# Impact Evaluation of the 2019 Existing Buildings Program

## Energy Trust of Oregon

June 1, 2021

SAFER, SMARTER, GREENER

*<Page intentionally blank>*

## Table of contents



## List of figures

# List of tables

# 0 EXECUTIVE SUMMARY

Energy Trust of Oregon (Energy Trust) hired DNV to complete an impact evaluation of Energy Trust's 2019 Existing Buildings program. This report presents the methods, results, and findings of the evaluation. The goal of the evaluation was to improve savings estimates and enhance the Existing Buildings program's effectiveness in delivering savings to customers.

## 0.1 Program overview

The Existing Buildings program began in March 2004 and is implemented by a program management contractor (PMC). ICF International was the PMC from January 1, 2013 – December 31, 2020, including the program year evaluated. In 2019, the program had four main tracks: Custom, Lighting (including standard, direct-install, and street lighting measures), Standard (calculated and prescriptive), and Strategic Energy Management (SEM).

## 0.2 Savings claimed

Table 0-1 shows the gross claimed program savings by track and fuel included in the program tracking data provided to DNV. The values shown are the site-level "working" savings listed in the data provided. These savings do not include adjustments for prior realization rates, net-to-gross, or transmission and distribution.

**Table 0-1: Program year 2019 claimed energy savings by fuel and track**

| Program Track | Unique Measure Lines | Claimed Electric Savings (kWh) | % of kWh Grand Total | Claimed Gas Savings (therms) | % of therms Grand Total |
|---|---|---|---|---|---|
| Lighting | 8,030 | 86,869,734 | 65% | 0 | 0% |
| Standard | 878 | 7,192,795 | 5% | 578,998 | 28% |
| Custom | 208 | 27,097,472 | 20% | 881,333 | 43% |
| **Capital Total** | **9,116** | **121,160,002** | **90%** | **1,460,331** | **71%** |
| Strategic Energy Management | 282 | 12,970,069 | 10% | 602,990 | 29% |
| **Grand Total** | **9,398** | **134,130,071** | **100%** | **2,063,321** | **100%** |

## 0.3 Evaluation results

Table 0-2 shows the evaluated savings by fuel and track. Table 0-3 provides the final program and track-level realization rates achieved. Table 0-4 provides a summary of the results for each track and primary sampling domain. The table shows the unweighted minimum, mean, and max realization rates (RR) for each track and domain.

**Table 0-2: Evaluated energy savings by fuel and track**

| Program Track | Evaluated Electricity Savings (kWh) 2019 | Evaluated Gas Savings (therms) 2019 |
|---|---|---|
| Lighting | 94,077,036 | |
| Standard | 7,053,360 | 457,991 |
| Custom | 24,352,887 | 755,154 |
| **Capital Measures Only** | **125,483,283** | **1,213,145** |
| Strategic Energy Management | 11,613,056 | 563,892 |
| **Grand Total** | **137,096,339** | **1,777,037** |

**Table 0-3: Program realization rates by fuel and track**

| Program Track | Electricity Realization Rates 2019 | Gas Realization Rates 2019 |
|---|---|---|
| Lighting | 108% | |
| Standard | 98% | 79% |
| Custom | 90% | 86% |
| **Capital Measures Only** | **104%** | **83%** |
| Strategic Energy Management | 90% | 94% |
| **Existing Buildings Program** | **102%** | **86%** |

**Table 0-4: Track and domain realization rate summaries, unweighted**

| Track / Primary sampling domain | Electric Results | | | | Gas Results | | | |
|---|---|---|---|---|---|---|---|---|
| | # Results | Min RR | Mean RR | Max RR | # Results | Min RR | Mean RR | Max RR |
| **Lighting** | **31** | **26%** | **98%** | **193%** | | | | |
| Direct Install | 12 | 26% | 81% | 141% | | | | |
| Standard Lighting | 16 | 90% | 110% | 193% | | | | |
| Street Lighting | 3 | 100% | 102% | 105% | | | | |
| **Standard** | **32** | **56%** | **97%** | **101%** | **43** | **15%** | **82%** | **136%** |
| Refrigeration | 13 | 84% | 99% | 100% | 5 | 100% | 100% | 100% |
| Others | 12 | 56% | 95% | 101% | 12 | 77% | 99% | 112% |
| Boiler | | | | | 13 | 15% | 41% | 115% |
| Food Equipment | 7 | 85% | 98% | 100% | 13 | 49% | 99% | 136% |
| **Custom** | **23** | **14%** | **98%** | **218%** | **18** | **0%** | **86%** | **145%** |
| Custom | 23 | 14% | 98% | 218% | 18 | 0% | 86% | 145% |
| **SEM** | **34** | **0%** | **133%** | **1,536%** | **26** | **44%** | **122%** | **311%** |
| Year-1 | 6 | 0% | 304% | 1,536% | 4 | 100% | 106% | 117% |
| Continuation | 28 | 0% | 96% | 292% | 22 | 44% | 125% | 311% |

## 0.4 Demand savings estimates

Energy Trust has developed summer and winter load factors, as a function of kWh savings, to estimate the demand savings achieved by each capital measure. Demand savings are not estimated for SEM measures. Energy Trust used this evaluation as an opportunity to see how the evaluation's adjustments to electric energy savings would impact Energy Trust's estimate of demand savings. Energy Trust provided DNV with the load factors to estimate the demand savings of any project based on the measure composition. Independent review of Energy Trust's load factors or site-specific adjustments to the factors themselves based on the data collected through this evaluation was out of the scope of this evaluation. DNV estimated the demand savings for each project in the sample frame based on the claimed energy savings and for every project evaluated using the evaluated savings. DNV then completed an expansion of demand savings from the sample to the program population. Table 0-5 compares the electric realization rates for the capital measures to the demand realization rates for the same measures by program track. The analysis shows that the demand realization rates are similar, but the application of the load factors does result in slightly different realization rates.

**Table 0-5: Comparison of electric energy and demand savings realization rates**

| Program | Electric Energy – | Summer kW - | Winter kW - |
|---|---|---|---|
| Track | RR | RR | RR |
| Lighting | 108.3% | 109.6% | 110.1% |
| Standard | 98.1% | 97.7% | 98.7% |
| Custom | 89.9% | 89.1% | 89.9% |
| **Capital Measures Only** | **103.6%** | **100.7%** | **106.0%** |

## 0.5 Historic capital measure performance

Table 0-6, Figure 1 and Figure 2 show historic program performance for capital measure tracks: lighting, standard, and custom. The table and charts do not include the SEM track, which was added to the Existing Buildings program impact evaluations in 2015.

**Table 0-6: Historic program performance, excluding SEM**

| Program Year | Verified Electric Savings (MWh) | Electric Realization Rate | Verified Gas Savings (therms) | Gas Realization Rate |
|---|---|---|---|---|
| 2008 | 41,887 | 99% | 746,564 | 87% |
| 2009 | 63,537 | 85% | 705,644 | 75% |
| 2010 | 91,884 | 107% | 1,486,729 | 86% |
| 2011 | 98,776 | 91% | 2,148,020 | 101% |
| 2012 | 86,911 | 95% | 1,174,676 | 79% |
| 2013 | 79,612 | 88% | 911,922 | 67% |
| 2014 | 82,699 | 81% | 973,143 | 72% |
| 2015 | 94,992 | 96% | 1,061,316 | 79% |
| 2016 | 104,962 | 92% | 1,228,416 | 87% |
| 2017 | 119,002 | 95% | 1,515,434 | 90% |
| 2018 | 134,660 | 104% | 915,956 | 73% |
| 2019 | 125,483 | 104% | 1,213,145 | 83% |

**Figure 1: Historic Non-SEM program electric savings and realization rates**

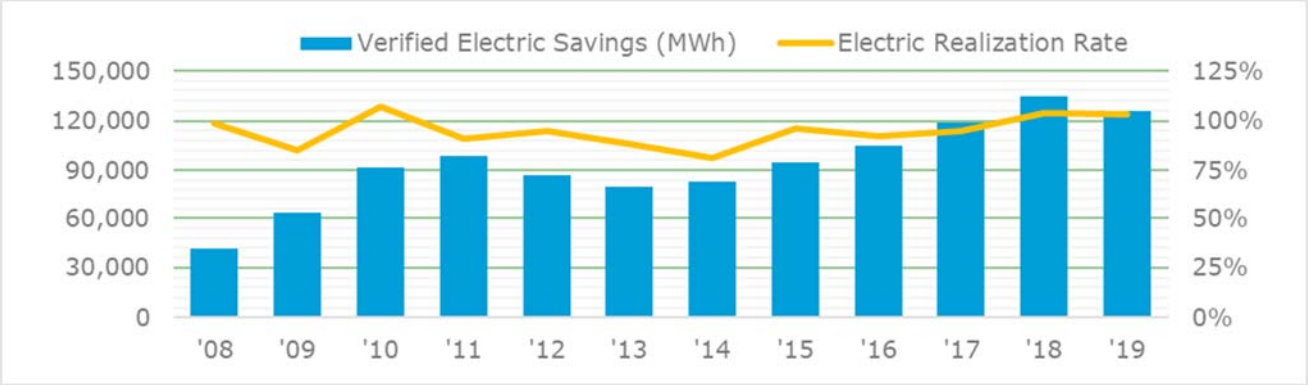
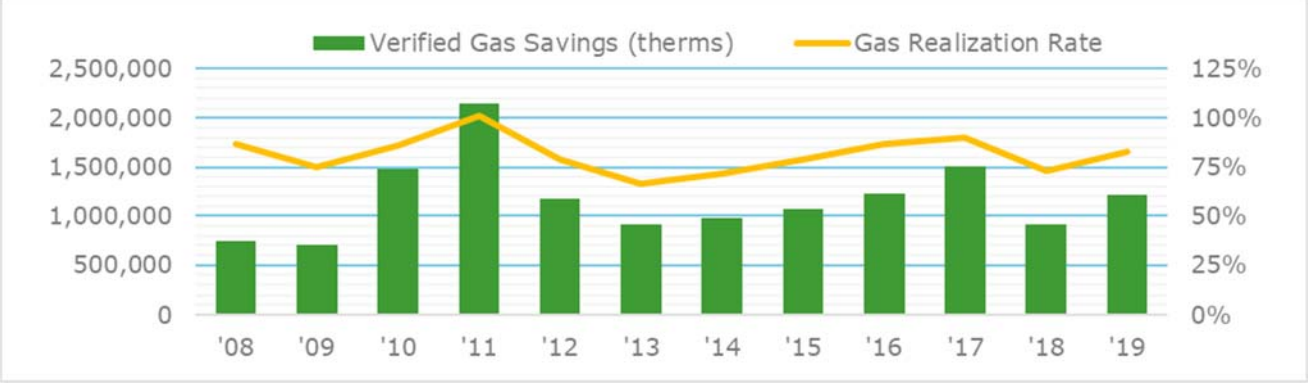



**Figure 2: Historic Non-SEM program gas savings and realization rates**



## 0.6 Historic SEM performance

Table 0-7, Figure 3 and Figure 4 show historic SEM performance over time.

**Table 0-7: Historic SEM program performance**

| Program Year | Verified Electric Savings (MWh) | Electric Realization Rate | Verified Gas Savings (Therms) | Gas Realization Rate |
|---|---|---|---|---|
| 2012 | 7,351 | 139% | -18,452 | -15% |
| 2013 | 8,988 | 103% | 174,390 | 47% |
| 2014 | 11,514 | 89% | 690,639 | 160% |
| 2015 | 9,217 | 89% | 446,946 | 83% |
| 2016 | 9,039 | 92% | 546,458 | 113% |
| 2017 | 5,540 | 92% | 137,968 | 66% |
| 2018 | 13,326 | 91% | 524,496 | 93% |
| 2019 | 11,613 | 90% | 563,892 | 94% |

**Figure 3: Historic SEM program electric savings and realization rates**

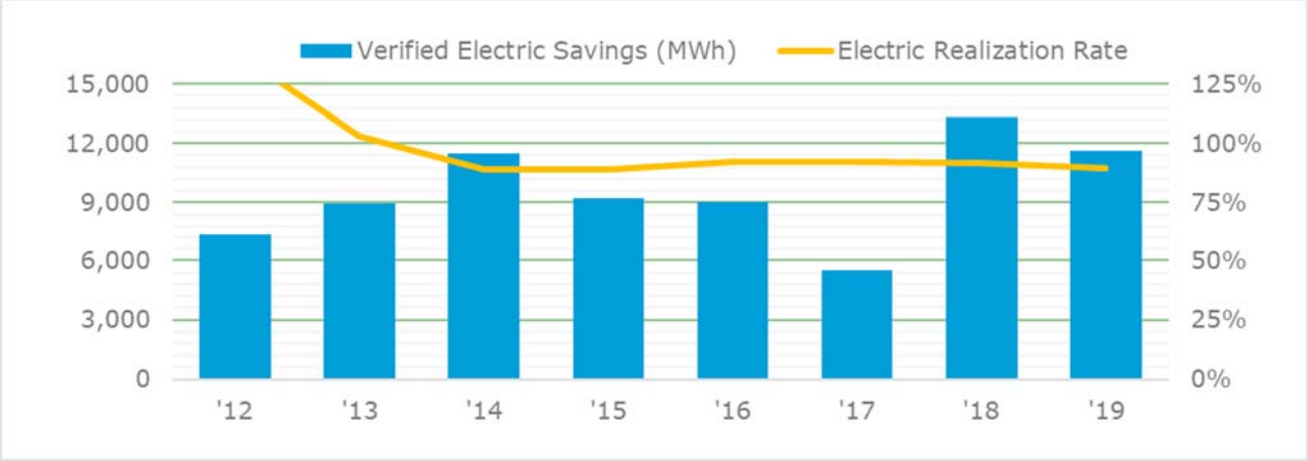
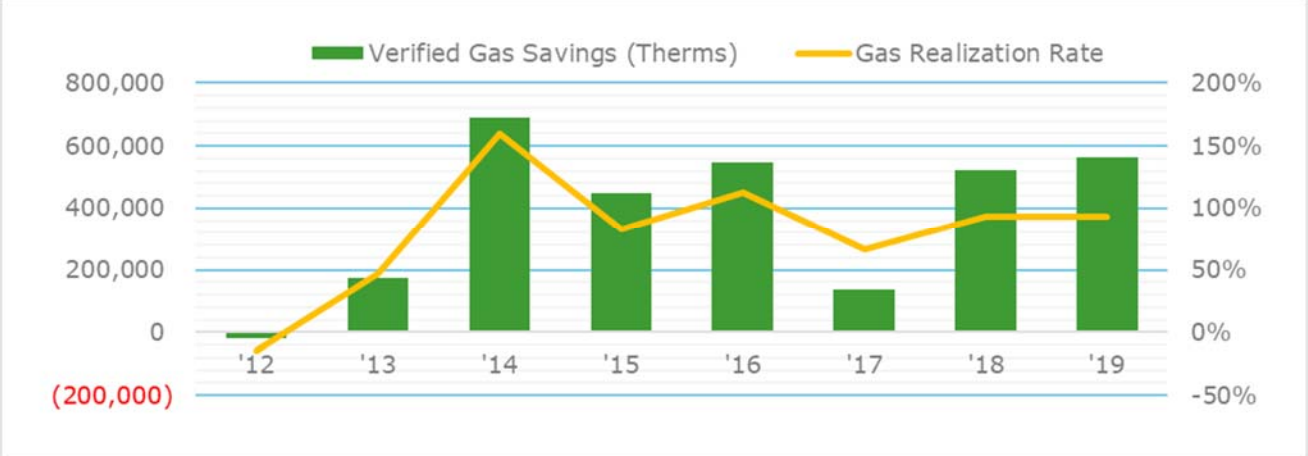



**Figure 4: Historic SEM program gas savings and realization rates**



## 0.7 Key evaluation findings and recommendations

This section provides key findings and recommendations resulting from this study. Additional findings are presented within each track-specific section.

### 0.7.1 Lighting track

**Finding** – We found that project documentation and program savings calculators were properly filled out and were sufficient for our evaluation needs. The program has done a good job of emphasizing the need for quality project documentation to trade allies and should continue to do so. We found no obvious errors in any of the assumptions used in the savings analysis.

For Standard Lighting measures we found good agreement between the reported operating parameters (lighting hours, quantities, and wattages) and the responses we received from site contacts.

Deemed savings values for Direct Install lighting projects tend to over-estimate the actual hours of operation, which results in a lower Direct Install evaluation realization rate. All DI measures assume 3,600 hours/year of operation per the regional mix from 2014 CBSA[1] data, regardless of business type or market.

- **Recommendation** – Based on results from PY2018 and PY2019 evaluation cycles, we suggest adopting an hours-of-use value more in line with actual consumption. 3,000 hours per year for lighting projects excluding controls would provide better alignment (large controls-only lighting projects have not typically over-estimated savings and need not be adjusted).

### 0.7.2 Standard track

**Finding** – For many projects completed under a corporate account (for example, this year's RTU Economizer / Demand Control Ventilation measures), it is often very difficult to track down a site contact specific to that location who is knowlegable enough to answer survey questions. The evaluation team was only able to speak with the Engineering Rebate Manager for the corresponding project, not the actual local site contact. This makes verification of site specific details, usage habits and other factors very uncertain, as Engineering Rebate Managers were only able to verify scope of installation at the time of install or before install. This echos a similar issue with a previous finding in the PY2018 impact evaluation.

- **Recommendation** – Limit the amount of projects applied through Engineering Rebate Managers, and/or consider revising the program application to have participants also provide a technically knowledgeable contact who is familiar with the installed measure (not just an administrative manager who applied for the measure incentive).

**Finding** – The evaluated savings for 12 of 13 sampled boiler projects are lower than reported savings. Seven of the sampled boiler projects were installed at schools, either at the primary, middle, or high school level. Boiler projects received considerable reduction in evaluated savings because of an error in assumptions used in the energy analysis that was uncovered as a result of the whole building gas consumption regression analysis.

- **Recommendation** – Projects utilizing the new boiler MAD UES values to estimate savings for high efficiency boiler installations should reduce this variance. If possible, review the total annual gas consumption of a school prior to finalizing savings in order to flag any sites for which savings is more

[1] Commercial Building Stock Assessment

than 20% of consumption. When flagged, adjust the savings claimed based on the consumption profile for the facility.

### 0.7.3 Custom track

**Finding** - Program models continue to estimate savings that suggest a significant reduction in annual consumption. DNV analyzed the actual change in facility consumption using the same regresssion methodology used for the Strategic Energy Management (SEM) evaluation. The COVID-19 impacts prevented this analysis from being used directly or as significantly as it has been used in past evaluations. However, the analysis did support the conclusion that two projects were not achieving any gas savings due to the measures installed.

- **Recommendation** – DNV continues to suggest that Energy Trust complete additional review of simulation inputs for sites expecting savings greater than 20% of consumption. Energy Trust should consider requiring ATACs to document in the technical analysis study (TAS) what simulation inputs are the largest drivers of savings for the project.

### 0.7.4 Strategic Energy Management

**Finding** – The Strategic Energy Management program has become a more complicated program over time, which has increased the cost to evaluate the program. The increase in complication is primarily driven by the increase in performance tracking tools (PTTs) used to estimate program savings. While it appears that improvement and consolidation of PTTs is occurring, there are still incidents where model inputs and information are located in inconsistent areas or are not appropriately accounted for in the model.

- **Recommendation** – DNV recommends that Energy Trust continue its efforts to create simplified and consistent PTT tools for program participants to use. DNV recommends the creation of a "Non-Routine Events" (NRE) log within the PTT that documents all capital projects (both those in the baseline and those during program years), any weather adjustments made, and any other NREs that are accounted for in the model (including baseline adjustments and gas curtailments). The log should state how the NRE is accounted for in the savings calculation.

# Memo

**To:** Board of Directors

**From:** Wendy Gibson, Sr. Program Manager – Existing Buildings
Kathleen Belkhayat, Program Manager – Commercial Energy Performance Management
Sarah Castor, Program Manager – Evaluation & Engineering

**cc:**

**Date:** August 13, 2021

**Re:** Staff Response to the Existing Buildings Program 2019 Impact Evaluation

The 2019 Existing Buildings program impact evaluation covered the program's four tracks: Custom, Lighting, Standard and Strategic Energy Management (SEM). As in past years, the evaluation found that the program is doing a good job of estimating savings for electric measures in all tracks, with an overall electric realization rate of 102%. Estimating gas savings proved more challenging, especially for Standard and Custom gas projects, and the overall realization rate was 86% for gas savings.

Energy Trust is committed to regularly updating the savings estimates and documentation for its standard measures, as recommended by the evaluator. The boiler measure was updated in late 2019 as recommended by the evaluator and an improvement in the realized savings for boilers should be seen in the impact of the 2020 program year. Energy Trust updated the measure analysis for direct install lighting in early 2021, before the recommendation to reduce assumed hours of use was received; this update included breaking out hours of use for different technologies rather than a single assumption for hours of use. When the measure is up for review again in early 2022, we will assess whether further changes are needed to improve savings accuracy.

In 2022, the program will begin developing a new performance tracking tool platform for SEM, which will make it easier for the program to assess engagement and consistently aggregate and analyze models to understand the correlation between actions and energy savings. Energy Trust will continue to track on evaluation results for K-12 public schools to identity ways to improve project performance for those buildings. We will also explore the possibility of collecting more relevant site contact information to facilitate evaluation.

As a health and safety precaution during the coronavirus pandemic, all data collection for this evaluation took place via telephone interviews or virtual site visits. These methods yielded the required data while also reducing travel costs associated with in-person site visits, and Energy Trust plans to continue using remote data collection methods in future evaluations to the greatest extent possible.

# 1 BACKGROUND

Energy Trust performs evaluations of its programs on a regular basis. DNV was selected to conduct an impact evaluation of Energy Trust's 2019 Existing Buildings program offering. This program offering is designed to deliver comprehensive energy efficiency options and services to commercial customers with existing buildings. The program offers incentives and technical support for the installation and operation of cost-effective energy efficiency measures for all major building end uses. This evaluation covers program year 2019. The goals of this evaluation were to:

- Develop estimates of Existing Buildings program gas and electric savings to establish realization rates for the 2019 program year. Information will be used for future program savings projections and budget developments and will be incorporated into Energy Trust's annual true-up of program savings.
- Report observations from the evaluation and make recommendations to help Energy Trust understand substantial deviations from claimed savings and to improve ex ante savings estimates and the effectiveness of future engineering studies and impact evaluations of Existing Buildings projects.

## 1.1 Energy Trust background

Energy Trust is an independent nonprofit organization, selected and overseen by the Oregon Public Utility Commission, to lead Oregon utility customers in benefiting from saving energy and generating renewable power. The services, cash incentives and solutions have helped participating customers of Portland General Electric, Pacific Power, NW Natural, Cascade Natural Gas, and Avista save more than $3.9 billion on their energy bills since 2002. The cumulative impact of their programs since 2002 has been a contributing factor in the region's low energy costs and in building a sustainable energy future. More information about Energy Trust's background, funding sources, strategic and action plans, policies and programs are available on their website at www.energytrust.org/about.

## 1.2 Program description

The Existing Buildings (EB) program began in March 2004 and is implemented by a program management contractor (PMC). ICF International was the PMC during program year 2019. The program had four main tracks in 2019: Custom, Lighting[2] (including standard, direct-install, and street lighting measures), Standard (simple calculations and prescriptive), and Strategic Energy Management (SEM). The program also maintained a few other tracks and pilots, which represented a small portion of program participants and savings. These small tracks were excluded from this evaluation. Custom track projects have their savings estimated through energy studies conducted by Allied Technical Assistance Contractors (ATACs). These studies may involve engineering calculations or energy simulation modeling. Standard Lighting track measures are installed directly by trade allies, while direct-install lighting measures were installed by a trade ally subcontractor to SmartWatt, under subcontract to the PMC. Standard track measures use savings estimates from reliable sources (including the Regional Technical Forum (RTF), ENERGY STAR, and others), as documented in Energy Trust measure approval documents (MADs). SEM savings are estimated based on a top-down analysis of building-level energy use and do not include savings from capital measures completed at the site through other program tracks during the SEM engagement. After completing a first

[2] As of calendar year 2021, Energy Trust has transitioned the Lighting track into its own program, separate from the Existing Buildings Program.

year of SEM, participants have the option of participating in Continuation, where they can claim additional savings and incentives for furthering their SEM activities.

## 1.3 Evaluation objectives

This evaluation was designed and completed to achieve the following primary objectives:

- Estimate the gas and electric savings achieved in program year 2019 (PY2019).
- Calculate gas and electric realization rates for PY2019.
- Provide savings and realization rates separately for SEM and non-SEM or Capital measures by fuel type.
- Provide realization rates to serve future program savings projections and budget developments.
- Report observations from the evaluation regarding program implementation and documentation, and compare assumptions regarding measure performance to actual performance.
- Provide recommendations to:
  - Understand substantial deviations from reported savings
  - Improve reported savings estimates
  - Improve effectiveness of future engineering studies and impact evaluations.

## 1.4 COVID-19 Adaptations

This evaluation was executed during a time of increased health and safety risk. DNV used remote methods for all data collection. Evaluated savings are not based solely on the as-found conditions during data collection. Evaluated savings are based both on as-found conditions and assumed post-installation normal conditions developed through participant interviews. This is grounded on our presumption that the operation conditions during the pandemic period are only temporary and do not substantially influence the lifetime performance and savings of the installed measures. Given the impact of COVID-19 on participant consumption, the evaluation does not directly use the results of any pre/post consumption analysis (IPMVP Option C) as the evaluated savings when the post period overlaps with the start of the COVID-19 pandemic in Oregon. Instead, the evaluation used the results of any pre/post consumption analysis as one piece of information available from which to estimate evaluated savings. In the case of Standard Boilers, the evaluation relied on pre-installation consumption instead of post-installation consumption to estimate facility heating load.

# 2 EVALUATION OVERVIEW

This section provides an overview of DNV's technical approach for the impact evaluation of this program. This section only describes the tasks used to determine the evaluated savings. Track-specific evaluation sections are provided following the overview. The track-specific sections discuss the actual activities and results for the program tracks.

## 2.1 Program database review

DNV reviewed the program tracking data provided by Energy Trust. This task helped DNV understand the measures and projects completed during the program year and begin to plan for the impact evaluation.

## 2.2 Sample design

DNV utilized stratified random sampling with certainty selection to identify the sample for this impact evaluation. Table 2-1 summarizes the final sample design implemented and the associated expected relative precision of the results. DNV intentionally oversampled for this evaluation due to an expectation that recruitment would face additional challenges during the COVID-19 pandemic. The sample and relative precisions shown are based on achieving the evaluation goal. All relative precisions are shown at the 90% confidence interval. The full sample design is discussed in Appendix A. The design for each track is discussed in the track-specific sections. The PY2019 evaluation sample frame included all measures supported by the four main tracks in Oregon during the year[3].

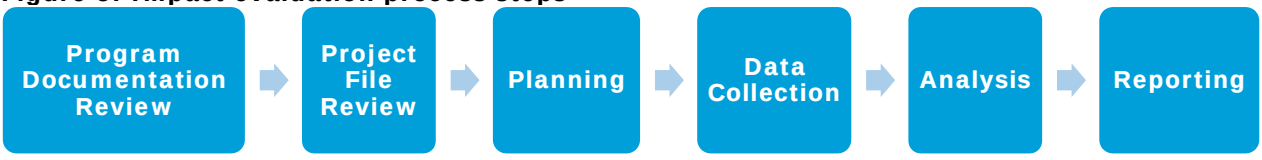
**Table 2-1: Sample summary**

| Program | Population | Evaluation Goal | % kWh | Electric Energy | Electric Demand | % therms | Gas Energy |
|---|---|---|---|---|---|---|---|
| Track | (N) | (n) | Sampled | Relative Precision | Relative Precision | Sampled | Relative Precision |
| Lighting | 2,822 | 33 | 7% | 14% | 21% | n/a | n/a |
| Standard | 749 | 68 | 44% | 14% | 19% | 45% | 10% |
| Custom | 159 | 34 | 48% | 12% | 16% | 57% | 13% |
| **Subtotal: Non-SEM** | **3,730** | **135** | **18%** | 10% | 15% | **52%** | 10% |
| SEM | 282 | 40 | 34% | 18% | n/a | 65% | 16% |
| **Total: All Tracks** | **4,012** | **175** | **20%** | **10%** | **15%** | **56%** | **10%** |

## 2.3 Site-specific evaluation

Site-specific impact evaluation was initiated after the final primary sample was identified. The site impact evaluation process steps used for this project are illustrated in Figure 5.

**Figure 5. Impact evaluation process steps**




The steps in this process were primarily applied at the track level and are discussed in the track-specific sections. A brief description of each step is provided below:

[3] The evaluation sample frame did not include NW Natural Washington or some small tracks (namely Cool Schools and Comprehensive Lighting Pilot), which were outside of the scope of this contract.

- Program Documentation Review: DNV reviewed a sample of project documentation to identify and understand what information is retained by Energy Trust to support compliance with the program's requirements and inform the estimate of savings for the project or measures. For sampled prescriptive measures, DNV also reviewed the measure approval documents.
- Project File Review: Our engineering team then conducted a thorough review of the project files for sampled projects, focused on the energy savings calculations and assumptions, feasibility study reports, and other supporting documentation. The review identified provided documentation, original calculation methodology, key uncertainty parameters and any concerns with the original savings estimation methods.
- Planning: Upon the completion of project document review and file review, DNV created a track, measure or site data collection and analysis plan based on the measures completed at each sampled site. This plan documented the project: the expected installed conditions, the data to be collected through the evaluation process, and the anticipated analysis method. In general, our plans followed the framework provided in the International Performance Measurement & Verification Protocol (IPMVP). However, there were times when the best evaluation approach was outside of the IPMVP framework. The following are the key elements that supplement the preparation of project evaluation plans:
  - **Evaluating Standard Calculated / Prescriptive Measures.** The measurement and verification (M&V) plan for standard calculated and prescriptive measures was the same across each measure selected for evaluation. The same information was gathered across all projects and the same analysis methodology employed unless project-specific circumstances required an alternative analysis method.
  - **Evaluating Complex Projects.** For projects with multiple interactive measures, the evaluation team reviewed all measures as one interactive system and estimated the achieved savings across all measures.
- Data Collection: Data collection occured through phone interviews and virtual site visits. The need for a site visit was determined based on the results of the program and project documentation review. Data collection activities verified equipment installation, verified operating conditions, and collected the information necessary to determine evaluated savings.
- Analysis: The ex-post savings analysis followed the M&V plan. DNV utilized the ex-ante savings estimation tools or their methodologies, unless the evaluators determined that there were major flaws in the ex-ante savings methodologies or determined that an alternative method provided a more reliable estimate of savings. For each sampled project, DNV produced estimates of evaluated electric and/or gas savings. DNV engineers also noted any opportunities for improvement in the accuracy of tracked savings estimates determined during the course of our analysis.
  - COVID-19 Analysis Adjustments: The current COVID-19 pandemic was considered when calculating a project's evaluated savings. The evaluated savings are based on post-installation normal conditions, not just as-found conditions. For each site, DNV assessed if the data collected directly represented normal post-installation loads, operation, and consumption or if adjustments are necessary to better represent normal consumption levels. Examples of inputs to savings estimates for which as-found conditions were not considered normal: tenant occupancy rates, space occupancy schedules, total building consumption, and current HVAC and lighting controls setpoints.

## 2.4 Sample extrapolation to track and program

DNV used a separate ratio estimator to obtain unbiased estimates of the total evaluated savings (either kWh or therms) for any group of interest. This estimator will yield, by design, unbiased estimates of some outcome measure, and is particularly beneficial when the outcome measure is correlated with something known for all members of the sample frame. In this case, the evaluated savings are logically correlated with claimed savings as listed in the tracking database. In general, the separate ratio estimator works as follows.

Suppose the indices:

$g$ = Application domains which are defined by track and fuel type (kWh or therms). For some outcome measures and domains of interest, strata had to be collapsed with one another during the estimation process. This occurred with $Y_g \neq 0$ but $\sum_{i \in Sample} w_{ig} y_{ig} = 0$ (these terms are defined below).

$i$ = Site.

And suppose:

$x_{ig}$ = Evaluated savings for site $i$ in group $g$.

$y_{ig}$ = Claimed savings for site $i$ in group $g$.

$w_{ig}$ = Sample weight for site $i$ in group $g$. This reflects the sample selection process that was used at the beginning of the study to select the original 202 sample points.

$Y_g$ = Population total claimed savings in group $g$. So $Y_g = \sum_{i \in Frame} y_{ig}$

$\hat{R}_g = \dfrac{\sum_{i \in Sample} w_{ig} x_{ig}}{\sum_{i \in Sample} w_{ig} y_{ig}}$ is the Ratio estimate for group $g$.

Then the separate ratio estimator that will yield the total evaluated savings is:

$$\hat{T} = \sum_g \left( Y_g \cdot \hat{R}_g \right)$$

And the ratio estimate of total modeled savings to total claimed savings is:

$$\hat{R} = \frac{\hat{T}}{\sum_g Y_g}$$

The procedure used for calculating ratio estimation by domains provides the correct standard error of the estimate for each domain and overall. The procedure also takes into account defined clusters of observations (customers) and stratification.

The standard error is calculated as drawn from a finite population: the measures completed within the analysis period with associated energy impacts in the program-tracking database. This calculation uses the Finite Population Correction (FPC) factor. This factor is a reduction to the calculated variance that accounts

for the fact that a relatively large fraction of the population of interest has been observed directly and is not subject to uncertainty. It is appropriate to apply precision statistics, such as confidence intervals, based on the standard error calculated in this manner when quantifying the results of the program during the study period only. The FPC factor reduces the calculated sampling error around the estimate more for smaller populations than for large.

# 3 LIGHTING TRACK EVALUATION

The lighting track evaluation includes three lighting delivery groups: Standard lighting, Direct Install lighting, and Street lighting. Table 3-1 shows the reported savings for lighting by delivery track. Table 3-2 shows the population frame for lighting measures. These measures represent about 65% of the electricity savings reported by the program.

**Table 3-1: Reported lighting track energy savings for 2019**

| Track | Electricity (kWh) |
|---|---|
| Direct Install Lighting | 8,911,893 |
| Standard Lighting | 74,256,806 |
| Street Lighting | 3,701,035 |
| **Lighting Total** | **86,869,734** |
| Existing Buildings program total | 134,130,071 |
| *Percent of Existing Buildings program savings* | 65% |

## 3.1 Sample design

DNV used stratified random sampling to select a representative sample of projects for evaluation designed to provide reliable savings estimates. Key elements of the design are:

- Creation of domains for Direct Install, Street Lighting, and Standard to ensure that all were represented in the evaluation sample.
- Stratification by size of savings reported (two size strata were used for Direct Install and Street Lighting domains; four size strata were used for the Standard domain) to increase the magnitude of savings evaluated and minimize the expected relative precision of evaluated savings.

Sampling occurred at the project level (Project ID). Table 3-2 summarizes the sample design for the lighting track. DNV intentionally oversampled for this evaluation due an expectation that recruitment would face additional challenges during the COVID-19 pandemic. DNV's goal for completed evaluated projects was 33 for this track. This design was expected to provide program year savings estimates with 13% relative precision at the 90% confidence level. Further detail on the sample design is available in the 2019 Existing Buildings Sampling Memo.

**Table 3-2: Lighting track sample design**

| Sub-Category | Fuel | Size Stratum | Population (N) | Primary Sample (n) |
|---|---|---|---|---|
| Direct Install Lighting | Electric | 1 | 407 | 7 |
| | | 2 | 126 | 6 |
| Standard Lighting | Electric | 1 | 1,574 | 6 |
| | | 2 | 432 | 6 |
| | | 3 | 204 | 6 |
| | | 4 | 71 | 6 |
| Street Lighting | Electric | 1 | 7 | 2 |
| | | Certainty | 1 | 1 |
| **TOTAL** | | | **2,822** | **40** |
| **EVALUATION GOAL** | | | | **33** |
| Percent of Reported kWh in primary sample | | | | 3% |

## 3.2 Lighting track evaluation methods

This section discusses the activities completed and associated findings of the impact evaluation.

### 3.2.1 Summary of approach

DNV completed these steps to evaluate this track:

- Documentation and file review: Review tracking data to identify savings reported, units reported, and measure codes used. Review of standard lighting calculator. File review to verify reported information through invoices and other provided documentation.
- Data Collection planning: Identification of the key input parameters for impact evaluation. Updates to impact evaluation data collection tool.
- Data collection: Phone interview of sampled participants using the instruments developed.
- Analysis: Estimate evaluated savings using the data collected to update key parameters.

### 3.2.2 Documentation and file review

DNV reviewed the project documentation provided for all projects included in the original sample design. There were two key findings from this review.

- Documentation was sufficient. The documentation for all standard and street lighting projects was comprehensive and included all relevant files. Documentation for several direct install projects was initially incomplete but was supplied by Energy Trust when requested.
- Calculation methodology reviewed. The program used a standard calculator (Excel workbook) to estimate project savings. No custom savings calculation workbooks were identified. The standard calculation tool was the same as the tool used in prior Existing Buildings evaluation with updates to wattages corresponding to Energy Trust's current standard calculator.

### 3.2.3 Data collection planning

DNV developed or updated data collection plans and tools to accomplish the impact evaluation.

The data collection plan focused on acquiring information to validate the accuracy of these key parameters used to estimate lighting energy savings:

1. **Annual hours of use** (*$Hours_{annual}$*) is the most uncertain savings parameter. Reducing uncertainty around this parameter is often the most beneficial outcome of lighting impact evaluations. The evaluation gathered information on:
    a. Self-reported facility or fixture schedules (by space)
    b. Lighting fixture controls by space (occupancy sensors, timers, photocell controllers, combination of controls)
    c. Behavioral changes due to change in lighting fixture or lighting controls

2. **Delta watts** *($\Delta W$)* is the difference between the pre-existing lighting fixture wattage and the installed lighting fixture wattage. Verification of $\Delta W$ included examination of:
    a. Pre-existing fixture types (including ballast type)
    b. Pre-existing fixture/lamp conditions (e.g., 4 lamp T8 fixtures but 20% of fixtures had 1 or more failed lamps)
    c. Pre-existing fixture wiring or behavioral usage (e.g., 3-lamp T8 fixture wired to turn on 1 lamp, 2 lamps, or all 3 lamps; users turned off half of the bay lights in the afternoons)
    d. Installed fixture types

e. Installed fixture wiring and replacement strategy (e.g., were installed fixtures wired the same as the fixtures they replaced? Were they installed on a 1:1 ratio?)

**3. Quantity**

a. Pre-existing fixture quantities (by space and/or fixture type)
b. Installed fixture quantities (by space and/or fixture type)
c. Quantity of fixtures added or removed since the original install date

**Interactive effects:** Current Energy Trust policy does not account for heating and cooling interactive effects on lighting measures.[4] DNV agrees with previous program evaluators[5] that interactive effects should be included to accurately estimate the value of the program. For this study, DNV estimated savings without interactive effects in order to directly assess the accuracy of the original savings calculations.

### 3.2.4 Data collection

In response to the COVID-19 pandemic, all data collection occurred via telephone interviews. We spoke with facility owners or operators to collect key parameter information.

### 3.2.5 Project level analysis

DNV developed a savings calculation workbook template that follows the methodology (flow and function) of Energy Trust's standard savings tool used in the lighting program for standard and street lighting projects. Savings that were claimed by Energy Trust and sampled by the evaluation were first re-created in the savings calculation workbook. Evaluated energy savings were calculated in the same workbook by adjusting the key savings parameters. The values used were determined from the most valid data source available.

**Key Savings Parameters -** The key savings parameters researched were:

- Annual hours of use ($Hours_{annual}$)
- Delta wattage (difference between pre-existing lighting fixture wattage and the installed lighting fixture wattage, ΔW)
- Quantity

Using these key savings parameters, direct annual energy (kWh) savings are very generally described as:

$$kWh_{savings} = \sum_{measures} \Delta W \times Hours_{annual} \times Quantity$$

As described in Section 3.2.3, we also included an interactive factor to estimate total evaluated savings for each project. The estimate based on interactive factors was not included in the results.

## 3.3 Lighting track evaluation results

This section presents the results of DNV's impact evaluation for this track.

### 3.3.1 Achieved sample

Table 3-3 shows the final sample (by number of projects) achieved across the entire lighting track. The final achieved evaluation sample differed from the sample design due to the following:

- Five sites never responded to emails or phone messages requesting interviews or site visits.
- One site went out of business before the evaluation.

[4] Heat is a byproduct of lighting. As lighting efficiency increases, the waste heat it generates decreases. This has an interactive effect on HVAC costs. During heating months, heating systems typically work harder to make up the heat that used to be generated by the lighting. In cooling months, the HVAC typically consumes less energy.

[5] Energy Trust of Oregon, Impact Evaluation of the 2013-2014 Existing Buildings Program, Prepared by ADM Associates Inc., 02/09/17. Available at: https://www.energytrust.org/wp-content/uploads/2017/02/EB_Impact_Evaluation_2013_2014.pdf

- Three sites were under new ownership or management at the time of the evaluation and had no knowledge of the program.

**Table 3-3: Final lighting track sample summary**

| Sub-Category | Size Stratum | Primary Sample (n) | Achieved Sample (n) | % Complete |
|---|---|---|---|---|
| Direct Install Lighting | 1 | 7 | 7 | 100% |
| | 2 | 6 | 5 | 83% |
| Standard Lighting | 1 | 6 | 4 | 67% |
| | 2 | 6 | 3 | 50% |
| | 3 | 6 | 4 | 67% |
| | 4 | 6 | 5 | 83% |
| Street Lighting | 1 | 2 | 2 | 100% |
| | 2 | 1 | 1 | 100% |
| **TOTAL** | | **40** | **31** | **78%** |
| **EVALUTION GOAL** | | **33** | **31** | **94%** |

## 3.3.2 Evaluated savings

Expansion from the sample to track-level results follows the methodology discussed in Section 2.4. Realization rates by installation method and for lighting as a whole are presented in Table 3-4. Overall, DNV estimates the evaluated lighting savings across all technologies and delivery channels to be 108% of the reported savings with a relative precision of 10.1% at the 90% confidence level.

**Table 3-4: Lighting track electric impact evaluation results by sub-category**

| Sub-Category | Completed Sample | Realization Rate | Standard Error | Rel. Precision @ 90% Confidence |
|---|---|---|---|---|
| Direct Install | 12 | 80% | 0.094 | 19.4% |
| Standard Lighting | 16 | 112% | 0.076 | 11.2% |
| Street Lighting | 3 | 100% | 0.006 | 1.0% |
| **Lighting** | **31** | **108%** | **0.067** | **10.1%** |

## 3.3.3 Savings variance

The Standard Lighting program had an overall gross realization rate (GRR) of 112%. The assumptions used to estimate reported savings appeared reasonable; twelve of the sixteen (75%) standard lighting projects evaluated had realization rates between 90% and 100%. The building type, reported savings, and evaluated savings for these projects are shown in Table 3-5, along with the main reasons for variance. The difference between the evaluated and reported hours of operation were a main variance driver in all 4 standard lighting projects with a GRR variance of more than 10%.

**Table 3-5: Standard Lighting variances by project (GRR variance > 10%)**

| # | Building Type | kWh Reported | kWh Evaluated | GRR | Primary Cause of Variance |
|---|---|---|---|---|---|
| 1 | Manufacturer | 8,141 | 9,173 | 113% | Actual hours are greater than reported |
| 2 | Retail | 55,037 | 61,729 | 112% | Actual hours slightly greater than reported; delta wattages slightly higher than claimed in calculation tool |
| 3 | Convention Center | 482,042 | 932,315 | 193% | (Controls measure) Occ censors in convention halls are keeping lights off much longer than 25% reduction projected; billing data supports result |
| 4 | Warehouse | 316,577 | 449,927 | 142% | Actual hours somewhat longer than reported |

The Direct Install Lighting program had an overall gross realization rate of 80%. Six projects achieved less than 90% GRR; two exceeded 110% GRR. The building type, reported savings, and evaluated savings are shown in Table 3-6, along with the main reason for variance.

**Table 3-6: Direct Install Lighting variances by project (GRR variance > 10%)**

| # | Building Type | kWh Reported | kWh Evaluated | GRR | Primary Cause of Variance |
|---|---|---|---|---|---|
| 1 | Automotive Repair | 20,406 | 13,301 | 65% | Actual hours shorter than deemed |
| 2 | Small Office | 3,636 | 1,796 | 49% | Actual hours shorter than deemed |
| 3 | Mini-mart | 4,892 | 6,885 | 141% | Actual hours longer than deemed |
| 4 | Small Retail | 14,326 | 4,965 | 35% | Actual hours shorter than deemed |
| 5 | Small Retail | 23,271 | 6,153 | 26% | Actual hours shorter than deemed |
| 6 | Automotive Repair | 27,648 | 15,231 | 55% | Actual hours shorter than deemed |
| 7 | Grocery | 29,525 | 35,091 | 119% | Actual hours longer than deemed |
| 8 | Manufacturing | 50,446 | 35,208 | 70% | Actual hours shorter than deemed |

## 3.4 Lighting track findings and recommendations

Our evaluation findings and recommendations specific to the lighting track are presented in this section. We will address the study's specific research questions first.

### 3.4.1 Track recommendations

#### 3.4.1.1 PY2019 Evaluation Research Questions

**Are there project files for every site and do those files contain complete information?**

**Evaluation Response**: As discussed in 3.2.2, we found that project documentation and program savings calculators were properly filled out and were sufficient for our evaluation needs. The program has done a good job of emphasizing the need for quality project documentation to trade allies and should continue to do so.

**Are there obvious errors in any of the assumptions used in the energy analysis?**

**Evaluation Response**: We found no obvious errors in any of the assumptions used in the savings analysis.

**Were there any post-installation changes in operating parameters or associated assumptions? If so, what were the consequent changes in energy savings?**

**Evaluation Response**: For Standard Lighting measures we found good agreement between the reported operating parameters (lighting hours, quantities, and wattages) and the responses we received from site contacts.

**What are the factors that result in large variances in energy savings from program estimates?**

**Evaluation Response**: Deemed savings values for Direct Install lighting projects tend to over-estimate the actual hours of operation, which results in a lower DI evaluation realization rate. All DI measures assume 3,600 hours/year of operation per the regional mix from 2014 CBSA[6] data regardless of business type or market.

- **Recommendation** – Based on results from PY2018 and PY2019 evaluation cycles, we suggest adopting an hours-of-use value more in line with actual consumption. 3,000 hours per year for lighting projects excluding controls would provide better alignment. (Large controls-only lighting projects have not typically over-estimated savings and need not be adjusted.) Any recommendations we make with respect to your lighting programs will, of course, be tempered by your annual assessment of EISA uncertainty.

**Evaluation Response**: Estimated kWh savings for four of the sixteen evaluated Standard Lighting sites differed by at least 10% from reported savings.

- **Recommendation** – We recommend that Energy Trust continue with current procedures. Program staff should continue to emphasize the importance of accurate estimates of operating hours during training for trade allies. DNV does not recommend any program change; any change would likely increase program complexity with no assurance that it would improve savings estimates.

#### 3.4.1.2 Other lighting findings and recommendations

**Finding** – As we have noted in previous evaluations, the Existing Buildings program does not account for the interactive effect of reduced lighting power on building HVAC systems. This has the potential to result in an oversimplified view of the societal value delivered by the program. The conclusions of the previous evaluator[7] are supported by DNV.

- **Recommendation** – We recognize the difficulty of designing a program that delivers both simplicity and rigor, but we continue to believe that Energy Trust should consider including estimates of interactive effects with HVAC systems when calculating the societal impact of their lighting programs. Energy Trust should consider changes to its savings calculation workbook (to include interactive effects) but should also continue to weigh the changes against increased workbook complexity. Future impact evaluations should continue to also estimate the impacts from interactive effects of lighting projects on all building systems in support of Energy Trust adopting the inclusion of interactive effects.

[6] Commercial Building Stock Assessment

[7] Energy Trust of Oregon, Impact Evaluation of the 2013-2014 Existing Buildings Program, Prepared by ADM Associates Inc., 02/09/17. Available at: https://www.energytrust.org/wp-content/uploads/2017/02/EB_Impact_Evaluation_2013_2014.pdf

# 4 STANDARD TRACK NON-LIGHTING EVALUATION

This section documents DNV's impact evaluation of savings reported through the standard non-lighting track (standard track). The standard track offered non-lighting prescriptive incentives for a large variety of electric and natural gas energy efficiency measures including refrigeration, cooking, HVAC, building shell, and office equipment.

The program estimates measure energy savings in this track using per-unit energy savings (UES) values that were either *stipulated values* or *calculated values* using a standard formula and equipment or site-specific measure characteristics. The standard track measures accounted for about 5% of the 2019 Existing Buildings program's reported electricity savings and 28% of the reported gas savings. Table 4-1 presents the energy use for the standard track measures and the overall Existing Buildings program.

**Table 4-1: Reported standard track energy savings for 2019**

| Track | Electricity (kWh) | Gas (Therms) |
|---|---|---|
| **Standard Non-Lighting** | **7,192,795** | **578,998** |
| Existing Buildings program total | 134,130,071 | 2,063,321 |
| *Percent of Existing Buildings program savings* | 5% | 28% |

## 4.1 Sample design

DNV used stratified random sampling to select an efficient representative sample of projects for evaluation designed to provide reliable savings estimates across program fuels. The sample design target included:

- Four technology subcategories determined based on the measure type before aggregation for sampling. These subcategories ensure that a variety of measures are selected for evaluation.
- Sample stratification and sampling at the project level, using unique Project IDs provided in the tracking data. All measures completed within a single sampled project of the same technology were therefore selected for evaluation.
- Stratification by primary fuel type to ensure the evaluation results include measures saving both gas and electricity.
- Stratification by size of savings reported to increase the magnitude of savings evaluated and minimize the expected relative precision of evaluated savings.

The sample design resulted in the selection of 83 projects for evaluation. DNV intentionally oversampled for this evaluation due an expectation that recruitment would face additional challenges during the COVID-19 pandemic. DNV's goal for completed evaluated project was 68 for this track. Sample goals were not set by stratum. Assuming 68 projects are evaluated, the sample was expected to provide program year savings estimates with 14% electric relative precision and 10% gas relative precision at the 90% confidence level. Further detail on sample design is available in Appendix A. Table 4-2 shows the design for this track.

**Table 4-2: Standard track sample design**

| Technology | Primary Fuel | Size Stratum | Population (N) | Primary Sample (n) |
|---|---|---|---|---|
| Refrigeration | Electric | 1 | 7 | 4 |
| | | Certainty 1 | 3 | 3 |
| | | Certainty 2 | 3 | 3 |
| | | Certainty 3 | 5 | 5 |
| Others | Electric | 1 | 155 | 7 |
| | | 2 | 30 | 6 |
| | Gas | 1 | 55 | 6 |
| | | 2 | 13 | 5 |
| | | 3 | 3 | 3 |
| Boiler | Gas | 1 | 17 | 4 |
| | | 2 | 9 | 4 |
| | | 3 | 6 | 4 |
| | | Certainty | 4 | 4 |
| Food Equipment | Electric | 1 | 59 | 6 |
| | Gas | 1 | 175 | 7 |
| | | 2 | 136 | 6 |
| | | 3 | 69 | 6 |
| **TOTAL** | | | **749** | **83** |
| **EVALUTION GOAL** | | | | **68** |
| Percent of Reported kWh goal | | | | 44% |
| Percent of Reported therms goal | | | | 45% |

## 4.2 Standard track evaluation methods

This section discusses the activities completed to evaluate this track.

### 4.2.1 Summary of approach

DNV used two approaches for the evaluation of standard track measures: *measure-specific* and *project-specific*. The following steps were completed in both approaches:

1) Documentation and file review: Reviewed tracking data to identify savings reported, units reported, and measure code used. Review of new and updated Measure Approval Document (MAD) to understand the eligibility requirements, savings algorithms, and savings values used to support reported savings. Reviewed project files to verify reported information through invoices and other provided documentation.
2) Data collection planning: Identified the key input parameters and stipulated values to research and how they should be verified (i.e. file review, phone interview, internet lookup, etc.). Then, created a list of interview questions.
3) Data collection: Interviewed sampled participants by telephone using the survey instruments developed for this purpose.
4) Analysis: Estimated evaluated savings using the data collected to update key parameters and/or map to the most correct MAD value. At this point the evaluation proceeded with either a measure-specific or a project-specific analysis as described next.
   a) Measure-specific: DNV used a more systematic and standardized measure-specific approach for measure types that occur three or more times in the sample. For each of these measures we created an Excel workbook that contains the relevant tracking data extract, and sequentially documents

each phase of our analysis including the file review, phone verification questions and responses, analysis of all the collected data, and the final evaluated results and dispositions. There is typically one workbook for each type of measure and some workbooks encompass multiple measure types.

b) Project-specific: A more customized, project-specific approach was used for measure types occurring fewer than five times in the sample, which were referred to as low-frequency measures. If DNV developed a measure-specific approach during a previous evaluation, then that approach was used. For other measures, a single file was used for a more free-form review of the available information, logging of verification questions and responses, and evaluation analysis results and findings. Additional materials and calculations were also used as needed to support the analysis. However, summarized findings for the file review, phone verification, analysis, and the final numeric evaluated results for all of these measures were also tabulated in an Excel workbook.

Table 4-3 shows all of the measure types for which savings were claimed in the standard track in 2019, notes which were sampled and not sampled in this evaluation and notes the evaluation approach type implemented.

**Table 4-3: Standard track sample design, count of unique projects by measure type and analysis approach**

| Measure Description | Project ID Count | | Analysis Approach |
|---|---|---|---|
| | Population | Sample | |
| Gas Fryer | 338 | 16 | Measure Type Approach |
| Boiler | 36 | 16 | |
| Cooler Doors | 13 | 12 | |
| Heat pump | 95 | 7 | |
| Ceiling insulation | 24 | 5 | |
| Lighting | 43 | 4 | |
| Dishwasher | 33 | 4 | |
| Tanked water heater | 29 | 3 | |
| Electric Convection Oven - Full Size | 19 | 2 | Project Specific Approach |
| Wall insulation | 11 | 2 | |
| Demand Control Ventilation | 8 | 2 | |
| Pipe insulation | 5 | 2 | |
| Steam traps | 3 | 2 | |
| New Refrigerated Cases | 2 | 2 | |
| Powerstrip | 31 | 1 | |
| Economizer | 12 | 1 | |
| Electric Combination Ovens | 8 | 1 | |
| Icemaker | 6 | 1 | |
| Motors | 6 | 1 | |
| Battery Charger | 6 | 1 | |
| Custom refrigeration | 5 | 1 | |
| Tankless water heater | 5 | 1 | |
| Electric Griddles | 3 | 1 | |
| Vent Hood - Gas Heat | 3 | 1 | |
| Strip Curtains | 1 | 1 | |
| Gas Convection Oven - Full Size | 25 | 0 | Not Sampled |
| Electric Hot Food Holding Cabinet, Any Size | 7 | 0 | |
| Generator Block Heater | 6 | 0 | |
| Gas Combination Ovens | 4 | 0 | |
| Gas-fired Conveyor Broilers, 22-28", Full Territory | 4 | 0 | |
| Server Closet Mini-split AC units | 3 | 0 | |
| Electric Hot Food Holding Cabinet - Half Size | 3 | 0 | |
| Controls | 2 | 0 | |

| Measure Description | Project ID Count | | Analysis Approach |
|---|---|---|---|
| | Population | Sample | |
| Radiant heating | 2 | 0 | |
| Custom welder | 2 | 0 | |
| HVAC | 2 | 0 | |
| Electric Vat Fryer | 2 | 0 | |
| Gas Steam Cooker | 2 | 0 | |
| Undercounter - high temp - Gas water heat | 2 | 0 | |
| Showerhead | 1 | 0 | |
| Faucet aerator | 1 | 0 | |
| Electric Steam Cooker | 1 | 0 | |
| Gas-fired Conveyor Broilers 22-28", Gas Only Territory | 1 | 0 | |
| Rack Oven - Double Rack | 1 | 0 | |

## 4.2.2 Documentation and file review

DNV reviewed the applicable MAD as well as site-specific project file documentation for the sampled measures. This section discusses the results of our review.

### 4.2.2.1 Measure Approval Documents

For Standard track measures, savings calculation approaches and values are provided in the MAD files. We received and reviewed new versions of MADs each with a corresponding MAD ID and MAD ID version, both of which are indicated in tracking data. The following documents DNV's review process and findings for the one additional MAD reviewed for this evaluation:

- Measure baseline condition: The assumed baseline condition was identified in most MADs. Evaluators need to understand the assumed baseline condition in order to assess the reliability of measure savings.
- Measure units: The unit basis for each measure is the denominator for each unit energy savings (UES) value. For example, MAD 42.1 for steam traps uses savings per failed trap while the newer version 42.2 uses savings per lb/hr capacity input of steam trap. This example illustrates why clearly identifying the unit basis in MADs is important. The unit basis was not clearly identified in some MADs and evaluators need to understand the unit basis in order to assess the reliability of measure savings, though there was improvement over the results of previous evaluations. We were able to determine the unit basis by reverse engineering the savings values for these older MADs[8]. However, the UES unit basis should be clearly defined in MAD tables, and both the UES and unit basis should be reported as part of the tracking data. The evaluation team continues to use this process to evaluate the PY2019 tracking data, as unit basis was not explicitly provided in tracking data.

### 4.2.2.2 Project file review

Project documentation for standard track projects was typically complete and extensive and included the application form, invoice, technical performance specification sheet, and ENERGY STAR (ES) documentation for ES measures. Overall, DNV found the project file documentation for the standard track was well organized, easy to access, consistent with the tracking data, and sufficient for independent verification. Table 4-4 summarizes the issues the team uncovered.

[8] Measures MADs lacking a clear indication of the unit basis include: PTHP, Dishwashers, Steam Trap (version 42.1), Commercial Insulation, and Strip Curtains.

**Table 4-4: Summary of Standard track file review results**

| | |
|---|---|
| **Signed App or End-User Agreement?** | There were no issues |
| **Folder Contains Signed 140P Form?** | There were no issues. |
| **Final Project Claimed Savings Match Total Value in Project Folder?** | There were no issues. |
| **Building Type Specified?** | Building types for all projects were specified. |
| **Models / Calculations in Folder?** | For the 2019 evaluation, high-level methodology and calculators used are noted in documentation, but the actual calculators were not provided. ENERGY STAR calculators are implied based on estimates provided in the food service measures MAD. |
| **Enough data to recreate savings?** | All project folders include enough data for us to provide an independent estimate of energy savings. |

### 4.2.3 Data collection

The only data collection method for standard track measures was a telephone interview due to the health and safety risks of a site visit associated with COVID-19. DNV followed a recruitment and communication protocol approved by Energy Trust for this project. The questions and overall evaluation approach for each measure were guided by the measure eligibility requirements, size and performance characteristics, complexity, available tracking data, and MAD savings approach (stipulated or calculated values). For all measures, at a minimum we verified installation and active operation, confirmed the business type, reviewed business hours, and asked about pre-retrofit conditions. All measures also included measure-specific parameter or condition questions.

### 4.2.4 Measure analysis

DNV estimated evaluated savings for all sampled measures with completed data collection. Inputs for the evaluated savings calculations were determined from the most valid data source including the telephone interview, tracking data, MAD file review, project file review, and other independent research. We did not typically revise the MAD algorithms, but used the collected data to either calculate a revised value or, more typically, map to a more correct MAD value. For ENERGY STAR equipment we used the latest version of the ENERGY STAR appliance calculator and combined that with our primary data. Excel workbooks were used to process and document the analysis and evaluated savings results and assumptions. Measure results are presented in Appendix D.

#### 4.2.4.1 Boiler Project Consumption Review

DNV completed a whole building gas consumption regression analysis for sampled boiler projects. DNV used a similar methodology to that used to evaluate Strategic Energy Management (SEM) savings. A baseline degree-day model was trained on 12-24 months of pre-project consumption. This regression was used to estimate the volume of gas used for weather sensitive heating during the 12-months prior to project installation and the gas likely consumed by other end uses. DNV estimated the boiler load and project savings using simplified boiler efficiency assumptions. DNV then reviewed the estimated consumption and savings along with information gathered during data collection and determined if the load assumption in the

savings calculation should be adjusted based on the post-installation consumption profile. In most cases, the consumption review resulted in a significant reduction in evaluated boiler savings. For many of the sampled boiler projects, the program estimated boiler savings are not supported by the metered gas consumption at the participating facilities. This is discussed further in the recommendations below and in Appendix D.

## 4.3 Standard track evaluation results

This section presents the track-level results of DNV's impact evaluation of the standard non-lighting track.

### 4.3.1 Achieved sample

Table 4-5 shows the final sample achieved across the entire standard track. DNV estimated evaluated savings for 69% of the sampled measure lines (81% of sampled projects). The final achieved evaluation sample differed from the sample design due to refusals and non-responses. This includes participants who could not be reached after exhausting our phone call protocol as well as a small number who refused to participate in the survey. Our protocol required calling up to 5 times at different times of the day. We also tried contacting the participant by email if they did not respond to phone calls.

**Table 4-5: Final standard track sample summary, count of projects**

| Technology | Primary Fuel | Size Stratum | Primary Sample (n) | Achieved Sample | % Complete |
|---|---|---|---|---|---|
| Refrigeration | Electric | 1 | 4 | 4 | 100% |
| | | Certainty 1 | 3 | 2 | 67% |
| | | Certainty 2 | 3 | 3 | 100% |
| | | Certainty 3 | 5 | 5 | 100% |
| Others | Electric | 1 | 7 | 5 | 71% |
| | | 2 | 6 | 6 | 100% |
| | Gas | 1 | 6 | 5 | 83% |
| | | 2 | 5 | 4 | 80% |
| | | 3 | 3 | 3 | 100% |
| Boiler | Gas | 1 | 4 | 4 | 100% |
| | | 2 | 4 | 2 | 50% |
| | | 3 | 4 | 4 | 100% |
| | | Certainty | 4 | 3 | 75% |
| Food Equipment | Electric | 1 | 6 | 3 | 50% |
| | Gas | 1 | 7 | 4 | 57% |
| | | 2 | 6 | 4 | 67% |
| | | 3 | 6 | 5 | 83% |
| TOTAL | | | 83 | 66 | 80% |
| **EVALUTION GOAL** | | | **68** | **66** | **97%** |
| Percent of Reported kWh in final sample | | | | 41% | |
| Percent of Reported therms in final sample | | | | 38% | |

### 4.3.2 Evaluated savings

Realization rates by sampling domain are shown in Table 4-6 below. The counts shown are the counts of unique projects within each fuel. If a project saved both fuels, then it is shown in both the top Electric and bottom Gas sections.

**Table 4-6: Standard track impact evaluation results by sampling technology and fuel saved**

| Technology | Fuel Saved | Projects Evaluated | Realization Rate | Standard Error | Rel. Precision @ 90% Confidence |
|---|---|---|---|---|---|
| Refrigeration | Electric | 13 | 100% | 0.001 | 0.1% |
| Other | Electric | 12 | 97% | 0.023 | 4.0% |
| Food Equipment | Electric | 7 | 94% | 0.040 | 7.1% |
| **Standard - Electric** | | **32** | **98%** | **0.012** | **2.1%** |
| Refrigeration | Gas | 5 | 100% | 0 | 0.0% |
| Other | Gas | 12 | 99% | 0.020 | 3.5% |
| Boiler | Gas | 13 | 30% | 0.034 | 19.2% |
| Food Equipment | Gas | 13 | 99% | 0.058 | 9.8% |
| **Standard - Gas** | | **43** | **79%** | **0.030** | **5.5%** |

Expansion from the sample to track-level results follows the methodology discussed in Section 2.4. Table 4-7 shows the overall electric standard track realization rates.

**Table 4-7: Standard track electric impact evaluation results**

| Track | Completed Sample | Realization Rate | Standard Error | Relative Precision at 90 % Confidence | Evaluated Savings (kWh) |
|---|---|---|---|---|---|
| Standard-2019, Projects Evaluated | 32 | 98% | 0.012 | 2.1% | 7,053,360 |

Table 4-8 shows the overall gas standard track realization rates. The gas realization rate is driven primarily by the evaluation results for gas fryer and space heating boiler measures.

**Table 4-8: Standard track natural gas impact evaluation results**

| Sub-Category | Completed Sample | Realization Rate | Standard Error | Relative Precision at 90 % Confidence | Evaluated Savings (therms) |
|---|---|---|---|---|---|
| Standard-2019, Projects Evaluated | 43 | 79% | 0.030 | 5.5% | 457,991 |

## 4.4 Standard track findings and recommendations

Our evaluation findings and recommendations specific to the standard track are presented in two sections, one that addresses overarching MAD file and tracking data issues and the other to address measure-specific findings.

### 4.4.1 PY2019 Research Questions

**Are there project files for every site and do those files contain complete information?**

**Evaluation Response**: Nearly all project files provided adequate information but could have provided more project scope, measure configuration detail, or relevant contact info.

- **Finding** – For many projects completed under a corporate account (for example, this year's RTU Economizer / Demand Control Ventilation measures), it is often very difficult to track down a site contact specific to that location who is knowlegable enough to answer survey questions. The

evaluation team was only able to speak with the Engineering Rebate Manager for the corresponding project, not the actual local site contact. This makes verification of site specific details, usage habits and other factors very uncertain, as Engineering Rebate Managers were only able to verify scope of installation at the time of install or before install. This also echos the same issue with a previous finding in PY2018 with project contacts having little-to-no technical knowledge about installed measures.

- **Recommendation** – Limit the amount of projects applied through Engineering Rebate Managers, and/or consider revising the program application to have participants provide a technically knowledgeable contact who is familiar with the installed measure (not just an administrative manager who applied for the measure incentive).

## Are there obvious errors in any of the assumptions used in the energy analysis?

**Evaluation Response**: The evaluation results for high efficiency condensing boilers demonstrate that the estimated heating load provided per unit of boiler capacity was incorrect for the equipment supported in 2019.

- **Finding** – The per unit savings for boilers was not supported by site consumptions levels for the projects sampled as part of this evaluation. However, the revised MAD which Energy Trust started using for PY2020 aligns with the evaluation result.
- **Recommendation** – Continue using the new boiler MAD UES values to estimate savings for high efficiency boiler installations. If possible, review the total annual gas consumption of a school prior to finalizing savings in order to flag any sites for which savings are more than 20% of consumption. When flagged, adjust the savings claimed based on the consumption profile for the facility.

## Were there any post-installation changes in operating parameters or associated assumptions? If so, what were the consequent changes in energy savings?

**Evaluation Response:** The operating parameters for standard measures are mostly operating hours for restaurant measures.

- **Finding** – Foodservice measures' operating hours were verified to differ slightly from MAD assumptions. Post-installation changes were mostly associated with operating hours or days, and equipment efficiency parameters. In most cases, the installed equipment was more efficient than assumed in the MAD (higher savings), but operated fewer hours than assumed in the calculator (lower savings). These two adjustments balanced each other and final evluated savings closely aligned with the original estimates.
- **Recommendation** – Continue to update deemed methods to reflect more up-to-date savings and assumptions through the current measure update process. The variances observed in this evaluation do not warrant any immediate changes to measure assumptions.

## What are the factors that result in large variances in energy savings from program estimates?

**Evaluation Response:** As previously mentioned, boiler measures are achieving significantly less savings than originally estimated.

- **Finding** – The evaluated savings for 12 of 13 sampled boiler projects are lower than reported savings. Seven of the sampled boiler projects were installed at schools, either at the primary, middle, or high school level. Boiler projects received considerable reduction in evaluated savings

because of the aforementoned error in assumptions used in the energy analysis that was uncovered as a result of the whole building gas consumption regression analysis.

- **Recommendation** – Projects utilizing the new boiler MAD UES values to estimate savings for high efficiency boiler installations should reduce this variance. If possible, review the total annual gas consumption of a school prior to finalizing savings in order to flag any sites for which savings is more than 20% of consumption. When flagged, adjust the savings claimed based on the consumption profile for the facility.

### 4.4.2 Measure Approval Document and tracking data recommendations

In this section, we present our findings and recommendations for the MAD and tracking data.

**Do the measure approval documents used by the program include sufficient information to estimate reliable savings, and if not, what specific changes should be made to improve them?**

**Evaluation Response**: For the PY2019 evaluation, DNV received an updated set of MADs, each with a corresponding MAD ID and MAD ID version. Both MAD ID and MAD ID version were indicated for each record of tracking data. This is a good practice as it eliminated any ambiguity on MAD versions for each measure claim. The evaluation team reviewed all of the MADs within the evaluated sample measures. There is still room for improvement in terms of transparency and traceability to source the savings.

DNV understands that Energy Trust has invested a great deal of effort into updating the format and content of these documents over time. While creating, maintaining, and updating prescriptive measure assumption documentation is a time-consuming process without a perfect solution, DNV recommends that Energy Trust continue to explore opportunities to improve the transparency, content, and application of its prescriptive measure supporting documentation system. The evaluation team continues to suggest the following for the contents of each MAD to ensure sufficient information for reliable savings estimation.

- Each MAD should clearly specify the unit basis for the unit energy savings (UES) in a consistent location within the documents. For example, vent hood savings are calculated as kWh or therms saved per motor horsepower, but this is not stated within the MAD. We were able to determine this by plotting MAD savings values against motor horsepower. The UES unit basis should be clearly defined in MAD tables and reported as part of the tracking data.
- Whenever possible, the MADs should show the methods and assumptions used to estimate savings in a simplified form. If possible, a one-line calculation should be provided showing the average values or range of values calculated. When the input assumptions vary based on application, the MAD should include the look-up table used. These look-up tables should be included in the document, either in-line or as an embedded Excel file.
- In all cases, the MAD should cite either the research that supports the assumptions used or the industry standards that support the assumed value. This applies to inputs to savings calculations, the baseline and installed equipment assumed, the measure life, and measure costs.
- When possible, the MAD or referenced supporting documentation should document the confidence level and relative precision of the input assumption or savings estimation used. These values provide a clear indication of savings reliability.

### 4.4.3 Additional Measure-level recommendations

Findings and recommendations for the measures with the largest impact on the overall electric and gas realization rates not already mentioned are listed below.

- **Finding** – Insulation measures. We found two sites that may have triggered building code update due to the purpose/nature of the installed project. One was a strip-mall with multiple tenant insulation upgrade, the other was a school seismic retrofit. However, we do not have enough evidence to confirm. Program claimed savings were credited as normal for these sites.
- **Recommendation**: Revise measure application such that the scope and background of the projects will be provided more explicitly and revise measure requirement to communicate above-code requirement more clearly. Projects should identify whether an insulation project triggered building code or not; a project that triggers building code requirements should use those code requirements rather than existing condition as the baseline for consumption.

# 5 CUSTOM TRACK EVALUATION

The 2019 custom track reported 208 unique measure lines through 159 unique projects providing over 27 GWh and 800,000 therms in annual energy savings. These savings accounted for 20% of the program's reported electricity savings and 43% of the program's reported gas savings. Table 5-1 shows the reported savings for custom projects.

**Table 5-1: Reported custom track energy savings for 2019**

| Track | Electricity (kWh) | Gas (Therms) |
|---|---|---|
| **Custom** | **27,097,472** | **881,333** |
| Existing Buildings program total | 134,130,071 | 2,063,321 |
| *Percent of Existing Buildings program savings* | 20% | 43% |

## 5.1 Sample design

DNV used stratified random sampling to select an efficient representative sample of projects for evaluation designed to provide reliable savings estimates. Key design elements were:

- Creation of domains based on the primary fuel saved, electricity or gas. This helped ensure sufficient results for both fuels.
- Stratification by size of savings reported (up to five size strata were used) and use of a certainty stratum to increase the magnitude of savings evaluated and minimize the expected relative precision of evaluated savings.

Sampling occurred at the project level (Project ID). DNV's sample design resulted in the selection of a primary sample of 39 projects. **DNV intentionally oversampled for this evaluation due to an expectation that recruitment would face additional challenges during the COVID-19 pandemic. DNV's goal for completed evaluated projects was 34 for this track.** Table 5-2 summarizes the sample design for the custom track. This design was expected to provide program year savings estimates with 15% relative precision at the 90% confidence level. Further detail on sample design is available in Appendix A.

**Table 5-2: Custom track sample design**

| Primary Fuel | Size Stratum | Population (N) | Primary Sample (n) |
|---|---|---|---|
| Electric | 1 | 63 | 5 |
| | 2 | 19 | 5 |
| | 3 | 10 | 5 |
| | Certainty | 1 | 1 |
| Gas | 1 | 33 | 5 |
| | 2 | 14 | 5 |
| | 3 | 9 | 5 |
| | 4 | 6 | 4 |
| | Certainty | 4 | 4 |
| **TOTAL** | | **159** | **39** |
| **EVALUTION GOAL** | | | **34** |
| Percent of Reported kWh in sample goal | | | 48% |
| Percent of Reported therms in sample goal | | | 57% |

## 5.2 Custom track evaluation methods

### 5.2.1 Summary of approach

DNV completed the following steps for the custom track impact evaluation:

- Project file review: Review of project files provided by Energy Trust.
- Data collection planning: Creation of project-specific measurement and verification plans.
- Data collection: Phone interviews with sampled participants.
- Analysis: Estimated evaluated savings using the data collected to update key input parameters.

### 5.2.2 Project file review

DNV reviewed each sampled project file for sufficient documentation, program savings methodology, and accurate savings reporting. This review included:

- Verification of the existence of signed application or participation agreement
- Identification of the building type
- Perform a web-based search to determine if the sampled commercial entity was operating normally, operating under modified conditions, or closed
- Determination if the file folder contained enough information for evaluation
- Verification of the existence of engineering calculations and/or energy simulation models with outputs that match the reported savings
- Identificantion of key building or system operation paremeters contributing the reported savings
- Identification of building operating parameters that may have been revised because of COVID-19
- Assessment of reported savings as the percentage of baseline energy consumption
- Assessment of the completeness of documentation

### 5.2.3 Measurement and verification planning

DNV created project-specific M&V plans to guide the data collection effort. These site-level M&V plans were created for each sampled site using DNV's project-specific M&V Plan template. These plans focused on the collection of information specific to the key research parameters identified. The study did not collect information on all drivers of end-use energy consumption.

### 5.2.4 Recruitment

Each sampled participant was called up to five times at different times of the day. DNV also tried contacting the participant by email if they did not respond to phone calls. Energy Trust's program implementation staff and contractors also assisted with recruiting and scheduling custom program participants.

The evaluation team successfully recruited program participants associated with 25 unique projects to participate in the evaluation. DNV was unsuccessful at recruiting 14 projects in the primary sample. Five sites refused to participate in the study, and six were nonresponsive (i.e., reached our call limit) to our engagement attempts. Two sites were under new ownership and unaware of their custom project details. One site permanently closed. The sampled sites that refused to participate cited a lack of resources or a lack of sufficient knowledge about the project, possibly because of staff turnover, as reasons for refusing to participate. Most of the sites DNV was unsuccessful in recruiting were in the sample strata for smaller savings claims. These rates of refusal or nonresponsiveness are considerably higher than in past evaluations

of this track within the Existing Buildings program. We believe a reason for nonresponsiveness may be attributed to staffing reductions and increased level of staff working remotely or heavily reduced hours due to the COVID-19 disruption. These nonresponsive sites were dropped after repeated attempts and support by the Existing Buildings implementation team.

Additional sites were not added to the recruitment pool to prevent placing unnecessary burden on customers already contacted for other sampled projects and the belief that the resources required to increase the number of evaluated projects did not justify the improvement in evaluated savings precision expected.

### 5.2.5 Data Collection

All data collection occurred remotely, either via telephone, videoconference, or virtual inspection. Data collection followed the M&V plan developed for each project. In many cases, screen shots of current setpoints and schedules were captured to document the as-found building controls sequences. No independent metering was completed for this evaluation. Our data collection also included gathering information on how COVID-19 has impacted the building's current and foreseeable occupancy schedule, HVAC operation schedule and anticipated changes in HVAC operation setpoints such as outside air flow rate.

### 5.2.6 Project analysis

DNV estimated evaluated savings for 25 of the 39 projects originally sampled. The final sample includes 23 projects with electric savings claimed and 18 projects with gas savings claimed. Whenever possible, DNV used the same calculation tool used by the program to estimate savings with revised inputs where necessary. If claimed or initial evaluated savings represented more than 10% of annual pre-project consumption, then DNV also completed a consumption review to inform the final evaluation results. Inputs for the evaluated savings calculations were determined from the most valid data source including participant interviews, site EMS data, schedules, setpoints, program project files, and utility meter data. Typically, adjustments were made to the post-installation analysis in order to model the conditions observed by the evaluation. Instead of evaluating against the conditions found at the time of data collection, the team assessed projects based on conditions prior to the COVID-19 disruption or against reasonable estimates of future expected conditions once the pandemic-driven disruption has abated. In some cases, the evaluation adjusted the pre-existing or baseline inputs based on interviews with the participants.

The site-specific EM&V files include identification of the prime reason(s) of variance between evaluated and program reported savings. Furthermore, the evaluation analysis categorized the primary discrepancy reason for the evaluated savings, such as baseline adjustment, tracking calculation error, calculation methodology, facility changes, etc. Finally, the evaluation analysis attributed the savings variance to either program-controllable or program-uncontrollable factors and assigned their percentage contributions to the total evaluated savings variance. These attributions were meant to assess the contribution factors that the program can address to improve. Project-specific results were provided to Energy Trust separately.

#### 5.2.6.1 Consumption Review

DNV completed a whole building regression analysis for all custom projects that installed new building controls or were expected to reduce facility consumption by more than 10%. Only monthly meter reads were available for this analysis. DNV used the same methodology for consumption review as used to evaluate Strategic Energy Management (SEM) savings. A baseline degree-day model was trained on 12-24 months of pre-project consumption. The baseline consumption for post-installation meter reads was determined by forecasting consumption using the degree-day model and weather data associated with each meter read. DNV compared the results of the analysis to the engineering calculations and information

gathered during data collection to determine the final evaluated savings for the project. For the PY2019 evaluation, DNV assumed that any consumption associated with a March 2020 meter read or later included the effects of the COVID-19 pandemic. In all cases, this prevented direct use of the forecast M&V method for evaluation savings due to lack of post-installation meter reads.

In two cases, DNV used this methodology to support an evaluated gas savings of zero therms. Both projects contained sufficient post installation meter reads during the first half of the 2019-2020 heating season (pre-COVID-19 effects) to show that no reduction in gas consumption had occurred. These conclusions were supported by DNV's concerns with the original simulation assumptions used to estimate and calibrate baseline facility consumption.

## 5.3 Custom track evaluation results

### 5.3.1 Achieved sample

Table 5-3 shows the final sample achieved across the entire custom track. DNV estimated evaluated savings for 74% of the original sample target. Fourteen sampled sites were not evaluated. See section 5.2.4 above more information on evaluation recruitment.

**Table 5-3: Final custom track sample summary**

| Primary Fuel | Size Stratum | Primary Sample (n) | Achieved Sample (n) | % Complete |
|---|---|---|---|---|
| Electric | 1 | 5 | 2 | 40% |
| | 2 | 5 | 4 | 80% |
| | 3 | 5 | 3 | 60% |
| | Certainty | 1 | 1 | 100% |
| Gas | 1 | 5 | 3 | 60% |
| | 2 | 5 | 3 | 60% |
| | 3 | 5 | 2 | 40% |
| | 4 | 4 | 4 | 100% |
| | Certainty | 4 | 3 | 75% |
| **TOTAL** | | **39** | **25** | **64%** |
| **EVALUTION GOAL** | | **34** | **25** | **74%** |
| Percent of Reported kWh in final sample | | | 36% | |
| Percent of Reported therms in final sample | | | 42% | |

### 5.3.2 Evaluated savings

Expansion from the sample to track-level results follows the methodology discussed in Section 2.4. Table 5-4 shows the overall electric custom track realization rate and Table 5-5 shows the overall gas custom track realization rate. The realization rates are driven by numerous factors, including changes to building operation or use, errors in the program analysis, and adjustments to simulation inputs. DNV captures our findings and recommendations in the sections that follow. Note that both fuels were evaluated, irrespective of the primary fuel sampling domain to which the project was originally assigned. Even with the recruitment challenges faced and lower-than-planned projects, Energy Trust can have confidence in the result.

**Table 5-4: Custom track electric impact evaluation results**

| Track | Completed Sample | Realization Rate | Standard Error | Relative Precision at 90 % Confidence | Evaluated Savings (kWh) |
|---|---|---|---|---|---|
| Custom-2019 | 23 | 90% | 0.072 | 13% | 24,352,887 |

**Table 5-5: Custom track natural gas impact evaluation results**

| Track | Completed Sample | Realization Rate | Standard Error | Relative Precision at 90 % Confidence | Evaluated Savings (therms) |
|---|---|---|---|---|---|
| Custom-2019 | 18 | 86% | 0.064 | 13% | 755,154 |

### 5.3.3 Custom School Evaluated Savings

Electric and gas savings were evaluated for seven K-12 schools and service districts in the sample. Of those seven sites, two sites evaluated had only electric savings, and five sites had both electric and gas savings. Table 5-6 shows the average unweighted electric realization rate for custom measures in schools. The electric realization rate is driven by controllable and uncontrollable factors; however, the largest contributor to the average unweighted discrepancy derived from a single site's calculation methodology error (a controllable discrepancy).

For this discrepancy, the baseline was adjusted by the evaluator to represent an HVAC system that had basic scheduling functionality. This resulted in no savings for the scheduling portion of the project. We determined that under this specific baseline scenario it was unreasonable to assume that the HVAC system (installed as an additional capacity) would have been implemented without basic scheduling functionality. This discrepancy accounted for 93% (-681,180 kWh) of the total average unweighted electric discrepancy for the seven evaluated school sites.

**Table 5-6: Custom schools electric impact evaluation results**

| Track | # Projects | Program Savings (kWh) | Evaluation Savings (kWh) | Realization Rate |
|---|---|---|---|---|
| Custom Track, Schools | 7 | 1,655,302 | 923,105 | 56% |

Table 5-7 shows the overall gas realization rate for custom measures in schools. The average unweighted gas realization rate is also largely impacted by a single site's discrepancy. For this site, the evaluated gas savings were zeroed out because the billing data do not support evidence for savings. Gas savings were estimated from a reduction in outside air ventilation; however, billing data and EMS data corroborate that the estimated outside air ventilation rate was not retained by the facility. This discrepancy accounted for 92% (-23,463 therms) of the total average unweighted gas discrepancy for the five evaluated school sites with gas savings.

**Table 5-7: Custom schools natural gas impact evaluation results**

| Track | # Projects | Program Savings (therms) | Evaluation Savings (therms) | Realization Rate |
|---|---|---|---|---|
| Custom Track, Schools | 5 | 78,021 | 52,509 | 67% |

### 5.3.4 Custom track findings and recommendations

Our evaluation findings and recommendations specific to the custom track are presented as responses to Energy Trust's key research questions.

**Are there project files for every site and do those files contain complete information?**

**Evaluation Response:** DNV received complete project files for all sites. The project documentation was organized in a consistent and recognizable manner. Most modeling files (e.g., eQUEST, Trane Trace) re-created reported savings and DNV spent an acceptable amount of time determining whether the modelling files were adequate and did not require follow up with Energy Trust. There was one project where the Trane Trace model did not perfectly re-create reported savings. Modeled electric and gas savings were 6.4% and 9.8% different from reported savings. We believe this discrepancy was due to difference in software versions.

- See the final recommendations bullets concerning Trane Trace models.

**Are there any obvious errors in any of the assumptions used in energy savings analyses, either in the original savings estimates or in verification of energy savings?**

**Evaluation Response:** DNV did not identify any systemic errors in the energy savings analysis, and very few significant calculation errors were identified during this project. The evaluation team did identify multiple simulation inputs that resulted in inaccurate estimates of energy savings. Below is one example:

- For one school project the program reported savings were about 165% of the baseline energy consumption. This was due to the program savings analysis assuming continuous baseline HVAC operation even though the facility manually turns off the system at the end of the day. Because of this error in baseline assumption, the program analysis inflated the baseline EUI to 108.39 kBtu/sq. ft. compared to the actual baseline EUI of 41.58 kBtu/sq. ft., i.e. an inflation of 261%.

**Were there any post-installation changes in operating parameters and associated assumptions? If so, what were the consequent changes in energy savings estimates for individual projects?**

**Evaluation Response:** DNV notes a couple projects where post-installation changes in operating parameters significantly affected the realization rate.

- One school site claimed significant outside air ventilation reductions that appear to have not been retained by the school operators. The project documentation claims that outside air dampers were reduced from the assumed baseline position of 25% open. However, the collected EMS screenshots indicate outside air damper positions that were greater than 25%. Additionally, billing data (pre- and post-COVID-19 periods) do not suggest savings materialized from reduced damper positions. We could not determine the reasons that lower outside air damper positions could not be retained, but the corroborating data (EMS screenshots and billing data) gave confidence to our determination of zero gas savings.
- A hot aisle containment project was not able to retain the proposed return air temperature of 87 °F. The site cannot currently have return air temperature above 77 °F because of an issue with their physical server configuration. The site contact could not confidently answer when this issue could be resolved, in part because the servers must be handled/administered by the equipment owners. This finding reduced the realization rate to 67% (reduced kWh savings by 598,499 kWh).

- **Recommendation:** These discrepancies are considered uncontrollable and participants' proposed projects cannot necessarily be screened to control for them. However, the school site claimed savings was significant (42% of baseline annual consumption); Energy Trust should consider implementing for high savings projects more rigorous pre-implementation screening and post-implementation verification protocol for ensuring that the project measures can reasonably achieve retention. This protocol may involve closer collaboration with facility staff and equipment operators.

**What factors result in large variances in measures savings (assumptions too conservative, incorrect hours of operation, loads differ from expectations, etc.)?**

**Evaluation Response:** Besides the errors listed above, there were some other factors that resulted in large variances in savings:

- Baseline parameter adjustment: There were four projects where DNV modified the baseline conditions/assumptions. Note these adjustments are different from baseline equipment efficiency or baseline scenario modifications; they pertain to pre-existing set points, schedules, etc. The assumed parameters varied by project and had positive and negative impacts on realization rates. One of these baseline factors related to the assumed chilled water supply temperature, which was verified to be a few degrees higher than what had been assumed in the reported calculation (this lowered the savings). Another project had assumed an existing roof insulation value of R-8, but after verifying the insulation composition with the site contact and estimating the R-value using the same tool as used by the program, the insulation was determined to be R-6.6 (this increased the savings).
- Changes in operating schedule: The evaluation updated building operating schedules based on data gathered during the evaluation. In some cases, these schedules differed from the operating schedules used in the reported savings analysis.
- Changes in operating setpoints: Similar with last year, the majority of the ex post revisions made were related to the control setpoints used in savings models (simulation or spreadsheet). The evaluation updated setpoints based on the data gathered during the evaluation. Most of the changes were related to: occupied/un-occupied cooling and/or heating setpoints, economizer high limit setpoint, chilled water and hot water plant operating setpoints and reset range, cooling and heating supply air temperature setpoints and reset range, and cooling and heating lock-out temperatures.

**Are there trends in savings realization by ATAC firm completing the energy study?**

**Evaluation Response:** DNV did not identify any clear trends by ATAC firm. All firms with multiple projects sampled over the past three evaluations had some projects with low variance and some projects with high variance. No systematic reasons for high or low variance were identified. DNV has shared this data with Energy Trust for their review. Additional review could be completed working with Energy Trust to determine a consistent list of ATAC names that can be used for every program year.

**Are the projects using the appropriate baseline (existing conditions or current market) to estimate savings and cost-effectiveness?**

**Evaluation Response:** The projects typically used the appropriate baseline; however, there was a notable exception where an inappropriate baseline scenario was used. The measure savings were derived from a hypothetical code baseline where equipment efficiency was equivalent to 2010 ASHRAE 90.1. However, the measure savings had assumed pre-existing scheduling conditions where the equipment operated 24/7. DNV

did not consider this operating condition to be appropriate. Code control requirements should have also been considered.

**Do you have any recommendations regarding energy savings analysis approaches and assumptions, or customer behavior or decision-making that would be helpful to Energy Trust in designing, implementing, or evaluating its programs in the future?**

**Evaluation Response:** DNV believes the following adjustments will improve Energy Trust's program:

- **Finding** – Evaluating savings based on Trane Trace simulation models continues to be more challenging than other methodologies. There were multiple cases for which the evaluation could not replicate the savings estimates using the models provided. Additionally, the Trane Trace models are more challenging to evaluate due to the required measure-by-measure modeling structure and difference between software versions.
- **Recommendation** – The PMC should keep the final models within their database and a record of the software version used to estimate final savings. This should save the time and budget needed to identify and locate the final models used for the project. DNV first made this recommendation in the PY2017 impact evaluation report and believes it was implemented during PY2019.
- **Recommendation** – DNV also recommends that Energy Trust implement the following modeling order for multi-measure simulation models; the baseline model first, followed by equipment replacement measures, then the revised operating schedule measures and finally, the control changes. This approach ensures that the baseline used represents the pre-project operation and individual measure savings are estimated over its previous operating condition. Increasing consistency in the modeling methods used will increase the reliability of program savings over time.
- **Finding** - Program models continue to estimate savings that suggest a significant reduction in annual consumption. DNV analyzed the actual change in facility consumption using the same regresssion methodology used for the Strategic Energy Management (SEM) evaluation. The COVID-19 impacts prevented this analysis from being used directly or as significantly as it has been used in past evaluations. The analysis did support the conclusion that two projects were not achieving any gas savings due to the measures installed.
- **Recommendation** – DNV continues to suggest that Energy Trust complete additional review of simulation inputs for sites expecting savings greater than 20% of consumption. Energy Trust should consider requiring ATACs to document in the technical analysis study (TAS) what simulation inputs are the largest drivers of savings for the project.

# 6 STRATEGIC ENERGY MANAGEMENT EVALUATION

The SEM track reported 282 unique measures at 282 sites providing over 12.9 million kWh and over 600,000 therms in annual energy savings in program year 2019. These savings account for 10% of the program's reported electricity savings and 29% of the program's reported gas savings. Table 6-1 shows the reported savings for SEM in program year 2019.

**Table 6-1: Reported SEM track energy savings for 2019**

| Track | Electricity (kWh) | Gas (Therms) |
|---|---|---|
| **SEM** | **12,970,069** | **602,990** |
| Existing Buildings program total | 134,130,071 | 2,063,321 |
| *Percent of Existing Buildings program savings* | 10% | 29% |

## 6.1 Sample design

DNV used stratified random sampling to select an efficient, representative sample of projects for evaluation designed to provide reliable savings estimates. Key design elements were:

- Creation of domains separating first year sites (Year1) from sites that have participated over multiple year (Year2+). This helped ensure representation of both participants types in the evaluation findings.
- Creation of domains based on the primary fuel saved, electricity or gas. This helped ensure sufficient results for both fuels.
- Stratification by size of savings reported and use of a certainty stratum to increase the magnitude of savings evaluated and improve the expected relative precision of evaluated savings.

Sampling occurred at the site level (CRM site number). DNV intentionally oversampled for this evaluation due an expectation that recruitment would face additional challenges during the COVID-19 pandemic. DNV's goal for completed evaluated sites was 40 for this track. Evaluation goals were not established for any specific domain or stratum. Table 6-2 summarizes the sample design for the SEM track. This design was expected to provide program year savings estimates with 20% relative precision at the 90% confidence level. Further detail on sample design is available in Appendix A.

**Table 6-2: SEM track sample design**

| Year | Primary Fuel | Size Stratum | Population (N) | Primary Sample (n) |
|---|---|---|---|---|
| Year1 | Electric | 1 | 8 | 4 |
| | | 2 | 4 | 2 |
| | Gas | Certainty 1 | 2 | 2 |
| | | Certainty 2 | 1 | 1 |
| Year2+ | Electric | 1 | 128 | 10 |
| | | 2 | 23 | 9 |
| | | Certainty | 1 | 1 |
| | Gas | 1 | 96 | 8 |
| | | 2 | 14 | 7 |
| | | Certainty | 5 | 5 |
| **TOTAL** | | | **282** | **49** |
| **EVALUATION GOAL** | | | | **40** |
| Percent of Reported kWh goal | | | | 34% |
| Percent of Reported therms goal | | | | 65% |

## 6.2 SEM track evaluation methods

### 6.2.1 Summary of approach

DNV completed the following steps for the SEM track impact evaluation:

- Doumentation review: Review of project files provided by Energy Trust for sufficient documentation.
- Project file review: Review of project files provided by Energy Trust for program savings methodology and accurate savings reporting
- Data collection planning: Creation of a track-specific measurement and verification plan and interview guide.
- Data collection: Phone interviews with recruited sampled participants to review program engagement, site energy drivers, actions taken, and necessary non-routine adjustments (NRAs).
- Measure analysis: Estimated evaluated savings using independent regression analysis of savings during the program year described in Appendix B.

### 6.2.2 Documentation review

DNV reviewed each sampled project file for sufficient documentation. This review included:

- Verification of the existence of a final participant report for the program year and a file documenting the estimation of energy savings achieved during the program year.
- Identification of the building type
- Assessment of the completeness of documentation.

### 6.2.3 Project file review

DNV reviewed each sampled project file for program savings methodology and accurate savings reporting. This review included the following steps:

- Verifying stated meter numbers and/or account numbers
- Identifying how many years the site has participated in SEM
- Identifying if the site was previously evaluated or reviewed
- Identifying use of non-weather variables, polynomials, or multiple degree-day variables in the regression model(s)
- Determining if the energy savings reported in the database is supported by a single or multiple regression models
- Extracting the utility consumption data used by the program, and non-weather independent variables (if used) for each facility and identifying if additional data is needed from Energy Trust
- Determining if any baseline adjustments occur in the model
- Identifying what capital projects are included in the model and extracting the associated savings values applicable to the sampled fuel and program year.

### 6.2.4 Measurement and verification planning

The SEM M&V Plan and interview guide focused on documenting the facility being evaluated, its consumption, reported SEM actions, and identified capital projects. A plan for each site was created using the guide and then used as part of the data collection interview process.

### 6.2.5 Data collection

Data collection was executed per the site M&V plan through an in-depth interview of facility personnel, completed onsite or via telephone. Through the in-depth interview, DNV staff captured information to:

- Verify engaged participation in the program during the sampled program year
- Verify the actions taken during the sampled program year to reduce energy consumption
- Determine if the standard modeling approach is sufficient for the site and what changes are required if not
- Determine what capital improvements or non-SEM activities impacted energy consumption during the sampled program year
- Identify any operating conditions or changes to the facility that may have affected the energy savings or the validity of the performance tracking tool (PTT) model; this includes capital projects installed during SEM engagement
- Identify known seasonal changes in facility use that might prevent modeling using weather only
- Understand basic occupancy, cooling, heating, process schedules and associated control sequences that should be reflected in consumption data, such as typical start and stop to heating and cooling seasons and use of free cooling.

### 6.2.6 Measure analysis

DNV estimated evaluated savings for 46 unique sites. The data collected through the interviews was used to develop an estimate of evaluated savings achieved during the program year. To estimate savings, DNV developed independent standard regression models using monthly utility meter data, weather data, and provided or collected data for other independent variables determined to be necessary.

Model development followed Energy Trust's "Commercial O&M Measurement and Verification Guideline for Energy Trust of Oregon's Commercial Strategic Energy Management (SEM) and Pay for Performance (PfP) offerings." Model validity was tested per the Statistical Criteria for Model Fitness.

## 6.3 SEM track evaluation results

### 6.3.1 Achieved sample

Table 6-3 shows the final sample achieved across the entire SEM track. DNV estimated evaluated savings for 94% of the primary sample and achieved 115% of the evaluation goal. Overall, the response rate for SEM participants was higher for this evaluation than during previous evaluations of this track. The final achieved evaluation sample exceeded the evaluation goal but did not achieve the entire sample because of the following:

- Refusals and Non-Responses: Incomplete sample strata were primarily due to those participants that could not be reached after exhausting our phone call protocol, and a small number who refused to participate in the evaluation. The refusals were associated with mutliple sampled sites with multiple years of participation in impact evaluation studies. DNV and Energy Trust accepted the refusals due to each organization's recent evaluation participation history. The customer's Energy Coach also tried contacting the participant if they did not respond to DNV's requests.

**Table 6-3: Final SEM track sample summary**

| Year | Primary Fuel | Size Stratum | Primary Sample (n) | Achieved Sample (n) | % Complete |
|---|---|---|---|---|---|
| Year1 | Electric | 1 | 4 | 4 | 100% |
| | | 2 | 2 | 2 | 100% |
| | Gas | Certainty 1 | 2 | 2 | 100% |
| | | Certainty 2 | 1 | 1 | 100% |
| Year2+ | Electric | 1 | 10 | 10 | 100% |
| | | 2 | 9 | 8 | 89% |
| | | Certainty | 1 | 1 | 100% |
| | Gas | 1 | 8 | 7 | 88% |
| | | 2 | 7 | 4 | 57% |
| | | Certainty | 5 | 7 | 140% |
| **TOTAL** | | | **49** | **46** | 94% |
| **EVALUTION GOAL** | | | **40** | **46** | **115%** |
| Percent of Reported kWh in final sample | | | | 32% | |
| Percent of Reported therms in final sample | | | | 59% | |

### 6.3.2 Evaluated savings

Expansion from the sample to track-level results follows the methodology discussed in Section 2.4. Table 6-4 shows the overall electric SEM track realization rate. The site-specific electric realization rates varied from 0% to 1,536% for PY2019. Table 6-5 shows the overall gas SEM track realization rate. The site-specific gas realization rates varied from 44% to 311% for PY2019.

The significant site-level savings variance is due to differences in baseline model form, formula errors within PTT files, sites assessed as disengaged in the program, incorrect accounting for non-routine events, and the impact a different model form and program year schedule can have on the incremental savings estimated for each year of participation.

**Table 6-4: SEM track electric impact evaluation results**

| Track | Completed Sample | Realization Rate | Standard Error | Relative Precision at 90 % Confidence | Evaluated Savings (kWh) |
|---|---|---|---|---|---|
| SEM-2019 | 34 | 90% | 0.06 | 11% | 11,613,056 |

**Table 6-5: SEM track natural gas impact evaluation results**

| Track | Completed Sample | Realization Rate | Standard Error | Relative Precision at 90 % Confidence | Evaluated Savings (therms) |
|---|---|---|---|---|---|
| SEM-2019 | 26 | 94% | 0.04 | 8% | 563,892 |

## 6.4 SEM track findings and recommendations

### 6.4.1 Energy Trust questions

This section provides responses to Energy Trust's SEM track research questions.

### Are the original SEM models and results well documented?

**Evaluation Response:** For PY2019, the original models continue to be sufficiently documented in that the evaluation could identify the independent variables used and the associated coefficients. The evaluation did identify documentation that consistently communicates the testing of independent variables and whether they passed fitness tests. Most of the sampled continuation sites had been re-baselined since their previous evaluation. The evaluation found the updated model files and versions easier to navigate.

### Were there any deviations from the SEM modeling guidelines, and if so, was there a satisfactory explanation, and were the deviations justified?

**Evaluation Response:** Yes, deviations from the modeling guidelines continued to exist in PY2019. DNV did observe a significant increase in re-baselined continuation sites which eliminated many previously observed deviations from the guidelines. Modelers continue to sometimes optimize the model fit by adding second heating degree-day (HDD) or cooling degree-day (CDD) terms. The evaluation rarely found an explanation discussing why the additional term, increased complexity, and reduced independence of the model is necessary. The evaluation did not include these second HDD or CDD terms in our models.

### How did the original baseline SEM models compare to the models used for evaluation?

**Evaluation Response:** Differences between the baseline evaluation and original models continue, but the number and magnitude of differences is less than for previous evaluations of the program.

- DNV did find sites that use a second HDD or CDD terms to improve model fit. DNV believes the use of a second degree-day (DD) term increases the model complexity and risks over fitting the model. Any sites using non-standard models should be required to document why the additional variable is necessary. The evaluation models use only one DD term.

### Were any important variables omitted from the original model?

**Evaluation Response:** The evaluation did not identify any sites for which the model omitted an independent variable that should have been included. There were a few cases where unnecessary variables were used in the original model and removed in the evaluation model.

### Were capital measures properly accounted for in the estimation of SEM savings?

**Evaluation Response:** The evaluation found capital projects to be properly accounted for in most projects. One site had included partial capital project savings for PY2019 involving a project that had not been fully implemented. With input from Energy Trust, DNV removed the capital project add-back resulting in positive electric incremental savings for PY2019.

### Are the SEM guidelines sufficient to guide the development of reasonable and robust models of savings?

**Evaluation Response:** Based on the realization results for PY2019 and previous years, DNV believes that the SEM guidelines are sufficient to continue using for development of models used in the SEM program. The guidelines mention important topics including M&V boundaries, model types, re-baselining, validation, and data collection. The guidelines also provide caveats and methods for managing anomalies like gaps, outliers, data shifting, and non-routine adjustments.

**What changes to the guidelines would improve model development and the accuracy of savings estimates?**

**Evaluation Response:** While DNV believes the guidelines are sufficient as is, there is opportunity to expand the guidelines with a set of practical examples within a supplemental appendix. The appendix could be a collection of anonymized projects that present modeling issues and solutions for optimizing the model's accuracy. Examples could include solutions and protocol for accommodating situations like fuel switching capital projects, gas curtailments, and savings from capital projects spanning several years or phases.

### 6.4.2 Other SEM findings and recommendations

Overall, the evaluation found the SEM program to be achieving over 90% of the energy savings claimed. The program is well documented with each site savings claim supported by an individual site model and customer report. DNV identified the following opportunities for improvement in the program that should increase the reliability of claimed savings and help mitigate the evaluation risk.

- **Finding** – The Strategic Energy Management program has become a more complicated program over time, which has increased the cost to evaluate the program. The increase in complication is primarily driven by the increase in performance tracking tool (PTTs) used to estimate program savings. While it appears that improvement and consolidation of PTTs is occurring, there are still incidents where model inputs and information are located in inconsistent areas or are not appropriately accounted for in the model.
  - **Recommendation** – DNV recommends that Energy Trust continue its efforts to create simplified and consistent PTTs for program participants to use. DNV recommends the creation of a "Non-Routine Events" (NRE) log within the PTT that documents all capital projects (both those in the baseline and those during program years), any weather adjustments made, and any other NREs that are accounted for in the model (including baseline adjustments and gas curtailments). The log should state how the NRE is accounted for in the savings calculation.
- **Finding** – DNV continued to find increased consistency in measurement periods for PY2019. However, DNV continues to find that sites' measurement start and end month vary.
  - **Recommendation** – Energy Trust should continue its efforts to get all sites on either the fall or spring schedule with the performance period starting and ending in the same month for all sites. This should continue to increase consistency and reduce the variance between claimed and evaluated savings.

# APPENDIX A. EVALUATION SAMPLE DESIGN MEMO

**Memo to:**
Sarah Castor, Energy Trust of Oregon

**Copied to:**
Andrew Wood, DNV

**Memo No.:** 1
**From:** Cameron Tuttle, DNV
**Date:** 6/25/2020
**Prep. by:** Thor Frantz, DNV
Cameron Tuttle, DNV
Benjamin Jones, DNV

**Commercial Existing Buildings Impact Evaluation Sampling Plan**

This memorandum summarizes DNV's draft sampling plan for the impact evaluation of the Energy Trust of Oregon's Commercial Existing Buildings program for program year (PY) 2019.

## Evaluation objectives

Existing Buildings program actions may target a site's electricity consumption, natural gas consumption, or both. The objectives of this evaluation considered in the development of this sampling plan are:

- Estimate achieved gas and electric savings for PY2019
- Develop separate gas and electric realization rates for PY2019 program true-up, for Strategic Energy Management (SEM) savings, and for Non-SEM savings
- Develop separate gas and electric realization rates for future program planning, if necessary.

## Sample Summary

This proposed sample is summarized in the table below. DNV believes the proposed sample and expected relative precision values are reasonable for this program and the results will achieve the study's objectives. The table also shows the sample sizes and relative precisions the evaluation is expected to achieve. The expected relative precision values are calculated utilizing error ratios informed by results from previous Energy Trust studies of the same program.

**Table A-1: Sample summary**

| Program<br><br>Track | Aggregated Track Population<br><br>(N) | Initial Sample Draw<br><br>(n) | Expected Evaluated Sample<br><br>(n) | Electric Energy kWh – Relative Precision | Electric Demand kW – Relative Precision | Gas – Relative Precision |
|---|---|---|---|---|---|---|
| Lighting | 2,822 | 40 | 33 | 14% | 21% | n/a |
| Custom | 159 | 39 | 34 | 12% | 16% | 13% |
| Standard | 749 | 83 | 68 | 14% | 19% | 10% |
| **Subtotal: Non-SEM** | **3,730** | **162** | **135** | 10% | 15% | 10% |
| SEM | 282 | 49 | 40 | 18% | n/a | 16% |
| **Total: All Tracks** | **4,012** | **211** | **175** | **10%** | **15%** | **10%** |

## Sample frame

Energy Trust provided DNV with the file "Measures 2019.xlsx" which shows energy efficiency measures completed during PY2019 through the Existing Buildings program. All pilot initiative measures were removed from the dataset by Energy Trust. The information in this file is considered the sample frame for this study and the savings listed under "working kWh" and "working therms" are considered the reported gross site-level savings.

DNV reviewed the sample frame file to confirm consistent measure classification. DNV did not reclassify any measures.

**Table A-2: Existing Buildings summary by Program Track and fuel, PY2019**

| Program Track | Unique Measure Lines | Working kWh | % of kWh Grand Total | Working therms | % of therms Grand Total |
|---|---|---|---|---|---|
| Lighting (includes Direct Install) | 8,030 | 86,869,735 | 65% | 0 | 0% |
| Custom | 208 | 27,097,473 | 20% | 881,332 | 43% |
| Standard | 878 | 7,192,794 | 5% | 578,999 | 28% |
| **Subtotal: Non-SEM** | **9,116** | **121,160,002** | **90%** | **1,460,331** | **71%** |
| SEM Cohort | 282 | 12,970,069 | 10% | 602,990 | 29% |
| **Grand Total** | **9,398** | **134,130,071** | **100%** | **2,063,321** | **100%** |

DNV converted the "working kWh" and "working therms" in the tracking file to "site Btu's." This conversion creates a single savings value to simplify stratification and the calculation of evaluation result weights. All aggregated evaluation results will be presented in kWh and therms. Only sampled electric measures will contribute to kWh results and only sampled gas measures will contribute to gas results. Tables in the appendix summarize the population. The following two equations illustrate the conversion from working kWh and working therms to Btu:

kWh_Btu = 3,412 * working_kWh

therms_Btu = 99,976 * working_therms

## Sampling Unit (Aggregation ID)

Measures are initially classified into the four Program Tracks listed below. The sampling unit varies based on the Track the project was completed under. The sampling unit recommendations are based on DNV's review of the program tracking data, specifically what types of measures are typically classified by project and site once initial Track classifications are completed. Reported savings are aggregated at the sampling unit level before size stratification and sample selection. The sampling units for the four Program Tracks are provided in the list below:

- Lighting – The sampling unit is the Project ID.
- Custom – The sampling unit is the Project ID
- Standard Non-Lighting – The sampling unit is the Project ID
- SEM – The sampling unit is the CRM Site Number listed as *et_sitenumber*

## Areas of Interest

DNV included the following areas of interest in the draft sample design:

- Lighting, Direct Install and Street Lighting – As with previous evaluations, unique sampling domains were created for these three sub-categories to ensure sufficient sample allocation within the lighting track.
- Standard, Refrigeration – A unique sampling domain was created to study this high impact category. Measures were identified by evaluationcode = "FRIDGE", "CUSTOMFRIDGE", "MOTOR", "CONTROLS", and "LIGHTING" (excluding re-lamping measures).
- Standard, Food Equipment – A unique sampling domain was created to study this high impact category. Measures were identified by evaluationcode = "FOODEQUIP", "DISHWASH", and "ICEMAKER".
- Standard, Boilers – A unique sampling domain was created to study this high impact category. Measures were identified by evaluationcode = "BOILER".
- SEM – DNV created two sampling domains for SEM based on the age of the baseline regression mode: One domain for models created in the last year (Year1) and one domain for older models (Year2+). Separate domains ensure the evaluation reviews a sample of the most recent models developed as there are many more sites in continuation than those that completed their first year.

## Stratification

Stratification is an important and commonly used design feature in most data collection efforts. Stratification refers to the process of partitioning the sample frame into distinct domains (or strata) and sampling is done independently within each domain. Stratification is often used to (1) improve precision of the final estimates and (2) control the sample size by subgroups of interest during the analysis. Precision is improved if strata are formed so that the population is relatively homogeneous within each stratum and relatively heterogeneous between strata.

Studies that involve analyzing data that could be highly variable between units often benefit by creating what is referred to as a **certainty stratum**. In this case projects or measures with the highest savings were placed in this stratum. This stratum is referred to as "certainty" because all frame units are selected for the data collection effort from this stratum. Since a census is being taken, the sampling variance associated with estimates created from this stratum is zero. A certainty stratum is suggested for this study. Figure A-6 below summarizes the domains used for this study.

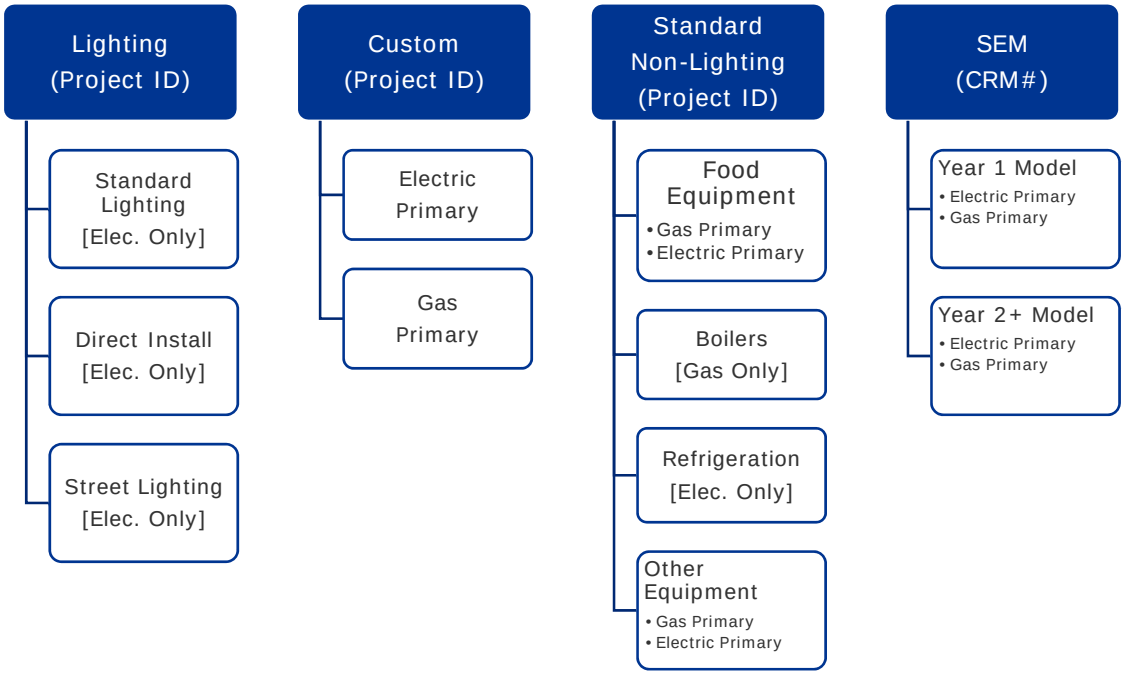
For this study, the sample will be selected independently within domains defined by the following:

- **Program Year**: 2019.
- **Program Track**: Lighting, Custom, Standard Non-Lighting (Standard), and SEM.
- **Track Sub-Category**: We applied additional categorization within Tracks. We used the field "ProductTrackDescription" to assign measures to a Program Track, and we used the fields "EvaluationCode" and/or "ProductCode" to further divide tracks into sub-categories.
    - Lighting: Standard Lighting, Direct Install, and Street Lighting
    - Custom: No sub-groups used
    - Standard Non-Lighting: Refrigeration, Boilers, Food Equipment, and Other
    - SEM: Year 1 and Year2+ (or "Continuation")

- **Fuel**: Electric and Gas classifications were used throughout the design. If an aggregated sampling unit saved both electric and gas, then the fuel classification was based on which fuel provided the majority of the site Btu savings.
  - *Exception*: All standard refrigeration projects were classified as electric, including cooler doors installed in spaces with gas heating.
- **Savings**: Additional size stratification was used within each track sub-category fuel domain to minimize the expected relative precision, ensure sample representation, and align with the evaluation's objectives.
  - *Certainty*: 30 sample points were selected at certainty.

Figure A-6 summarizes the domains used in this sample design before savings size stratification.

**Figure A-6. Domain assignments before size stratification**




## Sample Allocation to Strata

After the strata are formed, the next step was to allocate the sample to each stratum. The table below shows all strata in the sample design. The higher the size stratum value the larger the savings for the projects within the stratum.

**Table A-3: Stratification summary**

| Track | Sub-Category | Fuel | Size Stratum | Population (N) | Sample, Aggregation ID (n) |
|---|---|---|---|---|---|
| Lighting | Direct Install | Electric | 1 | 407 | 7 |
| | | | 2 | 126 | 6 |
| | Standard Lighting | Electric | 1 | 1574 | 6 |
| | | | 2 | 432 | 6 |
| | | | 3 | 204 | 6 |
| | | | 4 | 71 | 6 |
| | Street Lighting | Electric | 1 | 7 | 2 |
| | | | Certainty | 1 | 1 |

| Track | Sub-Category | Fuel | Size Stratum | Population (N) | Sample, Aggregation ID (n) |
|---|---|---|---|---|---|
| Custom | | Gas | 1 | 33 | 5 |
| | | | 2 | 14 | 5 |
| | | | 3 | 9 | 5 |
| | | | 4 | 6 | 4 |
| | | | Certainty | 4 | 4 |
| | | Electric | 1 | 63 | 5 |
| | | | 2 | 19 | 5 |
| | | | 3 | 10 | 5 |
| | | | Certainty | 1 | 1 |
| Standard | Refrigeration | Electric | 1 | 7 | 4 |
| | | | Certainty | 11 | 11 |
| | Others | Electric | 1 | 155 | 7 |
| | | | 2 | 30 | 6 |
| | | Gas | 1 | 55 | 6 |
| | | | 2 | 13 | 5 |
| | | | 3 | 3 | 3 |
| | Boiler | Gas | 1 | 17 | 4 |
| | | | 2 | 9 | 4 |
| | | | 3 | 6 | 4 |
| | | | Certainty | 4 | 4 |
| | Food Equipment | Electric | 1 | 59 | 6 |
| | | Gas | 1 | 175 | 7 |
| | | | 2 | 136 | 6 |
| | | | 3 | 69 | 6 |
| SEM Cohort | Year1 | Electric | 1 | 8 | 4 |
| | | | 2 | 4 | 2 |
| | | Gas | Certainty | 3 | 3 |
| | Year2+ | Electric | 1 | 128 | 10 |
| | | | 2 | 23 | 9 |
| | | | Certainty | 1 | 1 |
| | | Gas | 1 | 96 | 8 |
| | | | 2 | 14 | 7 |
| | | | Certainty | 5 | 5 |

## Sample Selection

Within each non-certainty strata, the measures or projects included in the evaluation were selected at random by assigning a random number to the sampling unit and sorting each stratum by this random number. Back-up sample points will be identified using these sorted lists. Within certainty strata, all projects are selected for evaluation.

## Expected Precision

DNV based the error ratios from the results of the PY 2015-2016, 2017, and 2018 Existing Buildings impact evaluations. The next table shows the error ratios assumed. DNV historically finds peak demand savings error ratios to be higher than energy error ratios. To provide Energy Trust with an estimate of the expected relative precision evaluated demand savings, if demand savings were evaluated for all electric sample points, DNV increased the assumed kWh ER by 0.2.

**Table A-4: Assumed error ratios**

| Track | Sub-Category | Assumed Electric Energy ER (kWh) | Assumed Electric Demand ER (kW) | Assumed Gas Energy ER (therms) |
|---|---|---|---|---|
| **Lighting** | Direct Install | 0.4 | 0.6 | n/a |
| | Street Lighting | 0.4 | 0.6 | n/a |
| | Standard Lighting | 0.4 | 0.6 | n/a |
| **Standard** | Refrigeration | 0.35 | 0.55 | 0.32 |
| | Others | 0.57 | 0.77 | 0.6 |
| | Food Equipment | 0.32 | 0.5 | 0.25 |
| | Boiler | n/a | n/a | 0.55 |
| **Custom** | Electric & Gas | 0.5 | 0.7 | 0.5 |
| **SEM** | Electric & Gas | 0.6 | n/a | 0.58 |

The next table summarizes the sample design and expected relative precision for various groups of interest. All "N" and "n" values are counts of the unique sampling units (Aggregation IDs) within each group. The relative precision values shown are calculated at the 90% confidence interval.

**Table A-5: Expected precision by track and fuel**

| Program<br>Track | Population<br>(N) | Expected Evaluated Sample<br>(n) | % kWh<br>Sampled | Electric Energy<br>Relative Precision | Electric Demand<br>Relative Precision | % therms<br>Sampled | Gas Energy<br>Relative Precision |
|---|---|---|---|---|---|---|---|
| Lighting | 2,822 | 33 | 7% | 14% | 21% | n/a | n/a |
| Custom | 159 | 34 | 48% | 12% | 16% | 57% | 13% |
| Standard | 749 | 68 | 44% | 14% | 19% | 45% | 10% |
| **Subtotal: Non-SEM** | **3,730** | **135** | **18%** | 10% | 14% | **52%** | 10% |
| SEM | 282 | 40 | 34% | 18% | n/a | 65% | 16% |
| **Total: All Tracks** | **4,012** | **175** | **20%** | **10%** | **14%** | **56%** | **10%** |

## Building Types

The following table shows the population and sample by building type. DNV aggregated all measure records within each program track by *et_sitenumber* to create this table. DNV used the *et_marketname* field from the tracking data. Highlighted cells which buildings types are sampled.. Stratification is used and therefore the sample is not expected to be perfectly representative of the building type distribution. DNV believes the current distribution of sample within each track is a fair representation of the different building types that participate in the program.

**Table A-6: Sample / Population by building type and track**

| | Custom | | Lighting | | SEM | | Standard Non-Lighting | |
|---|---|---|---|---|---|---|---|---|
| **Row Labels** | **Sample** | **Total** | **Sample** | **Total** | **Sample** | **Total** | **Sample** | **Total** |
| Arts, Entertainment and Recreation | 2 | 9 | 1 | 75 | 3 | 11 | 3 | 12 |
| Assisted Living | | | | | | 1 | | |
| Bank/Financial Institution | | | | 22 | 1 | 6 | | 1 |
| Campus Living | | | | | | 2 | | |
| Car Dealership/Showroom | | 1 | | 41 | | | 1 | 1 |
| Car Wash | | | | 6 | | | | |
| College/University | 4 | 18 | 2 | 34 | 7 | 20 | 2 | 7 |
| Commercial | 1 | 3 | | 33 | | 1 | 1 | 3 |
| Convenience Store | | 1 | 2 | 65 | | | 3 | 26 |
| Courthouse/Probation Office | | | | 3 | 1 | 3 | | 1 |
| Data Center | 2 | 4 | | | | | | 1 |
| Fire Station | 1 | 2 | | 22 | | 5 | | |
| Food Service | 1 | 14 | 3 | 97 | 1 | 2 | 29 | 389 |
| Gas Station | | | | 48 | | | | 1 |
| General Manufacturing | | | | 1 | | | | 1 |
| Greenhouse | | | | | | 1 | | |
| Grocery | | 1 | 2 | 53 | 3 | 4 | 17 | 39 |
| Gym/Athletic Club | 2 | 7 | 1 | 47 | 2 | 5 | | 4 |
| Healthcare | 2 | 5 | 1 | 78 | 4 | 41 | 1 | 7 |
| High School | 3 | 10 | | 26 | 1 | 11 | 6 | 17 |
| Hospital | 1 | 4 | 1 | 10 | 3 | 8 | 1 | 2 |
| Jail/Reformatory/Penitentiary | | | | 5 | 4 | 8 | 2 | 4 |
| K-12 School | 1 | 5 | 1 | 28 | | 1 | | 5 |
| Laundry/Dry Cleaner | | | | 17 | | | | |
| Library | 1 | 1 | | 6 | | 3 | | 4 |
| Lodging/Hotel/Motel | 2 | 8 | 1 | 74 | | | 8 | 90 |
| Meeting/Convention Center/Hall or Community Center | 1 | 4 | 1 | 35 | 1 | 5 | 1 | 11 |
| Middle School | 1 | 5 | 1 | 18 | 1 | 6 | 3 | 12 |
| Military (Armory, etc.) | | | | 4 | | | | 1 |
| Office | 9 | 43 | 3 | 328 | 10 | 71 | 4 | 22 |
| Parking Structure/Garage | | | 1 | 13 | | | | |
| Place of Worship | 1 | 2 | 1 | 121 | | | | 9 |
| Police | | | | 2 | | 4 | | 1 |
| Pre-K/Daycare | | 1 | | 24 | | | | |
| Primary School | 3 | 10 | | 44 | 1 | 22 | 4 | 34 |
| Printing and Related Support Activities | | | | 1 | | | | |
| Refrigerated Warehousing and Storage | | | | 1 | | 1 | | |
| Repair/Maintenance Shop | | 1 | 2 | 154 | | 1 | | 5 |
| Retail | | 4 | 12 | 474 | 4 | 29 | 1 | 21 |
| Single Family Home | | | | | 1 | 1 | | |
| Site Built Home | | | | 1 | | | | |
| Transportation Infrastructure (Tunnel, Roadway, Dock, etc.) | | | 1 | 226 | | | | |
| Utilities | | | | | | 1 | | |
| Vocational School/Community Classrooms | | | | 5 | | | | 1 |
| Warehousing and Storage | | 2 | 1 | 258 | 1 | 8 | 1 | 11 |
| Water Supply and Sewage Facilities | | | | 1 | | | | |
| **Grand Total** | **38** | **165** | **38** | **2,501** | **49** | **282** | **88** | **743** |

## Measure Types Sampled

The next table shows the number of sampled units by track and measure type. The measure type is the *evaluationdescription* provided in the tracking data.

**Table A-7. Sampled measure types**

| Track & Measure Type | Sampled Projects / Sites (n) | Program Projects / Sites (N) |
|---|---|---|
| **Custom** | **12** | **37** |
| Custom building controls | 25 | 88 |
| Custom Variable Frequency Drive | 6 | 25 |
| Retro Commissioning | 3 | 12 |
| Custom chiller | | 6 |
| Custom other measure | 3 | 6 |
| Custom boiler | 1 | 5 |
| Heat pump | 2 | 5 |
| Food equipment | 1 | 4 |
| Custom heat recovery | | 3 |
| Custom thermostat | 1 | 3 |
| Custom refrigeration | | 3 |
| Custom HVAC | 1 | 3 |
| Custom insulation | 2 | 3 |
| Custom fan | 1 | 2 |
| Air conditioning | | 2 |
| Custom ventilation | | 2 |
| Custom economizer | | 2 |
| Custom Variable Air Volume | 2 | 2 |
| Boiler | 1 | 1 |
| Ductless heat pump | | 1 |
| Custom energy management system | | 1 |
| Cooling tower | | 1 |
| Custom gas measure | | 1 |
| Custom pump | | 1 |
| Boiler tune-up | | 1 |
| Custom motor | | 1 |
| Domestic hot water measures | | 1 |
| **Lighting** | **49** | **3274** |
| Lighting | 39 | 2775 |
| Lighting controls | 3 | 261 |
| Custom lighting | 2 | 156 |
| Custom lighting control | 5 | 48 |
| Custom de-lamping | | 34 |
| **SEM Cohort** | **49** | **282** |
| Custom Operations & Maintenance | 49 | 282 |
| **Standard** | **87** | **777** |
| Food equipment | 22 | 414 |
| Heat pump | 7 | 95 |
| Boiler | 16 | 36 |
| Powerstrip | 1 | 31 |
| Tanked water heater | 3 | 29 |
| Dishwasher | 3 | 26 |
| Lighting | 2 | 25 |
| Ceiling insulation | 5 | 23 |
| Refrigerator | 14 | 15 |

| Track & Measure Type | Sampled Projects / Sites (n) | Program Projects / Sites (N) |
|---|---|---|
| Economizer | 1 | 12 |
| Wall insulation | 2 | 11 |
| Demand Control Ventilation | 2 | 8 |
| Motors | 1 | 6 |
| Generator Block Heater | | 6 |
| Battery Charger | 1 | 6 |
| Pipe insulation | 2 | 5 |
| Tankless water heater | 1 | 5 |
| Icemaker | 1 | 5 |
| Steam traps | 2 | 3 |
| Custom refrigeration | 1 | 3 |
| Server Closet Mini-split AC units | | 3 |
| Custom welder | | 2 |
| Controls | | 2 |
| Radiant heating | | 2 |
| HVAC | | 2 |
| Showerhead | | 1 |
| Faucet aerator | | 1 |

# APPENDIX B. SEM ANALYSIS METHODOLOGY

This appendix presents the methods used in this evaluation to develop gross SEM savings. The gross savings analysis relied on statistical energy consumption modeling using available historic energy consumption, weather data, and non-weather dependent variables expected to influence consumption at a sampled site. DNV primarily copied monthly facility energy consumption from the PTT files for the analysis. In some cases, Energy Trust provided the monthly consumption directly.

DNV applied one methodology to develop savings estimates for comparison with the claimed program achievements. DNV followed Energy Trust's Commercial O&M Measurement and Verification Guideline for Energy Trust of Oregon's Commercial Strategic Energy Management (SEM) and Pay for Performance (PfP) offerings. This guideline was provided to DNV by Energy Trust. This methodology primarily utilizes degree-day calculations to estimate baseline building performance during the program measurement period.

## Modeling background

### Modeling criteria

DNV considers statistical criteria and the appropriateness of the model when developing models for use in evaluation. In general, the strength of a model follows from its ability to tell a concise, consistent, and compelling story.

- Concise models are able to explain the appropriate amount of variation in the dependent variable under conditions experienced most frequently. There can be a large amount of variation in factors outside of weather that drive energy consumption. The intent of the energy consumption model is to best explain energy consumption as a function of weather and other predictor variables when those values are in the most common regions of their respective ranges.
- Consistent models have coefficient values with logical relationships. For example, a model should typically yield higher estimates of energy consumption as weather conditions become extreme or building occupancy or activity levels increase.
- Compelling models have a strong statistical fit. The probability that the coefficients are different than zero should generally be greater than 90%. Further, the overall model should account for a large amount of the observed variation in energy consumption. The adjusted R-squared statistic captures how much variation in the dependent variable (energy consumption) the model explains. Values greater than 0.8 denote a very strong statistical fit. Models that have an adjusted R-squared less than 0.5 are unable to explain half the variation in energy consumption.

To assess whether the models are consistent and concise, DNV assessed the available data on the drivers of energy consumption at SEM sites. Often, we did not have sufficient visibility into the energy drivers to assess if the models were well defined. For example, hospitals likely have factors other than weather that drive energy consumption. However, we did consider if the models made sense overall, adapting appropriately to the known variables:

- Was energy consumption predicted to change appropriately in response to the weather conditions?
- Were the predicted savings reasonable for the actions and measures implemented?

### Modeling vs. Fitting

One significant risk in statistical modeling is the trap of "over-fitting" to the available data when developing regression models. Curve-fitting tries to find an equation that fits well with the present data, while modeling tries to find an equation that represents the underlying data generator. Curve-fitting can be misleading and

can lead to over-fitting in the sense that the fitted curve may not accurately represent periods of time outside of what was used to create the curve; the classic example is always being able to fit an (n-1)th-degree polynomial to n data points. For these regression models, the energy consumption should be directly correlated with what actually drives usage. The DNV models are independent of any curve-fitting.

For this evaluation, DNV used adjusted R-squared values to assess the statistical fit. Adjusted R-squared is reduced when the model includes too many predictor variables. Increasing the number of variables may lead to a high R-squared value, but also can lead to interpretation issues, especially when the predictor variable is seemingly unrelated to energy consumption. The evaluation therefore limited the independent variables to weather-based variables and one non-weather variable.

## Site Baseline Modeling Approach

DNV utilized a standardized regression modeling approach for gas and electric usage to estimate annual energy consumption for each sampled site (or associated meter if multiple meters serve one site). DNV utilized HDD and/or CDD, rather than average temperature as used in many of the PTTs, to capture the underlying physical heating and cooling processes. If the program utilized a non-weather independent variable and the evaluation determined its use by reasonable, DNV used the same variable in its analysis. This standardized modeling approach serves to independently verify the claimed program savings. DNV developed the best model for each site based on the standard modeling criteria. In order to find the best model for each site, DNV tested several different models using various reference temperatures:

- Heating only - uses HDD term only. This model was used for all gas models.
- Cooling only – uses CDD term only.
- Single reference temperature – uses HDD and CDD calculated using the same reference temperature.
- Dual reference temperatures – uses HDD and CDD, where unique reference temperatures are calculated separately for cooling and heating.

### Model selection & development

**DNV developed 86 models using site-specific data from the baseline period** (consumption prior to the start of the program). DNV used the same months as the program for the baseline period unless sufficient data was unavailable, or a large capital project occurred during the baseline period. Model development for each site occurred in two stages:

**Stage 1, Determination of optimal model type reference temperatures:** The first stage determines the optimal reference temperature for each potential site model type. The temperature value that produced the highest adjusted R-squared value for a type was chosen to represent that type.

**Stage 2, Model type selection:** The best site model type of the four types listed above was the model type with the highest adjusted R-squared value. Table B- shows the model types used for the evaluation models developed. Nineteen (19) models also utilize a non-weather independent variable.

**Table B-1: Selected evaluation model types**

| Fuel | Temperature Response Model Type | Model Count |
|---|---|---|
| Electric | Constant | 1 |
| Electric | CDD Only | 16 |
| Electric | CDD & HDD, Single Reference Temperature | 17 |
| Electric | CDD & HDD, Dual Reference Temperature | 6 |

| Fuel | Temperature Response Model Type | Model Count |
|---|---|---|
| Electric | HDD Only | 4 |
| **Electric** | **Subtotal** | **44** |
| Gas | HDD Only | 42 |
| **All** | **Total** | **86** |

## Monthly Residuals

Energy savings for each month during the program are estimated as the difference between the modeled baseline energy and the actual energy consumption. This is referred to as the "monthly residual". This value is an estimate of the energy use avoided during the month due to all changes at the site. If the project installed a capital project after the baseline period, then any savings due to the capital project are included in the monthly residual.

## Program Year Savings

This section discusses how incremental program year savings are determined from monthly residuals.

### Program Year Assignment

Total program year energy savings are based on the sum of monthly residuals during the program year. DNV created a program year assignment schedule to determine which program year each monthly residual should be assigned to. The cohort schedule is based on the date of the participant's original cohort kick-off meeting or the end of the sites updated baseline period for re-baselined continuation sites. The standard analysis schedules are shown in Table B-2 at the end of this appendix. If participant enrolled additional sites in the program after the date of the kick-off meeting, the additional sites are assigned to a later cohort analysis schedule based on the either the end of the baseline period or the first year the program considered claiming savings. The assignments are selected to ensure that the first program savings year starts after the baseline concludes and is not earlier than the program assumed.

### Program year capital project savings

Individual capital measures associated with a sampled facility and fuel combination installed during the baseline or program year periods are included in this analysis. Concurrent capital project measure savings are accounted for by prorating the annual savings value per the measure installation date and cohort analysis schedule. For the program year under which the measure was initially installed, the measure savings are prorated by the number of days between the measure installation date and the end date for that program year, relative to 365 days for the full annual savings. For subsequent program years, the measure savings are prorated based on the number of days between the program year start and end dates, relative to 365 days for the full annual savings. Individual capital measure savings are then aggregated together for each facility to produce facility-level capital measures savings by program year and fuel type.

### Program year baseline adjustments

If the program used a baseline adjustment factor to adjust regression-based savings estimates, then the adjustment was reviewed through the evaluation. Similar to capital projects, baseline adjustments were included in each program year savings. Generally, the evaluation used the same methodology to calculate the adjustment as the program but used the outputs from the evaluation regression models.

### Program year SEM savings

Capital measure saving values are subtracted from the program year summation of monthly model residual savings values to arrive at the total SEM program savings achieved by program year and fuel type. Following

the program's guidelines, incremental savings are calculated as any SEM program savings that are greater than the SEM program savings claimed in previous years of program participation.

### Savings calculation summary

The following is a summary of the steps taken to estimate evaluated program year SEM savings:

1. Monthly Residuals: DNV calculated meter-level monthly energy savings as the difference between the estimated baseline consumption (using the regression model) and actual meter consumption. All calculations used monthly utility meter reads and daily weather data aggregated to each utility meter read period.
2. Program Year Assignment: DNV assigned each monthly residual to a program year based on schedules created for this evaluation.
3. Total Program Year Savings: DNV calculated the total savings achieved at each site by program year as the sum of monthly residuals assigned to each program year.
4. Program Year Capital Project Savings: DNV calculated program year capital savings based on the evaluation's estimate of annual capital project savings and the number of days in the assigned program year that the measure was installed.
5. Program Year Baseline Adjustment: DNV calculated program year baseline adjustment.
6. Total Program Year SEM Savings: DNV calculated the total SEM savings achieved in a program year as the difference between the Total Program Year Savings, the Program Year Capital Project Savings, and any Program Year Baseline Adjustment.
7. Incremental Program Year SEM Savings: DNV calculated Incremental Program Year SEM Savings as the difference between the Total Program Year SEM Savings for the program year and the maximum Total Program Year SEM Savings estimated for a previous program year.

**Table B-2: SEM program year assignment, standard cohort schedule**

| Month | Cohort 10 | Cohort 11 | Cohort 12 | Cohort 13 |
|---|---|---|---|---|
| Analysis Schedule Cohort | 10 | 11 | 12 | 13 |
| Kick Off Date | Jan-17 | Oct-17 | Jan-18 | Oct-18 |
| Jan-16 | BeforeSEM | BeforeSEM | BeforeSEM | BeforeSEM |
| Feb-16 | BeforeSEM | BeforeSEM | BeforeSEM | BeforeSEM |
| Mar-16 | BeforeSEM | BeforeSEM | BeforeSEM | BeforeSEM |
| Apr-16 | BeforeSEM | BeforeSEM | BeforeSEM | BeforeSEM |
| May-16 | BeforeSEM | BeforeSEM | BeforeSEM | BeforeSEM |
| Jun-16 | BeforeSEM | BeforeSEM | BeforeSEM | BeforeSEM |
| Jul-16 | BeforeSEM | BeforeSEM | BeforeSEM | BeforeSEM |
| Aug-16 | BeforeSEM | BeforeSEM | BeforeSEM | BeforeSEM |
| Sep-16 | BeforeSEM | BeforeSEM | BeforeSEM | BeforeSEM |
| Oct-16 | BeforeSEM | BeforeSEM | BeforeSEM | BeforeSEM |
| Nov-16 | BeforeSEM | BeforeSEM | BeforeSEM | BeforeSEM |
| Dec-16 | BeforeSEM | BeforeSEM | BeforeSEM | BeforeSEM |
| Jan-17 | BeforeSEM | BeforeSEM | BeforeSEM | BeforeSEM |
| Feb-17 | PY18 | BeforeSEM | BeforeSEM | BeforeSEM |
| Mar-17 | PY18 | BeforeSEM | BeforeSEM | BeforeSEM |
| Apr-17 | PY18 | BeforeSEM | BeforeSEM | BeforeSEM |
| May-17 | PY18 | BeforeSEM | BeforeSEM | BeforeSEM |
| Jun-17 | PY18 | BeforeSEM | BeforeSEM | BeforeSEM |
| Jul-17 | PY18 | BeforeSEM | BeforeSEM | BeforeSEM |
| Aug-17 | PY18 | BeforeSEM | BeforeSEM | BeforeSEM |
| Sep-17 | PY18 | BeforeSEM | BeforeSEM | BeforeSEM |

| Month | Cohort 10 | Cohort 11 | Cohort 12 | Cohort 13 |
|---|---|---|---|---|
| **Analysis Schedule Cohort** | **10** | **11** | **12** | **13** |
| **Kick Off Date** | **Jan-17** | **Oct-17** | **Jan-18** | **Oct-18** |
| Oct-17 | PY18 | BeforeSEM | BeforeSEM | BeforeSEM |
| Nov-17 | PY18 | PY18 | BeforeSEM | BeforeSEM |
| Dec-17 | PY18 | PY18 | BeforeSEM | BeforeSEM |
| Jan-18 | PY18 | PY18 | BeforeSEM | BeforeSEM |
| Feb-18 | PY19 | PY18 | PY19 | BeforeSEM |
| Mar-18 | PY19 | PY18 | PY19 | BeforeSEM |
| Apr-18 | PY19 | PY18 | PY19 | BeforeSEM |
| May-18 | PY19 | PY18 | PY19 | BeforeSEM |
| Jun-18 | PY19 | PY18 | PY19 | BeforeSEM |
| Jul-18 | PY19 | PY18 | PY19 | BeforeSEM |
| Aug-18 | PY19 | PY18 | PY19 | BeforeSEM |
| Sep-18 | PY19 | PY18 | PY19 | BeforeSEM |
| Oct-18 | PY19 | PY18 | PY19 | BeforeSEM |
| Nov-18 | PY19 | PY19 | PY19 | PY19 |
| Dec-18 | PY19 | PY19 | PY19 | PY19 |
| Jan-19 | PY19 | PY19 | PY19 | PY19 |
| Feb-19 | PY20 | PY19 | PY20 | PY19 |
| Mar-19 | PY20 | PY19 | PY20 | PY19 |
| Apr-19 | PY20 | PY19 | PY20 | PY19 |
| May-19 | PY20 | PY19 | PY20 | PY19 |
| Jun-19 | PY20 | PY19 | PY20 | PY19 |
| Jul-19 | PY20 | PY19 | PY20 | PY19 |
| Aug-19 | PY20 | PY19 | PY20 | PY19 |
| Sep-19 | PY20 | PY19 | PY20 | PY19 |
| Oct-19 | PY20 | PY19 | PY20 | PY19 |

# APPENDIX C. CUSTOM MEASURE RESULTS

This appendix provides summaries of the custom track evaluation results by measure category. Realization rates (RR) shown are mean across all measures evaluated (no weighting is applied). The first table is for PY2019 only and the second table combines the program years 2015-2019 evaluation results. DNV provided measure-specific results to Energy Trust separately.

**Table C-1: Custom track evaluation results by measure Category, PY2019 Only**

| Custom Evaluation Category & Measure Description | # Evaluated Electric Measures | # Evaluated Gas Measures | Electric GRR (%) | Gas GRR (%) |
|---|---|---|---|---|
| **HVAC** | **18** | **16** | **89%** | **79%** |
| Custom Building Controls | 1 | 1 | 91% | 83% |
| Custom VAV System | 3 | 2 | 102% | 99% |
| Custom Building Controls- DDC System | 11 | 11 | 89% | 75% |
| Custom Heat Pump- Air- to- Air (AAHP) | 1 | | 31% | |
| Custom Gas Rooftop Unit (RTU) | 2 | 2 | 98% | 78% |
| **Motors** | **7** | **2** | **93%** | **100%** |
| Custom VFD- Pump | 2 | | 48% | |
| Custom VFD Airside | 5 | 2 | 111% | 100% |
| **Other** | **6** | | **92%** | **0%** |
| Custom Other | 1 | | 67% | |
| Retrocommissioning (RCx) Control Sequences Optimization | 2 | | 92% | |
| Retrocommissioning (RCx)- Other | 1 | | 100% | |
| Custom Pool- Filtration | 2 | | 100% | |
| **Process Heating** | | **1** | **0%** | **100%** |
| Custom Boiler | | 1 | | 100% |
| **Fans** | **1** | | **93%** | **0%** |
| Custom Fans | 1 | | 93% | |
| **Weatherization** | **2** | **2** | **-162%** | **149%** |
| Custom Insulation- Roof | 2 | 2 | -162% | 149% |
| **Food Service** | **1** | **1** | **100%** | **100%** |
| Custom Kitchen Vent Hood- Fan VFD and Controls | 1 | 1 | 100% | 100% |
| **Grand Total** | **35** | **22** | **76%** | **89%** |

**Table C-2: Custom track evaluation results by measure category, PY2015 - PY2019**

| Custom Evaluation Category & Measure Description | # Evaluated Electric Measures | # Evaluated Gas Measures | Electric GRR (%) | Gas GRR (%) |
|---|---|---|---|---|
| **Controls** | **64** | **43** | **77%** | **71%** |
| Custom Building Controls | 60 | 43 | 77% | 71% |
| Custom EMS | 2 | | 84% | |
| EMS for BPTaC | 2 | | 100% | |
| **HVAC** | **56** | **32** | **101%** | **91%** |
| Custom Boiler | 1 | 2 | 100% | 98% |
| Custom Building Controls | 1 | 1 | 91% | 83% |
| Custom Chillers | 8 | 1 | 127% | 109% |
| Custom Demand Control Ventilation | 2 | 2 | 311% | 225% |
| Custom Economizers | 3 | | 93% | |
| Custom Gas | | 2 | | 100% |
| Custom HVAC | 23 | 8 | 87% | 74% |
| Custom VAV System | 4 | 3 | 92% | 96% |
| Custom Building Controls- DDC System | 11 | 11 | 89% | 75% |
| Custom Heat Pump- Air- to- Air (AAHP) | 1 | | 31% | |
| Custom Gas Rooftop Unit (RTU) | 2 | 2 | 98% | 78% |
| **Motors** | **36** | **6** | **79%** | **257%** |
| Custom Motors | 1 | | 88% | |
| Custom VFD Pump | 6 | | 77% | |
| Custom VFD- Pump | 2 | | 48% | |
| Custom VFDs | 22 | 4 | 75% | 336% |
| Custom VFD Airside | 5 | 2 | 111% | 100% |
| **Other** | **24** | **27** | **111%** | **129%** |
| Custom Other | 17 | 11 | 117% | 118% |
| Retrocommissioning (RCx) Control Sequences Optimization | 4 | 16 | 96% | 137% |
| Retrocommissioning (RCx)- Other | 1 | | 100% | |
| Custom Pool- Filtration | 2 | | 100% | |
| **Process Cooling** | **8** | **2** | **82%** | **36%** |
| Custom Chillers | 8 | 2 | 82% | 36% |
| **Process Heating** | **1** | **7** | **0%** | **60%** |
| Custom Boiler | | 6 | | 69% |
| Custom Heat Recovery | 1 | 1 | 0% | 8% |
| **Fans** | **1** | | **93%** | |
| Custom Fans | 1 | | 93% | |
| **Weatherization** | **2** | **2** | **-162%** | **149%** |
| Custom Insulation- Roof | 2 | 2 | -162% | 149% |
| **Food Service** | **1** | **1** | **100%** | **100%** |
| Custom Kitchen Vent Hood- Fan VFD and Controls | 1 | 1 | 100% | 100% |
| **Grand Total** | **193** | **120** | **86%** | **99%** |

# APPENDIX D. STANDARD MEASURE RESULTS

This appendix provides summaries of the standard track evaluation results by measure.

<table>
<tr><td>DNV Measure Description</td><td colspan="2">Boiler</td></tr>
<tr><td>Track: Evaluation Code</td><td colspan="2">Standard: BOILER</td></tr>
<tr><td>Measure Code(s)</td><td colspan="2">GFBOIL2500, GFBOIL300, GFBOIL3002500</td></tr>
<tr><td rowspan="3">This measure covers the installation of a gas-fired condensing boiler or a modulating burner on an existing boiler. Key evaluation parameters include equipment quantity, rated capacity, rated efficiency, replacement/add-on verification, return water temperature, steam end-use load and operational parameters.</td><td>RR: Avg.<br>(Min-Max)</td><td>27%<br>(15% - 115%)</td></tr>
<tr><td>Sample Target</td><td>16 Projects</td></tr>
<tr><td>Survey Completes</td><td>13 Projects</td></tr>
</table>

## Measure Information

**Program Data Review:** 53 unique measure lines were reported over the program year 2019. These lines accounted for 0% of electricity savings and 28.65% of gas savings reported for the Standard track.

**Program Delivery:** Standard

## Evaluation Summary

The evaluation primary sample goal was 16 sites. Interviews were completed for 13 sites. The evaluation developed two savings estimates for each site: a site-specific adjusted MAD savings estimate based on the data collected and a regression-based savings analysis. Final evaluated savings were determined based on a review of the data available and the evaluator's judgement regarding the sufficiency of the regression. A description of these two approaches is summarized below.

**Adjusted MAD Savings Approach: This approach uses adjustment factors based on key interview responses.**

1. The evaluation adjusted savings for boilers not operating in condensing mode. Condensing mode operation was assessed either by direct feedback from operators or by the stated or observed return water temperatures. When boilers were found to not operate in condensing mode, the savings were adjusted to reflect the reduction in operating efficiency. The average installed equipment efficiency is reduced to 88.9%, instead of the MAD document rated efficiency of 94%. For boilers that operate in condensing mode, the equipment's rated thermal efficiency was used. The table below provides a summary of the adjustments made. Only 1 out of the 13 sites evaluated showed evidence that the boilers do not operate in condensing mode.

2. Boilers that operated and condensing mode and had rated thermal efficiencies above the 94% referenced in the MAD were credited for the additional savings due to increased efficiency. These changes are represented in the "Operating Efficiency Scaling" column in the table below.

**Consumption Review: This element of our analysis uses pre-project consumption and expected efficiency gains to predict a reasonable savings value.**

1. Due to pandemic-related complications, the evaluation team elected not to employ utility bill regression analysis that takes into account post-project consumption data. Instead, pre-project annual consumption was adjusted for weather and then multiplied by the actual efficiency gain expected from replacing a non-condensing boiler with a condensing boiler. We used 75% for pre-project efficiency, 80% for baseline efficiency, and 94% for post-project efficiency. The result of this calculation divided by the reported savings in therms produced the load scaling factor in the table above.
2. Many of the sites evaluated for PY19 reported savings on the same order of magnitude as their annual consumption, which is not a reasonable expectation when the actual efficiency gain from installing a condensing boiler is about 20%. This consumption-based load scaling factor serves to adjust the savings for each project to more closely align with the actual gas consumption by the boilers at that site.
3. Site 11 is the only project for which this adjustment was not applied. There were too many pieces of equipment on the same gas meter to draw conclusions of boiler usage by looking at the billing data.
4. This adjustment is taking the place of the 75% adjustment factor that was given to sites with multiple boilers with lead/lag sequencing in previous evaluation years.

| Site | Annual Consumption Pre-project (therms) | Reported Savings (therms) | Reported Savings as % of Annual Consumption | Predicted Savings Based on Pre-project Consumption | Load Scaling Factor |
|---|---|---|---|---|---|
| Site 1 | 42,938 | 42,750 | 100% | 5,995 | 15% |
| Site 2 | 22,757 | 17,100 | 75% | 3,178 | 20% |
| Site 3 | 35,219 | 14,820 | 42% | 4,918 | 35% |
| Site 4 | 14,982 | 8,550 | 57% | 2,092 | 25% |
| Site 5 | 15,514 | 4,845 | 31% | 2,166 | 45% |
| Site 6 | 44,424 | 4,845 | 11% | 6,203 | 50% |
| Site 7 | 12,458 | 4,560 | 37% | 904 | 20% |
| Site 8 | 7,377 | 2,423 | 33% | 1,030 | 45% |
| Site 9 | 4,135 | 2,274 | 55% | 577 | 25% |
| Site 10 | 3,582 | 1,568 | 44% | 500 | 30% |
| Site 11 | N/A | 1,137 | N/A | N/A | 100% |
| Site 12 | 2,615 | 627 | 24% | 365 | 55% |
| Site 13 | 1,357 | 442 | 33% | 189 | 40% |
| **Total** | **207,358** | **105,940** | **51%** | **28,117** | **-** |

The following table combines the Efficiency Scaling developed from our interviews and the load scaling developed from the Consumption Review (except Site 11).

| Site | Site Boiler Quantity | Estimated Operating Efficiency | Operating Efficiency Scaling | Load Scaling | Evaluated UES, Therms | MAD UES, Therms |
|---|---|---|---|---|---|---|
| Site 1 | 3 | 94% | 99% | 15% | 0.42 | 2.85 |
| Site 2 | 2 | 96% | 114% | 20% | 0.65 | 2.85 |
| Site 3 | 2 | 96% | 113% | 35% | 1.12 | 2.85 |
| Site 4 | 1 | 97% | 117% | 25% | 0.83 | 2.85 |
| Site 5 | 2 | 97% | 118% | 45% | 1.51 | 2.85 |
| Site 6 | 2 | 94% | 100% | 50% | 1.43 | 2.85 |
| Site 7 | 1 | 96% | 113% | 20% | 0.64 | 2.85 |
| Site 8 | 1 | 89% | 67% | 45% | 0.85 | 2.85 |
| Site 9 | 2 | 98% | 124% | 25% | 0.88 | 2.85 |
| Site 10 | 1 | 94% | 99% | 30% | 0.85 | 2.85 |
| Site 11 | 1 | 97% | 115% | 100% | 3.28 | 2.85 |
| Site 12 | 1 | 95% | 106% | 55% | 1.66 | 2.85 |
| Site 13 | 1 | 95% | 106% | 40% | 1.21 | 2.85 |

This table compares the evaluated and reported savings for each project evaluated.

| Site | Evaluation Approach | Evaluated Savings (therms) | Reported Total (therms) | Realization Rate |
|---|---|---|---|---|
| Site 1 | MAD + load scaling adjustment | 6,373 | 42,750 | 15% |
| Site 2 | MAD + load scaling adjustment | 3,895 | 17,100 | 23% |
| Site 3 | MAD + load scaling adjustment | 5,846 | 14,820 | 39% |
| Site 4 | MAD + load scaling adjustment | 2,497 | 8,550 | 29% |
| Site 5 | MAD + load scaling adjustment | 2,572 | 4,845 | 53% |
| Site 6 | MAD + load scaling adjustment | 2,423 | 4,845 | 50% |
| Site 7 | MAD + load scaling adjustment | 1,028 | 4,560 | 23% |
| Site 8 | MAD + load scaling adjustment | 727 | 2,423 | 30% |
| Site 9 | MAD + load scaling adjustment | 703 | 2,274 | 31% |
| Site 10 | MAD + load scaling adjustment | 467 | 1,568 | 30% |
| Site 11 | MAD adjustment only | 1,308 | 1,137 | 115% |
| Site 12 | MAD + load scaling adjustment | 366 | 627 | 58% |
| Site 13 | MAD + load scaling adjustment | 188 | 442 | 42% |
| **Total** | - | **28,392** | **105,940** | **27%** |

## Evaluation Recommendations

The evaluation team found several sites with mutiple boilers operating in lead/lag type sequencing. In these cases boiler operators said that the lag boiler typically only operates under the coldest weather conditions. We were unable to collect specific runtimes or load of boilers, but it is likely that the lag boiler will operate much less than the MAD savings assume. MAD savings are for a single boiler providing the entire load. The evaluation team reviewed the MAD applicable to PY2020 and 2021. Energy Trust should continue to use the new MAD UES estimates as they align with these evaluation results.

The analysis of pre-project consumption data relative to reported savings shows that unrealistic estimates of savings can occur when prescriptive savings methods are used. This was especially the case in schools. DNV observed that there is significant extra heating capacity that is not typically utilized and therefore inflated predicted savings compared to what can be reasonably expected. In addition to using the new MAD UES savings, the evaluation team recommends that the MAD savings are compared to actual school consumption data to serve as a sanity check before finalizing an incentive and reporting savings.

## Adjustment to evaluation plan

We used a MAD adjusted savings approach, replacing the standardized load factor used in years past with a consumption-based scaling factor to account for sites with multiple boilers and oversized heating capacity.

| DNV Measure Description | Gas Fryers | |
|---|---|---|
| Track: Evaluation Code | Standard: FOODEQUIP | |
| Measure product code | GASFRY | |
| This measure covers the installation of new ENERGY STAR-compliant gas fryers. Fryers account for about 33.47% of Standard track gas savings. Key evaluation parameters include number of vats, vat capacity (Large or Standard), cooking energy efficiency, business hours, equipment operating hours, and pounds of food cooked per day. | **RR: Avg.** (Min-Max) | **97%** (42% - 138%) |
| | Sample Target | **21 fryers** |
| | Survey Completes | 15 fryers |

## Measure Information

**Program Data Review:** 344 unique measure lines were reported over the program year 2019. These lines accounted for 0% of electricity savings and 33.47% of gas savings reported for the Standard track.

**Program Delivery:** This is a standard prescriptive measure. Incentives are paid when the application, invoice/receipt, and ENERGY STAR certification are submitted.

## Evaluation Summary

The primary evaluation sample included 21 fryers over 16 sites for verification. Interviews were completed for 15 measure claims over 10 sites; the remaining sites in the sample are closed permanently.

- Telephone interviews were completed for data collection
- Verified tracking savings with MAD
- Pounds of food normally cooked per day are based on customer response, or on an adjusted-default value based on the operating hours if they are unable to estimate amount of food normally cooked.
- Operation was surprisingly resilient during pandemic; the main determinant of RR was equipment hours of use.
- All of the fryers incentivized by the program were more efficient than the standard Energy Star equipment, but almost none of the sites operated the fryers for as many hours as is assumed by the Energy Star calculator. Across the sample, these factors balanced each other out.
- Many sites were able to stay open for takeout during the pandemic while others closed completely. Overall, restaurant operations were slowed to some extent. Savings lost due to restaurant closures during Oregon's initial 6-week ban on in-person dining and slowed business throughout the year would have been captured by a conservative 1/12 reduction in savings for each site if the savings for this evaluation were to be assessed against the as-found conditions.

## Evaluation Recommendations

None

## Adjustment to evaluation plan

None

| DNV Measure Description | Wall Insulation, Roof/Ceiling Insulation and Pipe Insulation | |
|---|---|---|
| Track: Evaluation Codes | Standard: WALLINSULATE, CEILINGINSULATE, PIPEINSULATE | |
| Measure product codes | INSATTICEHP, INSATTICER, INSATTICG, INSROOFER, INSROOFG, INSROOFGR5R20, INSPIPELPS, INSPIPEMPS, INSWALLER, INSWALLG | |
| Insulation incentives are offered for wall, roof, attic, and pipe insulation. Two basic measures are “no existing insulation” and “some level of existing insulation.” Different heating system types are included: gas, electric resistance, and heat pump. Key evaluation parameters include building square footage, ceiling insulation square footage verification, wall insulation square footage verification, pipe insulation linear footage verification, building vintage, roof/attic, existing insulation verification, and space heating/space cooling verification. | **RR: Avg.** (Min-Max) | **Wall, Gas savings: 175%** (1 site only) **Roof/Ceiling, Gas savings: 100%** **Pipe, Gas savings: 100%** |
| | Sample Target | Wall: 2<br>Roof/Ceiling: 5<br>Pipe: 2 |
| | Survey Completes | Wall: 1<br>Roof/Ceiling: 3<br>Pipe: 2 |

## Measure Information

**Program Data Review:** 41 unique measure lines (12 for wall, 24 for roof/ceiling, 5 for pipe) were reported over the program year 2019. In total, these lines accounted for 0.68% of electricity savings and 12.93% of gas savings reported for the Standard track.

**Program Delivery:** This is a standard prescriptive measure.

## Evaluation Summary

The evaluation sample included 9 sites for verification. 6 interviews were completed. 3 sites did not respond to our requests.

- Telephone interviews and site visits (for some sites) were completed for data collection.
- Tracking savings were verified with MAD.
- A number of projects were suspected of triggering code updates due to their scope and nature. However, because of lack of evidence, by default these projects were fully credited.
- One driver of discrepancies is our verification that the insulation square-footage actually installed was more than the amount claimed in tracking.
- Measure Information:
  - MAD savings, based on square footage installed (wall/ceiling/attic insulation) or linear feet installed (pipe insulation), were verified.
  - Scope quantity was verified with site contact.
  - Pre-retrofit and post-retrofit R-values were verified with the site contact.

- Verified whether insulation project was a standalone upgrade or part of a larger gut-rehab/re-roofing effort.
- Verified heating/cooling uses for the space.
- Asked whether COVID-19 affected business operation, and hours of operation before/after the COVID-19 pandemic. Insulation measures are mostly affected indirectly by potentially reduced operating hours of certain businesses/organizations, especially schools.

**Evaluation Recommendations**

- Project folder should identify whether an insulation project triggered building code or not; a project that triggers building code requirements should use those code requirements rather than existing conditions as the baseline for consumption.
- COVID-19 implications: operating hours assumptions may be adjusted.

**Adjustment to evaluation plan**

MAD savings verification

| DNV Measure Description | Economizer control for RTUs, Demand Control Ventilation (DCV) for RTUs | |
|---|---|---|
| Track: Evaluation Codes | Standard: DCV, ECONO | |
| Measure product codes | RTUCTRLDCVGH, RTUCTRLDCVHPH, RTUCTRLECOGH | |
| This measure covers the installation of economizers and demand control ventilation (DCV) to rooftop units that are not required by code to include these features. This control feature must be implemented as new HVAC units are installed. Key evaluation parameters include heating system type, unit tonnage, supply air temperature, economizer low/high temp logic verification, CO2 sensor functionality verification, and additional control features (DCV and VFD) verification. | **RR: Avg.** (Min-Max) | **DCV: 100%** **Economizers: 100%** |
| | Sample Target | 3 (2 DCV, 1 Economizer) |
| | Survey Completes | 3 |

## Measure Information

**Program Data Review:** 3 unique measure lines were reported for program year 2019. These lines accounted for 0.83% of electricity savings and 0.23% of gas savings reported for the Standard track.

**Program Delivery:** This is a standard prescriptive measure. Incentives are paid when the application, invoice/receipt, and equipment specifications are submitted.

## Evaluation Summary

The evaluation sample included 3 sites for verification. 3 interviews were completed.

- Telephone interviews were completed for data collection
- Tracking savings were verified with MAD.
- MAD UES savings are on a per-ton basis, and the economizer measure is part of a group of rooftop unit control measures (economizers-DCV-VFD). These measures have a rolling-baseline requirement, i.e. in order to apply for DCV, the unit must also have an economizer; in order to apply for VFD, the unit must also have an economizer and DCV. This allows savings to be claimed for non-code-requirement measures even if they are combined with other code-requirement measures. The requirement for economizers was that the rooftop unit had to be new units with DX cooling and gas/heat pump heating.
- For economizer measure: economizer logic and tonnage were verified.
- For DCV measure: in addition to above verifications, we verified $CO_2$ sensor operation, ventilation damper settings, and economizer settings.
- All sampled site interviews were conducted with the Engineering Rebate Manager of the corresponding project, not the actual local site contact due to the corporate nature of sampled participants. The local site contact would have little to no technical knowledge on the measures implemented in this case.
- There were no direct COVID-19 impacts on this measure as evaluated sites were all essential businesses.

## Evaluation Recommendations

- Sites evaluated in the sample were all corporate participants where the evaluator was unable to speak to an actual local site contact knowlegable of the project, instead interviewing the remote Engineering Rebate Manager who was responsible for the project. We suggest the program require

participants to provide a “definitive technical contact” with direct knowledge of the operation of the equipment on which these measures were installed.

- We suggest the MAD include transparent documentation on the savings methodology. If a weighted average is calculated, the spreadsheet used should be embedded in the MAD.

**Adjustment to evaluation plan**

None, used MAD as the primary source for savings verification.

| DNV Measure Description | Packaged Terminal Heat Pumps | |
|---|---|---|
| Track: Evaluation Code | Standard: HEATPUMP | |
| Measure product codes | PTHEATPUMPHZ1, PTHEATPUMPHZ2 | |
| This measure covers the installation of packaged terminal heat pumps that replace electric resistance heat; it applies to multifamily buildings, assisted living facilities, dormitories, hotels, and motels. Key evaluation parameters include building location (climate zone), building type, installed quantity, typical setpoints, and occupancy rate. | **RR: Avg.** (Min-Max) | **99%** (98% - 100%) |
| | Sample Target | 7 |
| | Survey Completes | 6 |

## Measure Information

**Program Data Review:** 95 unique measure lines were reported over the program year 2019. These lines accounted for 33.31% of electricity savings and 0% of gas savings reported for the Standard track.

**Program Delivery:** This is a standard prescriptive measure. Incentives are paid when the application, invoice/receipt, and packaged terminal heat pump specifications are submitted.

## Evaluation Summary

The evaluation sample included 7 sites for verification. 6 interviews were completed. 1 site did not respond to our requests.

- Telephone interviews were completed for data collection.
- Tracking savings were verified with MAD.
- Quantity of HPs installed, average capacity per unit, average occupancy, and change in occupancy rate due to COVID-19 were verified.
- MAD savings are building-type and climate-zone dependent. 4 total savings categories are provided in MAD as a result of 2 climate-zones (CZ-1 and CZ-2 in Oregon) and 2 building types (hotels/motels and multifamily/assisted living/dorms)
- All sites verified are of hotels/motels/lodging building type.
- COVID-19 impacts: 4 out of 6 evaluated sites stated they were impacted by the COVID-19 pandemic, with 20% to 70% reduction in operation / occupancy rate since the pandemic. 2 sites stated they were not impacted by COVID-19, as they reported that they have the same occupancy rate as before the pandemic.

## Evaluation Recommendations

- Consider some COVID-19 adjustment factor due to reduction in operation.
- MADs should clearly state the baseline assumption used for the analysis and source.
- If a weighted average is calculated, the spreadsheet used should be embedded in the MAD.

## Adjustment to evaluation plan

None, used MAD as the primary source for savings verification.

| DNV Measure Description | Tank Water Heaters and Tankless Water Heaters | |
|---|---|---|
| Track: Evaluation Codes | Standard: TANKDHW, TANKLESS | |
| Measure product codes | DHWCONDNMF, CTWH199FS | |
| This measure covers the installation of condensing tank and tankless water heaters. Tank water heaters and tankless water heaters account for 3.62% of Standard track gas savings. Key evaluation parameters include building type, water temperature setpoint, and water heater quantity. | **RR: Avg.** (Min-Max) | **Tanked: 84%** (77.3% - 99.5%) **Tankless: 100%** (1 site only) |
| | Sample Target | **Tanked: 6** **Tankless: 1** |
| | Survey Completes | **Tanked: 6** **Tankless: 1** |

## Measure Information

**Program Data Review:** 43 (38 Tanked + 5 Tankless) unique measure lines were reported for program year 2019. These lines accounted for 0% of electricity savings and 3.62% of gas savings reported for the Standard track.

**Program Delivery:** This is a standard prescriptive measure. Incentives are paid when the application, invoice/receipt, and water heater equipment specifications are submitted.

## Evaluation Summary

The evaluation sample included 7 measure claims (6 tanked and 1 tankless) by 4 sites. 4 interviews were completed.

- Telephone interviews were completed for data collection
- Tracking savings were verified against the MAD. 3 measure claims were for a correctional facility. 2 measure claims were for a hospital. The third site was a restaurant and the fourth was a hotel.
- MAD savings are building-type dependent. Program savings followed building-type specific savings as listed on MAD following “Health care” / “Hotel/Motel” building types. (Also, the correctional facility was reported as “All Commercial Buildings.”)
- The hospital site had lower RR due to using the “All Commerical Building” UES of 2.2 therms instead of the “Health Care” UES of 1.7 therms.
- COVID-19 impacts: The restaurant had a 25% reduction in operations and the hotel saw a 50% reduction in operations. The correctional facility’s operation saw an increase in terms of personnel shifts and turnovers, but hot water demand / general operation level were not affected. Overall, water heater usage was affected by site operation levels.

## Evaluation Recommendations

- Use building-type specific savings as listed by MAD.

## Adjustment to evaluation plan

None, only used MAD as the source for savings verification.

| DNV Measure Description | Steam Traps | |
|---|---|---|
| Track: Evaluation Code | Standard: STEAMTRAP | |
| Measure product codes | STEAMTRAP, STMTRAPLPLU, STMTRAPMPHU | |
| This measure covers the installation of steam traps. These measures account for about 5.9% of Standard track gas savings. Key evaluation parameters include steam trap quantity, building type, typical system pressure/inlet pressure, boiler efficiency, and operating hour. | RR: Avg. (Min-Max) | 100% (100% - 100%) |
| | Sample Target | 2 |
| | Survey Completes | 2 |

## Measure Information

**Program Data Review:** 3 unique measure lines were reported for program year 2019. These lines accounted for 0% of electricity savings and 5.9% of gas savings reported for the Standard track.

**Program Delivery:** This is a standard prescriptive measure. Incentives are paid when the application and invoice/receipt are submitted.

## Evaluation Summary

The evaluation sample included 2 measures for verification. Both of these were surveyed. 1 site was an education site, the other was a correctional facility (which also installed water heater measures).

- Telephone interviews were completed for data collection
- Tracking savings and unit energy savings were verified against the MAD; savings categories were verified with hours-of-use and building type.
- COVID-19 impacts: operation impact are building-type dependent. The school maintained minimum operation despite doing distance learning (some staff continued to work at the facility.) For the correctional facility, there was a high infection rate among the population due to living conditions inherent within the facility. Staff were on rotating quarantines. However operation was not affected in terms of heating (steam) demand.

## Evaluation Recommendations

None

## Adjustment to evaluation plan

None

| DNV Measure Description | Dishwasher | | |
|---|---|---|---|
| Track: Evaluation Codes | Standard: DISHWASH, FOODEQUIP | | |
| Measure product codes | DWPPUHTEMPE, DWPPUHTEMPG, DWSTCONHTEMPE, DWSTDRUPHTEMPE, DWSTDRUPHTEMPG, DWSTDRUPLTEMPE, DWUCHTEMPE, DWSTDRUPLTEMPG | | |
| This measure covers the installation of new ENERGY STAR-compliant dishwashers. Dishwashers account for about 3.19% of Standard track electricity savings and 1.65% of gas savings. Key evaluation parameters include annual days of operation, racks washed per day, typical wash time, water use per rack, idle power draw, and use of water heater booster. | | **RR: Avg.** (Min-Max) | **Electric: 100%** (100% - 100%) **Gas: 100%** (100%-100%) |
| | | Sample Target | **5** |
| | | Survey Completes | 5 |

## Measure Information

**Program Data Review:** 44 unique measure lines were reported for program year 2019. These lines accounted for 3.19% of electricity savings and 1.65% of gas savings reported for the Standard track.

**Program Delivery:** This is a standard prescriptive measure. Incentives are paid when the application, invoice/receipt, and ENERGY STAR certification are submitted.

## Evaluation Summary

The evaluation sample included 5 sites for verification. 5 interviews were completed. Telephone interviews were completed for data collection.

- Verified tracking savings with MAD
- Business hours, racks washed per day, annual days of operation, and use of water heater booster were based on survey response
- COVID-19 Impacts: The school sites reported that dishwashers saw limited to no use due to the absence of students on campus.

## Evaluation Recommendations

None

## Adjustment to evaluation plan

None

| DNV Measure Description | LED Case Lighting | |
|---|---|---|
| Track: Evaluation Code | Standard: LIGHTING | |
| Measure product codes | LEDCLDRT12MED, LEDCLDRT8MED | |
| This measure covers the retrofit of refrigerated cases with LEDs. LED Case Lighting accounts for about 12.13% of Standard track electric savings. Key evaluation parameters include lighting controls, case length, case temperature, and replaced light type. | **RR: Avg.** (Min-Max) | **85%** (53% - 103%) |
| | Sample Target | 4 |
| | Survey Completes | 4 |

## Measure Information

**Program Data Review:** 43 unique measure lines were reported for program year 2019. These lines accounted for 12.13% of electricity savings and 0% of gas savings reported for the Standard track.

**Program Delivery:** This is a standard prescriptive measure. Incentives are paid on submission of application and invoice/receipts.

## Evaluation Summary

The evaluation sample included 4 measure claims at 2 sites for verification; each site had 2 measure claims. 2 interviews were completed.

- Telephone interviews were completed for data collection
- Verified tracking savings with MAD
- Business hours, case temperature, control type, and linear feet of case were confirmed via phone interview
- COVID-19 Impacts: Site contacts reported that operating hours were the same for their businesses before and during COVID-19 conditions.

## Evaluation Recommendations

None

## Adjustment to evaluation plan

None

| DNV Measure Description | Cooler Doors | |
|---|---|---|
| Track: Evaluation Code | Standard: FRIDGE | |
| Measure product codes | COOLDRETFITE, COOLDRETFITG, COOLDRETFITGOT, COOLDRETFITOEO, NRCWDLGOHEO, NRCWDMDGH | |
| These measures cover the installation of cooler door retrofits. Cooler doors account for 36.53% of electric savings and 7.13% of gas savings in the Standard track. Key evaluation parameters include linear feet of doors and cooling/heating system types. | **Electric RR: Avg.** (min-max) **Gas RR: Avg.** (min-max) | **100%** (100%-100%) **100%** (100%-100%) |
| | Sample Target | 13 |
| | Survey Completes | 12 |

## Measure Information

**Program Data Review:** 15 unique measure lines were reported over the program year 2019. These lines accounted for 36.53% of electricity savings and 7.13% of gas savings reported for the Standard track.

**Program Delivery:** This is a standard prescriptive measure. Incentives are paid on submission of the application and invoice/receipt.

## Evaluation Summary

The evaluation sample included 13 sites for verification. Interviews were completed for 12 sites. 1 site did not respond.

- Telephone interviews were completed for data collection
- Verified tracking savings with MAD
- Business hours, linear feet installed, and cooling/heating system types were obtained via phone interview.
- COVID-19 Impacts: All sites reported that operating hours were the same for their businesses before and during COVID-19 conditions.

## Evaluation Recommendations

None

## Adjustment to evaluation plan

None

## About DNV

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