Impact Evaluation of the 2018 Existing Buildings Program

Energy Trust of Oregon

July 21, 2020

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 GL to complete an impact evaluation of Energy Trust's 2018 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. ICF International has been the PMC since January 1, 2013. The program has four main tracks: Custom, Lighting (including standard, direct-install, and street lighting measures), Standard (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 GL. 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: 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	8,174	94,101,812	65%	0	0%
Standard	1,445	19,607,223	14%	741,222	41%
Custom	164	15,497,910	11%	509,471	28%
Capital Subtotal	**9,783**	**129,206,945**	**90%**	**1,250,694**	**69%**
Strategic Energy Management	291	14,569,986	10%	563,678	31%
Grand Total	**10,074**	**143,776,931**	**100%**	**1,814,372**	**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) 2018	Evaluated Gas Savings (therms) 2018
Lighting	102,469,850	
Standard	18,406,915	592,493
Custom	13,783,641	323,463
Capital Measures Only	**134,660,406**	**915,956**
Strategic Energy Management	13,326,261	524,496
Grand Total	**147,986,667**	**1,440,452**

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

Program Track	Electricity Realization Rates 2018	Gas Realization Rates 2018
Lighting	109%	
Standard	94%	80%
Custom	89%	63%
Capital Measures Only	**104%**	**73%**
Strategic Energy Management	91%	93%
Existing Buildings Program	**103%**	**79%**

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

Track / Primary sampling domain	Electric Results				Gas Results			
	Evaluation Results	Min RR	Mean RR	Max RR	Evaluation Results	Min RR	Mean RR	Max RR
Lighting	**27**	**4%**	**109%**	**488%**				
Direct Install	15	4%	94%	293%				
Standard Lighting	12	67%	129%	488%				
Standard	**35**	**12%**	**84%**	**199%**	**35**	**0%**	**128%**	**473%**
Boiler					17	40%	76%	144%
Others	20	12%	78%	199%	16	0%	186%	473%
Refrigeration	15	18%	93%	110%	2	100%	100%	100%
Custom	**26**	**2%**	**79%**	**140%**	**20**	**-130%**	**61%**	**165%**
Custom	26	2%	79%	140%	20	-130%	61%	165%
Strategic Energy Management	**20**	**0%**	**82%**	**245%**	**29**	**0%**	**76%**	**289%**
Year1	13	0%	84%	130%	20	0%	87%	289%
Year2+	7	0%	79%	245%	9	0%	53%	231%

0.4 Historic capital measure performance

Table 0-5, 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 0-5: 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%

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

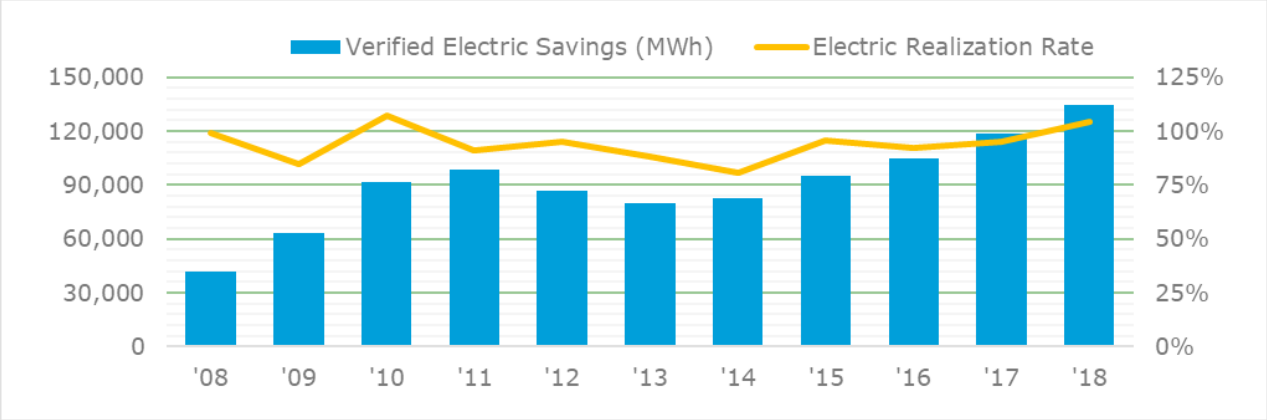
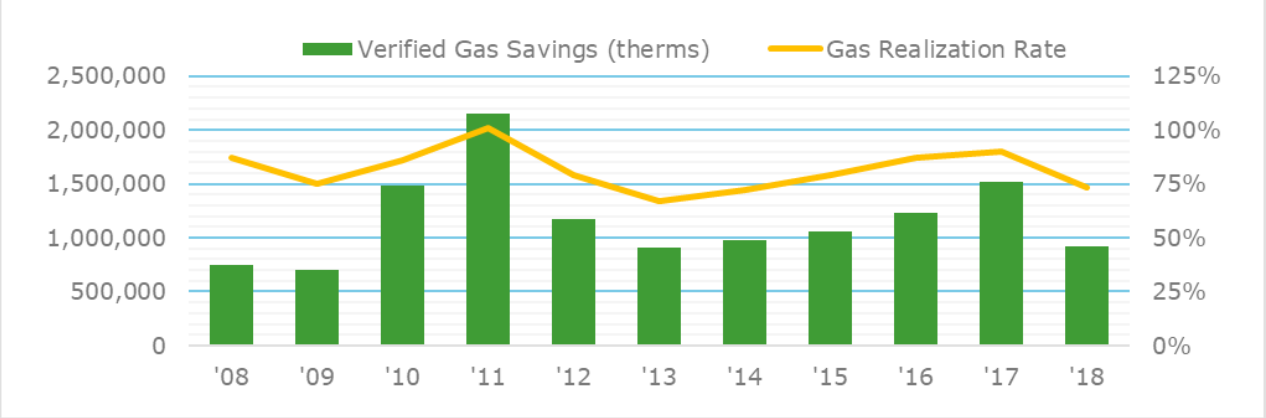


Figure 2: Historic Non-SEM program gas savings and realization rates



0.5 Historic SEM performance

Table 0-6, Figure 3 and Figure 4 show historic SEM performance over time.

Table 0-6: 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%

Figure 3: Historic SEM program electric savings and realization rates

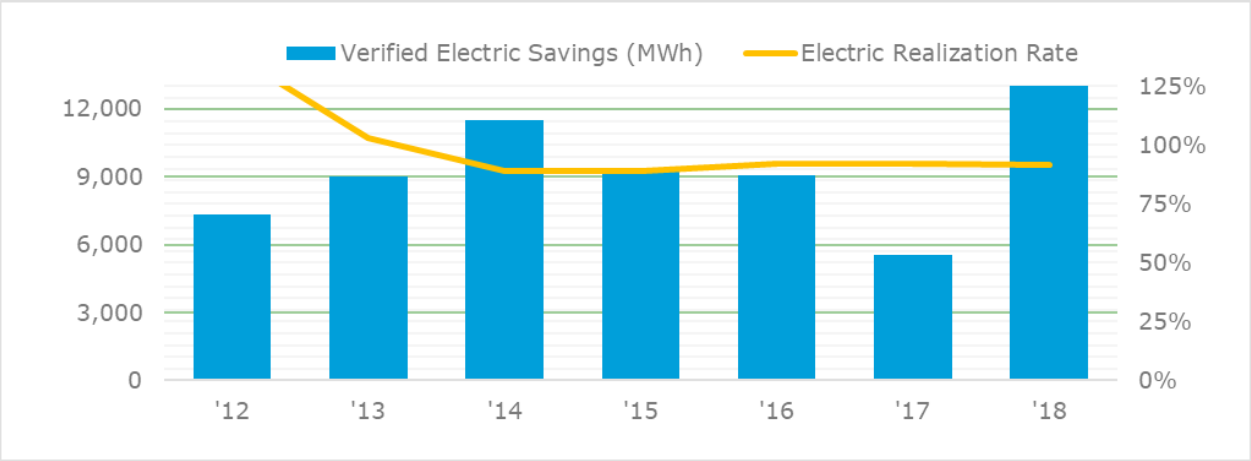
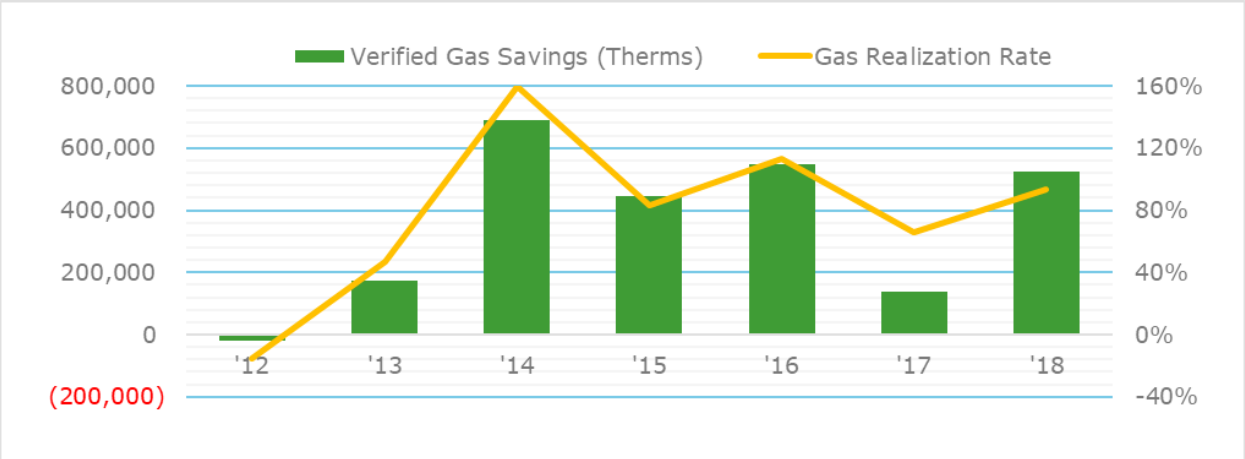


Figure 4: Historic SEM program gas savings and realization rates



0.6 Evaluation findings and recommendations

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

0.6.1 Lighting track recommendations

Finding – Deemed savings values for Direct Install lighting projects tend to over-estimate the actual hours of operation, which works to lower the realization rate. However, due largely to the under-estimated savings for a single site's controls measure (see row 13 in Table 3-6), our evaluation found an overall GRR of 116% for DI projects. 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** – We suggest a review of Measure Approval Document (MAD) 18.3, Small Commercial Direct Install for 2018, to ensure the assumptions in the program are still reasonable.

0.6.2 Standard track recommendations

DNV GL found standard measure savings claims to be sufficiently documented and well supported. We have identified opportunities for improvement in the measure approval documents and program processes. DNV GL believes these changes will increase the transparency, accuracy, and reliability of Energy Trust's standard track savings claims. Additional measure-specific recommendations are found in the standard track section of this report.

Finding - For the 2018 evaluation, DNV GL reviewed most of the MADs associated with the evaluated measures. As with the MADs we reviewed for previous program years, the evaluation team continues to find that the MADs do not provide sufficient transparency and traceability to support reliable savings estimates.

- **Recommendation** – DNV GL understands that Energy Trust has been 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 GL recommends that Energy Trust continue to explore opportunities to improve the transparency, content, and application of its prescriptive measure supporting documentation system.

Finding – Some assumptions in the MAD may be too general and did not include transparent methodology/reasoning for them, considering cases we have observed. Examples include: a tanked water heater measure using building-type weighted average savings without providing weights, and no definition of the two climate-zones used for packaged terminal heat pumps (PTHP).

- **Recommendation –** Energy Trust should continue to regularly update its MADs and improve documentation of the assumptions used in measure development. Energy should consider transitioning from a system with supporting documentation stored on internal servers to one that makes the methodologies, assumptions, and values used readily available to the public on the Energy Trust website.

0.6.3 Custom track recommendations

Overall, the evaluation found the custom project models developed by the program to be robust. DNV GL identified the following opportunities for improvement in model development that should increase the accuracy of individual project estimates.

[1] Commercial Building Stock Assessment

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 GL first made this recommendation in the program year 2017 (PY2017) impact evaluation report and believes it was implemented during PY2019.
- **Recommendation** – DNV GL 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; in some cases the claimed savings are as high as 70% of the baseline energy usage. DNV GL analyzed the actual change in facility consumption using the same regression methodology used for the Strategic Energy Management (SEM) evaluation. In some cases, the savings were found to exist. In other cases, the savings did not materialize.

- **Recommendation** – DNV GL continues to suggest that Energy Trust complete additional review of simulation inputs for sites expecting savings greater than 20% of consumption. DNV GL did not identify any evidence of further review such as discussions between ATACs and the PMC during this evaluation.
- **Recommendation** – Energy Trust should consider adjusting program implementation to complete the post installation verification (PIV) 3 to 4 months after project completion. This delay will allow the PIV process to also review post-installation consumption and assess if the significant reduction expected has materialized. If the reduction has not materialized, the PMC would have the opportunity to adjust the final savings claimed. This change would require adjustments to the incentive payment process. It may also adjust the calendar date by which such projects must be completed in order to achieve PIV before the end of the program year.

0.6.4 Strategic Energy Management 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. DNV GL 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 monitoring, tracking, and reporting (MTR) tools used to estimate program savings. There are now multiple tools and versions of those tools used by the program. As a result, the information supporting each savings claim is located in a different place within each tool.

- **Recommendation** – DNV GL recommends that Energy Trust continue its efforts to create simplified and consistent MTR tools for program participants to use. DNV GL recommends the creation of a "Non-Routine Events" (NRE) log within the MTR tool 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). The log should state how the NRE is accounted for in the savings calculation.

Finding – The SEM program is inconsistent in its treatment of campus facilities with central heating and/or cooling plants. For one campus, the program summed the measured savings (positive and negative) before adjusting for capital projects. For a different campus, the program used building-specific models and only summed savings after projects with negative incremental savings had been adjusted to zero. The impact of this difference becomes important when facility changes, especially program-claimed capital measures, installed in one building change the load seen at the central plant.

- **Recommendation** – Energy Trust should make one savings claim for campus participants with a central plant. The savings claim should be calculated by combining all building specific models and associated capital projects before determining if incremental savings have been achieved in the program year. Energy Trust should stop the practice of claiming savings for only the campus buildings that show positive incremental savings.

Finding – The site-specific realization rate for eight gas sites is below 5%. Seven of these eight sites achieved a site realization rate of 0%. Four of these seven 2018 sites were set to 0% by the evaluation team due to lack of engagement by the participant in the program. However, the total claimed savings across these sites was only 6,771 therms, ranging from 68 therms to 3,692 therms. DNV GL believes that the value of these savings does not support the cost of acquisition, cost of tracking and reporting, and the cost to evaluate.

- **Recommendation** – DNV GL continues to recommend that Energy Trust set a minimum threshold for savings claims from sites. If sites do not achieve the threshold for savings claims, then the incremental cumulative savings should not be claimed until a future program year when the savings are above the threshold. DNV GL recommends considering a threshold that prevents claiming savings below 1,000 therms. In PY2018, 73 of the 164 (44.5%) savings claims were below 1,000 therms, but represented only 4% of the gas savings claimed. Energy Trust could also consider a minimum threshold based on the percent reduction of consumption measured before capital project or other non-routine adjustments are made. DNV GL believes this change would reduce the number of claims associated with disengaged participants and improve savings reliability by ensuring the small changes in consumption persist over multiple years before being claimed.
- **Recommendation** – DNV GL also recommends that all participants consuming less than 50,000 therms per year be modeled using a standard heating degree day (HDD)-only baseline regression with the reference temperature optimized for model fit. There should be exceptions for critical non-weather independent variables. This change should increase the independence of the baseline regressions used, reduce the cost to evaluate, and better manage the program's evaluation risk.

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: November 12, 2020

Re: Staff Response to the Existing Buildings Program 2018 Impact Evaluation

The 2018 Existing Buildings program impact evaluation covered the program's four tracks: Custom, Lighting, Standard and Strategic Energy Management (SEM). 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 103%. Estimating gas savings proved more challenging, especially for Standard and Custom gas projects, and the overall realization rate was 79% for gas savings. SEM, where the gas realization rate tends to vary more by year than other tracks, performed well in 2018, achieving a 93% realization rate.

The Existing Buildings program has clarified rules around participation of projects on transport gas accounts and other rate schedules not eligible for Energy Trust participation; in order to receive an incentive offer, the project must transfer the account to an eligible rate schedule. The program plans to institute more checks on savings estimates for Custom track projects that claim to save more 20% of total building consumption, and to require parametric runs for building simulation models. The program also plans to explore ways to only claim SEM savings when SEM participants are engaged. In 2021, the program will begin developing a new performance tracking tool platform for SEM, which will make it easier for the program to aggregate and analyze models and understand the correlation between actions and energy savings.

Energy Trust is committed to regularly updating the savings estimates and documentation for its standard measures, as recommended by the evaluator. Since 2018, the program has updated measures for tanked water heaters and packaged terminal heat pumps, as suggested in the evaluation, along with many others.

With the transition of Existing Buildings program management and Commercial & Industrial Lighting program delivery in 2021, the program will be doing a broader review of its implementation and identifying additional ways to improve its estimation of savings.

1 BACKGROUND

Energy Trust performs evaluations of its programs on a regular basis. DNV GL was selected to conduct an impact evaluation of Energy Trust's 2018 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 2018. The goals of this evaluation were to:

- Develop estimates of Existing Buildings program gas and electric savings to establish realization rates for the 2018 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 $3.4 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 has been the PMC since January 1, 2013. The program has four main tracks: Custom, Lighting (including standard, direct-install, and street lighting measures), Standard (prescriptive), and Strategic Energy Management (SEM). The program also maintains a few other tracks and pilots, which represent 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 are 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 year of SEM, participants have the option of participating in Continuous SEM, 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 2018 (PY2018).
- Calculate gas and electric realization rates for PY2018.
- Provide savings and realization rates separately for SEM and non-SEM 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

2 EVALUATION OVERVIEW

This section provides an overview of DNV GL'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 GL reviewed the program tracking data provided by Energy Trust. This task helped DNV GL understand the measures and projects completed during the program year and begin to plan for the impact evaluation.

2.2 Sample design

DNV GL 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. The full sample design is discussed in Appendix A. The design for each track is discussed in the track specific sections. The PY2018 evaluation did not sample Standard track Gas Fryers or Street Lighting for evaluation as evaluation results from previous years were stable and the evaluation wanted to expand its review to other previously unevaluated measures with this year's sample. Gas Fryer and Street Lighting measures were separated into two unique domains to prevent sampling. The evaluation applied prior measure-level evaluation results to estimate PY2018 evaluated savings for these measures.

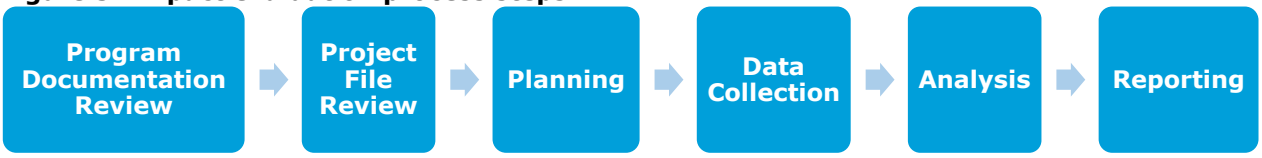
Table 2-1: Sample summary

Program Track	Population (N)	Sample (n)	% kWh Sampled	% therms Sampled	Electric – Relative Precision	Gas – Relative Precision
Lighting	2,684	30	3%	n/a	13%	n/a
Custom	123	38	57%	64%	12%	15%
Standard	1,041	72	22%	37%	14%	13%
Subtotal: Non-SEM	**3,848**	**140**	**12%**	**48%**	**10%**	**11%**
Strategic Energy Management	287	50	31%	53%	12%	12%
Total: All Tracks	**4,135**	**190**	**14%**	**50%**	**8%**	**8%**

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:

- Program Documentation Review: DNV GL 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 GL 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 GL 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/Prescriptive Measures.** The measurement and verification (M&V) plan for 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 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 GL 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 GL produced estimates of evaluated electric and/or gas savings. DNV GL engineers also noted any opportunities for improvement in the accuracy of tracked savings estimates determined during the course of our analysis.

2.4 Sample extrapolation to track and program

DNV GL 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 $Y_g = \sum_{i \in Frame} y_{ig}$

$\hat{R}_g = \frac{\sum_{i \in Sample} w_{ig} x_{ig}}{\sum_{i \in Sample} w_{ig} y_{ig}}$ 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 over 65% of the electricity savings reported by the program.

Table 3-1: Reported lighting track energy savings for 2018

Track	Electricity (kWh)
Direct Install	6,482,532
Standard Lighting	87,469,836
Street Lighting	149,444
Lighting Total	**94,101,812**
Existing Buildings program total	143,776,931
Percent of Existing Buildings program savings	65%

3.1 Sample design

DNV GL 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 and Standard to ensure that both were represented in the evaluation sample. **For PY2018, we did not sample Street Lighting; these measures were separated into a unique domain to prevent sampling.**
- Stratification by size of savings reported (three size strata were used for each 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. 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 Appendix A.

Table 3-2: Lighting track sample design

Sub-Category	Fuel	Size Stratum	Population (N)	Sample Target (n)
Direct Install Lighting	Electric	1	241	5
		2	99	5
		3	50	5
Standard Lighting	Electric	1	1,787	5
		2	399	5
		3	104	5
EVALUATION TOTAL			**2,680**	**30**
Percent of Reported kWh in 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 GL 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. Identification of data collection method - site visit or interview - for each site based on expected uncertainty. Updates to impact evaluation data collection tool.
- Data collection: Phone interview and/or onsite verification 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 GL reviewed the project documentation provided for all projects included in the original sample design. There were three key findings from this review.

- Documentation was sufficient. The documentation for the vast majority of lighting projects was comprehensive and included all relevant files.
- 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 from the 2018 Energy Trust calculation tool to Federal minimum wattages.

3.2.3 Data collection planning

DNV GL 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** *(ΔW)* is the difference between the pre-existing lighting fixture wattage and the installed lighting fixture wattage. Verification of ΔW 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 pre-existing; 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.[2] DNV GL agrees with previous program evaluators[3] that interactive effects should be included to accurately estimate the value of the program. For this study, DNV GL estimated savings without interactive effects in order to directly assess the accuracy of the original savings calculations.

3.2.4 Data collection

For Direct Install projects, data collection occurred via site visit for one site and via telephone interview for the remaining 15. For Standard Lighting projects, we interviewed 7 sites by telephone and performed site visits at 4. Whether via telephone or on site, we spoke with facility owners or operators to collect key parameter information. During the file reviews and initial recruitment, DNV GL flagged participants for possible site visits based on combinations of the following:

- Site contact, tenant, or ownership change. If the recruitment effort determined that the facility had changed owner or tenant, and the contact was not familiar with the incentivized project, the site might have been flagged for a site visit;
- The site was a high-priority data point for the stratum;
- Major renovation occurred or occupancy type changed; and
- Complex or custom lighting project that involved multiple measures or multiple space types.

3.2.5 Project level analysis

DNV GL developed a savings calculation workbook template that follows the methodology (flow and function) of Energy Trust's standard savings tool (Tab: Form 103L) 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

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

[3] 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

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 GL'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:

- Nine sites never responded to emails or phone messages requesting interviews or site visits.
- One site went out of business before the evaluation.
- One site refused to take part in evaluation efforts.
- One site had bad contact information; nobody with knowledge of the lighting project could be found.
- Where backup lighting projects were available, we attempted to recruit them. Even with a combination of telephone and email outreach and a minimum of four contact attempts for each project, we had particular trouble recruiting from the smallest standard lighting and largest Direct Install strata.

Table 3-3: Final lighting track sample summary

Sub-Category	Size Stratum	Sample Target (n)	Achieved Sample	% Complete
Direct Install Lighting	1	5	9	180%
	2	5	4	80%
	3	5	2	40%
Standard Lighting	1	5	2	40%
	2	5	6	120%
	3	5	4	80%
Grand Total		**30**	**27**	**90%**

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. The realization rate for Street Lighting from the PY2017 evaluation was used to calculate Street Lighting performance for PY2018. Overall, DNV GL estimates the evaluated lighting savings across all technologies and delivery channels to be 109% of the reported savings with a relative precision of 18.6% at the 90% confidence level.

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

Sub-Category	Projects Evaluated	Realization Rate	Standard Error	Rel. Precision @ 90% Confidence
Direct Install	15	116%	0.395	56.9%
Standard Lighting	12	108%	0.127	19.6%
Street Lighting	N/A	100%	0.000	0.0%
Lighting	**27**	**109%**	**0.121**	**18.6%**

3.3.3 Savings variance

The Standard Lighting program had an overall gross realization rate (GRR) of 108%. The assumptions used to estimate reported savings appeared reasonable; 6 of the 12 (50%) standard lighting projects evaluated had realization rates between 95% and 105%. 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 6 standard lighting projects with a GRR variance of more than 5%.

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

#	Building Type	kWh Reported	kWh Evaluated	GRR	Primary Cause of Variance
1	Large Retail (Exterior)	258,762	173,990	67%	All exterior lighting; pre-post assumed 12 hrs/day, but they actually use a combo of timer and photocell that results in fewer HOU than reported
2	Large Retail	37,830	28,316	75%	Actual hours shorter than reported
3	Assembly	90,375	82,264	91%	Actual hours for some areas were shorter than reported; a few items were not found during site visit
4	Large Retail	186,024	197,453	106%	Actual hours slightly longer than reported
5	Large Retail	647,942	701,462	108%	Actual hours slightly longer than reported
6	Hotel (Distributor Buy-down or DBD)	2,440	11,913	488%	Most lights are on 24/7 in a hotel lobby; buydown program assumes 3,796 hours/year (per DBD MAD)

The Direct Install Lighting program had an overall gross realization rate of 116%. Ten projects achieved less than 90% GRR; three exceeded 110% GRR. The building type, reported savings, and evaluated savings are shown in Table 3-5, 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	Retail	3,492	130	4%	Actual hours far shorter than deemed
2	Auto Services	5,937	3,077	52%	Actual hours shorter than deemed
3	Auto Services	43,440	23,247	54%	Some quantities incorrectly reported; actual hours shorter than deemed

4	Retail	5,148	3,141	61%	Actual hours shorter than deemed
5	Office	16,154	9,997	62%	Actual hours shorter than deemed
6	Retail	4,561	2,987	65%	Actual hours shorter than deemed
7	Retail	16,273	12,414	76%	No reported hours or invoices to confirm new equipment wattage
8	Veterinarian's Office	17,162	13,295	77%	Actual hours shorter than deemed
9	Retail	8,047	6,482	81%	Actual hours shorter than deemed
10	Retail	8,872	7,231	82%	Actual hours shorter than deemed
11	Grocery	25,181	37,885	150%	Actual hours longer than deemed
12	Gym/Athletic Club	9,087	14,468	159%	Actual hours longer than deemed
13	Retail	33,657	98,711	293%	Reported savings from controls were underestimated[4]

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

2018 High-level Research Questions:

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

Evaluation Response: As discussed in Section 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 should continue to emphasize the need for quality project documentation to trade allies.

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? (e.g., changes in operating hours for lighting)

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 (e.g. assumptions too conservative, incorrect hours of operation)?

[4] The high GRR for this stratum-3 site is due to control savings. The deemed savings for the Small Commercial Direct Install lighting measures assume a retail environment where occupancy sensors switch fixtures off for 5% of the operating hours. This site has (66) 6L TLED fixtures (75W/fixture) in a conditioned warehouse where the site contact estimated the fixtures were switched off "around a third of the time". (That is consistent with the 35% savings that the standard lighting calculator allows for warehouse spaces.)

Evaluation Response: Deemed savings values for Direct Install lighting projects tend to over-estimate the actual hours of operation, which works to lower the realization rate. However, due largely to the under-estimated savings for a single site's controls measure (see row 13 in Table 3-6), our evaluation found an overall GRR of 116% for DI projects. All DI measures assume 3,600 hours/year of operation per the regional mix from 2014 CBSA[5] data regardless of business type or market.

- **Recommendation** – We suggest a review of MAD 18.3, Small Commercial Direct Install for 2018, to ensure the assumptions in the program are still reasonable.

Evaluation Response: Estimated kWh savings for six of the twelve evaluated Standard Lighting sites differed by at least 5% from reported savings; three of these differed by more than 10%.

- **Recommendation** – Program staff should continue to emphasize the importance of accurate estimates of operating hours during training for trade allies. As in prior evaluations, DNV GL does not recommend any structural program change to address this; any change would likely increase program complexity with no assurance that it would improve savings estimates.

3.4.1.1 Other lighting findings and recommendations

Finding – As we have noted in previous evaluations, the Existing Buildings program does not account for the 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[6] are supported by DNV GL.

- **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 but should also continue to weigh the changes against increased workbook complexity. Future impact evaluations should continue to estimate the impact of lighting projects on all building systems.

[5] Commercial Building Stock Assessment

[6] 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 GL'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 standard track also included occupancy-sensor-controlled power strips, referred to as Power Strips or Smart Strips, which were purchased in bulk by the participant.

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 14% of the 2018 Existing Buildings program's reported electricity savings and 41% 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 2018

Track	Electricity (kWh)	Gas (Therms)
Standard Non-Lighting	**19,607,223**	**741,222**
Existing Buildings program total	143,776,931	1,814,372
Percent of Existing Buildings program savings	*14%*	*41%*

4.1 Sample design

DNV GL 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:

- Three technology subcategories determined based on the measure type before aggregation for sampling. These subcategories ensure that a variety of measures are selected for evaluation.
- For PY2018, the design prevented the selection of Gas Fryers by separating these measures into a unique technology domain. The evaluation chose to not sample this measure given the focus on them in previous evaluations, which yielded stable results. The prior evaluation results were used to estimate evaluated savings for PY2018 Gas Fryers.
- 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 were therefore selected for evaluation.
- Stratification by primary fuel type to ensure the evaluation results include measures savings 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 72 projects for evaluation and was expected to provide program year savings estimates with 13% gas and 14% electric relative precisions 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	Fuel	Size Stratum	Population (N)	Sample Target (n)
Refrigeration	Electric	1	88	6
		2	26	6
		3	22	6
		4	16	6
Others	Electric	1	331	5
		2	78	5
		3	23	5
	Gas	1	115	5
		2	43	5
		3	14	5
Boiler	Gas	1	17	5
		2	7	5
		3	5	4
		Certainty	4	4
EVALUATION TOTAL			**789**	**72**
Percent of Reported kWh in sample				22%
Percent of Reported therms in sample				37%

4.2 Standard track evaluation methods

This section discusses the activities completed to evaluate this track.

4.2.1 Summary of approach

DNV GL 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 one new 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:
   a) Interviewed sampled participants by telephone using the survey instruments developed for this purpose
   b) Interviewed sampled participants on-site if they were willing or if the site included a custom or SEM project that involved a site visit
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 GL used a more systematic and standardized measure-specific approach for measure types that occur five 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 GL 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 2018, 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 measure lines by measure type

Evaluation Approach	Measure Description	Measure ID Count	
		Population	Sample
Measure Type Approach	Electrically commutated motors (ECM) for refrigeration	160	44
	Boiler	57	38
	Anti-sweat heater controls	87	22
	Tanked water heater	61	9
	Heat pump	67	7
	Ceiling insulation	36	7
	LED case lighting	84	6
Project Specific Approach	Powerstrip	130	3
	Icemaker	74	3
	Economizer	5	3
	Dishwasher	47	3
	Refrigerator	23	2
	Dishwasher	38	2
	Vent hood	6	2
	HVAC	5	1
	Wall insulation	8	1
	Pipe insulation	1	1
	Oven	40	1
Not Sampled	Gas fryers[7]	322	
	Other food equipment	96	
	Virtualization	22	
	Battery charger	19	
	Generator block heater	12	
	Gas furnace	8	
	Custom refrigeration	7	
	Server closet mini-split AC units	5	
	Radiant heating	4	
	Compressed air	3	
	Demand control ventilation	3	
	Tankless water heater	3	
	Steam traps	3	

[7] Gas fryers were intentionally not sampled for the PY2018 impact evaluation.

Clothes washer	2	
Custom welder	1	
Other production efficiency	1	
Showerhead	1	

4.2.2 Documentation and file review

DNV GL 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 MADs not previously reviewed by DNV GL (i.e. RTU controls). All other relevant MADs were received and reviewed in the 2015-16 and 2017 program year evaluations. The following documents DNV GL's review process and findings for the one additional MAD reviewed for this evaluation:

- Measure baseline condition: The assumed baseline condition was not identified in the MAD and 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, vent hood savings are expressed as kWh saved per motor horsepower. The unit basis was not clearly identified in the MAD and evaluators need to understand the unit basis in order to assess the reliability of measure savings. We were able to determine the unit basis by reverse engineering the savings values in the MAD. 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 PY2018 tracking data, as unit basis were 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 GL 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. One project folder contained the 120P form from a different project for a different store under the same ownership. Finally, our file review revealed that one gas fryer project claimed savings for only one of the two fryers they installed. Table 4-4 summarizes the issues the team uncovered.

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 2018 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 primary data collection method for standard track measures was a telephone interview. In a few cases where DNV GL was already on site for measures sampled in a different track, data for standard track measures was collected in person. DNV GL 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 GL 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 B.

4.2.4.1 Whole Building Analysis

DNV GL completed a whole building regression analysis for sampled insulation and boiler projects. For insulation projects, DNV GL used the same methodology 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 consumptions using the degree-day model and weather data associated with each meter read. DNV GL 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 boiler projects, the regression was trained on the post installation facility gas consumption. The analysis was used to estimate the base and weather dependent heating consumption occurring at the facility since installation. DNV GL reviewed these values along with information gathered during data collection and determined if the load assumption in the savings calculation should be adjusted based the post-installation consumption profile.

4.3 Standard track evaluation results

This section presents the track-level results of DNV GL'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 GL 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	Fuel	Size Stratum	Sample Target (n)	Achieved Sample	% Complete
Refrigeration	Electric	1	6	4	67%
		2	6	4	67%
		3	6	2	33%
		4	6	5	83%
Others	Electric	1	5	3	60%
		2	5	4	80%
		3	5	5	100%
	Gas	1	5	4	80%
		2	5	6	120%
		3	5	4	80%
Boiler	Gas	1	5	5	100%
		2	5	4	80%
		3	4	4	100%
		Certainty	4	4	100%
Grand Total			**72**	**58**	81%

4.3.2 Evaluated savings

Realization rates by sampling domain are shown in Table 4-6 below.

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

Technology	Fuel	Projects Evaluated	Realization Rate	Standard Error	Rel. Precision @ 90% Confidence
Refrigeration	Electric	15	99%	0.025	4%
Others	Electric	19	71%	0.053	12%
Boiler	Gas	17	67%	0.016	4%
Others	Gas	16	127%	0.402	53%

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, PY2018 All Measures	34	94%	0.02	4%	18,406,915

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-2018, Projects Evaluated Only	35	94%	0.16	28%	458,357
Gas Fryers, PY2017 Evaluation Result	N/A	53%	0.14	44%	134,136
Standard, PY2018 All Measures	35	80%	0.13	24%	592,493

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 PY2018 Research Questions

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

Evaluation Response: Most project files provided adequate information but could have provided more project scope or measure configuration detail.

- **Finding** – Most project files provide complete information (signed application, equipment technical specifications, invoices, and in some cases a Post Installation Verification report). For the Vent Hood VFD measure one site only provided total HP as the measure quantifier, but verification revealed that this site had individual fans that added up to the total claimed HP. For insulation measures we found that invoices reflecting the total area of insulation are provided, but some sites were verified to have slightly different areas in the project scope than claimed.
  - **Recommendation** – For some measures that require additional project scope information for evaluators to understand where consumption is coming from (e.g. project scope area for insulation measures, fan configurations for VFD, RTU configurations for RTU controls, and boiler plants), consider requiring high-level project scope/system configuration.

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

Evaluation Response: There are some errors and potentially outdated analyses for tanked water heaters, PTHPs, and boilers.

- **Finding** – The MAD says that tanked water heater savings are based on 2003 CBECS data and 2008 EPRI studies. Analysis developed building-type-specific new installation savings and weighted average savings; the weighted average savings were used for the program savings. The weighted average savings is much lower than building-type-specific savings, especially for lodging and laundry, both of which are heavy users of hot water. We found that the PY2018 population includes a predominance of participants of the lodging building type, so using the weighted average underestimated savings.

- **Recommendation** – Review the weighted average against recent participation and future expectations, and consider the building type mix in the historical program portfolio. The savings assumed or weighting used should be adjusted to align with future program expectations.

Were there any post-installation changes in operating parameters or associated assumptions (e.g., changes in operating hours for lighting)? 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, return water temperature and end-use type for condensing boilers, and quantity of equipment/items for most other measures.

- **Finding** – Foodservice measures' operating parameters were verified to differ slightly from MAD assumptions. Post-installation changes were mostly associated with operating hours/days and equipment parameters. For dishwashers, we verified that fewer racks are washed per hour than assumed by the ENERGY STAR calculator. For ice machines, actual equipment ice harvest rate was slightly higher than the MAD assumed defaults. For hot food holding cabinents, equipment volume was lower than the default. For vent hoods, operating hours and days were slightly higher than the default.
- **Recommendation** – Continue to update deemed methods to reflect more up-to-date savings and assumptions through the current measure update process.
- **Finding** – The evaluation team continues to find condensing boiler measure sites (6 out of 17 sites) that do not show evidence of operating in condensing mode (return water temperature is not below 130°F.) This may be due to control strategies, weather restrictions, or space requirements. Boilers that are not operating in condensing mode had their efficiency reduced to 88.9%.
- **Recommendation** – Consider new methods to encourage boiler return water temperature requirements, for example requiring energy management system (EMS) screenshots or sequence of operation documentation before paying the incentive. Energy Trust should identify and consider additional methods to improve boiler controls and setpoints.

What are the factors that result in large variances in energy savings from program estimates (e.g. assumptions too conservative, incorrect hours of operation)?

Evaluation Response: We found in particular that insulation and tanked water heater measures mis-estimate actual savings.

- **Finding** – Insulation. Evaluated savings for a few insulation sites are lower than reported savings. The typical case involved a warehouse with associated conditioned office space where the measure savings were claimed for the entire space. Because a significant portion of the claimed area is either unconditioned or has different lower setpoints from the associated office space, the insulation will not be protecting against as much heat loss.
- **Recommendation** – Review insulation deemed savings, consider HVAC requirements and consider requesting information on space use by building zone. Energy Trust should also consider reviewing pre-installation electric and gas consumption for large insulation projects. If the pre-period consumption does not support the initial savings estimate, then an adjustment could be made.
- **Finding** – Tanked Water Heaters. We found that the predominant building types are hotels/motels and laundry services (both of which use large volumes of hot water), but the MAD UES value is a

weighted average value for savings across all building types. We found that the building-type-specific savings more accurately reflect the high consumption in lodging/laundry buildings.

  - **Recommendation** – The program should review the weighted average used against recent participation and future expectations. The program should adjust the savings assumed and weighting used to align with future program expectations. Consider building-type mix of the program year portfolio to determine weights. Finally, the program should update the MAD to provide the sources and methodology for determining the weighting.

- **Finding** – Hot food holding cabinents (HFHC) / Ice machines. We found that actual HFHC capacities and ice harvesting rates deviate from those assumed by the EnergyStar calculator, causing a lower realization rate.
  - **Recommendation** – Consider updating deemed assumptions to more closely reflect recent program activity.
- **Finding** – Dishwashers. We evaluated a high electric realization rate and low gas realization rate due to one site mistakenly reporting an electric water heater used for the dishwasher as gas-fueled. Also, one of the sites had their gas savings zeroed out because there was no evidence of a booster water heater upon verification.
  - **Recommendation** – We do not recommend action.
- **Finding** – Boiler MAD savings assumes most condensing boilers only operate during the heating season. However, evaluation findings suggest that condensing boilers are often used for domestic hot water (DHW) loads in lodging/laundry type buildings, or used to maintain a strict space setpoint temperature (e.g. in a swimming pool facility), neither of which are particularly weather-sensitive. This is one important reason why the evaluation team examined boiler savings in conjunction with site billing data.
  - **Recommendation** – Review MAD assumptions and possibly develop separate savings categories/tiers for different boiler end-uses.

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 PY2018 evaluation, DNV GL used the MADs provided for previous program year evaluations, plus any MADs updated for PY2018 or not previously reviewed by DNV GL. The evaluation team reviewed most of the MADs with the evaluated sample measures. As with the MADs we reviewed previously, the evaluation team continues to find that the MADs do not provide sufficient transparency and traceability to support reliable savings estimates.

DNV GL 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 GL 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). 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.
- The MAD should clearly specify the baseline condition for the measure, either pre-existing conditions (retrofit measures) or market practice (lost opportunity measures). The MAD should then provide the assumed efficiency of the baseline and the basis for the assumption.
- 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.

Below is a finding specific to the PY2018 evaluation.

Finding – Some assumptions in the MAD may be too general and do not include clear explanations of methodologies or reasoning. Examples include: a tanked water heater measure using building-type weighted average savings without providing weights, and no definition of the two climate-zones used for packaged terminal heat pumps (PTHP).

- **Recommendation –** Energy Trust should continue to regularly update its MADs and improve documentation of the assumptions used through this process. Energy should consider transitioning from a system with supporting documentation stored on internal servers to one that makes the methodologies, assumptions, and values used readily available to the public on the Energy Trust website.

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 are listed below.

- **Finding** – Space Heating Boilers. We found a number of sites claiming much higher boiler savings than their billing consumption data supports. These sites' evaluated savings are reduced to 10-50% of the claimed savings depending on the metered consumption. Conversely, there are sites that had claimed boilers as part of a larger system, serving a space with higher boiler demand (e.g. a swimming pool recreational facility), and their consumption data showed higher consumption and correspondingly higher savings. These sites' evaluated savings were increased proportionally.
- **Recommendation** – Consider boiler deemed UES to be building-type-dependent. Add savings modifiers to deemed UES for boilers based on different building types.

- **Finding** – Space Heating Boilers. We continue to find sites with multiple boilers operating in lead/lag sequencing. In these cases, boiler operators said that the lag boiler typically operates only in the coldest weather. We were unable to collect specific runtimes or loads of boilers, but we believe it is likely that the lag boiler will operate much less than the measure savings assume. The MAD for boilers used during PY2018 lacks the lead/lag assumptions used to estimate tracking savings. DNV GL's evaluation includes a review of utility meter consumption for all boiler installations.
  - **Recommendation A** – Adapt MAD savings to account for sites with multiple boilers that operate with lead/lag sequencing.
  - **Recommendation B** – Update program tracking and reporting to account for boilers installed as backups or sequenced as lag boilers. Consider identifying the quantity of incentivized boilers that will be primary versus backup or lead versus lag on the application. Add a field to the tracking data to capture the physical quantity of boilers. Regarding the baseline, the eligibility criteria should clearly state the baseline condition for existing buildings.
  - **Recommendation C** – Consider completing research on current practice for space heating boilers in Oregon. DNV GL believes that the current practice baseline efficiency for boilers is higher than the 80% assumed in the MAD. Recent research completed by our Massachusetts C&I evaluation team recommended increasing the assumed baseline for lost opportunity measures based on market activity in Massachusetts and recent Department of Energy rulemaking.[8,9]
  - **Recommendation D** – Consider completing a whole building degree-day regression analysis (similar to the Strategic Energy Management analysis) on recent or current boiler measure participants to identify the gas usage sensitive to changes in temperature. The results of this analysis and outputs from the simulation models referenced in the MAD could be used to more accurately estimate savings for this measure.

For Space Heating Boilers, DNV GL had two findings in the PY2017 evaluation that persisted into PY2018:

- **Finding** – Space Heating Boilers. As in PY2017, we found a number of boilers are providing functions other than space heating. The MAD requires that boilers provide space heating only, but the program application does not state that the incentive is only available to HVAC equipment.
  - **Recommendation** – The evaluation team continues to suggest identifying and developing savings estimates for non-space-heating uses.
  - **Recommendation** – Energy Trust should review program management processes to identify why a measure for which approved savings required a specific type of application continued to be installed in alternative applications.
- **Finding** – Space Heating Boilers. We continue to find boilers operating in conditions that made it unlikely that they typically operate in condensing mode. This reduces the operational efficiency and reduces savings. This observation persists in the PY2018 evaluation.
  - **Recommendation** - Have sites demonstrate that boilers will operate in condensing mode based on loading and estimated setpoints. As an alternative, the savings estimate could account for a percentage of boilers that do not operate in condensing mode.

[8] Gas Boiler Market Characterization Study Phase II - Final Report, Massachusetts Program Administrators and Energy Efficiency Advisory Council, March 1, 2017. http://ma-eeac.org/wordpress/wp-content/uploads/Gas-Boiler-Market-Characterization-Study-Phase-II-Final-Report.pdf

[9] Department of Energy, Commercial Packaged Boilers, final rule: https://energy.gov/sites/prod/files/2016/12/f34/CPB_ECS_Final_Rule.pdf
CPB webpage: https://www1.eere.energy.gov/buildings/appliance_standards/standards.aspx?productid=8

4.4.4 Future evaluation recommendations

- **Finding**: We continue to experience trouble contacting and then enlisting contacts at fast-food sites. They have irregular schedules, which makes reaching them difficult; they rarely respond to telephone messages; they resist scheduling site visits and interviews; and when reached, they often convey the impression that evaluations are a waste of time.
  - **Recommendation**: The evaluation team should engage program staff early and leverage their relationship with fast-food customers. Enlisting program staff early, rather than as a last resort, will allow staff to introduce the customer to the evaluation and establish the validity and importance of the evaluation.
- **Finding**: The evaluation team had trouble enlisting site contacts at grocery stores because grocery store employees are often reluctant to accept our interview request.
  - **Recommendation**: As with fast-food operators, the evaluation team should engage program staff early and leverage their relationship with grocery stores. Enlisting program staff early, rather than as a last resort, will allow them to introduce the customer to the evaluation and establish the validity and importance of the evaluation.
- **Finding**: The evaluation team found that many applications included contact information for people with little-to-no technical knowledge about the installed measures. At best, we could verify quantities for these claims. However, for high-rigor measures such as boiler and custom projects, a technically competent contact is necessary for an accurate and thorough evaluation.
  - **Recommendation**: Consider revising the program application to have participants provide a technically knowledgeable contact who is familiar with the installed measure. Also, as with the above recommendations, it is useful to establish the importance of the evaluation early with the participant/customer.

5 CUSTOM TRACK EVALUATION

The 2018 custom track reported 164 unique measure lines through 123 unique projects providing 15,497,910 kWh and 509,471 therms in annual energy savings. These savings account for 11% of the program's reported electricity savings and 28% 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 2018

Track	Electricity (kWh)	Gas (Therms)
Custom	**15,497,910**	**509,471**
Existing Buildings program total	143,776,931	1,814,372
Percent of Existing Buildings program savings	*11%*	*28%*

5.1 Sample design

DNV GL 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 four 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 GL's sample design included 38 custom projects that included 46 unique measures. 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

Fuel	Size Stratum	Population (N)	Sample Target (n)
Electric	1	44	6
	2	18	5
	3	10	5
	4	7	5
	Certainty	2	2
Gas	1	26	5
	2	9	4
	3	5	4
	Certainty	2	2
EVALUATION TOTAL		**123**	**38**
Percent of Reported kWh in sample			57%
Percent of Reported therms in sample			64%

5.2 Custom track evaluation methods

5.2.1 Summary of approach

DNV GL 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: Site visits and 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 GL 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
- 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
- Assessment of the completeness of documentation

5.2.3 Measurement and verification planning

DNV GL created project-specific M&V plans to guide the onsite data collection effort. These site-level M&V plans were created for each sampled site using DNV GL'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 Data collection

The evaluation team completed data collection for 32 projects. DNV GL was unsuccessful at recruiting 6 projects; one site refused to participate in the study. These projects were dropped after repeated attempts and support by the Existing Buildings implementation team.

5.2.5 Project analysis

DNV GL estimated evaluated savings for 32 of the 38 projects originally sampled. DNV GL used two analysis methods to estimate evaluated savings: the same calculation tool used by the program to estimate savings with revised inputs where necessary, or a whole building analysis. Inputs for the evaluated savings calculations were determined from the most valid data source including participant interviews, site observations, 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. However, in some cases the evaluation did adjust the pre-existing or baseline inputs based on interviews with the participants. Project-specific results were provided to Energy Trust separately.

5.2.5.1 Whole Building Analysis

DNV GL 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 GL used the same methodology used to evaluate Strategic Energy

Management (SEM) savings. A baseline degree-day model was trained on 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 GL 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. In some cases, DNV GL used this methodology to directly determine the evaluated savings for one project. In other cases, DNV GL compared the whole facility regression analysis results to evaluated savings calculated by modifying the original energy model developed. If the two methods provided significantly different results and the difference could not be attributed to other facility changes, then the engineering model was adjusted to improve alignment with the whole facility regression analysis. Details on the measure-specific evaluation methodologies used were provided to Energy Trust separately.

5.3 Custom track evaluation results

5.3.1 Achieved sample

Table 5-3 shows the final sample achieved across the entire standard track. DNV GL estimated evaluated savings for 84% of the original sample target. DNV GL was unsuccessful at recruiting and scheduling the six sampled projects dropped from the final analysis. These sites either did not respond to our or Energy Trust's requests for participation in the study. Our protocol required calling up to five times at different times of the day. We 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 scheduling custom program participants.

Table 5-3: Final custom track sample summary

Fuel	Size Stratum	Sample Target (n)	Achieved Sample	% Complete
Electric	1	6	3	50%
	2	5	4	80%
	3	5	3	60%
	4	5	5	100%
	Certainty	2	2	100%
Gas	1	5	5	100%
	2	4	4	100%
	3	4	4	100%
	Certainty	2	2	100%
Grand Total		**38**	**32**	**84%**

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 standard 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 GL 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.

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	26	89%	0.050	9%	13,783,641

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	20	63%	0.062	16%	323,463

5.3.3 Custom School Evaluated Savings

Electric and gas savings were evaluated for six K-12 schools and service districts in the sample. Of those 6 sites, 2 sites evaluated had only gas savings, and 4 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 numerous factors, including changes to building operation or use, errors in the program analysis, and adjustments to simulation inputs.

Table 5-6: Custom schools electric impact evaluation results

Track	# Projects	Program Savings (kWh)	Evaluation Savings (kWh)	Realization Rate
Custom Track, Schools	4	370,001	105,442	28%

Table 5-7 shows the overall gas realization rate for custom measures in schools. The gas realization rate is driven by numerous factors, including changes to building operation or use, errors in the program analysis, and adjustments to simulation inputs. Operational adjustments to calculations and billing analysis regression were utilized in the evaluation analyses for the sites.

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

Track	# Projects	Program Savings (therms)	Evaluation Savings (therms)	Realization Rate
Custom Track, Schools	6	85,408	49,996	59%

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 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 GL did not identify any systemic errors in the energy savings analysis, and very few significant calculation errors were identified during this project. However, DNV GL did identify one project that was ineligible for the gas savings claimed. This energy efficiency project reduced gas consumption on an account associated with a central plant that does not pay into the public purpose charge for energy efficiency. The customer has other gas accounts which do pay the energy efficiency charge; the project shifted gas consumption to these accounts, though the installed equipment did save energy over baseline equipment. According to program staff, the customer was planning to change the ineligible account to an eligible rate schedule, but as of March 2020, the customer still had not done so. While the project has a positive outcome for the customer, the savings should not have been claimed by Energy Trust.

- Energy Trust should review and update its eligibility requirements for customers with different account types and for projects that shift gas consumption from ineligible accounts to eligible accounts for the same customer.

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

Evaluation Response: The errors listed above resulted in large project-specific variances in savings. DNV GL also identified the following common parameters that resulted in large savings variances:

- Changes in operating schedule: The evaluation updated building operating schedules based on data gathered during the evaluation. In many cases, these schedules differed from the operating schedules used in the reported savings analysis.
- Changes in operating setpoints: The majority of the ex post revisions made were related to the control setpoints used in simulation models. 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. The source of the original setpoints assumed was typically unknown. The evaluation cannot therefore conclude if the setpoints were changed since the project was completed.

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

Evaluation Response: DNV GL did not identify any clear trends by ATAC firm. All firms with multiple projects sampled over the past two evaluations had some projects with low variance and some projects with high variance. No systematic reasons for high or low variance were identified. DNV GL 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.

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 GL 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 GL first made this recommendation in the PY2017 impact evaluation report and believes it was implemented during PY2019.
  - **Recommendation** – DNV GL 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; in some cases the claimed savings are as high as 70% of the baseline energy usage. DNV GL analyzed the actual change in facility consumption using the same regresssion methodology used for the Strategic Energy Management (SEM) evaluation. In some cases, the savings were found to exist. In other cases, the savings did not materialize.
  - **Recommendation** – DNV GL continues to suggest that Energy Trust complete additional review of simulation inputs for sites expecting savings greater than 20% of consumption. DNV GL did not identify any evidence of further review such as discussions between ATACs and the PMC during this evaluation.
  - **Recommendation** – Energy Trust should consider adjusting program implementation to complete the post installation verification (PIV) 3 to 4 months after project completion. This delay will allow the PIV process to also review post-installation consumption and assess if the significant reduction expected has materialized. If the reduction has not materialized, the PMC would have the opportunity to adjust the final savings claimed. This change would require adjustments to the incentive payment process. It may also adjust the calendar date by which projects must be completed in order to achieve PIV before the end of the program year.

6 STRATEGIC ENERGY MANAGEMENT EVALUATION

The SEM track reported 291 unique measures at 287 sites providing 14,569,986 kWh and 563,678 therms in annual energy savings in program year 2018. These savings account for 10% of the program's reported electricity savings and 31% of the program's reported gas savings. Table 6-1 shows the reported savings for SEM in program year 2018.

Table 6-1: Reported SEM track energy savings for 2018

Track	Electricity (kWh)	Gas (Therms)
SEM	**14,569,986**	**563,678**
Existing Buildings program total	143,776,931	1,814,372
Percent of Existing Buildings program savings	10%	31%

6.1 Sample design

DNV GL 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 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 GL's sample design included 48 unique site savings claims. 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	Fuel	Size Stratum	Population (N)	Sample Target (n)
Year1	Electric	1	83	6
		2	17	6
		3	10	5
	Gas	1	57	5
		2	12	5
		3	8	5
		Certainty	2	2
Year2+	Electric	1	46	4
		2	9	4
	Gas	1	38	4
		2	6	3
		Certainty	1	1
EVALUATION TOTAL			**289**	**50**
Percent of Reported kWh in sample				31%
Percent of Reported therms in sample				53%

6.2 SEM track evaluation methods

6.2.1 Summary of approach

DNV GL 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 project-specific measurement and verification plans
- Data collection: Sites visits and phone interviews with sampled participants
- 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 GL 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 GL 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

M&V Plans focused on documenting the facility being evaluated, its consumption, reported SEM actions, and identified capital projects. The plans were 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 GL 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 MTR 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 GL estimated evaluated savings for 44 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 GL 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 GL estimated evaluated savings for 88% of the sites sampled. The final achieved evaluation sample differed from the sample design due to 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 GL and Energy Trust accepted the refusals due to each organization's recent evaluation participation history. Our protocol required calling up to five times at different times of the day. We also tried contacting the participant by email if they did not respond to phone calls.

Table 6-3: Final SEM track sample summary

Year	Fuel	Size Stratum	Sample Target (n)	Achieved Sample	% Complete
Year1	Electric	1	6	6	100%
		2	6	4	67%
		3	5	3	60%
	Gas	1	5	5	100%
		2	5	5	100%
		3	5	4	80%

		Certainty	2	2	100%
Year2+	Electric	1	4	3	75%
		2	4	4	100%
	Gas	1	4	4	100%
		2	3	3	100%
		Certainty	1	1	100%
Grand Total	**Grand Total**		**50**	**44**	**88%**

6.3.2 Evaluated savings

Expansion from the sample to track-level results follows the methodology discussed in Section 2.4. The significant site-level savings variance is due to differences in baseline model form, formula errors within MTR 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 has on the incremental savings estimated for each year of participation.

Table 6-4 shows the overall electric SEM track realization rate. The site-specific electric realization rates varied from 0% to 245% for PY2018.

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	20	91%	0.13	23%	13,326,261

Table 6-5 shows the overall gas SEM track realization rate. The site-specific gas realization rates varied from 0% to 289% for PY2018.

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	29	93%	0.06	11%	524,496

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: 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 not identify any documentation that consistently communicates why one model was used instead of an alternative other than that it improved the model fit. However, the evaluation found that the increase in model file iterations and versions has significantly increased the complexity and cost of the SEM evaluation. The number of model file types also increased the risk that evaluators could not identify critical modeling parameters.

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 PY2018. DNV GL did observe a significant increase in re-baselined sites. However, some of the evaluated sites used models developed before the guidelines were provided. For models developed after the guidelines were developed, the evaluation did not find sufficient explanation for deviations. Modelers continue to optimize the model fit by adding second HDD or CDD terms that improve the fit. However, no explanation is provided discussing why the additional term, increased complexity, and reduced independence of the model is necessary.

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.

- The evaluation still found use of average temperature as an independent variable instead of degree-days in the original model. In some cases, average temperature was used in combination with a degree day variable. None of the evaluation models use average temperature.
- DNV GL found multiple sites that used a non-weather variable to account for a non-routine event in the baseline. In multiple cases, this non-weather variable could be avoided if the baseline length was reduced from 24 months to 12 months. In these cases, the evaluation used a reduced length baseline instead of a non-weather variable.
- DNV GL found multiple sites that use a second HDD term to improve the fit of the model. While these terms will improve the model fit for the baseline data available, DNV GL concluded that the term was always unnecessary and limited its models to only one HDD term. DNV GL believes the use of a second HDD 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.
- Only a few sites continue the use of average temperature squared. As documented in prior evaluations, the current modeling guidelines represent a significant improvement over these models. The use of a different baseline model did result in significant savings variance between the reported and evaluated savings. The evaluation does not use average temperature squared in any of its models.

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 multiple cases, as discussed above, where unnecessary variables were used in the original 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 project contained a formula error that omitted the capital project resulting in an overestimation of savings. Another project contained a formula error which resulted in an underestimation of savings.

- New for PY2018 was the use of realization rates to reduce to the savings deducted from the model due to capital savings. However, the application of realization rates was not consistent across the

program. DNV GL understands the theory supporting this additional adjustment, but questions whether the additional complexity is justified.

- DNV GL identified one campus site at which capital projects were not properly accounted for. The site is a campus with a central plant. The capital projects were known to impact both the equipment operating within each building and the load on the central plant. The program applied the capital savings to the buildings, but not to the central plant. Incremental savings were claimed for the central plant even though the cumulative savings for the buildings served by the central plant were negative. DNV GL recommends stopping this calculation practice.

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. DNV GL 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 monitoring, tracking, and reporting (MTR) tools used to estimate program savings. There are now multiple tools and versions of those tools used by the program. As a result, the information supporting each savings claim is located in a different place within each tool.

- **Recommendation** – DNV GL recommends that Energy Trust continue its efforts to create simplified and consistent MTR tools for program participants to use. DNV GL recommends the creation of a "Non-Routine Events" (NRE) log within the MTR tool 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). The log should state how the NRE is accounted for in the savings calculation.

Finding – The SEM program is inconsistent in its treatment of campus facilities with central heating and/or cooling plants. For one campus, the program summed the measured savings (positive and negative) before adjusting for capital projects. For a different campus, the program used building-specific models and only summed savings after projects with negative incremental savings had been adjusted to zero. The impact of this difference becomes important when facility changes, especially program-claimed capital measures, installed in one building change the load seen at the central plant.

- **Recommendation** – Energy Trust should make one savings claim for campus participants with a central plant. The savings claim should be calculated by combining all building-specific models and associated capital projects before determining if incremental savings have been achieved in the program year. Energy Trust should stop the practice of claiming savings for only the campus buildings that show positive incremental savings.

Finding – DNV GL found site savings continued to be calculated even when the participant was disengaged with the program.

- **Recommendation** – DNV GL continues to recommend to Energy Trust that participants exhibiting low engagement be classified under an inactive status, and the program not report savings from those participants. Energy Trust should also review program procedures to ensure that sufficient controls are in place to prevent energy coaches from reporting savings at disengaged participant

locations. The energy savings from the sites associated with these participants are often low compared to consumption or historic savings rates.

Finding – The site-specific realization rate for eight gas sites is below 5%. Seven of these eight sites achieved a site realization rate of 0%. In the program year 2017 evaluation, six gas sites achieved a site realization rate of 0%. Four of the 2018 sites were set to 0% by the evaluation team due to lack of engagement by the participant in the program. The total claimed savings across these sites was only 6,771 therms, ranging from 68 therms to 3,692 therms. DNV GL believes that the value of these savings does not support the cost of acquisition, cost of tracking and reporting, and the cost to evaluate. Table 6-6 divides the SEM gas savings claims into three groups based on the magnitude of savings claimed.

Table 6-6: SEM track natural gas savings claims, by magnitude group

Savings Range (therms)	Measure Line Claims	Claimed Savings (therms)	Percent of Claimed SEM Gas Savings
2,000 and up	61	498,052	88%
1,000 - 2,000	30	44,113	8%
0 - 1,000	73	21,513	4%
Totals	**164**	**563,678**	**100%**

- **Recommendation** – DNV GL continues to recommend that Energy Trust set a minimum threshold for savings claims from sites. If sites do not achieve the threshold for savings claims, then the incremental cumulative savings should not be claimed until a future program year when the savings are above the threshold. DNV GL recommends considering a threshold that prevents claiming savings below 1,000 therms. In PY2018, 73 of the 164 (44.5%) savings claims were below 1,000 therms, but represented only 4% of the gas savings claimed. Energy Trust could also consider a minimum threshold based on the percent reduction of consumption measured before capital project or other non-routine adjustments are made. DNV GL believes this change would reduce the number of claims associated with disengaged participants and improve savings reliability by ensuring the small changes in consumption persist over multiple years before being claimed.
- **Recommendation** – DNV GL also recommends that all participants consuming less than 50,000 therms per year be modeled using a standard HDD-only baseline regression with the reference temperature optimized for model fit. There should be exceptions for critical non-weather independent variables. This change should increase the independence of the baseline regressions used, reduce the cost to evaluate, and better manage the program's evaluation risk.

Finding - The level of activity documentation continues to vary broadly across participants in the program. Through the documentation review and data collection process, DNV GL observes a broad variation in the level of activity documentation in the tracking tool provided by the program. Some participants frequently document activities performed in the tracking tool, while others lack any considerable documentation of SEM-related activities.

- **Recommendation** – DNV GL recommends that Energy Trust continue to identify methods to track program engagement and energy management actions by participants. Documenting participant actions and program engagement is required to substantiate the existence of non-random energy savings.

Finding – DNV GL continued to find increased consistency in measurement periods for PY2018. However, DNV GL continues to find that sites' measurement start and end month vary.

- **Recommendation** – DNV GL recommends that all sites within a cohort start and end their measurement periods in the same month. This will increase consistency and reduce the variance between claimed and evaluated savings.

APPENDICES

APPENDIX A. EVALUATION SAMPLE DESIGN MEMO

Memo to:
Sarah Castor, Energy Trust of Oregon

Copied to:
Jennifer Canseco, DNV GL

Memo No.: 1
From: Andrew Wood, DNV GL
Date: 09/11/2019
Prep. by: Andrew Wood, DNV GL
Benjamin Jones, DNV GL

Commercial Existing Buildings Impact Evaluation Sampling Plan

This memorandum summarizes DNV GL's draft sampling plan for the impact evaluation of the Energy Trust of Oregon's Commercial Existing Buildings program for program year (PY) 2018.

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 PY2018
- Develop separate gas and electric realization rates for PY2018 to be used for program true-up.
- Develop separate gas and electric realization rates for PY2018 SEM savings
- Develop separate gas and electric realization rates for PY2018 Non-SEM savings
- Develop separate gas and electric realization rates for future program planning.

Sample Summary

This proposed sample is summarized in the table below. DNV GL 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 relative precisions expected in our response to the RFP. The expected relative precision values are based on error ratios determined in previous Energy Trust studies of the same program.

Table A-1: Sample summary

Program Track	Sample (n)	BTU – Relative Precision	Electric – Relative Precision	Gas – Relative Precision
Lighting	30	13%	13%	n/a
Custom	38	11%	12%	15%
Standard	72	13%	14%	13%
Subtotal: Non-SEM	**140**	9%	10%	11%
SEM	50	10%	12%	12%
Total: All Tracks	**190**	**7%**	**8%**	**8%**

Sample frame

Energy Trust provided DNV GL with the file "Measures 2018.xlsx" which shows energy efficiency measures completed during PY2018 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 GL reviewed the sample frame file to confirm consistent measure classification. DNV GL did not reclassify any measures.

Table A-2: Existing Buildings summary by fuel, PY2018

Program Track	Unique Measure Lines	Working kWh	% of kWh Grand Total	Working therms	% of therms Grand Total
Lighting (includes Direct Install)	8,174	94,101,812	65%	0	0%
Standard	164	15,497,910	11%	509,471	28%
Custom	1,445	19,607,223	14%	741,222	41%
Capital Total	**9,783**	**129,206,945**	**90%**	**1,250,694**	**69%**
SEM Cohort	291	14,569,986	10%	563,678	31%
Grand Total	**10,074**	**143,776,931**	**100%**	**1,814,372**	**100%**

DNV GL 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.

kwh_btu = 3,412 * working_kwh

therms_btu = 99,976 * working_therms

Sampling Unit (Aggregation ID)

Measures were 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 GL'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.

- 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

Areas of Interest

DNV GL included the following areas of interest in the draft sample design.

- Direct Install Lighting – As with previous evaluations, a unique sampling domain was created for this sub-program 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".
- Standard Boilers – A unique sampling domain was created to study this high impact category. Measures were identified by evaluationcode = "BOILER".
- SEM – DNV GL 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+)
- Gas Fryers & Street Lighting - For PY2018, we will not sample Gas Fryers or Street Lighting but will use previous results for these projects. These measures were separated into unique domains to prevent sampling.

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 6 summarizes the domains used for this study.

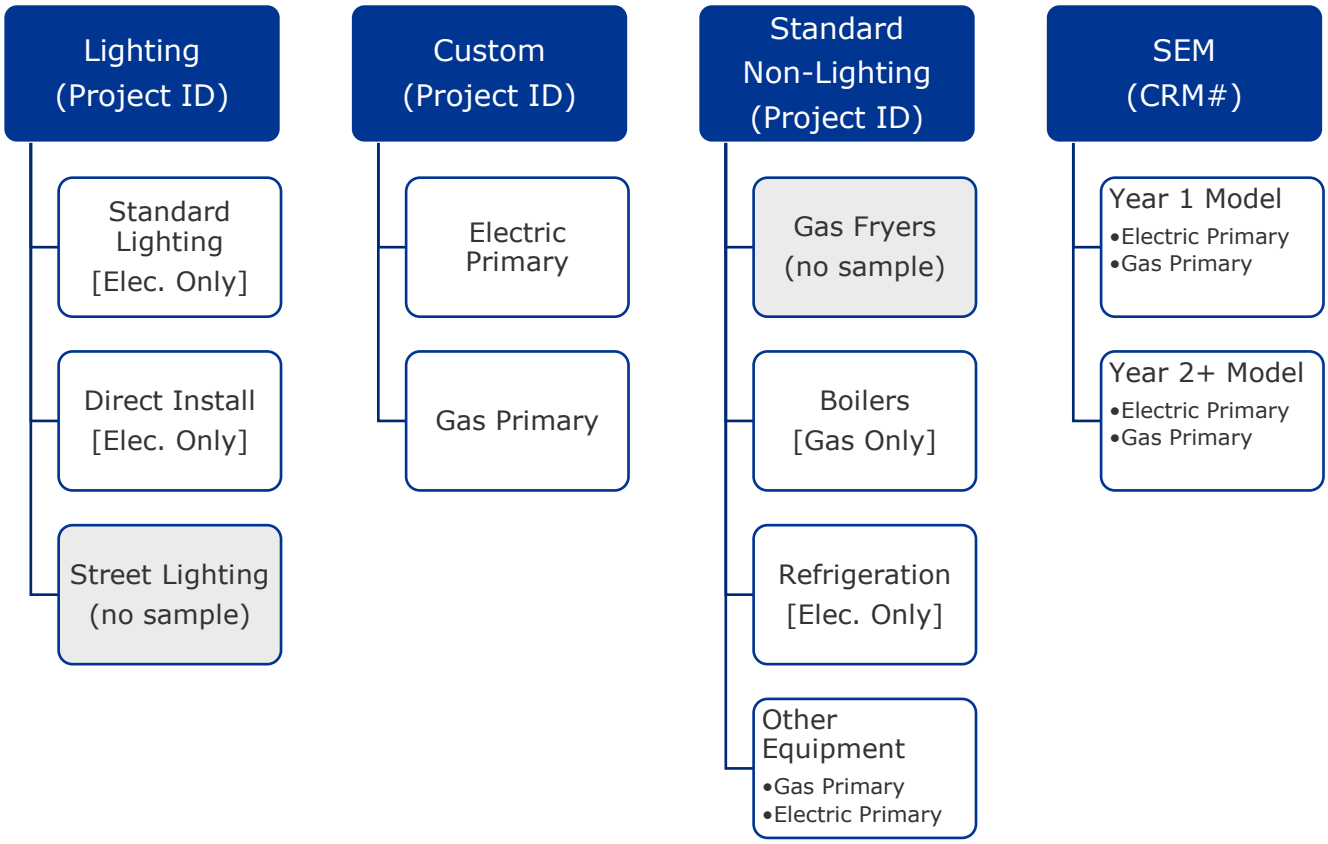
For this study, the sample will be selected independently within domains defined by the following:

- **Program Year**: 2018.
- **Program Track**: Lighting, Custom, Standard Non-Lighting (Standard), and SEM.
- **Track Sub-Group**: 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. For SEM we worked with Energy Trust to assign sub-groups to each measure line.
  - Lighting: Standard Lighting and Direct Install
  - Custom: No sub-groups used
  - Standard Non-Lighting: Refrigeration, Boilers, Other
  - SEM: Year 1 and Year2+
- **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*: 12 projects were selected at certainty.

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

Figure 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	241	5
			2	99	5
			3	50	5
	Standard Lighting	Electric	1	1787	5
			2	399	5
			3	104	5
Custom		Electric	1	44	44
			2	18	18
			3	10	10
			4	7	7
			Certainty	2	2
		Gas	1	26	26
			2	9	9
			3	5	5
			Certainty	2	2
Standard	Refrigeration	Electric	1	88	6
			2	26	6
			3	22	6
			4	16	6
	Others	Electric	1	331	5
			2	78	5
			3	23	5
		Gas	1	115	5
			2	43	5
			3	14	5
	Boiler	Gas	1	17	5
			2	7	5
			3	5	4
			Certainty	4	4
SEM Cohort	Year1	Electric	1	83	6
			2	17	6
			3	10	5
		Gas	1	57	5
			2	12	5
			3	8	5
			Certainty	2	2
	Year2+	Electric	1	46	4
			2	9	4
		Gas	1	38	4
			2	6	3
			Certainty	1	1

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 GL based the error ratios used on the results of the recent 2015-2016 and 2017 Existing Buildings impact evaluations. Table A-4 shows the error ratios assumed.

Table A-4: Assumed error ratios

Track	Sub-Category	ER Assumed	Sample (n)	Notes
Lighting	Direct Install	0.3	15	
	Standard Lighting	0.3	15	
Custom	Electric & Gas	0.6	38	
Standard	Refrigeration	0.5	24	
	Others	0.5	30	
	Boiler	0.75	18	
SEM	Electric & Gas	0.5	50	Assumes lower ER than prior results due to updated site models that follow guidelines.

Table A-5 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 level.

Table A-5: Expected precision by track and fuel

Program Track	Population (N)	Sample (n)	% kWh Sampled	% therms Sampled	Electric – Relative Precision	Gas – Relative Precision
Lighting	2,684	30	3%	n/a	13%	n/a
Custom	123	38	57%	64%	12%	15%
Standard	1,041	72	22%	37%	14%	13%
Subtotal: Non-SEM	**3,848**	**140**	**12%**	**48%**	10%	11%
SEM	287	50	31%	53%	12%	12%
Total: All Tracks	**4,135**	**190**	**14%**	**50%**	**8%**	**8%**

Building Types

The following table shows the population and sample by building type. DNV GL aggregated all measure records within each program track by site address to create this table. DNV GL used the *et_marketname* field from the tracking data. Highlighted rows are the top ten building types within each program. Stratification is used and therefore the sample is not expected to be perfectly representative of the building type distribution. DNV GL 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

	Lighting		Custom		Standard Non-Lighting		SEM	
Building Type	Program	Sample	Program	Sample	Program	Sample	Program	Sample
Amusement/Recreational	53	1	4	-	16	1	3	-
Assembly	2	-	1	-	4	-	2	-
Assisted Living Property	-	-	-	-	-	-	2	-
Auto Repair	74	1	-	-	5	2	-	-
Auto Services	92	3	3	-	8	-	1	-
Bank/Financial Institution	30	-	-	-	1	-	3	-
Beverage and Tobacco Product Manufacturing	1	-	-	-	-	-	-	-
Car Dealership/Showroom	48	1	-	-	3	-	-	-
Car Wash	16	-	-	-	1	-	-	-
College/University	51	-	8	4	8	1	18	3
Commercial	10	-	1	1	7	3	3	-
Convenience Store	47	-	-	-	28	4	-	-
Courthouse/Probation Office	5	-	-	-	1	-	4	1
Data Center	4	-	4	1	1	-	1	-
Education	13	-	-	-	2	-	1	-
Electric Power Generation, Transmission and Distribution	-	-	-	-	-	-	1	1
Enclosed Mall	4	-	1	-	-	-	-	-
Fire Protection	27	-	-	-	2	-	1	-
Gas Station	81	2	-	-	3	-	-	-
Grocery	73	4	-	-	129	20	2	-
Gym/Athletic Club	46	1	5	1	2	-	5	1
Health	6	-	-	-	1	-	2	1
High School	18	-	5	1	19	2	16	4
Hospital	19	-	5	1	3	-	17	3
Jail/Reformatory/Penitentiary	7	-	1	1	5	2	10	3
K-12 School	15	-	1	1	6	1	3	-
Laundry/Dry Cleaner	5	-	-	-	4	1	-	-
Library	6	-	4	1	-	-	1	-
Lodging/Hotel/Motel	216	1	7	3	83	8	-	-
Manufacturing	1	-	-	-	1	-	-	-
Medical Laboratory	-	-	-	-	-	-	1	1
Medical Office	50	-	2	-	10	1	32	6
Meeting/Convention Center/Hall or Community Center	20	-	2	-	22	-	1	1
Middle School	22	-	4	1	19	3	7	2
Military (Armory, etc.)	1	-	-	-	-	-	-	-

	Lighting		Custom		Standard Non-Lighting		SEM	
Building Type	Program	Sample	Program	Sample	Program	Sample	Program	Sample
Multifamily Unit	-	-	-	-	-	-	1	1
Museum	6	-	2	1	4	-	1	-
Non-Residential	1	-	-	-	1	-	7	1
Office	368	3	36	13	37	4	57	11
Parking Structure/Garage	18	-	-	-	2	-	-	-
Police	2	-	4	-	1	-	2	-
Primary School	28	-	5	2	22	5	23	4
Religious/Spiritual	119	1	-	-	24	3	-	-
Repair and Maintenance	2	-	-	-	1	-	-	-
Restaurant	78	-	8	2	502	9	-	-
Retail	447	9	8	2	41	2	20	1
Site Built Home	-	-	-	-	-	-	3	-
Super Center/Warehouse Club	6	-	-	-	-	-	-	-
Transportation Infrastructure (Tunnel, Roadway, Dock, etc.)	259	2	-	-	-	-	-	-
Unspecified Government/Public Sector	6	-	1	1	10	-	21	3
Veterinarian's Office	7	-	-	-	4	-	1	-
Warehousing and Storage	274	1	1	1	20	-	14	2

Public Schools

Public schools are of interest to Energy Trust for this evaluation. Table A-7 shows the number of sampled projects in public schools by track and measure type. No lighting projects were sampled directly and only one project occurred at a site (et_sitenumber) that was sampled for a different measure. Energy Trust can only expect anecdotal information on the performance of custom building controls and standard boilers at public schools in the results to this evaluation.

Table A-7. PY2018 Public school projects

Track & Measure Type	Program Projects/Sites (N)	Sampled Projects/Sites (n)	Not Sampled, but at a sampled site
LIGHTING	**83**		
Lighting	72		1
Lighting controls	9		
Custom lighting	2		
Custom de-lamping	0		
Custom lighting control	0		
CUSTOM	**15**		
Custom building controls	12	4	
Custom boiler	1		
Custom other measure	1		
Custom HVAC	1		
STANDARD	**66**		
Powerstrip	17	2	1
Boiler	10	6	
Ceiling insulation	9	2	
Food equipment	8		
Tanked water heater	8		2
Dishwasher	5		
Steam traps	3		
Wall insulation	2		
Server Closet Mini-split AC units	2		
Icemaker	1		
Virtualization	1		
SEM	**49**		
Custom Operations & Maintenance	49	10	
Grand Total	**213**	**24**	**4**

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 GL primarily copied monthly facility energy consumption from the MTR files for the analysis. In some cases, Energy Trust provided the monthly consumption directly.

DNV GL applied one methodology to develop savings estimates for comparison with the claimed program achievements. DNV GL 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 GL 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 GL 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 GL 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 GL models are independent of any curve-fitting.

For this evaluation, DNV GL 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 GL 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 GL utilized HDD and/or CDD, rather than average temperature as used in many of the MTRs, 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 GL used the same variable in its analysis. This standardized modeling approach serves to independently verify the claimed program savings. DNV GL developed the best model for each site based on the standard modeling criteria. In order to find the best model for each site, DNV GL 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 GL developed the models using site-specific data from the baseline period (consumption prior to the start of the program). DNV GL 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	4
Electric	CDD Only	19

Fuel	Temperature Response Model Type	Model Count
Electric	CDD & HDD, Single Reference Temperature	4
Electric	CDD & HDD, Dual Reference Temperature	4
Electric	HDD Only	4
Electric	**Subtotal**	**35**
Gas	HDD Only	33
All	**Total**	**68**

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. Prior to 2016, the SEM program would often estimate annual first-year savings from a measurement period less than 12 months. The second-year energy savings (or first continuation year) would then "true-up" savings by measuring 12-months starting from the end of the previous measurement period. DNV GL 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. For each cohort analysis schedule, where applicable, the following logic was applied to generate the schedule:

- SEM Year 1 – This is the first year for a participating facility and it contains 12 reads starting with the month following the Cohort Kick-Off workshop.
- SEM Year 2 – This is the second year for a participating facility and starts after Year 1 and ends after the following October. In most cases, this period contains less than 12 monthly reads.
- SEM Year 3+ or "Standard Year" – The Standard year contains the 12 reads from November – October. Every year except Year 1 and Year 2 is on the Standard Year schedule.

The standard analysis schedules are shown in Table B-2 at the end of this appendix (note that cohorts 7 & 8 have the same schedule). 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

The program used a baseline adjustment factor to adjust regression-based savings estimates at five sampled sites. Each 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 using 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 GL 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 GL assigned each monthly residual to a program year based on schedules created for this evaluation.
3. Total Program Year Savings: DNV GL 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 GL 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 GL calculated program year baseline adjustment.
6. Total Program Year SEM Savings: DNV GL 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 GL 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 1	Cohort 2	Cohort 3	Cohort 4	Cohort 5	Cohort 6	Cohort 7	Cohort 8	Cohort 9	Cohort 10	Cohort 11
Analysis Schedule Cohort	1	2	3	4	5	6	7	8	9	10	11
Kick Off Date	Nov-11	Jan-13	Oct-13	Jan-14	Oct-14	Jan-15	Oct-15	Oct-15	Oct-16	Feb-17	Oct-17
Oct-11	--	--	--	--	--	--	--	--	--	--	--
Nov-11	--	--	--	--	--	--	--	--	--	--	--
Dec-11	PY2012	--	--	--	--	--	--	--	--	--	--
Jan-12	PY2012	--	--	--	--	--	--	--	--	--	--
Feb-12	PY2012	--	--	--	--	--	--	--	--	--	--
Mar-12	PY2012	--	--	--	--	--	--	--	--	--	--
Apr-12	PY2012	--	--	--	--	--	--	--	--	--	--
May-12	PY2012	--	--	--	--	--	--	--	--	--	--
Jun-12	PY2012	--	--	--	--	--	--	--	--	--	--
Jul-12	PY2012	--	--	--	--	--	--	--	--	--	--
Aug-12	PY2012	--	--	--	--	--	--	--	--	--	--
Sep-12	PY2012	--	--	--	--	--	--	--	--	--	--
Oct-12	PY2012	--	--	--	--	--	--	--	--	--	--
Nov-12	PY2012	--	--	--	--	--	--	--	--	--	--
Dec-12	PY2013	--	--	--	--	--	--	--	--	--	--
Jan-13	PY2013	--	--	--	--	--	--	--	--	--	--
Feb-13	PY2013	PY2013	--	--	--	--	--	--	--	--	--
Mar-13	PY2013	PY2013	--	--	--	--	--	--	--	--	--
Apr-13	PY2013	PY2013	--	--	--	--	--	--	--	--	--
May-13	PY2013	PY2013	--	--	--	--	--	--	--	--	--
Jun-13	PY2013	PY2013	--	--	--	--	--	--	--	--	--
Jul-13	PY2013	PY2013	--	--	--	--	--	--	--	--	--
Aug-13	PY2013	PY2013	--	--	--	--	--	--	--	--	--
Sep-13	PY2013	PY2013	--	--	--	--	--	--	--	--	--
Oct-13	PY2013	PY2013	--	--	--	--	--	--	--	--	--
Nov-13	PY2014	PY2013	PY2014	--	--	--	--	--	--	--	--
Dec-13	PY2014	PY2013	PY2014	--	--	--	--	--	--	--	--
Jan-14	PY2014	PY2013	PY2014	--	--	--	--	--	--	--	--
Feb-14	PY2014	PY2014	PY2014	PY2014	--	--	--	--	--	--	--
Mar-14	PY2014	PY2014	PY2014	PY2014	--	--	--	--	--	--	--
Apr-14	PY2014	PY2014	PY2014	PY2014	--	--	--	--	--	--	--
May-14	PY2014	PY2014	PY2014	PY2014	--	--	--	--	--	--	--
Jun-14	PY2014	PY2014	PY2014	PY2014	--	--	--	--	--	--	--
Jul-14	PY2014	PY2014	PY2014	PY2014	--	--	--	--	--	--	--
Aug-14	PY2014	PY2014	PY2014	PY2014	--	--	--	--	--	--	--
Sep-14	PY2014	PY2014	PY2014	PY2014	--	--	--	--	--	--	--
Oct-14	PY2014	PY2014	PY2014	PY2014	--	--	--	--	--	--	--
Nov-14	PY2015	PY2015	PY2015	PY2014	PY2015	--	--	--	--	--	--
Dec-14	PY2015	PY2015	PY2015	PY2014	PY2015	--	--	--	--	--	--
Jan-15	PY2015	PY2015	PY2015	PY2014	PY2015	--	--	--	--	--	--
Feb-15	PY2015	PY2015	PY2015	PY2015	PY2015	PY2015	--	--	--	--	--
Mar-15	PY2015	PY2015	PY2015	PY2015	PY2015	PY2015	--	--	--	--	--
Apr-15	PY2015	PY2015	PY2015	PY2015	PY2015	PY2015	--	--	--	--	--
May-15	PY2015	PY2015	PY2015	PY2015	PY2015	PY2015	--	--	--	--	--

Month	Cohort 1	Cohort 2	Cohort 3	Cohort 4	Cohort 5	Cohort 6	Cohort 7	Cohort 8	Cohort 9	Cohort 10	Cohort 11
Analysis Schedule Cohort	1	2	3	4	5	6	7	8	9	10	11
Jun-15	PY2015	PY2015	PY2015	PY2015	PY2015	PY2015	--	--	--	--	--
Jul-15	PY2015	PY2015	PY2015	PY2015	PY2015	PY2015	--	--	--	--	--
Aug-15	PY2015	PY2015	PY2015	PY2015	PY2015	PY2015	--	--	--	--	--
Sep-15	PY2015	PY2015	PY2015	PY2015	PY2015	PY2015	--	--	--	--	--
Oct-15	PY2015	PY2015	PY2015	PY2015	PY2015	PY2015	--	--	--	--	--
Nov-15	PY2016	PY2016	PY2016	PY2016	PY2016	PY2015	PY2016	PY2016	--	--	--
Dec-15	PY2016	PY2016	PY2016	PY2016	PY2016	PY2015	PY2016	PY2016	--	--	--
Jan-16	PY2016	PY2016	PY2016	PY2016	PY2016	PY2015	PY2016	PY2016	--	--	--
Feb-16	PY2016	PY2016	PY2016	PY2016	PY2016	PY2016	PY2016	PY2016	--	--	--
Mar-16	PY2016	PY2016	PY2016	PY2016	PY2016	PY2016	PY2016	PY2016	--	--	--
Apr-16	PY2016	PY2016	PY2016	PY2016	PY2016	PY2016	PY2016	PY2016	--	--	--
May-16	PY2016	PY2016	PY2016	PY2016	PY2016	PY2016	PY2016	PY2016	--	--	--
Jun-16	PY2016	PY2016	PY2016	PY2016	PY2016	PY2016	PY2016	PY2016	--	--	--
Jul-16	PY2016	PY2016	PY2016	PY2016	PY2016	PY2016	PY2016	PY2016	--	--	--
Aug-16	PY2016	PY2016	PY2016	PY2016	PY2016	PY2016	PY2016	PY2016	--	--	--
Sep-16	PY2016	PY2016	PY2016	PY2016	PY2016	PY2016	PY2016	PY2016	--	--	--
Oct-16	PY2016	PY2016	PY2016	PY2016	PY2016	PY2016	PY2016	PY2016	--	--	--
Nov-16	PY2017	PY2017	PY2017	PY2017	PY2017	PY2017	PY2017	PY2017	PY2017	--	--
Dec-16	PY2017	PY2017	PY2017	PY2017	PY2017	PY2017	PY2017	PY2017	PY2017	--	--
Jan-17	PY2017	PY2017	PY2017	PY2017	PY2017	PY2017	PY2017	PY2017	PY2017	--	--
Feb-17	PY2017	PY2017	PY2017	PY2017	PY2017	PY2017	PY2017	PY2017	PY2017	PY2018	--
Mar-17	PY2017	PY2017	PY2017	PY2017	PY2017	PY2017	PY2017	PY2017	PY2017	PY2018	--
Apr-17	PY2017	PY2017	PY2017	PY2017	PY2017	PY2017	PY2017	PY2017	PY2017	PY2018	--
May-17	PY2017	PY2017	PY2017	PY2017	PY2017	PY2017	PY2017	PY2017	PY2017	PY2018	--
Jun-17	PY2017	PY2017	PY2017	PY2017	PY2017	PY2017	PY2017	PY2017	PY2017	PY2018	--
Jul-17	PY2017	PY2017	PY2017	PY2017	PY2017	PY2017	PY2017	PY2017	PY2017	PY2018	--
Aug-17	PY2017	PY2017	PY2017	PY2017	PY2017	PY2017	PY2017	PY2017	PY2017	PY2018	--
Sep-17	PY2017	PY2017	PY2017	PY2017	PY2017	PY2017	PY2017	PY2017	PY2017	PY2018	--
Oct-17	PY2017	PY2017	PY2017	PY2017	PY2017	PY2017	PY2017	PY2017	PY2017	PY2018	--
Nov-17	PY2018	PY2018	PY2018	PY2018	PY2018	PY2018	PY2018	PY2018	PY2018	PY2018	PY2018
Dec-17	PY2018	PY2018	PY2018	PY2018	PY2018	PY2018	PY2018	PY2018	PY2018	PY2018	PY2018
Jan-18	PY2018	PY2018	PY2018	PY2018	PY2018	PY2018	PY2018	PY2018	PY2018	PY2018	PY2018
Feb-18	PY2018	PY2018	PY2018	PY2018	PY2018	PY2018	PY2018	PY2018	PY2018	PY2019	PY2018
Mar-18	PY2018	PY2018	PY2018	PY2018	PY2018	PY2018	PY2018	PY2018	PY2018	PY2019	PY2018
Apr-18	PY2018	PY2018	PY2018	PY2018	PY2018	PY2018	PY2018	PY2018	PY2018	PY2019	PY2018
May-18	PY2018	PY2018	PY2018	PY2018	PY2018	PY2018	PY2018	PY2018	PY2018	PY2019	PY2018
Jun-18	PY2018	PY2018	PY2018	PY2018	PY2018	PY2018	PY2018	PY2018	PY2018	PY2019	PY2018
Jul-18	PY2018	PY2018	PY2018	PY2018	PY2018	PY2018	PY2018	PY2018	PY2018	PY2019	PY2018
Aug-18	PY2018	PY2018	PY2018	PY2018	PY2018	PY2018	PY2018	PY2018	PY2018	PY2019	PY2018
Sep-18	PY2018	PY2018	PY2018	PY2018	PY2018	PY2018	PY2018	PY2018	PY2018	PY2019	PY2018
Oct-18	PY2018	PY2018	PY2018	PY2018	PY2018	PY2018	PY2018	PY2018	PY2018	PY2019	PY2018

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 PY2018 only and the second table combines the program years 2015-2018 evaluation results. DNV GL provided measure-specific results to Energy Trust separately.

Table C-1: Custom track evaluation results by measure Category, PY2018 Only

Custom Evaluation Category & Measure Description	# Evaluated Electric Measures	# Evaluated Gas Measures	Electric GRR (%)	Gas GRR (%)
Controls	**14**	**12**	**64%**	**69%**
Custom Building Controls	14	12	64%	69%
HVAC	**5**	**2**	**93%**	**83%**
Custom HVAC	4	1	101%	75%
Custom VAV System	1	1	59%	90%
Motors	**7**	**2**	**82%**	**22%**
Custom VFD Pump	2		94%	
Custom VFDs	5	2	77%	22%
Other	**3**	**18**	**95%**	**133%**
Custom Other	1	2	85%	100%
Retrocommissioning (RCx) Control Sequences Optimization	2	16	100%	137%
Process Cooling	**6**	**2**	**77%**	**36%**
Custom Chillers	6	2	77%	36%
Process Heating		**4**		**52%**
Custom Boiler		4		52%
Grand Total	**35**	**40**	**77%**	**93%**

Table C-2: Custom track evaluation results by measure category, PY2015 - PY2018

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	**38**	**16**	**107%**	**102%**
Custom Boiler	1	2	100%	98%
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	1	1	59%	90%
Motors	**29**	**4**	**76%**	**336%**
Custom Motors	1		88%	
Custom VFD Pump	6		77%	
Custom VFDs	22	4	75%	336%
Other	**18**	**27**	**118%**	**129%**
Custom Other	16	11	120%	118%
Retrocommissioning (RCx) Control Sequences Optimization	2	16	100%	137%
Process Cooling	**8**	**2**	**82%**	**36%**
Custom Chillers	8	2	82%	36%
Process Heating	**1**	**6**	**0%**	**54%**
Custom Boiler		5		63%
Custom Heat Recovery	1	1	0%	8%
Grand Total	**158**	**98**	**89%**	**101%**

APPENDIX D. STANDARD MEASURE RESULTS

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

Measure specific approach: The initial pages contain tables summarizing the evaluation activities and documenting recommendations associated with specific measures sampled for evaluation.

Project specific approach: A single table follows the measure specific tables for the project specific approach measures. These are measures with 4 or less measures sampled for evaluation.

DNV GL Measure Description	Boiler	
Track: Evaluation Code	Standard: BOILER	
Product Codes	GFBOIL300, GFBOIL2500, GFBOIL3002500, MODBOILBURN	
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.	**RR: Avg.** (Min-Max)	**65%** (40% - 144%)
	Sample Target	23
	Survey Completes	17

Measure Information

Program Data Review: 57 unique measure lines were reported over the program year 2018. These lines accounted for 0% of electricity savings and 36% of gas savings reported for the standard track.

Program Delivery: Standard

Evaluation Summary

The original evaluation sample for PY2018 included 18 sites. The final sample included 37 boilers at 17 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. If multiple boilers were installed and the interview confirmed that the primary boiler meets the majority of heating loads throughout the year, then the savings were adjusted. The evaluation found that the lag boiler is still typically needed for heating during higher demand periods but not throughout the heating season. Specific loading times or weather correlations to building load were not available. To account for the reduction in operation at these sites, the savings are **multiplied by 75% for all boilers at the site**. The evaluators confirmed that the second boiler at these sites operates as a lag boiler and are not specifically a backup.

2. 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. 6 out of 17 sites evaluated have shown evidence that the boilers do not operate in condensing mode.

Table D-1. Site-level results for boilers

Site	Site Boiler Quantity	Estimated Operating Efficiency	Operating Efficiency Scaling	Multiple Boiler Scaling	Evaluated UES, Therms	MAD UES, Therms
Site 1	1	97%	115%	125%	4.09	2.85
Site 2	1	95%	106%	100%	3.02	2.85
Site 3	1	89%	67%	100%	1.92	2.85
Site 4	1	89%	67%	100%	1.92	2.85
Site 5	1	96%	113%	100%	3.22	2.85
Site 6	4	94%	100%	100%	2.85	2.85
Site 7	1	89%	67%	75%	1.44	2.85
Site 8	2	89%	67%	100%	1.92	2.85
Site 9	2	96%	112%	60%	1.92	2.85
Site 10	1	94%	100%	75%	2.14	2.85
Site 11	2	89%	67%	75%	1.44	2.85
Site 12	2	96%	109%	60%	1.86	2.85
Site 13	2	95%	104%	60%	1.77	2.85
Site 14	2	96%	114%	50%	1.62	2.85
Site 15	3	N/A (modulating burner)	N/A (modulating burner)	100%	0.80	0.80
Site 16	5	89%	67%	60%	1.15	2.85
Site 17	5	96%	112%	60%	1.92	2.85

Regression Analysis Approach: Estimating savings based on interview responses and gas meter consumption

The regression model is trained on billing data (therms) and heating degree days (HDDs) for the post period after the high efficiency boiler was installed. The regression calculates the optimal threshold temperature and uses this to calculate the HDDs for each billing period. The results can be assessed for how weather-sensitive the gas usage is and how well the billing data fits the model (adjusted R^2). The linear regression gas consumption is calculated for each billing period using the following form:

$$Therms = (Coefficiecnt)(HDD) + Constant$$

We collected the end-use descriptions that are supplied by the boiler and descriptions of other gas equipment that is associated with the gas meter. This information is used to determine if the boiler gas consumption can be isolated from the gas bills. The billing analysis utilizes MAD savings adjustments as a basis and is used to sanity-check and adjust evaluated savings depending on whether billed consumption justified reported consumption. For the sites where we had adequate billing data, we encountered the following scenarios:

- The boiler is only used for space heating and makes up 100% of the gas bill. In this case, we can use the regression savings approach.
- The boiler(s) are part of a multi-boiler plant with modulation, used for 75% DHW and 25% space heating, and operates 24/7. Billing regression results suggest the site has higher consumption than most. Combined with the modulation and smaller boiler size, the evaluator determined to have no load-scaling for these specific boiler(s). (Site 6)
- The boiler is part of a heating system serving a natatorium (recreational facility with indoor swimming pool). The system runs 24/7, and the indoor air temperature is kept at 87 ºF year-round. Main boiler loads include, but are not limited to, space heating (70%) and snow melting (30%) as the building has 5 HVAC units/boiler systems which all serve the pool space. Site billing consumption shows a much higher consumption than predicted boiler consumption, and site end-use justifies this increase. Therefore, a 1.25 load scaling adjustment was made. (Site 1)
- The boiler is used for space heating for more than 80% of its end-uses, however, site billing consumption does not support the savings claim. This is indicated when comparing "Total Annualized meter, therms" (billing data consumption annualized) and "Claimed annual boiler load, therms" (annual boiler consumption deduced from claimed savings and assumed efficiency difference) in the table shown below. (Site 9, 12, 13, 14, 16)

Using the isolated boiler load from the post period, we calculated a heating load based on the estimated operating efficiency. The operating efficiency is the same value that was estimated for the MAD adjusted savings approach above. The baseline energy use is calculated using the heating load and the baseline boiler efficiency of 80%. The therm savings is the difference between the baseline and post period therms.

Table D-2 provides a summary of the results for the sites that took the regression approach results into account when adjusting savings estimate.

Table D-2. Regression approach results for boilers

Site	Site 1	Site 6	Site 9	Site 12	Site 13	Site 14	Site 16
Adjusted R2	0.00	0.86	0.91	0.73	0.00	0.98	0.95
HDD Temperature Threshold, F	0	59	60	59	0	57	70
Constant	9,935.5	1,790.2	159.5	396.0	4,174.8	28.9	168.0
HDD Coefficient	-	2.5	2.4	9.9	-	8.0	19.9
Total Annualized meter, therms	105,497	27,599	9,187	34,391	14,702	22,974	139,680
Non-Weather Dependent, therms	-	21,483	1,755	4,356	-	347	2,184
Weather Dependent, therms	-	6,116	7,432	30,035	-	22,627	137,496
Percent of Meter Weather Dependent	24.7%	22.2%	80.9%	87.3%	89.8%	98.5%	98.4%
Claimed Annual Boiler Load, therms	4,264	21,375	12,825	42,750	32,063	55,575	160,313

The high adjusted R^2 shows that the billing data fits this model. For Sites 1 and 13 the best regression model for the data was a flat curve, therefore no curve-fitting statistics were produced. The boiler at Site 1 is used 75% for space heating and 25% snow melting. The boilers at Sites 6 and 17 have a large portion of boiler load classified as DHW (laundry, kitchen, etc.) Site 6 is a truck stop travel-center with a variety of facilities. Site 17 is a prison and its boilers were operating at low percentage at the time of visit. Table D-3 shows the approach and savings results for each site.

Table D-3. Evaluation approach and results for boiler sites

Site	Evaluation Approach	Evaluated Savings Therms	Reported Total therms	Realization Rate
Site 1	MAD + billing regression adjustment	1,632	1,137	144%
Site 2	MAD adjustment	601	567	106%
Site 3	MAD adjustment	764	1,137	67%
Site 4	MAD adjustment	1,628	2,423	67%
Site 5	MAD adjustment	2,417	2,138	113%
Site 6	MAD + billing regression adjustment	5,700	5,700	100%
Site 7	MAD adjustment	2,874	5,700	50%
Site 8	MAD adjustment	2,874	4,275	67%
Site 9	MAD + billing regression adjustment	2,308	3,420	67%
Site 10	MAD adjustment	6,413	8,550	75%
Site 11	MAD adjustment	3,592	7,125	50%
Site 12	MAD + billing regression adjustment	7,454	11,400	65%
Site 13	MAD + billing regression adjustment	5,316	8,550	62%
Site 14	MAD + billing regression adjustment	8,421	14,820	57%
Site 15	MAD adjustment	16,800	16,800	100%
Site 16	MAD + billing regression adjustment	17,242	42,750	40%
Site 17	MAD adjustment	57,706	85,500	67%
Total	-	**143,741**	**221,991**	**65%**

The evaluation team recommends ensuring that the gas meter number specifically assigned to the boiler is listed on the application and captured in program tracking data. Additionally, collecting the boiler end use (space heating, reheating, DHW, pool heating, etc.) could help this type of assessment in the future.

Evaluation Recommendations

A number of sites had boilers servicing equipment other than space heating. The MAD savings assume that boilers provide space heating only. DNV GL continues to suggest that the boiler load end-uses be identified and alternatetive savings estimates be developed.

Some boilers were found to have conditions that made it unlikely that they typically operate in condesing mode. Specifically the return water temperature did not meet the threshold that reflects condensing mode. This reduces the operational efficiency as well as savings. It may be beneficial to have sites demonstrate that boilers will operate in condensing mode based on loading, estimated setpoints and return water temperature. As an alternative, the savings estimate could account for a percentage of boilers that will not operate in condesing mode. This observation persists for PY2018 program participants.

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. We recommend that savings account for sites with mutiple boilers.

Adjustment to evaluation plan

We used both a billing data regression analysis and MAD adjusted savings approach, depending on the data available.

<table>
<tr><th>DNV GL Measure Description</th><td colspan="2">Wall Insulation, Roof/Ceiling Insulation and Pipe Insulation</td></tr>
<tr><th>Track: Evaluation Code</th><td colspan="2">Standard: WALLINSULATE, CEILINGINSULATE, PIPEINSULATE</td></tr>
<tr><th>Product Codes</th><td colspan="2">INSWALLEHP, INSWALLG, INSATTICER, INSROOFER, INSATTICEHP, INSROOFG, INSROOFGR5R20, INSATTICG, INSPIPEDHW180</td></tr>
<tr><td rowspan="3">Insulation is offered for wall insulation, roof insulation or attic insulation, and pipe insulation. Two basic measures are no existing insulation and some level of existing insulation. Different heating system types are covered: gas, electric resistance, and heat pump. Key evaluation parameters include building area, ceiling insulation area verification, wall insulation area verification, pipe insulation linear footage verification, building vintage, roof/attic, existing insulation verification, and space heating/space cooling verification.</td><td>RR: Avg.
(Min-Max)</td><td>Wall: 100%
(1 site only)
Roof/Ceiling, Gas savings: 54%
(8% - 100%)
Roof/Ceiling, Electric savings: 100%
(1 site only)
Pipe: 67%
(1 site only)</td></tr>
<tr><td>Sample Target</td><td>Wall: 2 (1 sample, 1 backup)
Roof/Ceiling: 9 (7 sample, 2 backup)
Pipe: 1</td></tr>
<tr><td>Survey Completes</td><td>Wall: 1
Roof/Ceiling: 6
Pipe: 1</td></tr>
</table>

Measure Information

Program Data Review: 45 unique measure lines (8 for wall, 36 for roof/ceiling, 1 for pipe) were reported over the program year 2018. In total, these lines accounted for 0.63% of electricity savings and 12.90% of gas savings reported for the Standard track.

Program Delivery: This is a standard prescriptive measure.

Evaluation Summary

The evaluation sample included 12 sites for verification. 8 interviews were completed. 4 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, and 2 to 3 sites were cross-checked with the billing data regression estimate.
- Measure Information:
    - MAD savings are square-footage dependent.
    - 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 relative to the space.
    - Verified billing consumption for 4 sites. Savings are reduced for 3 of these sites. A summary of the reduced sites is provided in the table below.

Site	Measure	Evaluated UES, Therms	MAD UES, Therms
Site 1	Insulation Roof Gas	0.02 therm/sqft	0.25 therm/sqft
Site 2	Insulation Roof Gas	0.19 therm/sqft	0.25 therm/sqft
Site 3	Pipe Insulation - DHW	3.8 therm/linear ft pipe	5.7 therms/linear ft pipe

- Site 1 is a warehouse that consumes about 3,000 therms per year, but claimed 12,000+ therms in savings. The unit savings (0.25 therm/sqft) scales with sqft with the MAD assumption of 72°F occupancy setpoint. However, 90% of the building is warehouse space, which is kept at 52°F. Based on a billing regression analysis similar to ones done for boilers, about 10% of the claimed savings was estimated. Therefore the savings is reduced to about 10%.
- Both Site 2 (school) and Site 3 (church) show billing consumption that do not support their savings claim. Their savings are reduced to 75% and 67%, respectively.

Evaluation Recommendations

- Savings based on building area alone may not be accurate as HVAC systems do not serve zones/spaces uniformly for some building types. Consider revising measure requirement and/or adjusting incentive rates/categories based on percentage space covered by HVAC system and building setpoints.
- MADs should be updated to include measure theory, such as "retrofit", "early replacement", "lost opportunity", or terms specifically used by Energy Trust.
- 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

MAD savings verification for most sites and billing verification for some sites

DNV GL Measure Description	HVAC Occupancy Sensors	
Track: Evaluation Code	**Standard: HVAC**	
Product Codes	HVACINROOM	
This measure covers the installation of occupancy-based PTHP/PTAC controllers, which incorporate an occupancy sensor integrated with the room thermostat to provide occupancy-controlled heating and/or cooling. Key evaluation parameters include quantity verification, standard setpoint of thermostat, occupancy rate, and setback strategy implementation if any.	**RR: Avg.** (Min-Max)	**84%**
	Sample Target	1
	Survey Completes	1

Measure Information

Program Data Review: 5 unique measure lines were reported over the program year 2018. These lines accounted for 0.27% of electricity savings and 0% of gas savings reported for the Standard track.

Program Delivery: This is a standard prescriptive measure.

Evaluation Summary

The evaluation sample included 1 site for verification. 1 interview was completed.

- Telephone interviews were completed for data collection.
- Tracking savings were verified with MAD.
- MAD savings feature a single estimated UES based on a 2007 BC Hydro study and is adjusted based on average 60% occupancy rate.
- Site level RR is 84% and verified occupancy rate is 50%, so evaluated UES is de-rated to 50% of the original MAD estimate instead of 60%.

Evaluation Recommendations

- MADs should be updated to include measure theory as "retrofit", "early replacement", or "lost opportunity".
- 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 GL Measure Description	Packaged Terminal Heat Pumps	
Track: Evaluation Code	Standard: HEATPUMP	
Product Codes	PTHEATPUMPHZ1, PTHEATPUMPHZ2	
This measure covers the installation of packaged terminal heat pumps that replace electric resistance heat, and it applies to multifamily buildings, assisted living facilities and dorms, as well as hotels and motels. Key evaluation parameters include building location (climate zone), building type, installed quantity, typical setpoints and occupancy rate.	**RR: Avg.** (Min-Max)	**68%** (56% - 105%)
	Sample Target	7
	Survey Completes	6

Measure Information

Program Data Review: 67 unique measure lines were reported over the program year 2018. These lines accounted for 7.23% 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.
- 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/assist living/dorms).
- Climate-zone definition from MAD: CZ-1 is West, CZ-2 is East.
- All sites verified are of hotels/motels/lodging building type. Savings are adjusted based on heating degree-days (HDD base-60) of each site's location and existing MAD savings (with corresponding degree-day values).
-

Evaluation Recommendations

- MADs should be updated to include measure theory as "retrofit", "early replacement", or "lost opportunity".
- MADs should clearly state the baseline assumption used for the analysis and source, including climate-zone sources and references
- 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 GL Measure Description	Power Strips	
Track: Evaluation Code	Standard: POWERSTRIP	
Product Codes	OCCPLUGSTRIP, OCCPLUGSTRIPSI, LOADPLUGSTRIP	
The smart strip was a TrickleStar Motion Sensor Power Strip. Key evaluation parameters include quantity verification, end-use space/building type verification, and the use of motion-sensing feature of the equipment.	**RR: Avg.** (Min-Max)	**91%** (83% - 100%)
	Sample Target	4
	Survey Completes	3

Measure Information

Program Data Review: 130 unique measure lines were reported over PY2018. These lines accounted for 1.22% of electricity savings and 0% of gas savings reported for the Standard track.

Program Delivery: This is a standard prescriptive measure.

Evaluation Summary

The evaluation sample included 4 sites for verification. 3 interviews were completed. Project data was never received for the last site.

- Telephone interviews were completed for data collection; powerstrip quantities were verified verbally.
- Tracking savings were verified with MAD.
- Two sites are schools and one site is small retail. One school site says they have a small portion of claimed powerstrips that are unused.
- Schools tend to buy powerstrips in bulk and deploy them as needed (not all purchased are used immediately; some are stored).

Evaluation Recommendations

- MADs should be updated to include measure theory as "retrofit", "early replacement", or "lost opportunity".
- 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 GL Measure Description	Economizer control for RTUs	
Track: Evaluation Code	Standard: ECONO	
Product Codes	RTUCTRLECOGH	
This measure covers the installation of economizers to rooftop units which are not required by code. 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, and additional control features (DCV and VFD) verification.	**RR: Avg.** (Min-Max)	**100%** (100% - 100%)
	Sample Target	3
	Survey Completes	3

Measure Information

Program Data Review: 5 unique measure lines were reported over the program year 2018. These lines accounted for 0.04% of electricity savings and 0% of gas savings reported for the Standard track.

Program Delivery: This is a standard prescriptive measure.

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; that is, in order to qualify for DCV incentives a unit must also have an economizer and in order to qualify for VFD incentives the unit must be equipped with both 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 a new unit with DX cooling and gas/heat pump heating.
- Economizer logic and tonnage were verified. For 1 site, evidence of DCV features were found, but these were found to be required by code after researching the specific building specifications. Therefore, only the economizer claim was valid for this site.

Evaluation Recommendations

- MADs should be updated to include measure theory as “retrofit”, “early replacement”, or “lost opportunity”.
- MADs should clearly state the baseline assumption used for the analysis and source.
- Provide transparent documentation on 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 GL Measure Description	Tank Water Heaters	
Track: Evaluation Code	Standard: TANKDHW	
Product Codes	DHWCOND	
This measure covers the installation of condensing tank water heaters. Tank water heaters account for about 2.62% of gas savings. Key evaluation parameters include building type, water temperature setpoint, and water heater quantity.	**RR: Avg.** (Min-Max)	**415%** (383% - 473%)
	Sample Target	7
	Survey Completes	5

Measure Information

Program Data Review: 61 unique measure lines were reported over the program year 2018. These lines accounted for 0% of electricity savings and 2.62% 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 sites for verification. 5 interviews were completed. 2 sites did not respond to our requests.

- Telephone interviews were completed for data collection.
- Tracking savings were verified with MAD. 3 sites were motel/hotels. 2 sites are laundry services.
- MAD savings are building-type dependent. The program savings value is a much lower weighted average of building type specific savings. We determined that using building-type-specific savings would be more accurate in this case, as most participants for this measure are lodging type in PY2018.
- The lodging and laundry building types use higher amounts of hot water than other building types, therefore savings adjustment is justified.
- The building-type-specific UES for dormitories is 2.21, which implies that the UES for lodging should be higher than the existing-building weighted average of 1.37.

Evaluation Recommendations

- Consider the building-type mix of program claims to revise the savings rate of tank water heaters.
- The program should review the weighted average used against recent participation and future expectations. The savings assumed or weighting used should be adjusted to align with future program expectations. Alternately consider revising the application to have building-type dependent incentive/savings rates.
- MADs should be updated to include measure theory as "retrofit", "early replacement", or "lost opportunity".
- Provide transparent documentation on the weighted average UES and sources of all building type weights. The calculation/methodology for these weights should be embedded in the MAD.

Adjustment to evaluation plan

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

DNV GL Measure Description	Vent Hood	
Track: Evaluation Code	**Standard: FOODEQUIP**	
Product Codes	VENTHOODG	
This measure covers the installation of new ENERGY STAR-compliant vent hoods. Vent hoods account for about 0.16% of gas savings and 0.04% of electric savings. Key evaluation parameters include motor horsepower, business hours, and annual days of operation.	**RR: Avg.** (Min-Max)	**Electric: 148%** (119% - 199%) **Gas: 224%** (199% - 239%)
	Sample Target	2
	Survey Completes	**2**

Measure Information

Program Data Review: 4 unique measure lines were reported over the program year 2018. These lines accounted for 0.04% of electricity savings and 0.16% 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 2 sites for verification. 2 interviews were completed.

- Telephone interviews and site visits were completed for data collection.
- Verifying tracking savings with MAD.
- Business hours and annual days of operation were obtained from participant surveys. Equipment specifications including motor horsepower were obtained from project documentation.
- Both sites additionally had higher hours of use and days of operation than were assumed.

Evaluation Recommendations

Ensure incentivized equipment UES are consistent with deemed savings value on MAD document.

Adjustment to evaluation plan

None

DNV GL Measure Description	Hot Food Holding Cabinet	
Track: Evaluation Code	Standard: FOODEQUIP	
Product Codes	HOTFOODCABHALF	
This measure covers the installation of new ENERGY STAR-compliant hot food holding cabinets. Half-size hot food holding cabinets account for about 0.85% of electric savings in the standard track. Key evaluation parameters include cabinet size (interior volume), idle energy rate, and hours of use.	**RR:**	4.7%
	Sample Target	1
	Survey Completes	1

Measure Information

Program Data Review: 35 unique measure lines were reported over the program year 2018. These lines accounted for 0.85% 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 ENERGY STAR certification are submitted.

Evaluation Summary

The evaluation sample included 1 site for verification. 1 interview was completed. This project was promoted from the backups toward the end of the data collection phase.

- A telephone interview was completed for data collection.
- Verifying tracking savings with MAD. This project had only half-size cabinets.
- Typical hours of use are based on customer response. Interior volume and idle energy use was taken from product specifications.
- Energy Star calculator was used to calculate savings based on parameters mentioned above.
- Low RR as a result of reduced evaluated interior volume of the hot food cabinet based on the provided equipment specifications compared to the default (3 cubic feet versus 19 cubic feet).

Evaluation Recommendations

Use the ENERGY STAR calculator for traceability and transparency.

Provide information in MAD document about what constitutes “half-size” or “full-size” cabinets.

Adjustment to evaluation plan

None

DNV GL Measure Description	Dishwasher	
Track: Evaluation Code	Standard: DISHWASH, FOODEQUIP	
Product Codes	DWSTDRUPLTEMPG, DWSTDRUPLTEMPE DWSTDRUPHTEMPG, DWSTDRUPHTEMPE DWUCHTEMPE	
This measure covers the installation of new ENERGY STAR-compliant dishwashers. Dishwashers account for about 1.6% of gas savings and 1.9% of electricity savings. Key evaluation parameters include annual days of operation, racks washed per day, average daily operation, typical wash time, water use per rack, idle power draw, and use of water heater booster.	**RR: Avg.** (Min-Max)	**Electric: 105%** (13% - 43%) **Gas: 2.9%** (0%-11%)
	Sample Target	**5**
	Survey Completes	5

Measure Information

Program Data Review: 58 unique measure lines were reported over the program year 2018. These lines accounted for 1.9% of electricity savings and 1.6% 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 and site visits were completed for data collection

- Verifying 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.
- All of the measures had significantly lower daily hours of use and racks washed than is used in the default ENERGY STAR calculator. This was the primary case for low realization rates.
- One measure lost additional gas savings due to water heater booster savings being zeroed out. The absence of water heater booster was determined by response to our survey.
- The reason for the high electric realization rate and low gas realization rate is because the project that reported the largest savings was incorrectly reported as having gas as a fuel type. Both file review and phone survey indicated that the equipment was electric.

Evaluation Recommendations

Ensure incentivized equipment quantities are consistent with claimed savings value (one site had a typo on claimed savings from tracking data).

Adjustment to evaluation plan

None

DNV GL Measure Description	Ice Machine	
Track: Evaluation Code	**Standard: ICEMAKER**	
Product Codes	ICESCUSMT2, ICEIMHSMT1, ICEIMHLGT1, ICESCULGT1	
This measure covers the installation of new ENERGY STAR-compliant icemakers. Icemakers account for about 0.32% of electric savings. Key evaluation parameters include equipment type, ice harvest rate, and annual days of operation.	**RR: Avg.** (Min-Max)	**69%** (29% - 118%)
	Sample Target	3
	Survey Completes	5

Measure Information

Program Data Review: 74 unique measure lines were reported over the program year 2018. These lines accounted for 0.32% 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 ENERGY STAR certification are submitted.

Evaluation Summary

The evaluation sample originally included 3 sites for verification; 2 more were added from the backups toward the end of the data collection phase. 5 interviews were completed.

- Telephone interviews were completed for data collection.
- Verifying tracking savings with MAD document.
- Equipment type and ice harvest rate were obtained from documentation in the project files. Annual days of operation were obtained from participant surveys. The major discrepancy in savings was due to equipment having higher ice harvest rates than was assumed in the ENERGY STAR calculator.

Evaluation Recommendations

Use the ENERGY STAR calculator for traceability and transparency.

Adjustment to evaluation plan

None

DNV GL

Driven by our purpose of safeguarding life, property and the environment, DNV GL enables organizations to advance the safety and sustainability of their businesses. Our origins stretch back to 1864, and our reach today is global. Operating in more than 100 countries, we are dedicated to helping our customers make the world safer, smarter and greener.