. This is due to the site with 154 measures that has been mentioned previously. (The Brightline savings estimate was more than twice the NGRID estimate for this site, and this site produced more savings than any others in the final gas analysis file.)

Table 20: Results by Percent Savings Bin – Gas

Percent Savings Bin	# of Premises	Statistic	NGRID (therms)	Brightline (therms)	Savings Ratio (%)
5% - 10%	11	Average	51,185	123,261	22
		Median	2,818	8,338	234
10% - 20%	12	Average	8,365	665	18
		Median	6,852	4,559	113
20% - 30%	3	Average	14,407	-2,956	46
		Median	9,820	8,139	83
30% - 40%	4	Average	30,985	4,978	33
		Median	21,965	3,716	60
40% - 50%	0	Average	--	--	--
		Median	--	--	--
50% - 60%	4	Average	8,150	-4,299	-58
		Median	5,889	-2,347	-83
60% - 70%	0	Average	--	--	--
		Median	--	--	--
70% - 80%	0	Average	--	--	--
		Median	--	--	--
80% - 90%	0	Average	--	--	--
		Median	--	--	--
90% - 100%	0	Average	--	--	--
		Median	--	--	--

2.3.1.5.2 Results by Number of Measures Installed

A number of premises installed multiple retrofit measures.[21] For such measures, our billing analysis treated the pre-installation period as the period before any measures were installed. The post-installation period was the period after all measures were installed. Thus, the savings estimate derived from the billing analysis is an estimate of the combined savings for premises that installed multiple measures. Using three bins, Table 21 shows the results by the number of measures installed. The bins are: just one measure installed, more than one measure installed, and 154 measures installed. The second bin ("more than one measure" does not include the premise with 154 measures.) The binned results are largely in line with the overall results – the Brightline savings estimates are smaller than the NGRID savings estimates with the exception of the site with 154 measures, which has been noted previously. The Brightline savings estimate is more than twice the NGRID savings estimate for this site.

Table 21: Results by Number of Measures – Gas

# of Measures	# of Premises	Statistic	NGRID (therms)	Brightline (therms)	Savings Ratio (%)
Just one	14	Average	5,077	251	-38
		Median	2,176	1,338	101
More than one	19	Average	16,728	11,493	41
		Median	8,070	3,168	38
154	1	Average	474,269	1,135,812	239
		Median	474,269	1,135,812	239

2.3.1.5.3 Results by Measure Type

Some premises installed only custom measures, some premises installed only prescriptive measures, and some premises installed a mixture of custom and prescriptive measures. Table 22 compares National Grid savings with Brightline savings for these three bins. There were no clear takeaways from this breakdown.

[21] We're defining measures as unique rows in the tracking data. It's possible that one single tracking data entry could incorporate multiple technologies. Such cases would be counted as one measure in this analysis.

Table 22: Results by Measure Type – Gas

Type of Measure	# of Premises	Statistic	NGRID (therms)	Brightline (therms)	Savings Ratio (%)
Only Custom	27	Average	10,854	5,380	56
		Median	6,181	1,359	85
Only Prescriptive	3	Average	131	-667	-528
		Median	131	-678	-518
Mixture of measures	4	Average	142,434	301,581	139
		Median	44,011	33,671	139

2.3.1.5.4 Results by Annual Consumption Bins

Table 23 shows the results by customer size (based on pre-retrofit annual consumption). For this breakout, three annual consumption bins were defined: less than 25,000 therms, between 25,000 and 100,000 therms, and greater than 100,000 therms. No patterns related to annual consumption bin are present. The Evergreen team ran the same comparison using non-weather-normalized savings and found nearly identical results. These results are shown in Table 24. Greater discussion on non-weather-normalized savings is presented in Section 2.3.2.3.

Table 23: Results by Annual Consumption – Weather-Normalized

Annual Consumption	# of Premises	Statistic	NGRID (therms)	Brightline (therms)	Savings Ratio (%)
< 25,000 therms	13	Average	2,052	639	-104
		Median	1,112	-463	-84
25,000 - 100,000 therms	13	Average	7,498	5,694	128
		Median	7,215	9,221	161
> 100,000 therms	8	Average	92,378	159,421	23
		Median	36,129	-8,892	-19

Table 24: Results by Annual Consumption – Non-Weather-Normalized

Annual Consumption	# of Premises	Statistic	NGRID (therms)	Brightline (therms)	Savings Ratio (%)
< 25,000 therms	13	Average	2,052	622	-82
		Median	1,112	-382	-84
25,000 - 100,000 therms	13	Average	7,498	5,178	119
		Median	7,215	8,716	145
> 100,000 therms	8	Average	92,378	168,312	35
		Median	36,129	-6,604	-10

2.3.2 Avoided Energy Use

2.3.2.1 Approach

The avoided energy use approach removes the notion of a "typical weather year" from the problem. That is, instead of estimating what savings would be in a typical weather year, we estimate savings observed at the meter. This approach mimics the viewpoint of the customer, as the customer only sees bills (not what their bills would be under a different set of weather conditions).

Similar to the weather-normalized approach, the avoided energy use approach begins by separating the billing records into pre-installation and post-installation periods. From there, we use the pre-period billing records to create a mathematical model that relates consumption and the relevant temperature variables (CDD and HDD):

$$Avg. Daily\ Consumption = Intercept + \beta_1 \times (Avg. Daily\ CDD) + \beta_2 \times (Avg. Daily\ HDD)$$

The intercept and slope terms (β_1 and β_2) will vary from premise to premise. After developing the model, the relationship is cast over the weather conditions observed in the post-installation period (as opposed to TMY conditions). This provides an estimate of what post-installation consumption would have been if the pre-installation relationship between consumption and temperature had carried into the post period. Avoided energy use is the difference between estimated post-installation consumption and actual post-installation consumption (then annualized as necessary).

2.3.2.2 Results – Electric

The estimates of avoided energy use were largely in line with our estimates of weather-normalized savings. Figure 20 compares these two estimates for the 298 premises that remained in the electric analysis file. There is some minor variation but there is a positive linear correlation between the savings estimates for the two methods.

Figure 20: Comparison of Methods – Electric

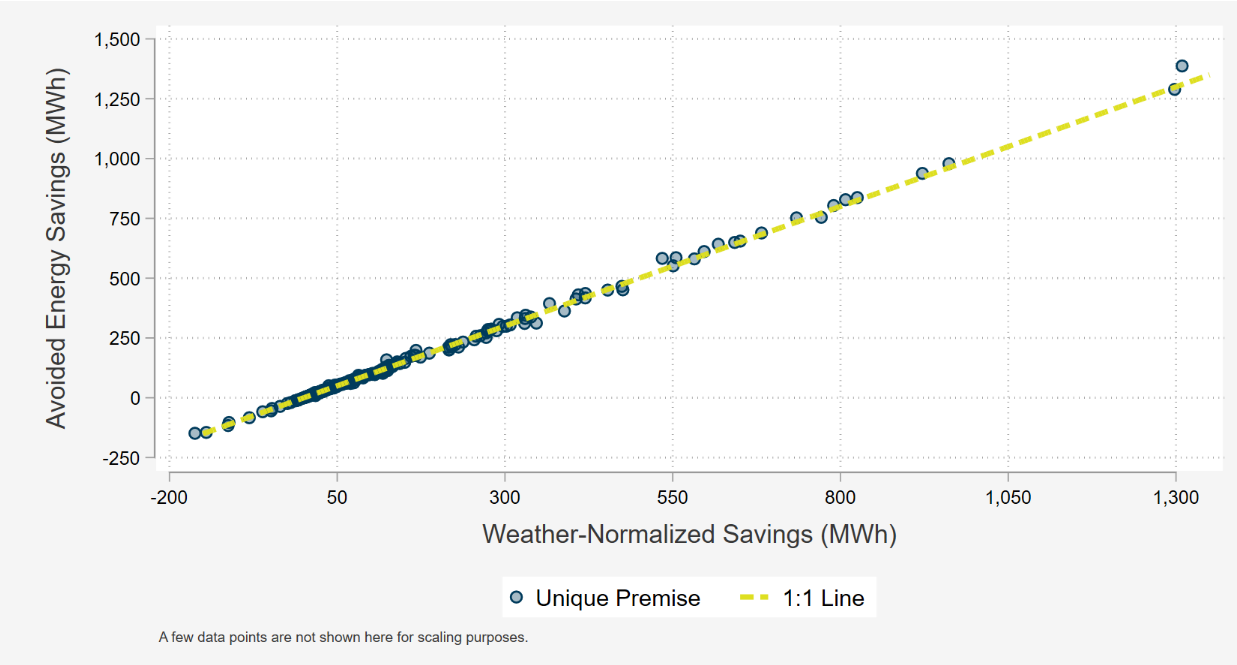


Table 25 shows the average and median savings values using the avoided energy approach. The same values are shown for the NGRID tracking data savings estimates and the weather-normalized savings estimate derived via billing analysis. Average savings using the avoided energy use approach are slightly lower than average savings using the weather-normalized approach, but median savings using the avoided energy use approach are slightly higher. Additional breakdowns (by percent savings, by building type, by number of measures installed, by measure type) are not shown in this section since the avoided energy use estimates were so strongly correlated with the weather-normalized savings estimates (Figure 20). The results would essentially be identical.

Table 25: Annual kWh Savings by Method

Metric	Number of Premises	NGRID Savings	Weather-Normalized Savings	Avoided Energy Use
Average	298	191,687 kWh	130,139 kWh	130,514 kWh
Median	298	53,328 kWh	39,733 kWh	41,197 kWh

2.3.2.3 Results – Gas

Like with electric savings, the estimates of avoided energy use for gas were largely in line with our estimates of weather-normalized gas savings. Figure 21 compares these two estimates for the 34 premises that remained in the final gas analysis file. There is some minor variation, but the overall trend is quite clear. A few data points with large savings are not shown for scaling purposes.

Figure 21: Comparison of Methods

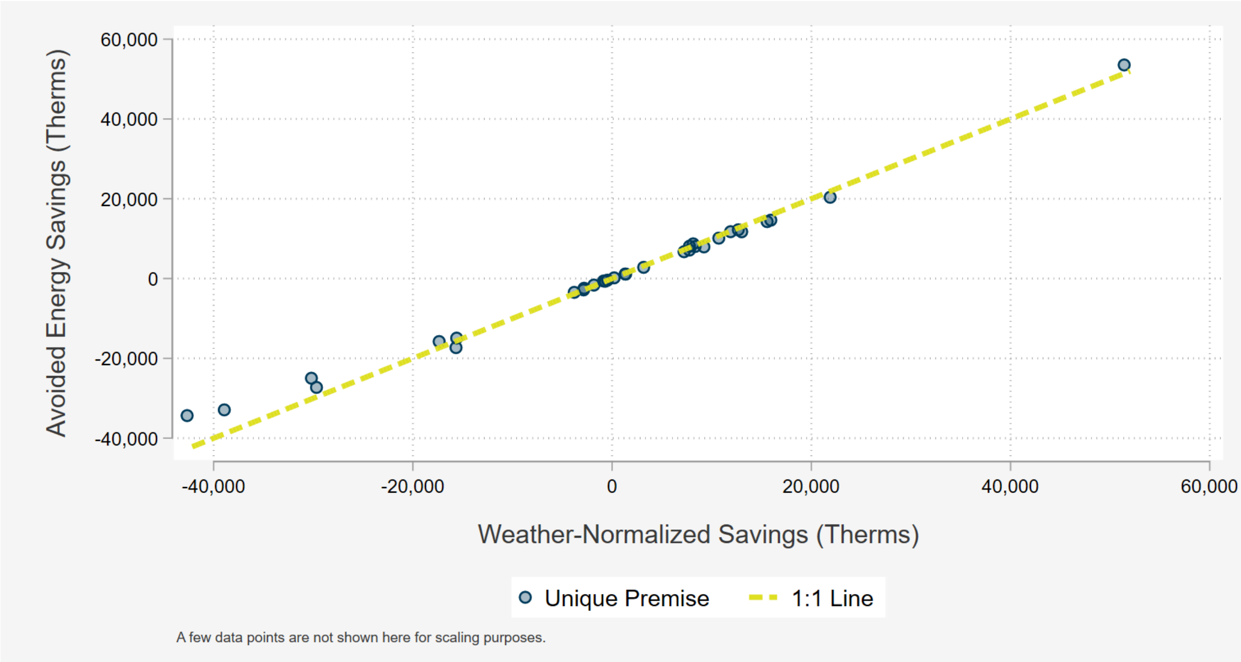


Table 26 shows the average and median savings values using the avoided energy approach. The same values are shown for the NGRID tracking data savings estimates and the weather-normalized savings estimate derived via billing analysis. Average savings using the avoided energy use approach are slightly higher than average savings using the weather-normalized approach, but median savings using the avoided energy use approach are slightly lower. Additional breakdowns (by percent savings, by number of measures installed) are not shown in this section since the avoided energy use estimates were so strongly correlated with the weather-normalized savings estimates (Figure 21).

Table 26: Annual Therms Savings by Method

Metric	Number of Premises	NGRID Savings	Weather-Normalized Savings	Avoided Energy Use
Average	34	25,388 therms	39,932 therms	41,820 therms
Median	34	6,335 therms	2,264 therms	1,984 therms

2.3.2.4 Potential Issues with a Billing Analysis

As has been noted, a billing analysis is not the best tool for measuring savings for all premise/measure combinations. Predictable load patterns will facilitate the precise measurement of savings, as will relatively larger effect sizes. If a site's load pattern is highly volatile and cannot be linked to other independent variables, or if the effect of the ECM is expected to be small relative to annual consumption, a billing analysis may not return a reliable or precise estimate of savings. The scope of this research was undeniably ambitious. This section discusses some potential threats to the validity of the findings from a billing analysis. Some of the factors discussed impact the precision of the savings estimate and others may lead to bias.

2.3.2.4.1 Non-routine Events

Any non-routine events, such as a temporary closure in the pre-installation or post-installation period, confound the energy impact of the ECM if not properly accounted for. During the data cleaning process, our team took steps to tag potential outliers. Such records were not included in the analysis. Still, it's possible that some non-routine events remain unaccounted for. The effect a non-routine event would have on the savings estimate depends on when the non-routine event occurs and the directional impact on consumption. For example, if a non-routine event occurs in the post-installation period and lowers consumption, then consumption in the post period will be less than what it would have been absent the non-routine event. This discrepancy will feed into our regression coefficients – they will under-estimate the true relationship between consumption and weather. Thus, the weather-normalized estimate of post period consumption will be understated. This means the savings estimate would be overstated.

In general, there is no blanket statement regarding the directional affect a non-routine event has on the savings estimate derived from a billing analysis. Such an event may result in overestimating savings or underestimating savings – it depends on when the event occurs and the type of non-routine event.

2.3.2.4.2 Economic Factors

A pre/post billing analysis generally assumes that the only differences between the pre-installation period and the post-installation period are the weather and the measure itself. In reality, there are other factors that play into how a business uses energy. One such factor for some businesses is the performance of the economy. For example, a manufacturing warehouse that was hit hard by the recently imposed tariffs may reduce the number of operating days or shifts. Such a scenario could lead to instances where the tracking data savings estimate and the billing analysis savings estimate are very different, as the change in consumption is unrelated to weather or the ECM. Directionally, this could affect the savings estimate in multiple ways. In the example posed in this paragraph, suppose the reduction in operating shifts occurs as the post period begins. Without adding information regarding operational changes to the regression model, the model will attribute the decrease in energy use associated with the operational change to the ECM. Thus, the savings estimate would be overstated. In the example in the next paragraph, the billing analysis would underestimate savings (supposing the expansion occurs primarily in the post period.)

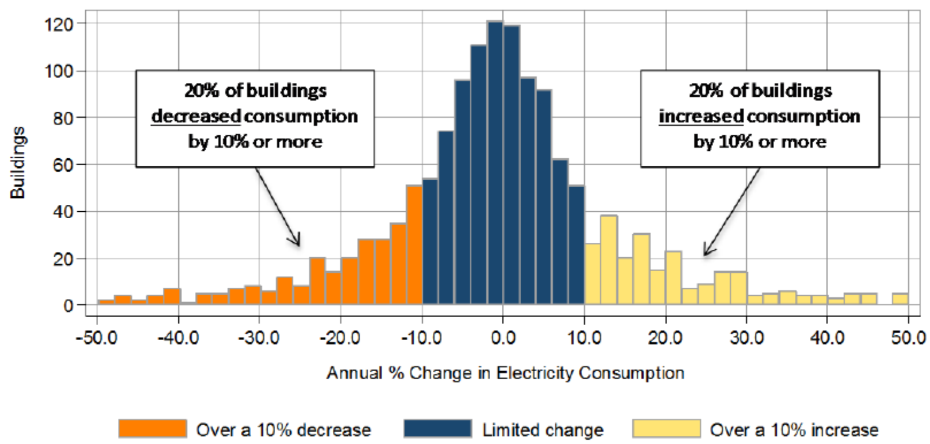
Another example would be a business that is expanding. Such a business may have more incentive to participate in the program, as participation would reduce their per-unit energy costs and also potentially increase their production capacity (depending on the equipment installed). Energy efficiency aside, expansion likely means that energy consumption at the premise is increasing. This increase would be captured by the billing analysis, but the context around the increase would not be understood without more information.

Like with non-routine events, there is no blanket statement describing the directional affect such scenarios have on the savings estimate produced via billing analysis. The direction of the effect depends on how the operational change alters energy use (increased or decreased consumption?) and which analysis period is affected (pre period, post period, or both?).

2.3.2.4.3 Unexplained Variation

Some of the variation in energy use from year to year simply cannot be explained by variables that are readily available. Even absent energy efficiency measures, buildings can experience year to year changes in electricity consumption that are greater than 10% in either direction.[22] This conclusion is drawn from an ACEEE paper that sought to determine how much energy consumption varies from year to year without ECMs. Figure 22, which is drawn from said paper, shows the percent change in annual electricity consumption for a sample of commercial buildings. Importantly, the distribution is centered at zero, but 40% of buildings saw changes in excess of 10% in magnitude.

Figure 22: Year to Year Changes in Energy Consumption Absent EE



Similarly, in a 2013 paper that examined commercial building energy consumption profiles by Phillip Price and David Jump[23], the authors concluded: "In most buildings and most years, the largest source of year-to-year change in energy use is neither energy conservation measures nor year-to-year variation in weather, it is changes in characteristics of building operation and occupant behavior such as operating hours, thermostat settings, the number of occupants, the type of activities performed in the building, and so on." Hence, much of the variation in energy consumption from one year to the next cannot be explained by variables that are readily available.

On average, we would not expect the unexplained variation to result in systematically overstated or understated savings estimates, as the average annual change in electricity consumption in Figure 21

[22] Bode, J, Caririllo, L., Basarkar, M. 2014. Whole Building Energy Efficiency and Energy Savings Estimation: Does Smart Meter Data with Pre-screening Open up Design and Evaluation Opportunities?

[23] Price, P., Jump, D, Granderson, J. Sohn, M. Addy, N. 2013. Commercial Building Energy Baseline Modeling Software: Performance Metrics and Method Testing with Open Source Models and Implications for Proprietary Software Testing

was around 0% – increases in consumption and decreased in consumption are equally likely. Case by case, the effect of year-to-year changes in energy consumption depends on the nature of the change. Consider a hypothetical site that participated in an energy efficiency program but would have fallen in the orange region in Figure 21 even if they hadn't participated. Without including characteristics of building operation and/or occupant behaviors in the regression models, the billing analysis would attribute the decrease in consumption to the ECM. Thus, the billing analysis savings estimate would be too high, as it's truly an estimate of the effect of the ECM plus a decrease in consumption.

2.3.2.4.4 Meter Matching

We can only request billing data for accounts/meters that are represented in the tracking data. This can create issues, as we are blind to other accounts/meters that should potentially be aggregated for the billing analysis. For example, consider a case where there are five meters at one site and the ECM affects two or three of the meters. Further, suppose the billing data we receive is for one of the meters affected by the ECM. The billing analysis would provide an estimate of savings observed just at that meter rather than all savings associated with the ECM. In this example, the billing analysis underestimates savings. In general, for any premises where there are meter matching imperfections, we'd expect the meter matching issues to create a downward bias in the savings estimates produced by this billing analysis (i.e., savings underestimated).

2.4 Objective One: Billing Analysis Conclusions and Recommendations

2.4.1 Key Takeaways

The weather-normalized annual kWh savings produced by the electric billing analysis were positively correlated with the savings estimates stored in the NGRID tracking data, but the estimates derived via the billing analysis were generally lower. When looking at the results site by site (i.e., ignoring the magnitude of the savings), Brightline estimates were about 80% of NGRID estimates on average. A 95% confidence interval for this 80% figure is: 70.3%, 89.8%. This result is not wholly unexpected, as National Grid will adjust gross savings using realization rates and in-service rates – adjustment factors that typically reduce the savings estimate based on findings from prior evaluations. (The unadjusted savings values were the point of comparison in this analysis.) Recall again that the meter matching issues, if present, would produce a downward bias in the billing analysis savings estimate. If such issues were removed (assuming their presence), this would help to bridge the gap between the billing analysis savings estimates and the tracking data savings estimates. The differences in the estimates could also be explained by a number of other factors: business expansion resulting in greater energy use, operational changes leading to increased/decreased hours of operation, random year-to-year variations in energy use, data entry errors (e.g., incorrect project completion dates in the tracking data), faulty or outdated TRM assumptions, etc.

On the gas side, the sample size was significantly smaller than the sample size for the electric billing analysis (34 and 298, respectively), which makes it harder to draw conclusions. The weather-normalized savings estimates produced by the billing analysis were largely uncorrelated with the savings estimated stored in the NGRID tracking data (Figure 17). Our team thought one possible explanation for the differences between tracking data savings estimates and billing analysis savings estimates could be electric retrofits that led to increased gas consumption. In investigating this hypothesis, we found a few premises that installed electric retrofit measures which led to increased gas consumption (per the tracking data), but this did not help explain the differences. For a number of premises, our team saw significant increases in consumption in the post-installation period. A review of the raw billing data confirmed that consumption did indeed increase at these premises (Figure 18). For such premises, it may be the case that the increased consumption was a reflection of an expanding business. This is something that is not captured by our billing analysis, as there are no metrics of production or expansion included in the models. Meter matching issues bay also explain discrepancies.

2.4.2 Recommendations

Based on the key take aways noted above, the BrightLine team recommended customer interviews for the second component of Task 3. These interviews sought to understandthe variances in the energy savings shown in National Grid's tracking database and the energy savings calculated by the BrightLine team through the billing analysis. We will also sought to understand customers' expectations for energy and bill savings, the source of these expectations, and perceptions of the changes to their electric and gas bills based on the projects implemented. The activities and outcomes of the customer interviews are presented in the following section.

2.4.3 Output Files

Several output files have been provided in tandem with this report. These files include:

- The gas and electric analysis files
- A results workbook that shows regression output from our billing analysis regression models
- A summary file that shows the findings for each premise. Relevant fields include:
    - The tracking data savings estimate
    - The billing analysis savings estimate
    - The margin of error around the billing analysis savings estimate
    - Fractional savings uncertainty (relative precision)
    - An estimate of the percent savings
    - Pre-installation annual load at the premise
    - The number of pre-installation months and post-installation months used in the analysis

3 Objective Two: Follow-up Customer Interviews and Project Documentation Review

The second key sub-component objective for Task 3 was to conduct follow-up customer interviews (and potentially site visits) to understand any large discrepancies between the gross savings estimates stored in the tracking data and the savings estimates derived from the billing analyses. To achieve this objective, the BrightLine team first contacted a select group of customers to participate in telephone interviews with a member of the BrightLine team. The interviews sought to gain an understanding for variances in the energy savings shown in National Grid's tracking database and the energy savings calculated by the BrightLine team through the billing analysis. The interviews also sought to understand customers' expectations for energy and bill savings, the source of these expectations, and perceptions of the changes to their electric and gas bills based on the projects implemented.

Based on the results of the billing analysis, the BrightLine team's sample frame and strategy for customer interviews included the following:

- For gas projects we attempted to contact all customers in the billing analysis sample. The sample contained approximately 36 premises, of which 23 contained unique customer information.
- For electric projects we targeted 45 completed interviews using a random sample of customers with a greater than 25% discrepancy between tracking savings and bill savings. The sample pool of electric customers was approximately 168 unique premises with contact information.
- Project completion dates for sampled projects ranged from 2015-2018.
- The BrightLine team attempted to interview customers on both sides of the discrepancy (i.e., those for whom the billing analysis showed more savings than expected, and those for whom the billing analysis showed less savings than expected). Initially, we reached out to customers who implemented projects in 2017 and 2018 as we expected better recall with more recent projects. In the end, it was necessary to use the entire sample and interview respondents include customers who implemented projects in 2015, 2016, 2017, and 2018.
- In order to support higher recruitment for the interviews, we offered all customers who agreed to participate in the interview a $20 gift card for their time.

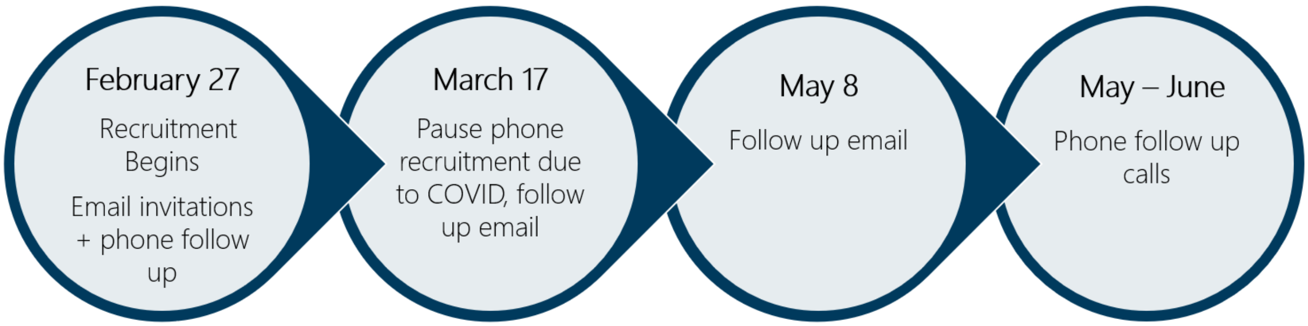
3.1 Interview Recruitment Summary

The BrightLine team began recruitment in late February with an email invitation to all customers for which we had email addresses, followed by telephone recruitment. Due to the Coronavirus pandemic, we paused phone recruitment on March 17th.[24] On May 8th, we resumed recruitment, which continued

[24] We sent an email to all customers who had not previously responded to our inquiry stating that we were pausing activities due to the pandemic but if anyone wished to participate, they could still reach out to us.

throughout May and June. After reaching out to each contact with up to 3 email messages and 3 voice messages the BrightLine team stopped to avoid over-burdening customers with additional outreach efforts. See Figure 23 for a summary of the recruitment activities.

Figure 23. Schedule of recruitment activities



The BrightLine team completed 34 interviews, 27 with customers who implemented electric projects, five with customers who implemented gas projects, one whose projects included both electric and gas measures and one with a fuel switching customer.[25] The final number of completes fell short of the target, which we partially attribute to the pandemic, partially to the turnover of staff at the organizations who participated in the program, and partially to the tracking data including contact information for contractors and not customers, making it difficult to interview a person knowledgeable with the project's post-installation experience. Table 27 outlines the interview sample disposition and response rate and Table 28 outlines the disposition for the incompletes. Table 29 summarizes the count of completed responses by year that the projects were implemented.

[25] The customer who spoke about their experience with both electric and gas measures was originally included among the electric records, and their billing analysis completed based on their electric measure. The customer who spoke about their experience with a fuel switching measure (switching from electric to gas water heating) was originally included among the electric records, and their measure is accurately recorded as a DHW fuel switch.

Table 27. Interview sample disposition – summary of completes relative to targets

Fuel Type	Available Records	Completes	Response Rate	Target Completes	% of Target Achieved
Electric	145	27	19%	45	60%
Gas	23	5	22%	32	16%
Other (electric + gas; fuel switch)[26]	N/A	2	N/A	N/A	N/A
Total	168	34	20%	77	44%

Table 28. Interview sample disposition – summary of incompletes

Incompletes	Count
No response	45
List error[27]	25
Refused	24
Wrong contact (e.g., contractor, non-technical rep)	15
Staff turnover – new contact unfamiliar	12
Duplicate contacts for facility	8
Connected but failed to complete	3
Partial response	2
Total	134

[26] Targets were set and National Grid data records are organized according to "gas" and "electric" categories only. However, through the interview process BrightLine Group gathered more specific information about the nature of the projects. In the interest of clarity that more detailed categorization is presented in this summary table.

[27] A list error is a case in which the contact information available was incorrect (e.g., phone number was invalid, disconnected, or no longer in service, and we were not successful in attempts to contact the organization through alternate means like connecting to a main office line).

Table 29. Project year for interview respondents

Project Year	Count
2015	7
2016	9
2017	10
2018	8

3.2 Customer Interview Findings

A common theme observed in the analysis of responses was that few respondents track energy savings; 14 respondents stated that they 'don't know' if the energy savings they are realizing from their energy efficient project aligns with their expectations. Table 30 summarizes the billing analysis observations (conducted as the first activity for Task 3) compared to respondent observations. Among the 20 respondents who are monitoring their energy savings at some level:

- Only six (6) observed that their savings is misaligned with expectations.
- Of those six, five respondents' observations aligned with the billing analysis findings (i.e., they believed they were saving more or less than expected, and the BrightLine team's billing analysis had the same conclusion).

Several respondents noted that the installation of additional energy efficiency measures (beyond the measure we were asking the respondent about) makes it difficult for them to track the energy savings associated with any one project. In some of these instances, respondents noted an overall sense that "we're doing the right thing and we must be saving energy."

Table 30. Summary of billing analysis findings vs. respondent observations and consistency across observations

BrightLine team billing analysis / Respondent observation	Count	Consistency across billing analysis / respondent observations
Billing Analysis found:		
Less savings than expected	27	4
More savings than expected	7	1
Respondent observed:		
Less savings than expected	4	4
More savings than expected	2	1
Think savings align with expectations	14	N/A
Don't know	14	N/A

Another key theme that emerged from the interviews is that facility and operational changes likely account for a significant amount of the misalignment in energy savings. Over two-thirds of respondents reported that changes in their facility or operations have occurred since project completion, or that they may be experiencing equipment failure or installation errors. Of those respondents, 21 described circumstances that may account for the findings of the billing analysis. For example, one respondent explained that their process efficiency improvements included changing some production methods, and this enabled them to double the production capacity of their plant. This was consistent with the billing analysis which found the plant was using more electricity than expected following the efficiency improvements. As another example, a respondent described that they had taken a large AC unit offline since completing the project. That was consistent with the billing analysis findings which showed lower energy consumption than expected following the project. Table 31 summarizes the facility and/or operational changes that were reported.

Table 31. Facility and operational issues that may account for misalignment of savings with expectations

Types of Facility / Operations Issues Reported	Reports of facility or operational change by respondents who…	
	Had less savings than expected	Had more savings than expected
Change in facility loads / volume of production / hours of operation	7	3
Additional energy efficiency projects / equipment improvements	0	3
Meter sharing	3	0
Potential equipment / installation failure	5	0
Discovered deficiency in equipment when installing efficiency improvement; repair increased load (e.g., some equipment had been offline)	2	0

We asked respondents if they recalled receiving an audit of their facility prior to project implementation and the influence of the audit report in their decision to proceed with the project. Table 32 shows that 20 respondents recalled receiving an audit, and of those, 13 felt the audit was important in their decision to implement the project.

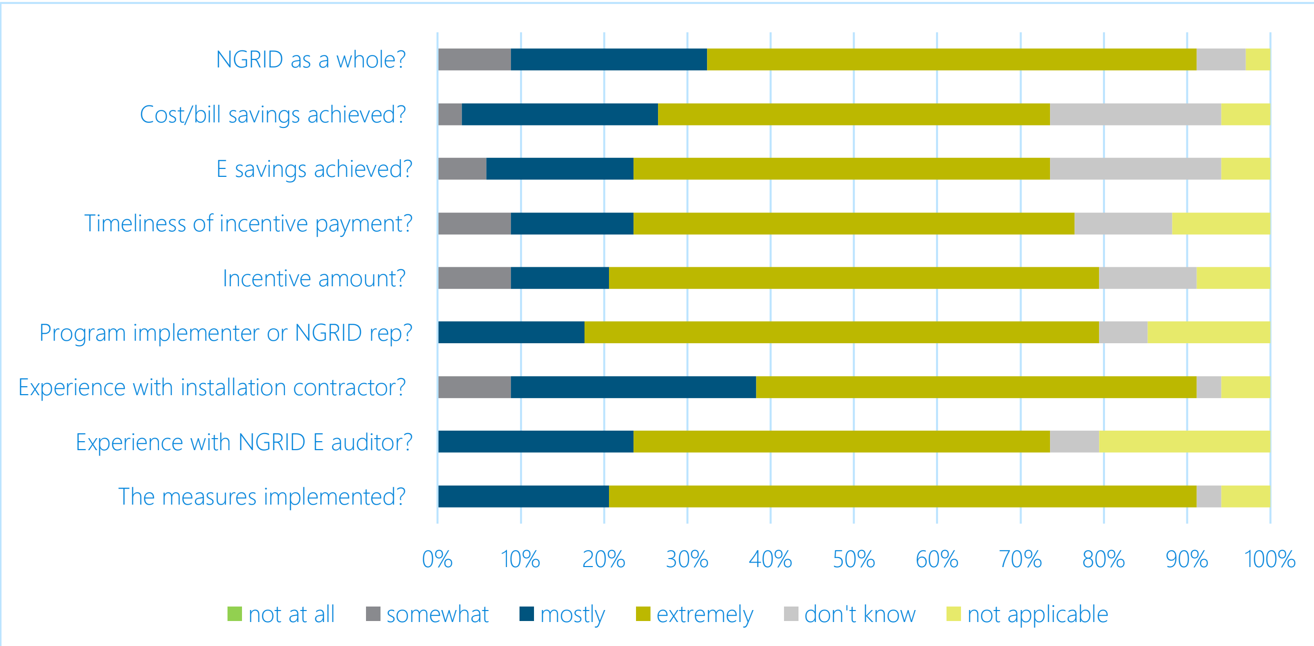
Table 32: Respondents Who Recalled Receiving Audit Report

Audit Question	Count of 'yes' responses
Did you have an energy audit through National Grid's program?	20
If yes, did this energy audit directly recommend the implementation of the measure(s) you intalled through the program?	15
Were the audit report's savings estimates important information for convincing your management to invest in energy efficiency measures?	13
What was the source of this (energy savings) expectation – (i..e energy audit report, auditor, installation contractor, other?)	5 - responses noted that energy audit was source of savings expectation

In additional to these findings, one respondent commented that the audit report may have been inaccurate and overstated savings expectations.

We asked questions regarding respondents' level of satisfaction with various aspects of the project implementation, including satisfaction with National Grid and installation contractors. Figure 24 summarizes the responses. Overall, everyone was "mostly" or "extremely" satisfied with their program experience and with National Grid as a whole, including the energy and cost savings achieved. The "don't knows" and "not applicable" responses are valuable as well, as some respondents were not familiar enough with the projects to be able to report on their satisfaction in some areas given the role they had with the project.

Figure 24. Program satisfaction



3.3 Recommendations for Next Steps

The BrightLine team identified five (5) respondents for which follow-up project documentation review, follow-up phone interviews, and/or virtual visits were conducted to better understand the discrepancies between National Grid's tracking savings and the savings estimated through the billing analysis. All the sites had projects completed in either 2017 or 2018, which increased the likelihood that respondents would have helpful recall of project details.

Of the five sites recommended for follow up, two sites reported concerns with the quality of equipment or installation. The three remaining sites were identified because the site contact is engaged and trying to monitor energy savings, and based on the initial interviews we were not able to discern apparent causes for the mismatch in National Grid's reported savings and the observed savings from the billing analysis. Three of the five projects recommended for follow up were predominately made up of lighting measures. One project is a custom HVAC measure, and one is a custom refrigeration measure.

4 Objective Two: Project Documentation Review

The objective for conducting the follow-up project documentation review was to verify data collected during the interview and to allow for a more detailed investigation of the project by collecting additional data points such as the project application, invoices, audit report, etc. We reviewed project documentation to get a clearer understanding of the types and quantities of measures installed, to review assumptions used by National Grid to calculate reported energy savings, to independently calculate estimated energy savings for the project based on data provided, and to assess whether conducting follow-up phone interviews and/or virtual verifications would be a worthwhile effort.

The initial project documentation request was submitted to National Grid on September 25, 2020 and documentation was received October 5, 2020. A second data request was submitted to National Grid on November 9, 2020 and was fulfilled on January 13, 2021. Project documentation requested for each project included:

- Project application
- Project Invoice
- Manufacturer specification (spec) sheets
- Audit report (if audit was conducted)
- Savings calculations including assumptions used to estimate project energy savings
- Any pre or post inspection reports (if applicable)

4.1 Project Documentation Review Findings

The BrightLine team conducted project documentation for all five projects originally identified. Table 33 outlines the information provided by National Grid for each project. Note that audits and pre and/or post inspections were not necessarily conducted by National Grid or their program vendors for all projects, and in these cases, reports associated with audit or inspections could not be provided. In addition, audit reports and savings calculations are not required for prescriptive lighting projects.

Table 33: Project Documentation Review, Summary of Information Provided

Project	Measure	Project application	Project invoice	Specification sheets	Audit report	Savings calculations	Any pre- or post-inspection reports
A	Lighting (4ft retrofit tube kits)	√	√	√	X	√	√
B	Lighting - LED High Bay, Low Bay, & Occ Sensors	√	√	√	X	√	√
C	Lighting - Interior & Occupancy Controls	√	√	√	X	X	X
D	Custom HVAC	√	√	√	X	√	√
E	Refrigeration Equipment and Controls	√	√	√	√	√	√

Table 34 summarizes the findings from each of the project documentation review activities, including the electric savings reported by National Grid, the savings found from the billing analysis and the electric savings calculated based on the project documentation provided to the BrightLine team. As outlined, the BrightLine team found that for three of five projects reviewed, National Grid's calculated savings values are reasonable and appropriate. One lighting project was miscalculated because National Grid mistakenly calculated savings using bulb wattages instead of fixture wattages. For the fifth project, it is BrightLine's assessment that the hours of use of the facility were overstated and therefore the electric savings value was overestimated. We understand that in many instances National Grid utilizes the hours of use value reported directly by the client. However, it is recommended that the self-reported value be checked for reasonableness by National Grid prior to reporting energy savings to ensure that the estimates are as accurate as possible.

One of the main objectives of this activity was to identify reasons for the variance in the billing analysis savings and National Grid's reported energy savings value. As noted in Table 34, the BrightLine team was able to identify the reasons for the variance for four of five projects reviewed. No further project reviews or customer interviews are recommended and BrightLine's overall assessment is that National Grid's reported electric energy savings are reasonable and that reasons for variances between National Grid's reported value and the savings calculated as part of the billing analysis are explainable.

Table 34: Project Documentation Review Summary of Findings

Project	Measure	NGRID Savings (kWh)	Brightline Billing Analysis Savings (kWh)	BrightLine Documentation Review Savings (kWh)	Summary	Potential Reason for Discrepancy
A	Lighting (4ft retrofit tube kits)	106,235	68,079	63,682	NGrid did not include the correct number of lamps in the installed case increasing the savings above the correct value. Additionally, NGrid's hours of use were likely overstated for this facility. These two adjustments aligned the savings calculated through algorithms with the billing analysis results, within 6%.	Error in NGrid savings calculation and variation in actual hours of use.
B	Lighting - LED High Bay, Low Bay, & Occ Sensors	102,965	87,111	88,254	NGrid's wattage and fixture counts were accurate, but HOU could be overstated for this type of facility. Reducing the HOU aligned algorithm results with billing analysis finding.	NGrid's wattage and fixture counts were accurate, but HOU could be overstated for this type of facility. Reducing the HOU aligned algorithm results with billing analysis finding.
C	Lighting - Interior & Occupancy Controls	35,348	29,495	37,084	NGrid's savings found to be reasonably calculated.	According to customer interview, new BMS system was installed (not incentivized) that may have changed operations of other equipment.
D	Custom HVAC	30,171	-112,634	N/A	The customer claims that after installation of the new VRF system, he did not have a gas bill and the electric bill increased. NGrid calculated savings using appropriate baseline inputs from MA & RI TRMs for new equipment purchase with a standard efficiency baseline, thus did not try to capture the large increase in electricity usage from the fuel conversion from gas to electric. This project should not have been classified as a retrofit.	Conversion from natural gas heating to electric heat pump explains negative billing analysis results.
E	Refrigeration Equipment and Controls	190,102	-111,352	209,268	BrightLine confirmed algorithm savings within 10% of NGrid's gross reported savings using CA eTRM savings methodologies for floating head & floating suction savings.	Customer reported satisfaction with the equipment and level of energy savings. Cause for discrepancy could not be determined.

5 Appendix – Alternative Billing Analysis Results

While a percent savings value greater than 100% is theoretically impossible (i.e., a site cannot save more than they consume), a percent savings value north of 50% does seem improbable. Sites with percent savings greater than 100% were removed from the original billing analysis. What if that threshold was changed to 50%? This section presents the results with this new threshold. The complete set of filters is shown in Table 35. All filters other than Filter 6 are identical to the filters shown in Table 4. Note that results from the avoided energy use analysis are not reproduced in this appendix since those results closely mirror results from the primary approach (weather-normalized savings).

Table 35: Number of Premises Removed by Adjusted Filtering

Filter #	Filter	Electric Premises		Gas Premises	
		Removed	Remaining	Removed	Remaining
--	Original population of retrofit premises	---	1,575	---	421
1	Premise not found by National Grid	24	1,551	1	420
2	Premise/account relationship is not 1:1	130	1,421	31	389
3	Zero consumption in pre or post period	6	1,415	1	388
4	Remove premises with SEM measures	3	1,412	0	388
5	Less than 12 months of pre/post data	224	1,188	76	312
6	Percent savings < 5% or > 50%	518	669	133	181
7	Relative precision > ±50%	412	257	145	36
8	Savings expire during analysis window[28]	1	256	2	34
9	Savings ratio outliers	6	250	2	30
	Final analysis data set	1,325	250	391	30

[28] For these premises, it was the case the multiple measures were installed across a few years. For example, a measure with a useful life of three years was installed in January 2015 and a measure with a useful life of five years was installed in July 2017. The savings from the first measure would expire in the middle of the post-installation period

5.1 Results – Electric

All impact estimates produced by the Brightline billing analysis were found to be statistically significant at the 95% confidence level.[29] Table 36 shows average and median annual kWh savings for the premises that remained in our electric analysis file. Both NGRID tracking data savings and the savings estimate from the Brightline billing analysis are shown. For both the average and the median, the Brightline estimate comes in around 80% of the NGRID estimate. The median estimates are lower than the average estimates because there are a few premises with very high savings values that inflate the average. The median is unaffected by these values and is likely the better measure of central tendency in this case.

Table 36: Savings Comparison – Electric

Fuel	Measure	# of Premises	NGRID	Brightline
Electric	Sum	250	40,481 MWh	33,635 MWh
Electric	Average	250	161,923 kWh	134,539 kWh[1]
Electric	Median	250	52,129 kWh	39,733 kWh

[1] A 95% confidence interval for the average savings value was calculated to be (98,189, 170,889).

Note that the tracking data savings values shown in Table 36 are not adjusted with realization rates and in-service rates.[30] The savings values that National Grid actually claims have these adjustments factored in. Table 37 compares unadjusted savings values, adjusted savings values, and the savings values calculated via the billing analysis. The remainder of the figures, tables, and comparisons in this report use the unadjusted savings values.

Table 37: Comparison with Adjusted Savings Values

Source	# of Premises	Total Savings (MWh)	% Difference Relative to Unadjusted Savings
NGRID – Unadjusted	250	40,481	---
NGRID – Adjusted	250	36,440	-10%
Brightline Analysis	250	33,635	-17%

[29] This is as expected. Filter 7 in Table 35 removes premises where uncertainty exceeds the savings estimate which would render the estimate statistically insignificant.

[30] A realization rate represents the percentage of claimed savings that are actually realized. Consider a lighting upgrade in a building where hours of use are believed to be 5,000 hours annually. If lights in this building are actually on for 4,500 hours annually, then the actual savings value will be less than the expected savings value since the initial assumption overstated use by 500 hours. An in-service rate represents the percentage of program-supported equipment that is installed.

Figure 25 and Figure 26 show the distribution of savings per premise derived via the billing analysis and in the tracking data, respectively.

Figure 25: Distribution of Electric Savings – Billing Analysis

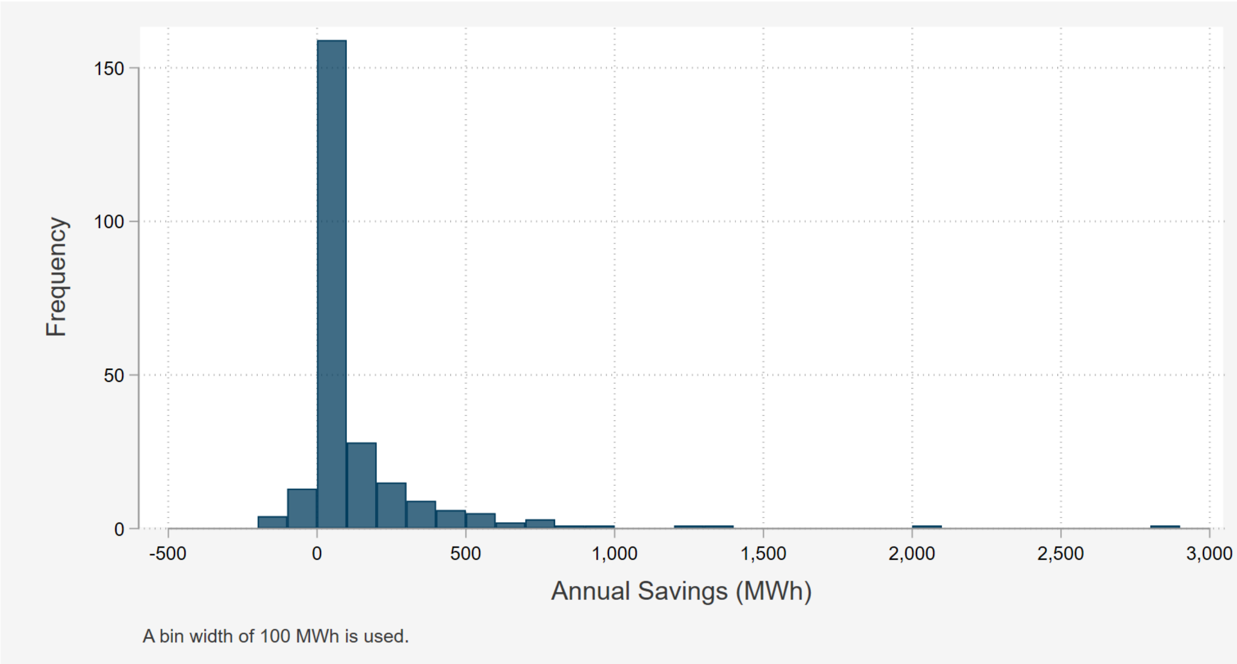
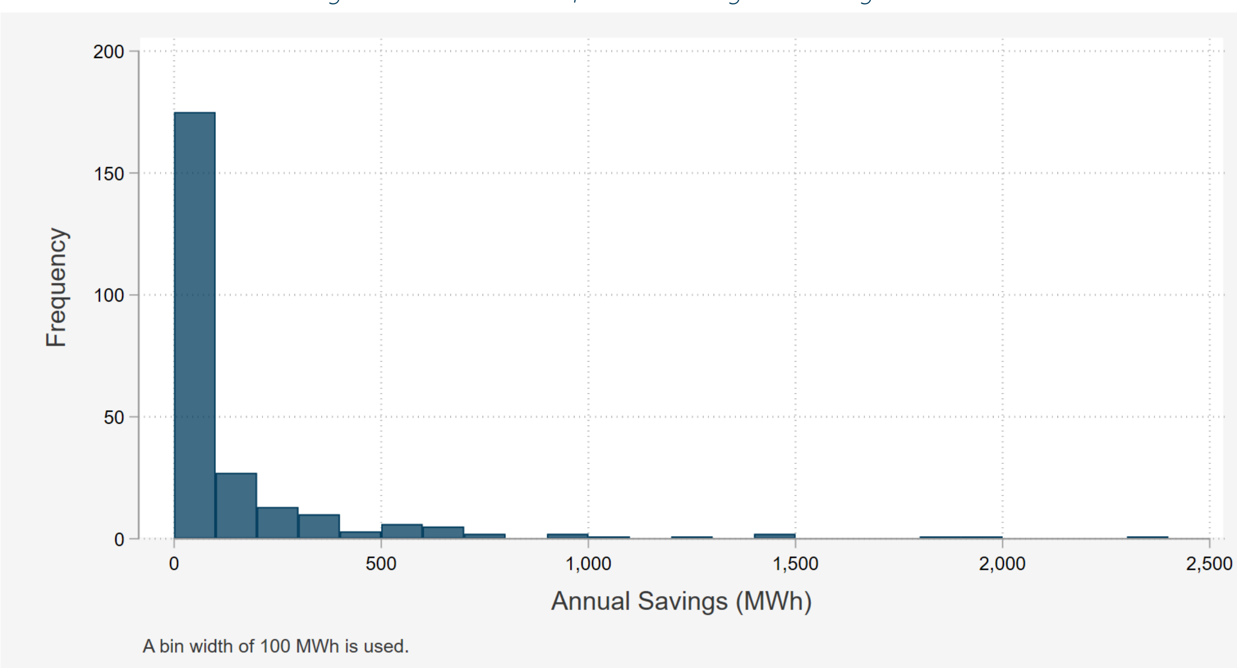
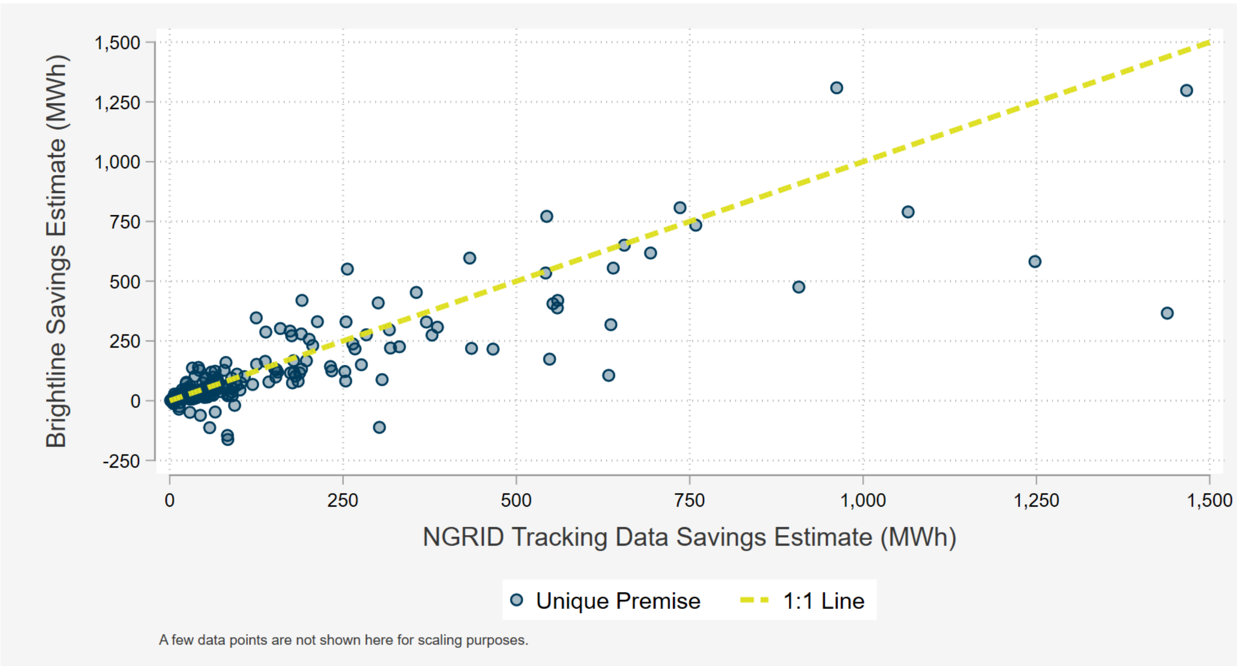


Figure 26: Distribution of Electric Savings – Tracking Data



By premise, Figure 27 shows the Brightline electric savings estimates plotted against the NGRID tracking data savings estimates. A few data points were not included for scaling reasons. The figure certainly shows a positive linear trend, though there are a number of premises for which the Brightline billing analysis produced a negative savings estimate (i.e., increased consumption in the post period). The negative values may be indicative of premises that are undergoing expansion (e.g., producing more widgets per day) or some other change between the pre period and post period other than the ECM itself and weather. Recall that one key assumption used herein is that the only differences between the two periods are the ECM and the weather. Any other changes that effect energy consumption will get confounded with the savings estimate produced via the billing analysis.

Figure 27: Comparison of MWh Savings



Due to inherent noise in the data, there is some error in the savings estimates produced by the billing analysis. For this reason, the Brightline team also calculated a 95% savings confidence interval for each premise. Such an interval is a range of values which we expect to contain the true savings value (assuming the only differences between the pre-installation and post-installation periods are the ECM itself and weather conditions). We compared the NGRID savings values with these intervals. Results are shown in Table 38. Most commonly, the NGRID savings value fell above the confidence interval calculated by the Brightline team. Of course, there is also uncertainty associated with the National Grid tracking estimates. Although they are reported as integer values, TRM assumptions and engineering rules-of-thumb are generally based on averages from studies that also have a margin of error.

Table 38: Confidence Interval Comparisons – Electric

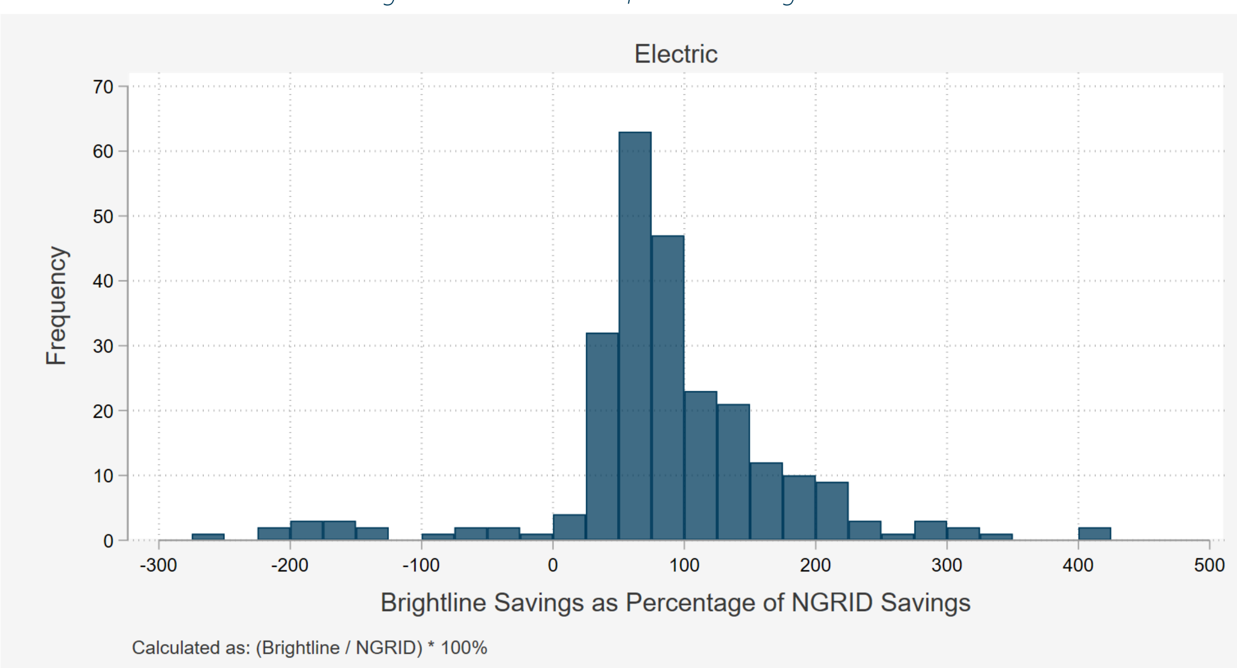
Result	Count	Average NGRID Savings (kWh)	Average Brightline Savings (kWh)
NGRID Estimate Below Savings Interval	50	135,472	222,594
NGRID Estimate Within Savings Interval	82	169,750	163,626
NGRID Estimate Above Savings Interval	118	167,692	77,015

The Brightline team also looked at savings ratios. For each premise, the Brightline team calculated a savings ratio as:

$$Savings\ Ratio = \frac{Brightline\ Savings\ Estimate}{NGRID\ Savings\ Estimate} * 100\% \quad (3)$$

In cases where the Brightline and NGRID savings estimates were virtually equal, this ratio would be near 100%. In cases where the Brightline savings estimate was much larger, the ratio would be much larger than 100%. Similarly, in cases where the Brightline savings estimate was much lower than the NGRID estimate, the ratio would be much lower than 100%. Figure 28 shows the distribution of these ratios for the electric billing analysis. A negative ratio indicates that the Brightline billing analysis produced a negative savings estimate (e.g., consumption increases after the ECM is installed). The average and median of the distribution are 87% and 81%, respectively – this is approximately where the distribution in Figure 28 peaks. A 95% confidence interval for the average ratio was calculated to be (76.1%, 98.7%).

Figure 28: Distribution of Electric Savings Ratios



The Brightline team reviewed the results by percent savings bins, number of measures installed, building type, and measure type (custom or prescriptive). Results for these breakouts are discussed below.

5.1.1 Results by Percent Savings

The calculation of percent savings was discussed in 2.3.1.2. Importantly, percent savings were calculated using the NGRID savings estimate in the numerator rather than savings estimated calculated by the Brightline team. Our team divided premises into 10 bins based on percent savings – 5%-10%, 10%-20%, 20%-30%, and so on.[31] Summary stats for each bin are shown in Table 39. The table also shows average and median savings ratios for each bin. (Equation (3) shows how savings ratios are calculated at the premise level.) In the first two bins, the estimates produced by the Brightline team exceed the estimates in the NGRID tracking data. For the larger percent savings bins, the difference between NGRID savings estimates and Brightline savings estimates grows. In the larger bins, there is an increased likelihood that there is a data mismatch between the premise number in the tracking data and the premise (or premises) where the savings occur. Or there may be some non-measure, non-program issue going on with the site that has not been identified. In such cases, we would expect the billing analysis to underestimate savings, as the billing analysis will only measure savings for one premise.

Table 39: Results by Percent Savings Bin – Electric

Percent Savings Bin	# of Premises	Statistic	NGRID (kWh)	Brightline (kWh)	Savings Ratio (%)
5% - 10%	34	Average	137,138	169,860	164
		Median	42,132	76,054	162
10% - 20%	74	Average	148,016	163,333	107
		Median	51,345	51,249	106
20% - 30%	86	Average	135,425	98,630	61
		Median	50,162	34,277	70
30% - 40%	31	Average	226,046	135,704	69
		Median	78,435	36,530	68
40% - 50%	25	Average	248,436	123,357	40
		Median	46,593	28,442	52

[31] Recall that premises with percent savings less than 5% were filtered out of the analysis data set. This is why the first bin runs from 5% to 10% rather than 0% to 10%.

5.1.2 Results by Number of Measures Installed

A number of premises installed multiple retrofit measures.[32] For such premises, our billing analysis treated the pre-installation period as the period before any measures were installed. The post-installation period was the period after all measures were installed. Thus, the savings estimate derived from the billing analysis is an estimate of the combined savings, not the savings for any particular measure.

The majority of sites installed three measures or fewer, but there were a number of premises with a measure count in the 20s and one with a measure count above 100. We reviewed results by the number of measures installed, though we took a binary approach here. The comparison bins were "just one measure" and "more than one measure." Table 40 shows the results. The table also shows average and median savings ratios for each bin. (Equation (3) shows how savings ratios are calculated at the premise level.) For sites with just one measure installed, the Brightline savings estimates came out slightly higher than the NGRID estimates, on average. For sites with multiple measures installed, Brightline savings estimates were lower than the NGRID estimates, on average.

Table 40: Results by Number of Measures – Electric

# of Measures	# of Premises	Statistic	NGRID (kWh)	Brightline (kWh)	Savings Ratio (%)
Just one	40	Average	78,425	84,282	95
		Median	34,707	29,618	126
More than one	210	Average	177,827	144,112	86
		Median	52,635	43,399	78

5.1.3 Results by Building Type

Results by building type are shown in Table 41. Building type assignments were drawn from the tracking data. Not all premises were assigned a building type. Tables entries are sorted by savings ratio (using the median value) from least to greatest. For building types at the top of the table, there was not much agreement between the tracking data savings estimate and the billing analysis findings. Towards the bottom of the table, there was more agreement between the savings estimates. Hospitals and small offices produced some of the largest discrepancies, though the sample sizes here are exceptionally small.

[32] We're defining measures as unique rows in the tracking data. It's possible that one single tracking data entry could incorporate multiple technologies, such as a lighting upgrade combined with occupancy sensors. Such an example would be counted as one measure in this analysis.

Table 41: Results by Building Type – Electric

Building Type	# of Premises	Statistic	NGRID (kWh)	Brightline (kWh)	Savings Ratio (%)
Hospital	1	Average	83,791	-162,039	-193
		Median	83,791	-162,039	-193
Small Office	6	Average	32,433	-12,440	-42
		Median	18,314	3,766	-67
Refrigerated Space	2	Average	87,652	41,444	48
		Median	87,652	41,444	48
Multifamily low-rise	2	Average	133,522	66,336	60
		Median	133,522	66,336	60
Secondary School	2	Average	66,161	43,106	65
		Median	66,161	43,106	65
Multi Story Retail	3	Average	199,789	198,452	86
		Median	179,166	117,974	66
University	3	Average	248,996	166,792	79
		Median	174,009	164,835	67
Fast Food Restaurant	66	Average	47,475	35,168	79
		Median	47,327	32,610	70
Light Industrial	10	Average	362,141	408,552	76
		Median	187,858	155,124	77
Primary School	11	Average	98,291	57,540	63
		Median	80,458	62,303	77
Large Office	14	Average	174,814	180,500	84
		Median	54,571	42,774	79
Other	45	Average	251,241	151,123	79
		Median	59,900	30,587	80
Heavy Industrial	7	Average	395,410	391,745	128
		Median	254,254	237,539	90
Warehouse	6	Average	151,311	160,239	54
		Median	39,744	49,984	96
Grocery	21	Average	300,965	283,928	109
		Median	213,024	287,489	99
Automobile	5	Average	73,849	71,643	86
		Median	46,593	46,805	100
Multifamily high-rise	4	Average	192,579	172,322	121
		Median	221,016	186,344	108
Hotel	7	Average	322,064	293,141	116
		Median	177,027	123,580	128
Big Box Retail	12	Average	238,552	236,201	166
		Median	83,891	119,218	132
Small Retail	20	Average	21,135	24,749	125

Building Type	# of Premises	Statistic	NGRID (kWh)	Brightline (kWh)	Savings Ratio (%)
		Median	16,255	25,592	133
Religious	3	Average	5,584	3,973	93
		Median	5,994	9,657	175

5.1.4 Results by Measure Type

Some premises installed only custom measures, some premises installed only prescriptive measures, and some premises installed a mixture of custom and prescriptive measures. Table 42 compares National Grid savings with Brightline savings for these three bins. The table also shows average and median savings ratios for each bin. (Equation (3) shows how savings ratios are calculated at the premise level.) Savings ratios were highest for premises that only installed prescriptive retrofit measures (i.e., no custom measures).

Table 42: Results by Measure Type – Electric

Type of Measure	# of Premises	Statistic	NGRID (kWh)	Brightline (kWh)	Savings Ratio (%)
Custom measures	78	Average	102,313	87,306	82
		Median	50,704	35,182	72
Prescriptive measures	138	Average	134,547	96,864	94
		Median	43,516	36,186	90
Mixture of measures	34	Average	409,789	395,816	75
		Median	231,219	216,445	77

5.1.5 Results by Technology

Lighting measures accounted for approximately 80% of the records in the electric tracking data and approximately 70% of the savings. To investigate whether the two savings estimates were better aligned for lighting measures, we put each premise in one of three bins: no lighting measures, some lighting measures, and only lighting measures. Premises in the first bin did not have any lighting measures in the tracking data, while premises in the other bins had at least one lighting measure in the tracking data. Premises in the "only lighting measures" did not show any non-lighting measures in the tracking data. Comparisons are shown in Table 43. In each bin, average and median estimates from the billing analysis were less than average and median estimates from the tracking data. The table also shows average and median savings ratios for each bin. (Equation (3) shows how savings ratios are calculated at the premise level.) On average, savings ratios were better in the "no lighting upgrades" bin compared to the other two bins.

Table 43: Results by the Presence of Lighting Upgrades

Lighting Bin	# of Premises	Statistic	NGRID (kWh)	Brightline (kWh)	Savings Ratio (%)
No Lighting Upgrades	33	Average	129,917	112,242	113
		Median	102,634	74,711	94
Some Lighting Upgrades	96	Average	213,842	192,472	70
		Median	53,238	39,733	71
Only Lighting Upgrades	121	Average	129,460	94,658	94
		Median	37,690	30,160	94

5.1.6 Results by Annual Consumption Bins

Table 44 shows the results by customer size (based on pre-retrofit annual consumption). For this breakout, five annual consumption bins were defined: less than 50 MWh, between 50 and 100 MWh, between 100 and 500 MWh, between 500 and 1,000 MWh, and greater than 1,000 MWh. Savings ratios were highest in the 50 - 100 MWh bin. Overall, no patterns related to annual consumption bin are present.

Table 44: Results by Annual Consumption – Electric

Annual Consumption	# of Premises	Statistic	NGRID (kWh)	Brightline (kWh)	Savings Ratio (%)
< 50 MWh	21	Average	6,890	3,842	43
		Median	5,994	3,763	82
50 - 100 MWh	22	Average	19,109	17,694	112
		Median	15,892	16,793	101
100 - 500 MWh	122	Average	48,292	34,208	80
		Median	46,770	32,353	72
500 - 1,000 MWh	28	Average	156,713	120,384	115
		Median	154,709	109,673	73
> 1,000 MWh	57	Average	519,931	449,488	96
		Median	318,411	318,206	88

5.2 Results – Gas

All impact estimates produced by the Brightline billing analysis were found to be statistically significant at the 95% confidence level.[33] Table 45 shows average and median annual therm savings for the premises that remained in our gas analysis file. Both NGRID tracking data savings and the savings estimate from the Brightline billing analysis are shown. The average savings estimate for the Brightline billing analysis is higher than the average from the tracking data, but this is solely due to the premise with the largest savings value – the Brightline savings estimate for this premise was more than twice as large as the NGRID estimate. This premise had 154 measures installed (153 steam trap measures accounting for 452,063 annual therms saved and one custom insulation measure account for 22,206 annual therms saved). Without this premise, the average savings values would be 12,886 therms and 8,244 therms for NGRID and Brightline respectively. The median is not affected by this one premise and is a better measure of central tendency for this collection of premises.

Table 45: Savings Comparison – Gas

Fuel	Statistic	# of Premises	NGRID	Brightline
Gas	Sum	30	830,575 therms	1,374,886 therms
Gas	Average	30	27,686 therms	45,830 therms[1]
Gas	Median	30	6,335 therms	7,493 therms

[1] A 95% confidence interval for the average savings value was calculated to be (-32,765, 124,424).

Note that the tracking data savings values shown in Table 45 are not adjusted with realization rates and in-service rates.[34] The savings values that National Grid actually claims have these adjustments factored in. Table 46 compares unadjusted savings values, adjusted savings values, and the savings values calculated via the billing analysis. The bottom three rows remove the site with 154 measures from the totals. The remainder of the figures, tables, and comparisons in this report use the unadjusted savings values.

[33] This is as expected. Filter 7 in Table 35 removes premises where uncertainty exceeds the savings estimate which would render the estimate statistically insignificant.

[34] A realization rate represents the percentage of claimed savings that are actually realized. Consider a lighting upgrade in a building where hours of use are believed to be 5,000 hours annually. If lights in this building are actually on for 4,500 hours annually, then the actual savings value will be less than the expected savings value since the initial assumption overstated use by 500 hours. An in-service rate represents the percentage of program-supported equipment that is installed.

Table 46: Comparison with Adjusted Savings Values

Source	# of Premises	Total Savings (therms)	% Difference Relative to Unadjusted Savings
NGRID – Unadjusted	30	830,575	---
NGRID – Adjusted	30	769,005	-7%
Brightline Analysis	30	1,374,886	+66%
NGRID – Unadjusted	29	356,306	---
NGRID – Adjusted	29	297,178	-17
Brightline Analysis	29	239,074	-33%

Figure 29 and Figure 30 show the distribution of savings derived via the billing analysis and in the tracking data, respectively.

Figure 29: Distribution of Gas Savings – Billing Analysis

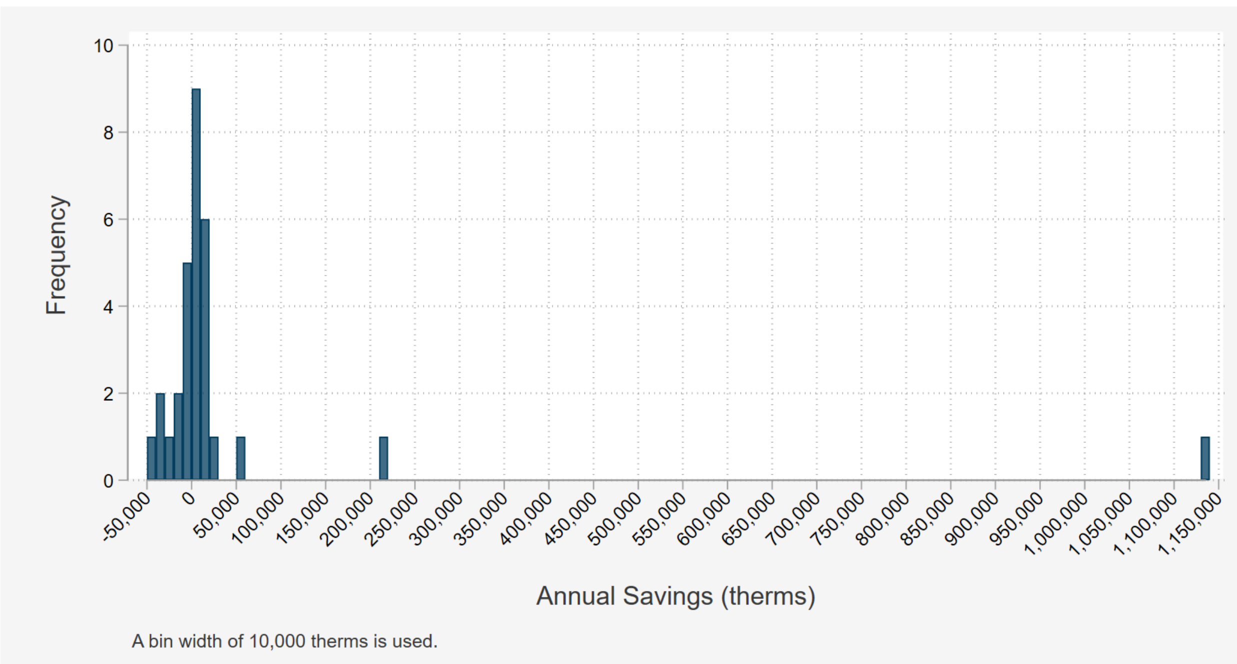
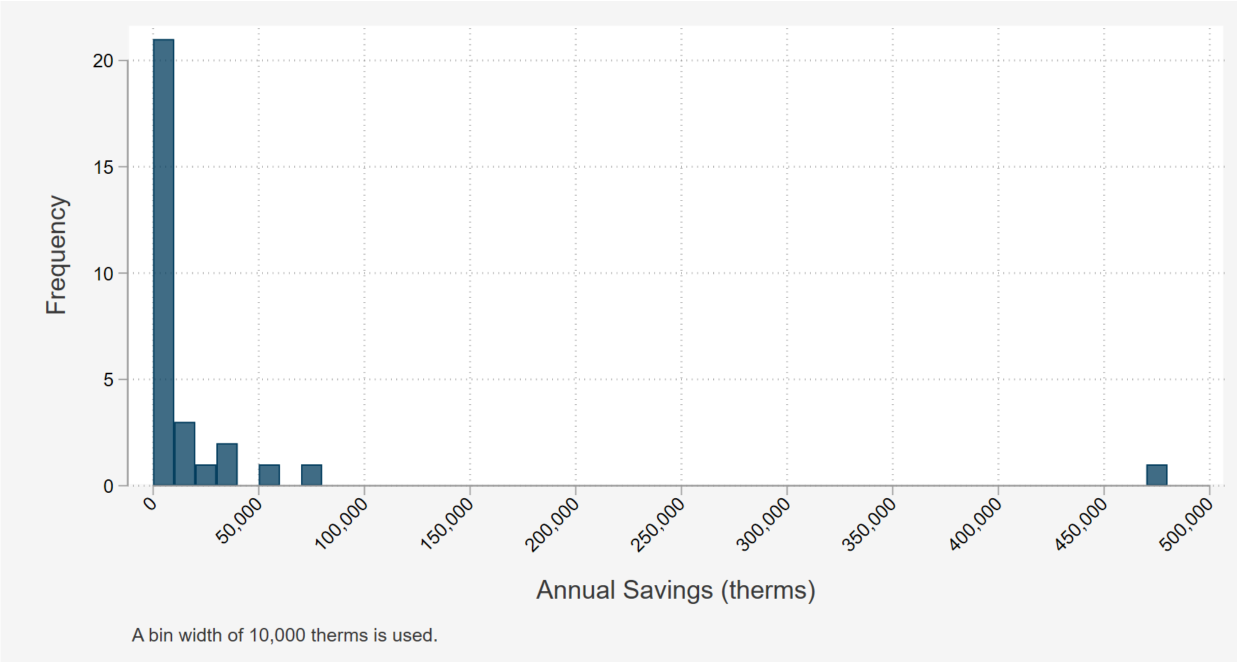
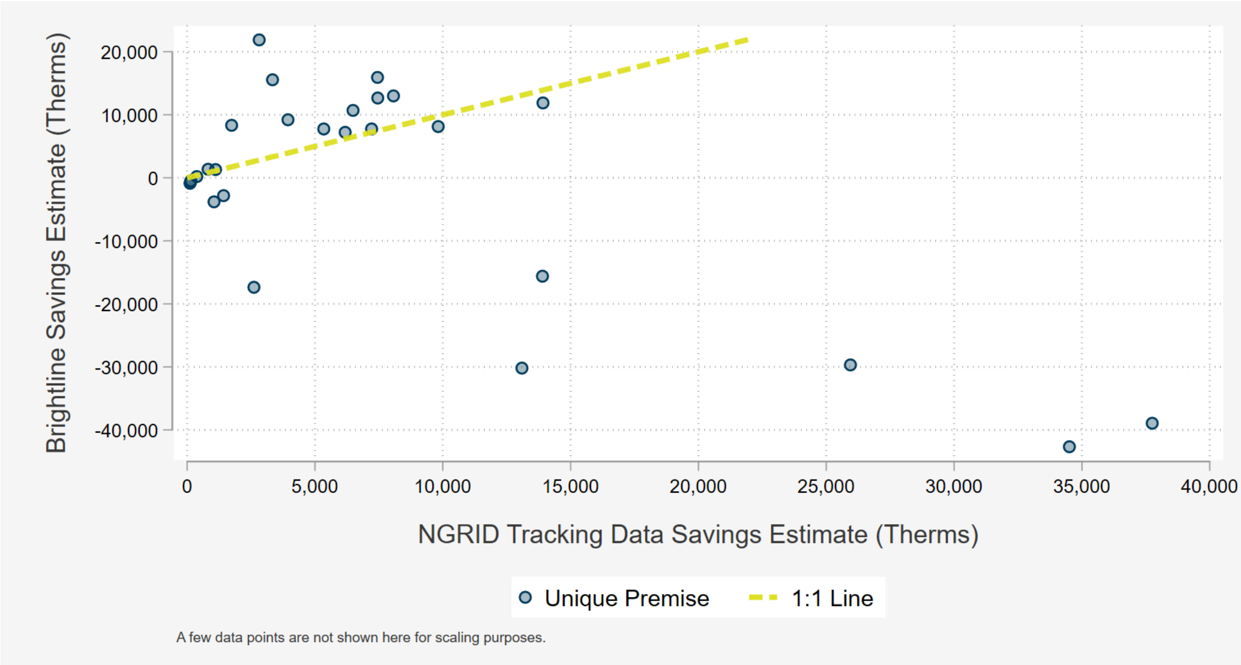


Figure 30: Distribution of Gas Savings – Tracking Data



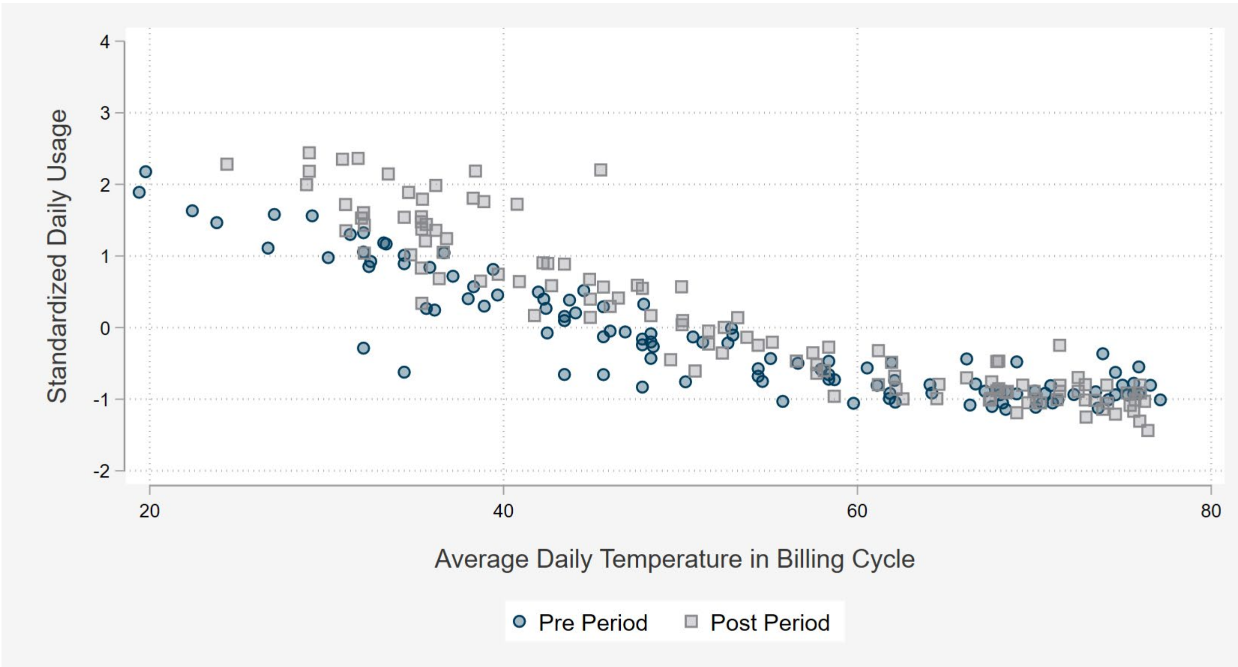
By premise, Figure 31 shows the Brightline gas savings estimates plotted against the NGRID tracking data savings estimates. A few data points were not included in the figure for scaling reasons. Unlike the trend in the figure showing kWh impacts (Figure 27), there is not much of a trend between Brightline gas savings estimates and NGRID gas savings estimates. Additionally, a higher percentage of the Brightline gas savings estimates are negative (compared to the electric savings estimates). As with the negative savings estimates for electric, it might be the case that the negative values are indicative of premises that are undergoing expansion (e.g., producing more widgets per day) or other non-routine events.

Figure 31: Comparison of Therms Savings



Our team reviewed the raw billing data for any premise for which the Brightline savings estimate was less than -10,000 therms per year. Indeed, the raw billing data for these premises showed an increase in average daily consumption in the post-installation period relative to the pre-installation period. Figure 32 shows the results using standardized consumption values as a function of outdoor temperature. "Standardizing" is a transformation of the data that puts all of the premises on a common consumption scale (i.e., it removes differences in magnitude and spread), which makes it easier to compare trends across premises.

Figure 32: Standardized Daily Gas Consumption for Premises with Negative Savings



Due to inherent noise in the data, there is some error in the savings estimates produced by the billing analysis. For this reason, the Brightline team also calculated a 95% savings confidence interval for each premise. Such an interval is a range of values which we expect to contain the true savings value (assuming the only differences between the pre-installation and post-installation periods are the ECM itself and weather conditions). We compared the NGRID savings values with these intervals. Results are shown in Table 47. Most commonly, the NGRID savings value fell above the confidence interval calculated by the Brightline team.

Table 47: Confidence Interval Comparisons – Gas

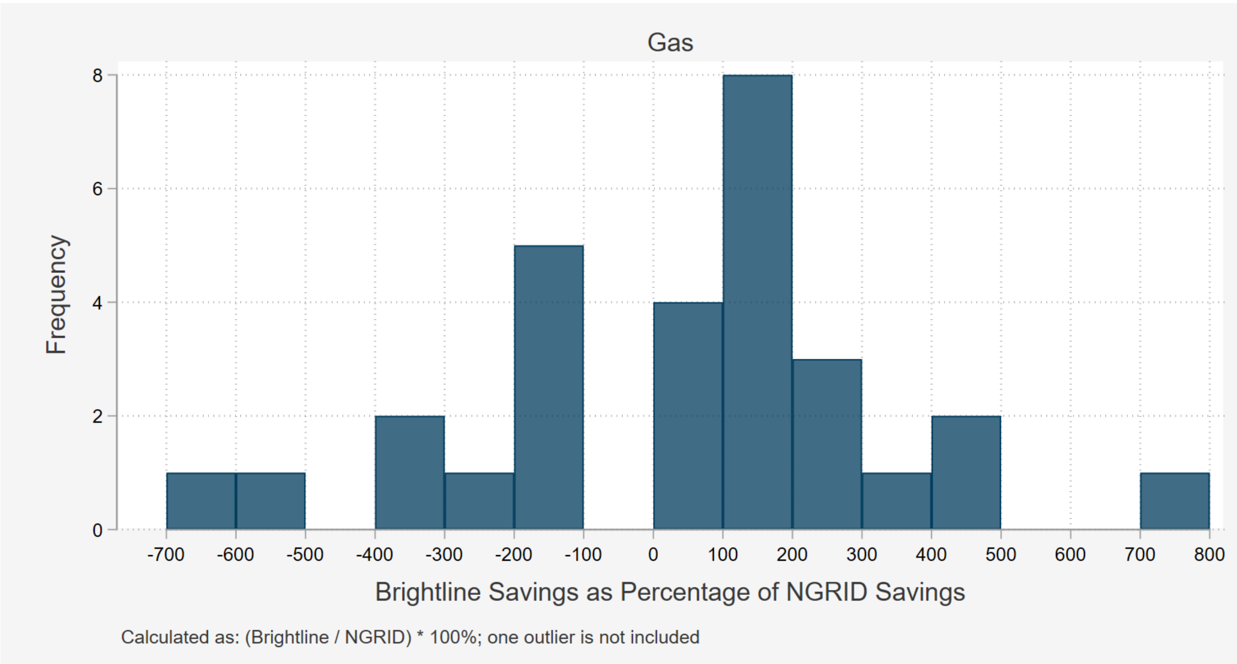
Result	Count	Average NGRID Savings (therms)	Average Brightline Savings (therms)
NGRID Estimate Below Savings Interval	9	63,045	160,872
NGRID Estimate Within Savings Interval	8	6,560	7,308
NGRID Estimate Above Savings Interval	13	16,207	-10,109

The Brightline team also looked at savings ratios. For each premise, the Brightline team calculated a savings ratio as:

$$Savings\ Ratio = \frac{Brightline\ Savings\ Estimate}{NGRID\ Savings\ Estimate} * 100\% \qquad (4)$$

In cases where the Brightline and NGRID savings estimates were virtually equal, this ratio would be near 100%. In cases where the Brightline savings estimate was much larger, the ratio would be much larger than 100%. Similarly, in cases where the Brightline savings estimate was much lower than the NGRID estimate, the ratio would be much lower than 100%. Figure 33 shows the distribution of these ratios for the gas billing analysis. A negative ratio indicates that the Brightline billing analysis produced a negative savings estimate (e.g., consumption increases after the ECM is installed). There is a considerable amount of spread in the figure, with ratios ranging from -700% up to 800%. This implies there is not much agreement between the tracking data savings estimate and the billing analysis savings estimate.

Figure 33: Distribution of Gas Savings Ratios



The Brightline team reviewed the results by percent savings bins and by number of measures installed. (Building types were sparsely populated for gas measures, so that breakout was not examined for gas savings.)

5.2.1 Results by Percent Savings

The calculation of percent savings was discussed in 2.3.1.2. Importantly, percent savings were calculated using the NGRID savings estimate in the numerator rather than savings estimated calculated by the Brightline team. Our team divided premises into bins based on percent savings – 5%-10%, 10%-20%, 20%-30%, and so on. Summary stats for each bin are shown in Table 48. The table also shows average and median savings ratios for each bin. (Equation (4) shows how savings ratios are calculated at the premise level.) In each bin, both median and average Brightline estimates fall below the NGRID estimates with one exception – the Brightline average for the 5% – 10% bin. The NGRID average savings estimate is much higher in this bin. This is due to the site with 154 measures that has been mentioned previously. (The Brightline savings estimate was more than twice the NGRID estimate for this site, and this site produced more savings than any others in the final gas analysis file.)

Table 48: Results by Percent Savings Bin – Gas

Percent Savings Bin	# of Premises	Statistic	NGRID (therms)	Brightline (therms)	Savings Ratio (%)
5% - 10%	11	Average	51,185	123,261	22
		Median	2,818	8,338	234
10% - 20%	12	Average	8,365	665	18
		Median	6,852	4,559	113
20% - 30%	3	Average	14,407	-2,956	46
		Median	9,820	8,139	83
30% - 40%	4	Average	30,985	4,978	33
		Median	21,965	3,716	60
40% - 50%	0	Average	--	--	--
		Median	--	--	--

5.2.2 Results by Number of Measures Installed

A number of premises installed multiple retrofit measures.[35] For such measures, our billing analysis treated the pre-installation period as the period before any measures were installed. The post-installation period was the period after all measures were installed. Thus, the savings estimate derived from the billing analysis is an estimate of the combined savings for premises that installed multiple measures. Using three bins, Table 49 shows the results by the number of measures installed. The bins are: just one measure installed, more than one measure installed, and 154 measures installed. The second bin ("more than one measure" does not include the premise with 154 measures.) The binned results are largely in line with the overall results – the Brightline savings estimates are smaller than the NGRID savings estimates with the exception of the site with 154 measures, which has been noted previously. The Brightline savings estimate is more than twice the NGRID savings estimate for this site.

[35] We're defining measures as unique rows in the tracking data. It's possible that one single tracking data entry could incorporate multiple technologies. Such cases would be counted as one measure in this analysis.

Table 49: Results by Number of Measures – Gas

# of Measures	# of Premises	Statistic	NGRID (therms)	Brightline (therms)	Savings Ratio (%)
Just one	14	Average	5,077	251	-38
		Median	2,176	1,338	101
More than one	15	Average	19,015	15,704	67
		Median	8,070	8,139	83
154	1	Average	474,269	1,135,812	239
		Median	474,269	1,135,812	239

5.2.3 Results by Measure Type

Some premises installed only custom measures, some premises installed only prescriptive measures, and some premises installed a mixture of custom and prescriptive measures. Table 50 compares National Grid savings with Brightline savings for these three bins. There were no clear takeaways from this breakdown.

Table 50: Results by Measure Type – Gas

Type of Measure	# of Premises	Statistic	NGRID (therms)	Brightline (therms)	Savings Ratio (%)
Only Custom	24	Average	11,201	7,239	74
		Median	6,335	7,493	112
Only Prescriptive	3	Average	131	-667	-528
		Median	131	-678	-518
Mixture of measures	3	Average	187,117	401,051	173
		Median	79,634	51,409	214

5.2.4 Results by Annual Consumption Bins

Table 51 shows the results by customer size (based on pre-retrofit annual consumption). For this breakout, three annual consumption bins were defined: less than 25,000 therms, between 25,000 and 100,000 therms, and greater than 100,000 therms. No patterns related to annual consumption bin are present.

Table 51: Results by Annual Consumption – Weather-Normalized

Annual Consumption	# of Premises	Statistic	NGRID (therms)	Brightline (therms)	Savings Ratio (%)
< 25,000 therms	10	Average	1,310	983	-121
		Median	936	-128	-71
25,000 - 100,000 therms	12	Average	6,538	7,474	145
		Median	6,852	9,963	163
> 100,000 therms	8	Average	92,378	159,421	23
		Median	36,129	-8,892	-19