DNV·GL

Impact Evaluation of the 2017 Existing Buildings Program

Energy Trust of Oregon

Date: 8/8/2019

<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 2017 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	Working kWh	% of kWh Grand Total	Working therms	% of therms Grand Total
Lighting	6,675	80,527,411	61%		
Standard	1678	20,127,512	15%	900,864	48%
Custom	218	24,452,156	19%	780,488	41%
Capital Measures Only	**8,571**	**125,107,079**	**95%**	**1,681,352**	**89%**
SEM Cohort	166	6,014,681	5%	209,043	11%
All Existing Buildings	**8,737**	**131,121,760**		**1,890,395**	

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. Note that the evaluated savings for Custom Gas is not equal to the realization rate times the claimed energy savings. Due to information learned through this evaluation, one large project was removed from the sample frame and will be evaluated separately.

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

Program Track	Electricity Savings (kWh)	Gas Savings (therms)
	2017	2017
Lighting	79,302,959	
Standard	17,711,325	950,155
Custom	21,987,514	565,279
Capital Measures Only	**119,001,799**	**1,515,434**
Strategic Energy Management	5,539,687	137,968
Grand Total	**124,541,486**	**1,653,402**

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

Program Track	Electricity Realization Rates	Gas Realization Rates
	2017	2017
Lighting	99%	
Standard	88%	105%
Custom	90%	87%
Capital Measures Only	**95%**	**90%**
Strategic Energy Management	92%	66%
Existing Buildings Program	**95%**	**87%**

0.4 Historic capital measure performance

Table 0-4, Figure 1 and Figure 2 show historic program performance for capital measures. The table and charts do not include the SEM track, which was added to the Existing Buildings program impact evaluations in 2015.

Table 0-4: 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%

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

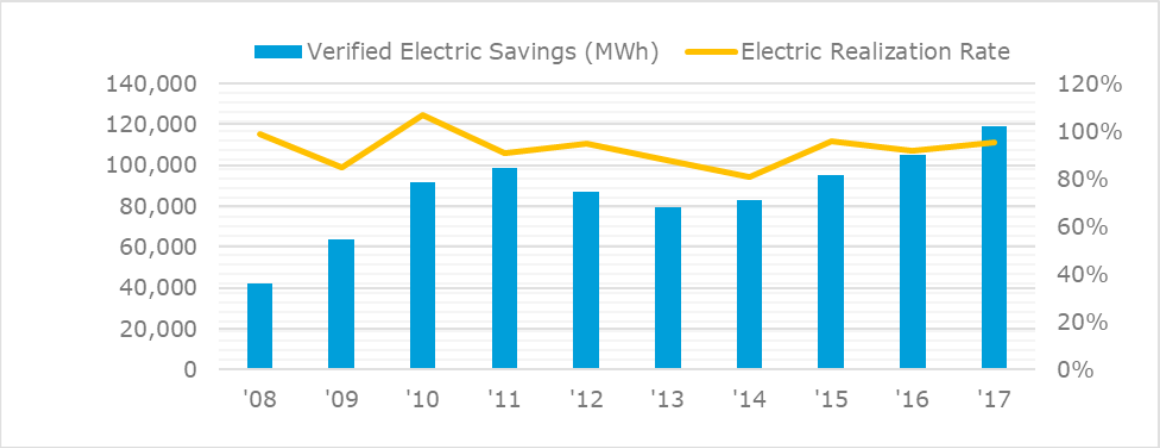
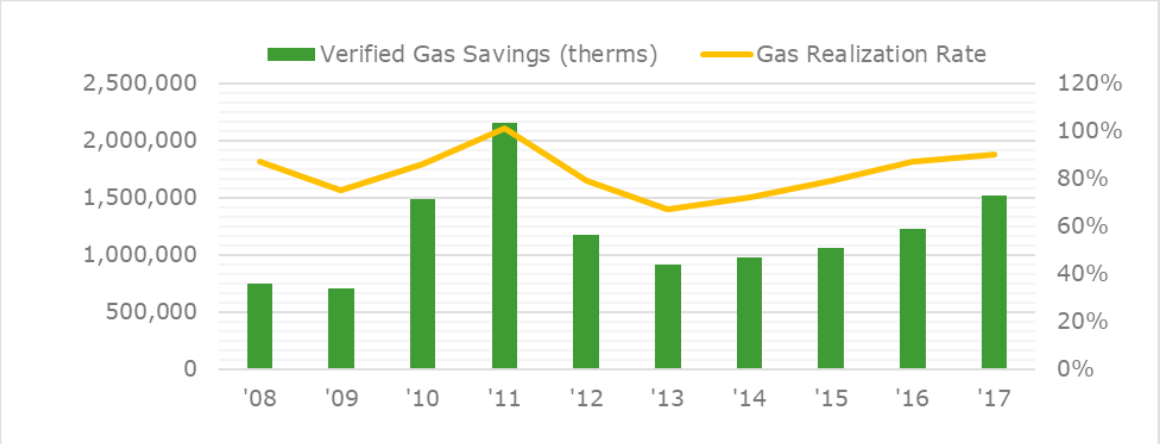


Figure 2: Historic Non-SEM program gas savings and realization rates



0.5 Historic SEM performance

Table 0-5, Figure 3 and Figure 4 show historic SEM performance over time.

Table 0-5: 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%	128,402	66%

Figure 3: Historic SEM program electric savings and realization rates

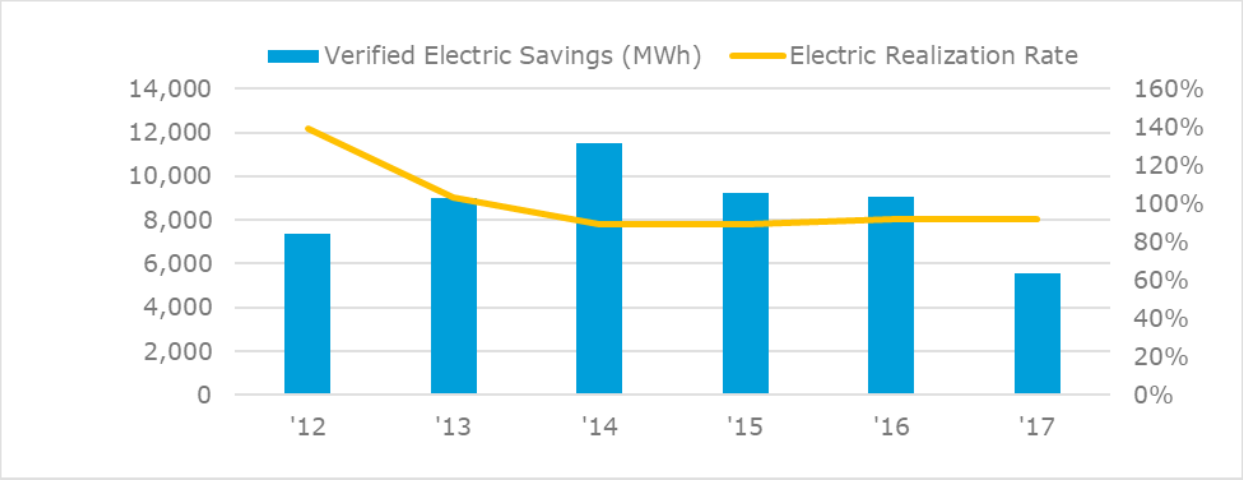
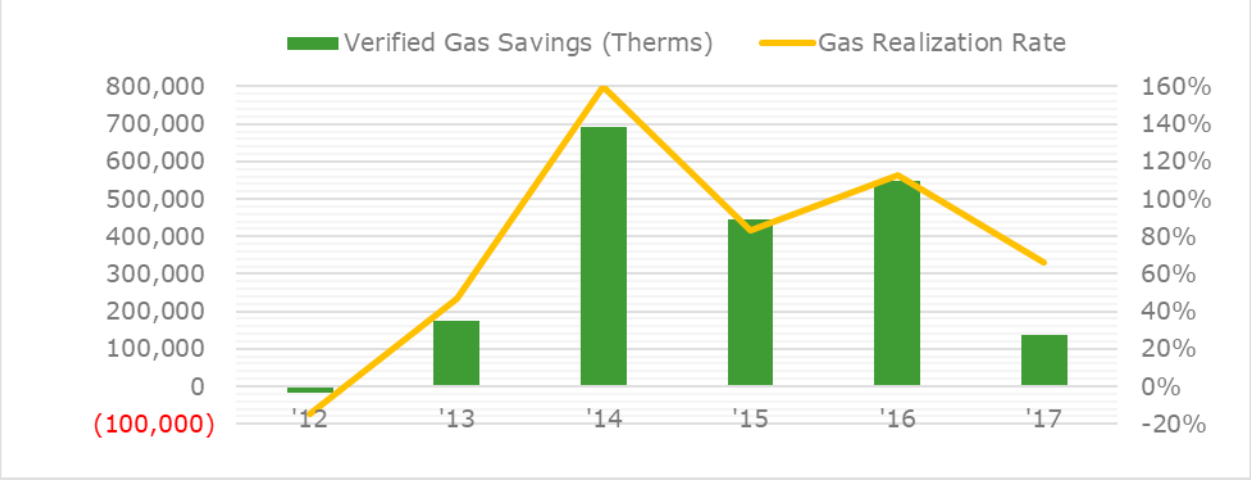


Figure 4: Historic SEM program gas savings and realization rates



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 recommendations

- **Finding** – The estimated kWh savings differed by at least 10% from the reported savings for only four lighting projects. Two projects had lower evaluated savings than reported savings; two had higher. Overall, the program accurately estimated lighting end-use energy savings.
  - **Recommendation** – Program staff should continue to emphasize the importance of accurate estimates of operating hours during training for trade allies.
- **Finding** – General satisfaction with tubular LEDs (TLEDs) is high and performance issues are minimal. The twenty-three survey participants gave their TLEDs an average rating of 4.7 out of 5, with all but one giving either a 4 or 5. One participant gave a 3 rating, saying his new TLED fixtures were bright enough but he wasn't sure he was saving money on his electric bill. Only one customer indicated that he'd had a problem with any of his TLEDs (a fixture stopped working), and he had not yet contacted his contractor to resolve it. No respondents had removed any lamps or fixtures since the retrofit, another indication of high satisfaction with lighting system performance.
  - **Recommendation** – Continue supporting the installation of TLEDs. No systematic concerns were identified. DNV GL recommends not including these TLED specific questions in future impact evaluations of this program.

0.6.2 Measure Approval Documents recommendations

- **Finding -** For the 2017 evaluation there was only one MAD we had not already reviewed, the 2014 MAD for variable-speed drives on vent hoods. As with the MADs we reviewed for the 2015-16 program years, we found that the MAD does 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.

0.6.3 Standard measure recommendations

- **Finding** – *Space Heating Boilers.* The evaluation team found several sites with multiple boilers operating in lead/lag 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 we believe it likely that the lag boiler will operate much less than the measure savings assume. Measure savings are currently for a single boiler providing the entire load.
  - **Recommendation** – Measure savings should be adjusted to assume that most sites with mutiple boilers will operate with lead/lag sequencing and the lag boiler load will be signficantly less than the lead load. Measure savings documentation should be updated to transpartenly communicate the basic assumptions and structure used to estimate measure savings.
- **Finding** – *Space Heating Boilers.* DNV GL found a number of boilers are providing functions other than space heating. The measure savings documentation assumes that boilers provide space heating only.

- **Recommendation** – Consider identifying and developing savings estimates for non-space-heating uses or adjust the program design for non-space-heating boiler applications to improve the accuracy of savings estimation.

- **Finding** – *Space Heating Boilers.* DNV GL found boilers operating in conditions that made it unlikely that they operate in condensing mode. Operating outside the condensing zone reduces the operational efficiency of the installed boiler and reduces savings.
  - **Recommendation** – The program could require sites to demonstrate that boilers will operate in condensing mode based on loading and estimated setpoints on the application. Any verification activities completed on boiler installations should include a review of the operating setpoints.

0.6.4 Custom 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.

- **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.
  - **Recommendation** – Energy Trust should require the PMC to 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.
- **Finding** – Program models continue to estimate savings that suggest a significant reduction in annual consumption. In some cases, the savings were found to exist. In other cases, the savings did not materialize.
  - **Recommendation** – Energy Trust should complete additional review of simulation inputs for sites expecting savings greater than 20% of consumption.

0.6.5 Strategic Energy Management recommendations

- **Finding –** The site specific realization rate for eight gas sites is below 20%. Six of these sites achieved a site realization rate of 0%. These results are the primary driver of the 66% gas realization rate for this track. These sites did not have capital project adjustments and only one has a baseline/other adjustment. In most cases, these sites are achieving cumulative savings over the baseline, but no incremental savings were achieved in program year 2017. DNV GL believes cases like this will continue to exist until all sites have baseline models meeting the current guidelines.
  - **Recommendation** – DNV GL recommends that Energy Trust continue its efforts to re-baseline continuation participants with average mean temperature baselines. Reducing differences in the baseline modelling approach will reduce this variance in continuation participants.
  - **Recommendation** – DNV GL also recommends that Energy Trust consider not claiming continuation savings that are a small percent (less than 2%) of total consumption for participants in their 3rd year or later if the baseline model does not meet the current guidelines. Based on this evaluation, the degree-day baseline modelling approach is more likely to not support the savings claim than to support the claim.
- **Finding** – Participants continue to value energy coaches and peer-to-peer learning. Participants cite benefits from the insights provided by working closely with energy coaches to identify and execute

operational and capital improvement opportunities. Participants also commented on perceiving value from the peer-to-peer information exchanges with participants of a similar facility type. These learning exchanges provide participants with practical 'case study' examples to draw upon, as well as benchmarking and competitive motivation across organizations with similar facilities.

- **Recommendation** – DNV GL recommends that Energy Trust continue to identify program improvements that allow energy coaches to spend more time working with participant staff to support energy conservation opportunities.

Memo

To: Board of Directors

From: Jay Olson, Sr. Program Manager – Commercial
Kathleen Belkhayat, Program Manager – Commercial
Sarah Castor, Evaluation Sr. Project Manager

cc:

Date: September 19, 2019

Re: Staff Response to the Impact Evaluation of the 2017 Existing Buildings Program

The 2017 Existing Buildings program impact evaluation confirmed that the program is doing a good job of estimating electric and gas savings from capital measures, with savings realization rates of 95% and 90%, respectively. This finding is important given that the program has been increasing its activity and savings steadily since 2013.

The impact evaluation included the program's four tracks: custom, lighting, standard and Strategic Energy Management (SEM). The SEM portion of the 2017 impact evaluation included SEM continuation participants only, those participants that were in their second or later year of engagement in SEM, and demonstrated a good realization rate for electric savings (92%) and a lower gas realization rate than expected or found in recent program years (66%). A key factor in the SEM gas realization rate was the fact that many of the savings models used for participating sites were out of date and do not conform to current Energy Trust guidelines for modeling energy savings. For the past two years, the Existing Buildings program has been replacing older savings models with new ones that conform to program guidelines, which should help improve savings estimation. A review of SEM projects for the 2018 Existing Buildings impact evaluation, currently in progress, revealed that 65% of projects had models less than a year old and these new models account for more than 70% of the electric and gas savings claimed by SEM in 2018.

Energy Trust is committed to regularly updating the savings estimates and documentation for its standard measures, as recommended by the evaluator. In 2019, Existing Buildings program staff updated standard measures for boilers and boiler burners, grocery refrigeration, and various lighting technologies, in additional to many other measures. The updates will take effect with the 2020 program year and address many of the suggestions made by the evaluator with respect to these measures.

Interviews with participants who installed tube light-emitting diodes (TLEDs) in 2017 confirm the findings from a similar investigation in the 2015-2016 program year evaluation: TLED participants are satisfied with the performance of their lighting and have experienced almost no issues with the technology. This finding is reassuring given the rapid growth in the installation of TLEDs over the last four years, and Energy Trust does not see a need to continue to collect in-depth information on satisfaction with TLEDs in future impact evaluations.

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 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 2017. The goals of this evaluation are to:

- Develop estimates of Existing Buildings program gas and electric savings to establish realization rates for the 2017 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 $2.3 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 2017 (PY2017).
- Calculate gas and electric realization rates for PY2017.
- 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.

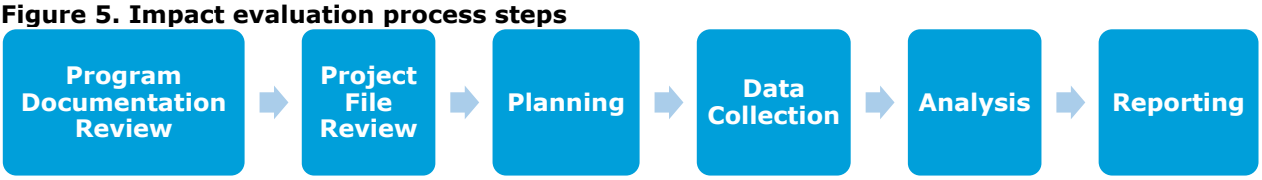
Table 2-1: Sample summary

Program Track	2017 Electric				2017 Gas			
	% of Savings in Sample	N	n	Relative Precision (@ 90% CI)	% of Savings in Sample	N	n	Relative Precision (@ 90% CI)
Lighting	10%	2,366	48	17%				
Standard	10%	796	36	22%	17%	682	39	18%
Custom	45%	157	32	19%	49%	84	18	20%
Subtotal: Non-SEM	**17%**	**3,319**	**116**	**12%**	**32%**	**766**	**57**	**13%**
SEM	64%	117	36	17%	66%	82	28	19%
Total (All Measures)	**19%**	**3,436**	**152**	**11%**	**36%**	**848**	**85**	**12%**

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 is 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 is outside of the IPMVP framework. The following are the key elements that supplement the preparation of project EM&V plans:
  - **Evaluating Standard/Prescriptive Measures.** The 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 determine 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 60% of the electricity savings reported by the program.

Table 3-1: Reported lighting track energy savings by delivery

Track	Electricity (kWh)
Lighting	69,293,686
Direct Install	4,724,333
Street Lighting	6,509,392
Lighting Total	**80,527,411**
Existing Buildings program total	131,121,760
Percent of Existing Buildings program savings	61%

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 a technology domain for Street Lighting due to the unique attributes of these projects and the magnitude of savings reported per project. Further stratification was performed within Street Lighting to separate one larger city's projects from all other street lighting projects.
- Creation of domains for Direct Install and Standard to ensure that both were represented in the evaluation sample.
- Stratification by size of savings reported (up to three 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). Table 3-2 summarizes the sample design for the lighting track. This design was expected to provide program year savings estimates with 17% relative precision at the 90% confidence interval. Further detail on 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	250	7
		2	80	7
		3	38	6
Street Lighting: City 1	Electric	1	4	2
Street Lighting: Non-City 1	Electric	1	7	1
		Certainty	1	1
Standard Lighting	Electric	1	1,528	8
		2	350	8
		3	108	8
EVALUATION TOTAL			**2,366**	**48**
Percent of Reported kWh in sample	Direct Install			10%
	Street Lighting			70%
	Standard Lighting			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. Creation of an impact evaluation data collection tool. Update of our TLED interview guide from the 2015-16 program year evaluation.
- **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. Analyze responses to TLED interview questions.

3.2.2 Documentation and file review

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

- **Documentation was sufficient.** The documentation for the 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 the 2015-16 Existing Buildings evaluation.

3.2.3 Data collection planning

DNV GL developed data collection plans and tools to achieve both the impact evaluation and, for projects including TLEDs, participant feedback objectives of this evaluation. The TLED interview guide is available in Appendix B. The objectives of the interview along with the results are discussed in section 3.4.

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.[1] DNV GL agrees with previous program evaluators[2] 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 and street lighting projects, data collection occurred exclusively via telephone interview. For Standard projects, we interviewed 9 sites by telephone and performed site visits at 11. 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.

Interviews with program participants who installed TLED measures included additional questions specific to TLEDs. We asked the same questions about TLED fixtures that we asked in the 2015-16 evaluation.

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 project. 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.

[1] 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.

[2] 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

Key Savings Parameters - The key savings parameters researched were:

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

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

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

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

3.3 Lighting track evaluation results

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

3.3.1 Achieved sample

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

- One site, a manufacter/distributor, agreed to a site visit but when our engineer arrived nobody available was familiar with either the visit or the lighting project.
- Another site, a supermarket, canceled without notice.
- Another supermarket never responded to our multiple requests for a site visit.

Table 3-3: Final lighting track sample summary

Sub-Category	Size Stratum	Sample Target (n)	Achieved Sample	% Complete
Direct Install Lighting	1	7	5	71%
	2	7	7	100%
	3	6	6	100%
Street Lighting: City 1	1	2	2	100%
Street Lighting: Non-City 1	1	1	1	100%
	Certainty	1	1	100%
Standard Lighting	1	8	7	88%
	2	8	8	100%
	3	8	5	63%
Grand Total		**48**	**42**	88%

3.3.2 Evaluated savings

Expansion from the sample to track-level results follows the methodology discussed in Section 2.4. Realization rates by technology and for lighting as a whole are presented in Table 3-4. Overall, DNV GL estimates the evaluated lighting savings across all technologies to be 98% of the reported savings with a relative precision of 4.4% at the 90% confidence interval.

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

Sub-Category	Completed Sample	Realization Rate	Standard Error	Relative Precision at 90 % Confidence
Standard - Lighting	20	98%	0.031	5.2%
Direct Install	18	103%	0.031	5.1%
Street Lighting	4	100%	0.000	0.0%
Lighting Total	**42**	**98%**	**0.026**	**4.4%**

3.3.3 Savings variance

The assumptions used to estimate reported savings were reasonable for most projects: 37 of the 42 (88%) projects we evaluated had realization rates between 90% and 110%. Actual vs. reported hours of operation were a factor in all three projects that did not achieve at least 90% realization rates. For the two projects with realization rates above 110%, operating hours were a factor in just one. Table 3-5 shows the five projects with realization rates below 90% or above 110%.

Table 3-5: Lighting measures variance summary by building type

Building Type	Reported kWh	Evaluated kWh	Realization Rate	Reason for Variance
Large Office Building	16,583	9,526	57%	Some tenant spaces remain unoccupied following installation so evaluated savings for those spaces are zero.
Auto Repair (DI)	7,969	10,349	77%	Actual operating hours are only 77% of reported hours
Manufacturer	59,728	48,716	82%	Operating hours for one measure are lower than reported; installation rate for one measure is lower than reported; incorrect savings factor for controls was reported
Religious Assembly	5,257	6,275	119%	Operating hours are longer than reported
Bank Branch	82,974	108,314	131%	Reported kWh didn't account for emergency lighting hours

3.4 TLED interview results

The program provides incentives for the three types of TLEDs.[3] As part of the impact evaluation effort, DNV GL addressed numerous TLED research questions related to four broad topics:

- General measure performance
- TLED removal and replacement after the initial installation

[3] KNOW BEFORE YOU BUY: TLED BASICS, https://www.energytrust.org/wp-content/uploads/2017/05/EB_FS_TLED.pdf

- Controls
- Participant decision-making regarding program participation

Table 3-6 provides a description of the different TLED installation types and some of the key differences between them. Table 3-7 links these research topics to specific research questions. The full interview guide is available in Appendix B.

Table 3-6: TLED types

Type	Description	Key Differences
Type A	Uses either an existing or a new fluorescent-lamp ballast	Lowest installation cost of the three but is susceptible to ballast failure.
Type B	Requires that the fixture be modified to connect the TLED directly to 120/277V sources	No fluorescent-lamp ballast results in line voltage (120-227V) across the lamp. Although this type avoids issues with ballasts, maintenance staff need to be careful to replace with Type B TLEDs and not fluorescent tubes.
Type C	Requires that the existing fluorescent ballast be replaced with a low-voltage driver in the fixture to supply power to the TLED	Installation costs are highest for this type, but lamp efficiency is higher and ballast failure is not a concern.

Table 3-7. TLED research topics and survey questions

Topic	Research Question
General Performance	Are the TLEDs installed through the program in 2017 operating well? Are participants satisfied with their performance?
	Are there any differences in satisfaction by TLED installation type (Type A, B, or C), space use or customer type?
	Are participants experiencing any issues with TLEDs, like buzzing, flickering, early failure, etc.?
	For participants who changed the quantity and/or placement of lighting fixtures, are they satisfied with their overall lighting system design?
Removal and Replacement	Have any participants removed any of the incentivized TLEDs installed in 2017?
	If a participant removed any incentivized TLEDs, when did they remove them?
	If a participant removed any TLEDs, why did they remove them?
	If a participant removed any TLEDs, with what type of lighting did they replace them?
Controls	What control strategies are being used with TLEDs?
	For participants who installed TLED products and controls,[4] were the controls installed before, at the same time as, or after the TLEDs were installed?
	For participants who installed TLED products and controls, are they satisfied with how the controls are operating, or are there issues to be addressed?
	Do participants intend to install new or additional controls for TLEDs within the next 12 months? If so, what type(s)?

[4] In this context, "controls" means any control scheme other than a simple on/off switch.

Topic	Research Question
Participant Decision-Making[5]	Was the interview participant involved in the decision to install the incentivized TLEDs?
	Did the participant consider other TLED types in addition to the incentivized type(s)?
	What motivated the participant to install the specific type(s) of incentivized TLEDs (A, B, C)?

We addressed these research questions using telephone surveys with 23 program participants who received incentives for TLED measures. The subsections that follow provide detailed survey results. Table 3-8 shows the number of survey participants for each major TLED type along with the program participation and reported savings. Of the 23 survey participants, only 1 changed the quantity and/or placement of lighting fixtures as part of their TLED retrofits, so the majority of TLED survey responses represent participants who replaced fixtures and/or lamps but did not redesign their lighting systems.

Table 3-8. TLED survey participation by type

TLED Type	Number of Respondents	Number of Projects	Reported Savings (kWh)
Type A (new or existing fluorescent ballast)	17	770	16,868,529
Type B (line voltage, no ballast)	5	572	6,339,070
Type C (remote driver)	1	72	451,733
Total	**23**	**1,414**	**23,659,332**

3.4.1 TLED measure performance

General satisfaction. Participants rated their general satisfaction with TLED measures on a scale of 1 to 5 where 1 means "not at all satisfied" and 5 means "very satisfied." Of 23 participants, 22 reported that they were "satisfied" or "very satisfied" with their TLED measures. One respondent, with Type A TLEDs, provided a rating of 3 (moderate satisfaction). This respondent said he was happy with the output of the TLEDs but hadn't noticed a change in his electrical consumption. There were no noteworthy patterns in satisfaction based on TLED type, space usage, or customer type—not surprising given the high satisfaction overall.

Equipment performance. When asked directly whether they had experienced any performance issues with their TLEDs (such as buzzing, flickering, early failure, and so on), only one of the participants we surveyed reported an issue. One of his Type A TLED fixtures had "stopped working" but he had not yet contacted his contractor for repair. There were no other issues, complaints, or concerns about TLEDs from any respondent.

Type B TLEDs. Of the 23 TLED survey participants, 5 received incentives for Type B TLEDs. Type B TLEDs require installers to remove ballasts and directly wire sockets to line voltage. Three of the Type B participants recalled specifically being told by the installers to make sure they replaced these bulbs with the same type on failure.

[5] Questions regarding decision-making were asked in instances in which the interview respondent was the decision-maker—i.e., we did not attempt to find a separate contact to address these questions.

3.4.2 TLED removal and replacement

Interviewers asked TLED survey participants whether they removed any of the incentivized TLED fixtures or lamps since participating in the program. Of the 23 survey respondents who answered the question, none had removed any incentivized TLED fixtures or bulbs. These results underscore the high levels of satisfaction with program-incentivized TLED measures among participants.

3.4.3 TLED controls

Interviewers read TLED participants a list of control types and asked which types they used. Of the 23 TLED survey participants who answered the question, all but one reported that they controlled their TLEDs with on/off switches.

- A single user reported using occupancy sensors to control his TLEDs. On a scale of 1 to 5, he rated his satisfaction with the control scheme a 2 because "They don't activate all the time and they turn off too soon." He had not discussed the issue with his contractor.

Installation plans. We asked TLED survey participants whether they plan to install any new or additional controls for their TLEDs within the next 12 months. Of the 23 who answered the question, none planned to install additional controls.

3.4.4 TLED participant decision-making

Five TLED survey participants reported that they were involved in the decision to install the TLEDs for which they received incentives through the Existing Buildings program (n=23). We asked these participants:

- Whether they considered other TLED types in addition to the type(s) they installed
- What motivated them to install the specific type(s) they installed (A, B, C)

None of these five participants reported that they considered other TLED types in addition to the type or types they installed.

We asked these five participants what had motivated them to select the type or types of TLEDs they selected. All five cited "contractor recommendation" as the top reason. One also cited price as a motivation.

3.5 Lighting track findings and recommendations

Our evaluation findings and recommendations specific to the lighting track are presented in this section.

3.5.1 Track recommendations

- **Finding** – In the case of four standard lighting projects and one direct install project, the estimated kWh savings differed by at least 10% from the reported savings. Two projects had lower evaluated savings than reported savings; two had higher.
  - **Recommendation** – Program staff should continue to emphasize the importance of accurate estimates of operating hours during training for trade allies. 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 estimates of savings.
- **Finding** – 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 overestimation of the societal value delivered by the program. The conclusions of the previous evaluation report[6] are supported by DNV GL.

[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

- **Recommendation** – Energy Trust should consider including estimates of interactive effects when calculating the societal value of this program. Energy Trust should consider changes to its savings calculation workbook but weigh the changes against the added workbook complexity required. Future impact evaluations should continue to estimate the impact of lighting projects on all building systems.

3.5.2 TLED specific recommendations

- **Finding** – General satisfaction with TLEDs is high and performance issues are minimal. The twenty-three survey participants gave their TLEDs an average rating of 4.7 out of 5, with all but one giving either a 4 or 5. One participant gave a 3 rating, saying his new TLED fixtures were bright enough but he wasn't sure he was saving money on his bill. Only one customer indicated that he'd had a problem with any of his TLEDs (a fixture stopped working), and he had not yet contacted his contractor to resolve it. No respondents had removed any lamps or fixtures since the retrofit, another indication of high satisfaction with lighting system performance. These findings are consistent with the findings for TLEDs in the 2015-2016 Existing Buildings impact evaluation.
  - **Recommendation** – Continue supporting the installation of TLEDs. No systematic concerns were identified.
- **Finding** – "Recommendation from the contractor" was the leading reason participants selected the TLED type they installed.
  - **Recommendation** – Since program attribution is high, we have no recommendation to change this program.

4 STANDARD TRACK NON-LIGHTING EVALUATION

This section documents DNV GL's impact evaluation of non-lighting savings acquired 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 15% of the 2017 Existing Buildings program's reported electricity savings and 48% 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 2017

Track	Electricity (kWh)	Gas (Therms)
Standard Non-Lighting	**20,127,512**	**900,864**
Existing Buildings program total	131,121,760	1,890,395
Percent of Existing Buildings program savings	15%	48%

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. DNV GL sampled at the measure level, using unique Measure IDs in the data. The sample design target included 68 points spread across five sub-categories. Additional strata used include primary fuel type and size of savings claim (up to three size strata were used). This design was expected to provide program year savings estimates with 18% relative precision at the 90% confidence interval. Further detail on sample design is available in Appendix A.

Table 4-2: Standard track sample design

Technology	Fuel	Size Stratum	Population (N)	Sample Target (n)
Powerstrips	Electric	1	120	6
Refrigeration	Electric	1	130	6
		2	37	6
Others	Electric	1	346	6
		2	67	5
		3	26	5
	Gas	1	162	3
		2	61	3
		3	17	2
Boiler	Gas	1	24	5
		2	9	5
Gas Fryers	Gas	1	164	6
		2	132	5
		3	79	5
EVALUATION TOTAL			**1,374**	**68**
Percent of Reported kWh in sample				6%
Percent of Reported therms in sample				12%

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:

- **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.
- **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.
- **Data collection**: Interviewed sampled participants by telephone using the survey instruments developed.
- **Analysis**: Estimated evaluated savings using the data collected to update key parameters and/or map to the most correct MAD value.

The two approaches created to complete the evaluation were:

- ***Measure-specific:*** DNV GL used a more systematic and standardized *measure-specific* approach for measure types[7] 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.
- ***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*. For each of these measures, a single Word document 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 in the standard track, notes which were sampled and not sampled in this evaluation, and notes the evaluation approach type implemented.

[7] Measure type is based on a specific field called evaluationcode in the program databased that is used to identify similar measures.

Table 4-3: Standard track sample design by measure type

Measure Type	% of Track Savings		Measure ID Count		Evaluation Approach
	%kWh	%therms	Population	Sampled	
Food Equipment	5.4%	41%	610	23	Measure Type Approach
Refrigeration	69.5%	18%	295	18	
Boiler	0.0%	22%	46	14	
Powerstrip	1.8%	0%	228	8	
Controls	6.1%	0%	37	5	
Ceiling Insulation	1.9%	7%	42	4	Project Specific Approach
Icemaker	0.4%	0%	97	2	
Motor	1.4%	0%	24	1	
Wall Insulation	0.1%	0%	10	1	
AC Mini Server	0.0%	0%	1	0	Not Sampled
Block Heater	0.3%	0%	3	0	
Clothes Washer	0.0%	0%	1	0	
Compressed Air	0.0%	0%	2	0	
Custom Refrigeration	3.4%	0%	16	0	
Custom Welder	0.0%	0%	1	0	
Economizer	0.0%	0%	1	0	
Gas Furnace	0.0%	1%	40	0	
Gas Unit Heater	0.0%	0%	9	0	
Heat Pump	3.2%	0%	42	0	
HVAC	0.2%	0%	5	0	
Lighting Controls	7.9%	0%	1	0	
Pipe Insulation	0.0%	5%	9	0	
Radiant Heat	0.0%	1%	5	0	
Showerhead	0.0%	0%	1	0	
Steam Trap	0.0%	4%	7	0	
Tank Water Heater	0.0%	1%	42	0	
Tankless Water Heater	0.0%	0%	1	0	

4.2.2 Documentation and file review

DNV GL reviewed the applicable MAD as well as site-specific project file documentation for all of 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 one additional MAD (vent hoods); all other relevant MADs were received and reviewed in the 2015-16 program year evaluation. As we noted in the 2015-16 evaluation, supporting documentation was referred to within the MAD but was only available from links to a drive internal to Energy

Trust. 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 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.

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?	Project P00001240733 had project P00001240724's 120P form. There were no other issues.
Folder Contains Signed 140P Form?	There were no issues.
Final Project Claimed Savings Match Total Value in Project Folder?	One gas fryer project (P00001273225) appears to claim savings for only one of the two installed fryers. There were no other issues.
Building Type Specified?	Building types for all projects were specified.
Models / Calculations in Folder?	As was the case with the 2015-16 evaluation, there were few savings calculators in the documentation, but that was not a problem for prescriptive measures in this evaluation.
Enough data to recreate savings?	All sixty-eight 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, when DNV GL was already on site for a measure sampled in a different track, data was collected in person for standard track measures. 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 C.

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 72% of the measures sampled. 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

Technology	Fuel	Size Stratum	Sample Target (n)	Achieved Sample	% Complete
Powerstrips	Electric	1	6	6	100%
Refrigeration	Electric	1	6	6	100%
		2	6	2	33%
Others	Electric	1	6	4	67%
		2	5	2	40%
		3	5	1	20%
	Gas	1	3	3	100%
		2	3	2	67%
		3	2	2	100%
Boiler	Gas	1	5	4	80%
		2	5	4	80%
Gas Fryers	Gas	1	6	5	83%
		2	5	4	80%
		3	5	4	80%
Grand Total			**68**	**49**	72%

4.3.2 Evaluated savings

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

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

Technology	Fuel	Projects Evaluated	Realization Rate
Powerstrip	Electric	6	55%
Refrigeration	Electric	8	84%
Others	Electric	8	98%
Others	Gas	7	54%
Boiler	Gas	8	80%
Gas Fryers	Gas	13	53%
Refrigeration	Gas	2	368%

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
Standard Non-Lighting	22	88%	0.06	11%

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

Track	Completed Sample	Realization Rate	Standard Error	Relative Precision at 90 % Confidence
Standard Non-Lighting	30	105%	0.40	64%

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 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 2017 evaluation there was only one MAD we had not already reviewed, the 2014 MAD for variable-speed drives on vent hoods. As with the MADs we reviewed for the 2015-16 program years, we found that the MAD does not provide sufficient transparency and traceability to support reliable savings estimates.

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. Below are our thoughts on what each MAD should include to ensure sufficient information for reliable savings estimation.

- The 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 method and/or 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 assumption varies 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 imbedded Excel file.
- The MAD should clearly specify the baseline condition for the measure, either pre-existing conditions (retrofit measures) or current 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 interval and relative precision of the input assumption or savings estimation used. These values provide a clear indication of savings reliability.

Below are additional findings based on this evaluation:

- **Finding** – There is no direct link to the MAD file used for each measure in the tracking data, nor to the measure names used in the MAD file. Identifying the correct MAD file was a multi-step effort: first finding the files with a similar measure name, then calculating the unit energy savings (UES) values from the tracking data from the total kWh or therms and the quantity, then finding the best match in the MAD doc to the tracking data measure name and the calculated UES value.
  - **Recommendation A** – Create consistent and traceable file names. The MAD file and tracking data should use the same measure names and/or use unique measure codes so that there is direct traceability to the source of the savings approach and UES values for every measure in the tracking data. If new measure codes are created, then the MAD should be updated.
  - **Recommendation B** – Create a way to directly identify the applicable MAD file used for each measure. One possible solution is to create a mapping table of Measure Code to MAD file name. Another is to insert and populate a MAD file name field in the tracking data.
  - **Recommendation C** – Include program years and programs that the savings documented in the file are approved to be used in. Include a measure history table similar to the table in policy documents that shows the last revision date, revisions made, and date of next review. Do not update the measure savings within a program year unless necessary. These changes should improve the transparency of the MAD update process and traceability of savings claims to supporting documentation.

- **Recommendation D** – Include the UES value, the units according to the unit basis for the measure, and the physical count of units installed in the tracking data. Future impact evaluation should research the accuracy of these values in the tracking data.

- **Finding** – The supplemental or supporting references for assumed values and approaches were not properly referenced (title, author, date) and only a link to an internal server location for the studies was provided.
  - **Recommendation** – For supplemental or supporting documents, studies, reports or calculators, consider embedding those files in the Word doc, attaching to a created PDF file, and/or providing a proper and complete reference so that an external or web search can easily find the reference. Reference sources should ideally be publicly and readily available, and the savings values and methods used in the MAD files should be traceable and transparent.
- **Finding** – The MAD files use a variety of structures and formats, including some that appear to be long narratives from emails. The variety of formats used makes it difficult to ensure that the information for a measure is completely and consistently documented across measures.
  - **Recommendation** – Create a template with the structure that can encompass all the measures and all the information needed to characterize a prescriptive measure, then phase that template in as measures are updated (targeting high-impact measures) or complete a separate project to update all MAD files.
- **Finding** – *ENERGY STAR measures.* The ENERGY STAR calculators are readily available, relatively easy to use and defaults are very transparent, citable, and version controlled. For many of the measures, Energy Trust uses the basic ENERGY STAR algorithms, but makes adjustments to some of the parameters by referencing sources that could not always be validated.
  - **Recommendation A** – Consider using the ENERGY STAR calculator directly and as part of the MAD documentation. Furthermore, consider using the ENERGY STAR calculator for every participant, but use realistic operating hours to reflect each business. One alternative to consider is the creation of a measure for non-restaurant business types (for example, caterers) that assumes lower usage. If created, the measure should be tested for cost-effectiveness.
  - **Recommendation B –** Conduct research on current practice to validate the baseline assumptions for incentivized equipment in Oregon.
- **Finding –** The savings methodologies and assumptions are not easily available to the public and should be. A public reference manual, measure database, or work paper library improves the transparency of program assumptions, methods, and savings estimations.
  - **Recommendation –** Energy Trust should develop and implement a plan to transition 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.

4.4.2 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** – *Smart Power Strips.* We looked only at smart power strips purchased by participants. We did not look at leave-behind smart power strips for this program year. The sample included 6 measure claims spread across 2 buyers, both of which are public schools. We found that one purchaser had deployed their entire stock for use on computers and peripherals, while the other purchaser was

deploying them as teachers and staff request them. At the time of the interview, this purchaser said he had ordered the maximum incentivized quantity and had deployed "about a quarter" of them.

- **Finding** – *Refrigerated case LED lighting.* The evaluation team was unable to contact any participants with knowledge of this measure. The contact data for this measure, and for grocery chains generally, included a central contact who was typically involved in either purchasing the equipment or submitting the incentive request but who had no knowledge of installation specifics and often did not know who best to reach on site to answer our questions.
  - **Recommendation –** Project folders should include an on-site technical contact for each installation site who can respond to questions about equipment operation.
- **Finding** – *Cooler case measures.* Interactive HVAC savings are claimed for gas heating systems but not for electric heating systems, and the reported heating system was incorrect for one of the verified sites.
  - **Recommendation** - Improve identification of the space heating system type, adjust the electric savings for electric space heating systems, or consider dropping the interactive gas savings and only claiming the direct kWh savings. DNV GL believes that Energy Trust should consider the full benefits and costs for all fuel types of the measure but recognizes that tracking interactive effects adds complexity. At a minimum, interactive effects should be accounted for consistently across the Existing Buildings program.
- **Finding** – *Gas Fryers.* There is an ENERGY STAR calculator for this measure. Energy Trust did not use the calculator directly; it used the calculation approach, some of the defaults, and some revised parameters. Phone response values for verified pounds of food and hours of fryer operation tended to be lower than assumed in Energy Trust's calculations. This measure is aimed at restaurants, but two sites – a catering service and a fraternal organization – had significantly shorter operating periods than the measure assumes.
  - **Recommendation** – As mentioned above, Energy Trust should consider the creation of measures for non-restaurant business types that use cooking equipment. If created, the measure should be tested for cost-effectiveness. Use the ENERGY STAR calculator for traceability and transparency. Review and consider revising the assumed pounds of food and fryer hours per day; using participant estimates would be best. Consider stopping incentives to fast-food chains or other 7/24/365 sites where ENERGY STAR equipment is already common practice.
- **Finding** – *Space Heating Boilers.* We found that a number of sites had boilers providing functions other than space heating. The MAD assumes that boilers provide space heating only.
  - **Recommendation** – Consider identifying and developing savings estimates for non-space-heating uses.
- **Finding** – *Space Heating Boilers.* We found some boilers operating in conditions that made it unlikely that they typically operate in condensing mode. This reduces the operational efficiency and reduces savings.
  - **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.
- **Finding** – *Space Heating Boilers.* The evaluation team found several sites with multiple boilers operating in lead/lag 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

we believe it 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.

- **Recommendation A** – Adapt MAD savings to account for sites with mutiple 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.

- **Finding** - *Roof/Attic Insulation.* The low electric realization rate is primarily due to evaluated savings of zero at two projects with gas heating systems that were identified as remodel/rehab by the participant, and as such were considered alterations under the energy code and subject to required minimum R-values for roof insulation without incentives.
  - **Recommendation** – Ensure that projects are true retrofits and not complete remodel-rehabs of the space or building that are subject to compliance with energy codes and minimum insulation R-value requirements. If remodel-rehabs are allowed, then only incentivize insulation above the code minimum - but savings and cost-effectiveness will diminish rapidly, especially for gas heating.

4.4.3 Future evaluation recommendations

- **Finding:** The team often had 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; 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 them to introduce the customer to the evaluation and establish the validity and importance of the evaluation.

[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

5 CUSTOM TRACK EVALUATION

The 2017 custom track reported 218 unique measures providing 24,452,156 kWh and 780,488 therms in annual energy savings. These savings account for 20% of the program's reported electricity savings and 46% 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 by delivery

Track	Electricity (kWh)	Gas (Therms)
Custom	**24,452,156**	**780,488**
Existing Buildings program total	125,107,079	1,681,352
Percent of Existing Buildings program savings	20%	46%

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 interval. 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	73	7
	2	23	7
	3	12	6
	Certainty	4	4
Gas	1	43	5
	2	13	4
	3	8	4
	Certainty	1	1
EVALUATION TOTAL		**177**	**38**
Percent of Reported kWh in sample			44%
Percent of Reported therms in sample			39%

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**: Sites 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 made onsite verification at 15 project sites and conducted interviews with the facility personnel for 23 project sites. One site was dropped because of no response received from the facility contact.

5.2.5 Project analysis

DNV GL estimated evaluated savings for 36 of the 38 projects originally sampled. DNV GL used two analysis methods: 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 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. DNV GL used this methodology to determine the evaluated savings for three projects:

1. Large Office, Direct Digital Controls (DDC) installation: The regression analysis confirmed that the facility is saving energy compared to consumption prior to the project. However, the monthly savings rate does not support the savings originally estimated. DNV GL engineers had identified issues with the original analysis methodology that introduced additional uncertainty. DNV GL decided to use the regression analysis results to estimate final evaluated savings for the site.
2. Small Office, Advanced RTU Controls: The regression analysis confirmed that the facility is saving energy compared to consumption prior to the project. However, the monthly savings rate does not support the savings originally estimated. DNV GL concluded that the regression analysis is the best method for estimating savings at this site and used the results of the analysis as the evaluated savings.
3. Municipal Building, DDC installation: The regression analysis did not support the achievement of energy savings at this facility. DNV GL engineers had identified calculation errors in the original analysis and believed that significant uncertainty existed regarding the baseline modeling assumptions. DNV GL concluded that there is not sufficient evidence of energy savings at the facility. The evaluated savings result for this facility is zero savings.

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 95% of the original sample target. The final achieved evaluation sample differed from the sample design due to the following:

- **Dropped site:** The one large gas certainty site was dropped from the sample and population frame after the initial interview. DNV GL learned that the site had experienced issues with the performance of the installed equipment and had recently stopped using it until a solution was found. Additionally, the equipment was part of a larger project with similar measures claimed in 2018. Energy Trust decided to evaluate the 2017 project and the related 2018 measures seperately once the site implemented a solution and true-up program year savings accordingly.
- **Non-Responsive site:** One site 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.

Table 5-3: Final custom track sample summary

Primary Fuel	Sample Target	Achieved Sample	% Complete
Electric	24	23	96%
Gas	14	13	93%
Combined	**38**	**36**	**95%**

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 for the track. 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. 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 the project was originally assigned to.

Table 5-4: Custom track electric impact evaluation results

Track	Completed Sample	Realization Rate	Standard Error	Relative Precision at 90 % Confidence
Custom-2017	31	90%	0.051	9%

Table 5-5 shows the overall gas standard track realization rate for the track. The gas realization rate was determined after the dropped certainty project was removed from the population frame. 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. DNV GL captures our findings and recommendations in the sections that follow.

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

Track	Completed Sample	Realization Rate	Standard Error	Relative Precision at 90 % Confidence
Custom-2017	18	87%	0.085	16%

5.3.3 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 aspects of the models used in the energy savings analyses by the PMC or program allies that may be of concern to Energy Trust?

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

- **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 estimating using the models provided.
  - **Recommendation** – Energy Trust should require the PMC to 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.
- **Finding** – Program models continue to estimate savings that suggest a significant reduction in annual consumption. In some cases, the savings were found to exist. In other cases, the savings did not materialize.
  - **Recommendation** – Energy Trust should complete additional review of simulation inputs for sites expecting savings greater than 20% of consumption.

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 calculation errors were identified during this project.

- One project contained errors specific to fan energy calculations.
- One project incorrectly modeled the type of economizers on site.

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 both 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 setpoints, occupied/un-occupied heating setpoints, economizer high limit setpoint, chilled water and heating 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.

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:

- *Increase documentation of changes to building controls*: For many control upgrade or modification measures, the evaluation team found little or no information available to support the measure changes and the inputs used in simulation modeling. DNV GL recommends that pre- and post-project control setpoints be documented either through facility EMS screenshots, plots of EMS trends, or a text narrative.

6 STRATEGIC ENERGY MANAGEMENT EVALUATION

The SEM track reported 166 unique measures at 164 sites providing 6,014,681 kWh and 209,043 therms in annual energy savings in program year 2017. These savings account for 5% of the program's reported electricity savings and 11% of the program's reported gas savings. Table 6-1 shows the reported savings for SEM in program year 2017.

Table 6-1: Reported SEM track energy savings by delivery

Track	Electricity (kWh)	Gas (Therms)
SEM Cohort	**6,014,681**	**209,043**
Existing Buildings program total	131,121,760	1,890,395
Percent of Existing Buildings program savings	5%	11%

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 across program years. 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 interval. Further detail on sample design is available in Appendix A.

Table 6-2: SEM track sample design

Fuel	Size Stratum	Population (N)	Sample Target (n)
Electric	1	57	7
	2	19	6
	3	11	6
	Certainty	5	5
Gas	1	45	7
	2	14	7
	3	9	6
	Certainty	4	4
EVALUATION TOTAL		**164**	**48**
Percent of Reported kWh in sample			50%
Percent of Reported therms in sample			37%

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 the data collected to update key modeling parameters.

6.2.2 Documentation review

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

- Verification of the existence of signed application or end-user 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.

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 the program
- Identifying if the site was previously evaluated or reviewed
- Identifying use of non-weather variables, polynomials, or multiple degree-day variables
- o Determining if the energy savings reported in the database is supported by a single or multiple regression models
- Extraction of the monthly utility consumption data, and non-weather independent variables (if used) for each facility and identify 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 extract 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 completed on-site or via telephone. The evaluation team completed 8 interviews onsite discussing 18 unique sites and performed phone interviews with 11 facility personnel regarding 18 unique sites. Twelve sites were dropped because of no response received from the facility contact, the facilities did not have a suitable person who could answer the evaluator's queries, or the facility refused to participate in the evaluation interview. 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 MT&R 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 36 sites. The data collected through the interviews was be 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 75% 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 refusal 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

Primary Fuel	Sample Target	Achieved Sample	% Complete
Electric	24	15	63%
Gas	24	21	88%
Overall	**48**	**36**	**75%**

6.3.2 Evaluated savings

Expansion from the sample to track-level results follows the methodology discussed in Section 2.4. Table 6-4 shows the overall electric SEM track realization rate. The electric realization rate is driven primarily by the change to baseline models that comply with the modeling guidance document and by zeroing out savings for sites which exhibit a lack of any substantial program engagement. The site-specific electric realization rates varied from 0% to 154% for PY2017.

Table 6-4: SEM track electric impact evaluation results

Track	Completed Sample	Gross Realization Rate	Standard Error	Relative Precision at 90 % Confidence
SEM Cohort-2017	15	92%	0.06	11%

Table 6-5 shows the overall gas SEM track realization rate. The gas realization rate is driven primarily by nine sites found to have achieved 20% or less of the reported savings for the program year. The significant savings variance is due to the differences in baseline models and how the different model form changes the incremental savings estimated for each year of participation. In most cases, these sites are achieving cumulative savings over the baseline, but no incremental savings were achieved in PY2017. The site-specific gas realization rates varied from 0% to 202% for PY2017.

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

Tracks	Completed Sample	Gross Realization Rate	Standard Error	Relative Precision at 90 % Confidence
SEM Cohort-2017	21	66%	0.11	22%

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 were 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.

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 existed in the PY2017 evaluation sample as many of the evaluated sites used models developed before the guidelines were provided. For models developed after the guidelines were developed, the evaluation did not identify a satisfactory explanation for deviations.

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

Evaluation Response: There are significant differences between the baseline evaluation and original models. All sampled participants were continuation participants and most models were created before the modeling guidelines were developed. The majority of sites evaluated used average temperature as an independent variable instead of degree-days in the original model. Many sites also used 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.

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.

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

Evaluation Response: The evaluation found capital projects to be properly accounted for.

6.4.2 Other SEM findings and recommendations

Our evaluation findings and recommendations specific to the SEM track are presented in this section.

- **Finding –** The site-specific realization rate for eight gas sites is below 20%. Six of these sites achieved a site realization rate of 0%. These results are the primary driver of the 66% gas realization rate for this track. These sites did not have capital project adjustments and only one has a baseline/other adjustment. In most cases, these sites are achieving cumulative savings over the baseline, but no incremental savings were achieved in PY2017. DNV GL believes cases like this will continue to exist until all sites have baseline models meeting the current guidelines.
  - **Recommendation** – DNV GL recommends that Energy Trust continue its efforts to re-baseline continuation participants with average mean temperature baselines. Reducing differences in the baseline modelling approach will reduce this variance in continuation participants.
  - **Recommendation** – DNV GL also recommends that Energy Trust consider not claiming continuation savings that are a small percent (less than 2%) of total consumption for participants in their 3rd year or later if the baseline model does not meet the current guidelines. Based on this evaluation, the degree-day baseline modelling approach is more likely to not support the savings claim than to support the claim.
- **Finding** – Participants continue to value energy coaches and peer-to-peer learning. Participants cite benefits from the insights provided by working closely with energy coaches to identify and execute operational and capital improvement opportunities. Participants also commented on perceiving value from the peer-to-peer information exchanges with participants of a similar facility type. These learning exchanges provide participants with practical 'case study' examples to draw upon, as well as benchmarking and competitive motivation across organizations with similar facilities.
  - **Recommendation** – DNV GL recommends that Energy Trust continue to identify program improvements that allow energy coaches to spend more time working with participant staff to support energy conservation opportunities.
- **Finding** – In recent years, participants were transitioned from the smaller cohorts used for each engagement cycle to one "mega" cohort that includes all participants from the early cohorts. Participant

reaction to the "mega" cohort are mixed. Some participants feel that the larger cohort has provided signifciant learning opportunities. Other participants feel that the meetings started to feel less relevant to them and therefore they were less motivated to attend, participate or pay attention. DNV GL found some evidence that the largest institutional participants find the least value in the mega cohort, but provide the most value to participants with smaller building footprints.

  - **Recommendation** – Based on this finding, DNV GL recommends that Energy Trust rely on its Energy Coaches to identify participants that are not finding sufficient value in the cohort workshops to engage and then assess if an alternative engagement approach better fits those organizations or would increase the perceived value of the mega cohort. An opportunity may exist to recognize large organizations for the value of the information they provide to smaller organizations.

- **Finding** – DNV GL found site savings continued to be calculated even when the participant was disengaged with the program. In one PY2017 case, the energy coach updated the MTR calculations at the end of the program year instead of the participant and neglected to include the non-weather variable inputs. If the non-weather inputs had been included, no savings would have been claimed.
  - **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.
- **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 found that many participants were discussing the establishment of new baselines for their facilities with Energy Trust. The evaluation previously recommended updating baselines and this year's evaluation confirmed that actions were taken to address the recommendation.
  - **Recommendation** - DNV GL continues to recommend that Energy Trust work with participants to regularly update baselines and establish a clear understanding of the baseline update process at the start of engagement with new program participants.
- **Finding** – DNV GL found increased consistency in measurement periods for PY2017. This finding demonstrates that the program has taken actions to address previous evaluation recommendations.
  - **Recommendation** – None.

APPENDICES

APPENDIX A. EVALUATION SAMPLE DESIGN MEMO

Memo to:
Sarah Castor, Energy Trust of Oregon

Copied to:
Jennifer Canseco, DNV GL

Memo No.: 002
From: Andrew Wood, DNV GL
Date: 05/22/18
Prep. by: Andrew Wood, DNV GL
Santosh Lamichhane, 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, program year 2017.

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 PY2017
- Develop separate gas and electric realization rates for PY2017 to be used for program true-up.
- Develop separate gas and electric realization rates for PY2017 SEM savings
- Develop separate gas and electric realization rates for PY2017 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

		2017 Draft Sample				Proposal
Measure Type	**Fuel**	% of Reported Savings in Draft Sample	N Frame	n Sample	Relative Precision (@ 90% CI)	Relative Precision (@ 90% CI)
Capital (Non-SEM)	Electric	17%	3,319	116	12%	11%
	Gas	32%	766	57	13%	14%
SEM	Electric	64%	117	36	17%	20%
	Gas	66%	766	28	19%	19%
ALL (Capital + SEM)	Electric	19%	3,436	152	11%	10%
	Gas	36%	848	85	12%	12%

Sample frame

Energy Trust provided DNV GL with the file "Measures2017.xlsx" which shows energy efficiency measures completed during PY2017 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: Sample frame summary by fuel, PY2017

Program Track	Unique Measure Lines	Working kWh	% of kWh Grand Total	Working therms	% of therms Grand Total
Lighting	5,402	69,293,686	53%		
Direct Install	1,261	4,724,333	4%		
Street Lighting	12	6,509,392	5%		
Standard	1678	20,127,512	15%	900,864	48%
Custom	218	24,452,156	19%	780,488	41%
Capital Total	**8,571**	**125,107,079**	**95%**	**1,681,352**	**89%**
SEM Cohort	166	6,014,681	5%	209,043	11%
Grand Total	**8,737**	**131,121,760**	**100%**	**1,890,395**	**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. This includes Standard Lighting, Direct Install, and Street Lighting
- Standard Non-Lighting – The sampling unit is the Project ID
- Custom – 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 – A unique sampling domain was created for this sub-program to ensure sufficient sample allocation.

- Purchased Power Strips - A unique sampling domain was created for this sub-program to ensure sufficient sample allocation. Measures were identified by productcode = "OCCPLUGSTRIP".
- Standard Refrigeration – A unique sampling domain was created to study this high impact category. Measures were identified by evaluationcode = "FRIDGE" or "CUSTOMFRIDGE".
- Standard Boilers – A unique sampling domain was created to study this high impact category. Measures were identified by evaluationcode = "BOILER".
- Standard Gas Fryers - A unique sampling domain was created to prevent oversampling this measure. Measures were identified by productcode = "GASFRY".

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. So the sampling variance associated with estimates created from this stratum is zero (since a census is being taken). A certainty stratum is suggested for this study. For this study, the sample will be selected independently within domains defined by the following:

- **Program Year**: 2017. If any comparison to previous results is requested.
- **Program Track**: Custom, Lighting, Standard Non-Lighting, and SEM.
- **Track Sub-Category**: Additional categorization was used within tracks.
    - Lighting: Direct Install, Standard Lighting, Street Lighting-City 1, Street Lighting-Non-City 1.
    - Standard Non-Lighting: Purchased Powerstrips, Refrigeration, Gas Fryers, Boilers, Other
- **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*: 15 projects were selected at certainty.

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)	Sample, Measure Lines	Avg. Working kWh	Avg. Working therms.
Lighting	Direct Install	Electric	1	250	7	15	5,055	
			2	80	7	31	19,298	
			3	38	6	45	50,441	
	Non-City 1	Electric	1	7	1	1	52,626	
			Certainty	1	1	1	2,671,864	
	City 1	Electric	1	4	2	2	867,287	
	Standard Lighting	Electric	1	1,528	8	18	10,677	
			2	350	8	25	63,569	
			3	108	8	97	284,535	
Standard	Powerstrips	Electric	1	120	6	6	2,703	
	Refrigeration	Electric	1	130	6	12	61,529	44
			2	37	6	6	180,554	4,323
	Others	Electric	1	346	6	6	3,357	0
			2	67	5	8	22,358	5
			3	26	5	5	77,722	21
		Gas	1	162	3	3	788	338
			2	61	3	3	5,042	1,025
			3	17	2	2	497	6,526
	Boiler	Gas	1	24	5	6	0	3,189
			2	9	5	8	0	13,639
	Gas Fryers	Gas	1	164	6	6	0	569
			2	132	5	5	0	750
			3	79	5	5	0	1,448
Custom		Electric	1	73	7	7	350,475	55,328
			2	23	7	10	1,494,578	229,929
			3	12	6	6	2,926,691	487,695
			Certainty	4	4	8	5,430,915	1,357,729
		Gas	1	43	5	5	7,212	18,950
			2	13	4	4	286,093	72,153
			3	8	4	5	477,664	261,092
			Certainty	1	1	1	0	134,370
SEM Cohort		Electric	1	57	7	7	86,641	14,003
			2	19	6	6	274,669	50,973
			3	11	6	7	725,620	120,119
			Certainty	5	5	5	2,152,311	430,462
		Gas	1	45	7	7	29,612	2,571
			2	14	7	7	24,467	7,674
			3	9	6	6	132,555	17,090
			Certainty	4	4	5	397,457	99,364

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 Existing Buildings impact evaluation. Table A-4 shows the error ratios assumed.

Table A-4: Assumed error ratios

Track	Sub-Category	ER Assumed	Sample Size (n)
Lighting	Direct Install	0.5	20
	Standard Lighting	0.5	24
	Street Lighting	0.10	4
Standard	Power Strips	0.5	6
	Refrigeration	0.5	12
	Others	0.58	24
	Boiler	0.63	10
	Gas Fryers	0.63	16
Custom		0.58	38
SEM		0.75	48
Total			**202**

Table A-5 summarize 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 assume that DNV GL estimates evaluated savings for 90% of the original project's samples.

Table A-5: Expected precision by track and fuel

Program Track	2017 Electric				2017 Gas			
	% of Savings in Sample	N	n	Relative Precision (@ 90% CI)	% of Savings in Sample	N	n	Relative Precision (@ 90% CI)
Lighting	10%	2,366	48	17%				
Standard	10%	796	36	22%	17%	682	39	18%
Custom	45%	157	32	19%	49%	84	18	20%
Subtotal: Non-SEM	**17%**	**3,319**	**116**	**12%**	**32%**	**766**	**57**	**13%**
SEM	64%	117	36	17%	66%	82	28	19%
Total (All Measures)	**19%**	**3,436**	**152**	**11%**	**36%**	**848**	**85**	**12%**

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 have more than 100 sites across all four tracks. The sample is not expected to be perfectly representative of the building type distribution since stratification is used.

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

	Lighting		Standard Non-Lighting		Custom		SEM	
Building Type	**Program**	**Sample**	**Program**	**Sample**	**Program**	**Sample**	**Program**	**Sample**
Affordable Multifamily Property	1							
Amusement/Recreational	28		14	1	7	1	5	1
Assembly	10		4	1				
Assisted Living Property							1	
Auto Repair	73	6	3		2			
Auto Services	72	1	8		1		1	
Bank/Financial Institution	41	1			1		2	2
Brewery			1					
Cafeteria	1							
Campus Living Property							4	2
Car Dealership/Showroom	24		1	1				
Car Wash	9	1						
College/University	30		5	2	12	3	16	6
Commercial	21	1	16	1	9	5	3	
Convenience Store	88		38	1	1			
Courthouse/Probation Office	2		2		2		1	1
Data Center	1				7	3		
Education	4		9					
Enclosed Mall	1				1			
Fire Protection	22	1	1		4	1	2	
Funeral/Cremation	1							
Gas Station	68		4					
Grocery	52	2	174	17	1			
Gym/Athletic Club	35	1	12	1	7	3	3	2
Health	6		2					
High School	12	1	26	1	2	1	2	1
Hospital	11		2		5	1	7	2
Hospitality			1					
Indoor Agriculture	1							
Industrial							1	
Jail/Reformatory/Penitentiary	8	1	4	1	2	1	5	3
K-12 School	7		21	1	1		1	
Laundry/Dry Cleaner	6		3		2	1		
Library	4						1	1
Lodging/Hotel/Motel	59	2	64	1	7	2		
Medical Laboratory	1						2	2
Medical Office	42		12	1	2		11	3
Meeting/Convention Center/Hall or Community Center	16		14	1	5		1	
Middle School	2		24	1	3		4	1
Military (Armory, etc.)			1					
Museum	8	1	2		1			
Office	244	3	68	2	40	9	36	10
Parking Structure/Garage	17	1			1	1	1	
Police	2							
Primary School	20		77	2	5		7	1
Printing and Related Support Activities	1							

	Lighting		Standard Non-Lighting		Custom		SEM	
Building Type	**Program**	**Sample**	**Program**	**Sample**	**Program**	**Sample**	**Program**	**Sample**
Religious/Spiritual	87	6	22				1	
Repair and Maintenance	1							
Restaurant	123	1	548	23	11	4		
Retail	396	5	37	1	12	1	13	2
Single Family Home			1					
Site Built Home	1		3					
Super Center/Warehouse Club	2							
Transportation Equipment Manufacturing	2							
Transportation Infrastructure (Tunnel, Roadway, Dock, etc.)	504	6						
Unspecified Government/Public Sector	15	1	6		4		22	6
Veterinarian's Office	4		3		1			
Warehousing and Storage	184	6	9		1	1	3	1

APPENDIX B. TLED INTERVIEW GUIDE

Energy Trust of Oregon PY 2017 Tubular LED (TLED) Telephone Interview

Draft: 7/9/2018

Key Research Questions

The interview will capture information on participants' operations, controls, and satisfaction with incentivized TLED products. The table below links the key research questions for this effort to the specific interview questions that address them.

Topic	Research Question	Relevant Interview Question(s)
General Performance	Are the TLEDs installed through the program in 2017 operating well? Are participants satisfied with their performance?	P1, P1A, W1, W1A
	Are there any differences in satisfaction by TLED installation type (at line voltage, with existing or new ballast, etc.), space use or customer type?	N/A[1]
	Are participants experiencing any issues with TLEDs, like buzzing, flickering, early failure, etc.?	P2, P2A, P3, P4, P4A, P4B, P5, P5A, P5B, P6, P6A, P7, P8
	[For participants who changed the quantity and/or placement of lighting fixtures] Are participants satisfied with their overall lighting system design?	P9, P9A, P9B, P10, P10A
Removal and Replacement	Have any participants removed any of the incentivized TLEDs installed in 2017?	R1, R2
	[If participant removed any incentivized TLEDs] When did participants remove the incentivized TLEDs?	R3
	[If participant removed any TLEDs] Why did they remove the incentivized TLEDs?	R4A, R4B
	[If participant removed any TLEDs] With what type of lighting did participants replace the incentivized TLEDs?	R5, R5A
Controls	What control strategies are being used with TLEDs?	C1
	[For participants who installed TLED products and controls[2]] Did these controls receive incentives from the EB program? If not, why not?	C2, C2A
	[For participants who installed TLED products and controls[2]] Were the controls installed before, at the same time as, or after the TLEDs were installed?	C3
	[For participants who installed TLED products and controls[2]] Are participants satisfied with how the controls are operating, or are there issues to be addressed?	C4, C4A
	Do participants intend to install new or additional controls for their TLEDs within the next 12 months? If so, what type(s)?	C5, C5A
Decision-making[3]	Was interview participant involved in decision to install the incentivized TLEDs?	D1
	Did participant consider other TLED types in addition to the incentivized type(s)?	D2, D2A
	What motivated the participant to install the specific type(s) of incentivized TLEDs (A, B, C)?	D3A, D3B, D3C
	Did participant consider LED fixtures in addition to the incentivized type(s)?	D4, D4A

[1] Installation type and customer type are tracking data variables, and the verification interview will determine space usage. DNV GL will compare interview results regarding satisfaction by each of these variables.
[2] Note that "controls" in this context refers to any control type other than an on/off switch.
[3] Note that research questions regarding decision-making are not in the original scope of work for this study, but we added questions on this topic for instances in which the interview respondent is the decision-maker—i.e., we will not attempt to find a separate contact to address these questions.

Database Variable Definitions

Variable Name	Definition[1]
Customer_Name	Respondent contact name
Installeddt	Date on which equipment was installed
Productdescription	Measure name (brief description)
Program	Terms the program the contact is likely to recognize Most likely: "Energy Trust Incentives" Also: "Energy Trust Existing Buildings" For Direct Install: "Small Business Energy Savings" or "SmartWatt". SmartWatt is the Direct Install contractor.
Site Address	Equipment installation address

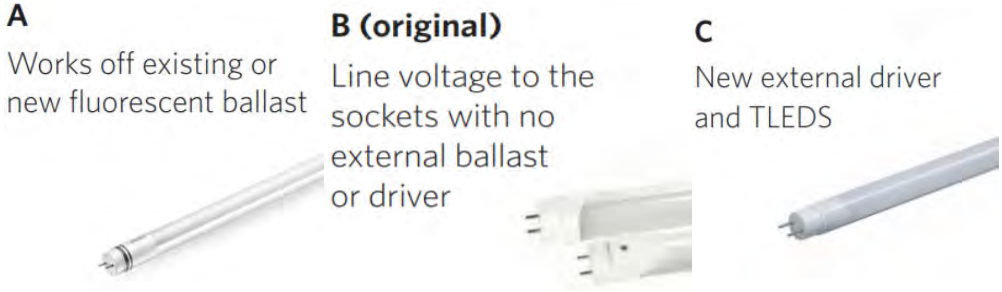
[1] Unless otherwise noted, the database can contain more than one value for each variable for each respondent.

Instructions to Interviewers

- Do not read response options unless instructed to do so in questionnaire (["READ LIST"]). Never read response options for don't know (98) or refused (99).
- If more than one TLED installation type occurred at the site, prioritize the type with higher reported savings in the instructions regarding which measure(s) should be the focus of the interview.
- Responses must be recorded in Excel response file.
- Commence interviewing once you have identified the appropriate respondent. This should be someone familiar with the incentivized equipment and its operation.
- Prioritize your impact questions. Then complete this interview. If respondent asks how much time, estimate 15 extra minutes.
- If asked what the purpose of the interview is, state something like:
    - "Linear or Tubular LEDs are being installed and incentivized more and more. Energy Trust wants to know if participants like you are satisfied with the technology and if you've had any issues, and they want to understand how you are controlling the lights."
- If asked, have others had issues:
    - "I am not aware of issues. Asking you questions today is part of Energy Trust's effort to make sure there aren't issues"

Technology Information for Interviewers

Name	Definition
TLEDS Drivers	Tubular Light Emitting Diodes A "driver" or "remote driver" is used when the fluorescent lamp and ballast are removed (type C lamps)
TLED **Type A** lamp	**Retrofitted to existing fixtures.** These types of tubes replace T12, T8, and T5 lamps and operate using an internal driver that is powered directly from an existing linear fluorescent ballast. Type A installations do not require any modifications to the existing fixture and many people refer to this type of lamp as 'plug-and-play.' Type A tubes have reduced efficiency due to power loss from the existing ballast and limited dimming and control capabilities.
TLED **Type B** lamp	**Wired directly to the line voltage.** These types of tubes operate using an internal driver and are powered directly from the main voltage that is supplying the fixture. The existing fixture must be modified for these tubes which has led to calling this type of lamp 'ballast bypass.' These lamps require installation by a certified electrician and have limited dimming and control capabilities.
TLED **Type C** lamp	**Supplied with dedicated LED drivers.** These types of tubes operate using a remote driver, which replaces the existing fixture's ballast. These types of tubes require modifying the existing fixture, but the power being directed to the sockets are low-voltage, not AC Mains. Type C tubes are more efficient than the other types.
Links for more info:	http://cltc.ucdavis.edu/sites/default/files/files/publication/LED_Retrofit_Options_Linear_Fluorescent_FINAL.pdf https://www.energytrust.org/wp-content/uploads/2017/05/EB_FS_TLED.pdf http://www.lutron.com/TechnicalDocumentLibrary/TLED_Lighting_Scenarios_for_Retro_App_Whitepaper.pdf



Introduction

I'd like to ask you a few questions regarding the TLED lighting equipment that was incentivized through the Existing Buildings program on or around <Installeddt>.

General Performance

P1. On a scale of 1 to 5 where 1 means "not at all satisfied" and 5 means "very satisfied," how satisfied are you with the general performance of the <Measuredesc>?

1	1 – not at all satisfied
2	2
3	3
4	4
5	5 – very satisfied
98	Don't know
99	Refused

P1A. [IF SATISFACTION IS NOT LOW (IF P1 ≠ 1, 2, OR 3), SKIP TO P2] Why do you say that? [IF NECESSARY: What could be improved?]
[RECORD VERBATIM RESPONSE]

P2. Since installing <Measuredesc> have you observed any performance issues such as buzzing, flickering while dimming, or any light quality issues?

1	Yes
2	No
98	Don't know
99	Refused

Article I. **P2A.** [IF NO PERFORMANCE ISSUES (IF P2 ≠ 1), SKIP TO P9] What performance issues have you observed? [ACCEPT MULTIPLE RESPONSES]

1	Buzzing
2	Flickering
3	Poor dimming
4	Incompatibility with dimming controls
5	Poor/low light level
6	Poor light quality
7	Fixture failure
8	Ballast failure
9	Lamp failure (burned out)
10	Failure (unspecified)
97	Other [SPECIFY]
98	Don't know
99	Refused

[IF SATISFACTION IS NOT LOW (IF P1 ≠ 1, 2, OR 3) AND NO PERFORMANCE ISSUES (P2 ≠ 1), SKIP TO P6 for Type B, P9 for Type A and C]

[IF <Measuredesc> ≠ TYPE A FIXTURE, SKIP TO P6 for Type B, P9 for Type A and C]]

Type A Questions Only

P3. Did an electrician or contractor inspect any of the ballasts for any of the incentivized TLED fixtures where your satisfaction was low or you experienced poor performance?

1 Yes
2 No
98 Don't know
99 Refused

P4. Did an electrician or contractor replace any of the ballasts for any of the incentivized TLED fixtures where your satisfaction was low or you experienced poor performance?

1 Yes
2 No
98 Don't know
99 Refused

P4A. [IF ELECTRICIAN OR CONTRACTOR DID NOT REPLACE BALLASTS (P4 ≠ 1) SKIP TO P5] Did replacing the ballast(s) correct the performance issues?

1 Yes
2 No
98 Don't know
99 Refused

P4B. [IF BALLAST REPLACEMENT FIXED THE ISSUE (P4A = 1), SKIP TO P5] What was the problem after the electrician or contractor replaced the ballast(s)?

[RECORD VERBATIM RESPONSE]

P5. Did your electrician or contractor tell you anything regarding what to expect about ballast performance?

1 Yes
2 No
98 Don't know
99 Refused

P5A. [IF ELECTRICIAN OR CONTRACTOR DID NOT SAY ANYTHING ABOUT BALLAST PERFORMANCE (P5 ≠ 1) SKIP TO P5B] What did the electrician or contractor say about ballast performance?

[RECORD VERBATIM RESPONSE]

P5B. [IF P5A RESPONSE MENTIONS EARLY FAILURE, SKIP TO P6] Did your electrician or contractor tell you that the ballasts may fail before the lamps fail or burn out?

1 Yes
2 No
98 Don't know
99 Refused
99 Refused

[IF <Measuredesc> ≠ TYPE B, SKIP TO P9]

Type B Questions Only

P6. Did your electrician or contractor mention anything about safety when he or she installed the incentivized TLEDs? [IF NECESSARY, SPECIFY TYPE B TLED]

1 Yes
2 No
98 Don't know
99 Refused

P6A. [IF ELECTRICIAN OR CONTRACTOR DID NOT MENTION SAFETY (P6 ≠ 1), SKIP TO P7] What did your electrician or contractor mention about safety?

[RECORD VERBATIM RESPONSE]

P7. [IF P6A RESPONSE DOES NOT MENTION HIGH VOLTAGE WIRING] Did your electrician or contractor mention the potential shock hazard that may exist with the type of incentivized TLEDs you installed? [IF NECESSARY, SPECIFY TYPE B TLED]

1 Yes
2 No
98 Don't know
99 Refused

P8. Are there any safety labels on the incentivized TLED fixtures?

1 Yes
2 No
98 Don't know

If respondent is unaware of safety concerns, shock hazard, or does not have safety labels, notify Data Collection Lead and Project Manager after the interview. Inform respondent that we will notify Energy Trust so they can follow up.

All Types Questions

P9. When the incentivized <Measuredesc> were installed, did you change the quantity or placement of the lighting fixtures?

1 Yes, changed quantity of fixtures
2 Yes, changed placement of fixtures
3 Yes, changed quantity and placement of fixtures
4 No
98 Don't know
99 Refused

Article II. **P10.** [IF NO CHANGE TO QUANTITY/PLACEMENT (P9 ≠ 1, 2, or 3), SKIP TO R1] On a scale of 1 to 5 where 1 means "not at all satisfied" and 5 means "very satisfied," how satisfied are you with the overall design of your new lighting system?

Article III. 1 1 – not at all satisfied
2 2
3 3
4 4
5 5 – very satisfied
98 Don't know
99 Refused

P10A. [IF SATISFACTION IS NOT LOW (P10 ≠ 1, 2, or 3), SKIP TO R1] Why do you say that? [IF NECESSARY: What could be improved?]
[RECORD VERBATIM RESPONSE]

[IF <Measuredesc> ≠ TYPE B, SKIP TO R1]

Additional Type B Questions

P11. When your electrician or contractor installed the incentivized TLEDs, did he or she mention anything about how to replace the lamps if they fail? [IF NECESSARY, SPECIFY TYPE B TLED]
1 Yes
2 No
98 Don't know
99 Refused

P11A. [IF ELECTRICIAN OR CONTRACTOR DID NOT MENTION HOW TO REPLACE FAILED LAMPS (P11 ≠ 1), SKIP TO P12] What did your electrician or contractor mention about how to replace the lamps if they fail?
[RECORD VERBATIM RESPONSE]

P12. [IF P11A RESPONSE DOES NOT MENTION REPLACING FAILED LAMPS WITH THE SAME TYPE] Did your electrician or contractor mention that, if a lamp fails, you should only replace it with another Type B lamp?
1 Yes
2 No
98 Don't know
99 Refused

Since Installation, Removal and Replacement (All Types)

R1. Since installing the TLED lighting equipment, have you removed any of the incentivized TLED fixtures or bulbs?
1 Yes, removed bulbs
2 Yes, removed fixtures
3 Yes, removed fixtures and bulbs
4 No
98 Don't know
99 Refused

R2. [IF DID NOT REMOVE ANY TLEDs (R1 ≠ 1, 2, OR 3), SKIP TO C1] In total, how many did you remove?
[RECORD QUANTITY OF PROGRAM-DISCOUNTED FIXTURES]
[RECORD TOTAL QUANTITY OF PROGRAM-DISCOUNTED BULBS]

R3. In what month and year did you remove them?
[RECORD MONTH AND YEAR FOR FIXTURES]
[RECORD MONTH AND YEAR FOR BULBS]

R4A. [IF DID NOT REMOVE ANY FIXTURES (R1 ≠ 2 OR 3), SKIP TO R4B] Why did you remove the incentivized fixtures?

100 Flickering
101 Poor dimming
102 Incompatibility with dimming controls
103 Poor/low light level
104 Poor light quality
105 Fixture failure
106 Ballast failure
107 Other [SPECIFY]
108 Don't know
109 Refused

R4B. [IF DID NOT REMOVE ANY BULBS (R1 ≠ 1 OR 3), SKIP TO R5] Why did you remove the incentivized bulbs?

0 Because I removed the fixtures
110 Flickering
111 Poor dimming
112 Incompatibility with dimming controls
113 Poor/low light level
114 Poor light quality
115 Fixture failure
116 Ballast failure
117 Lamp failure (burned out)
118 Other [SPECIFY]
119 Don't know
120 Refused

R5. Did you replace the incentivized TLED lighting you removed?

1 Yes
2 No
98 Don't know
99 Refused

R5a. [IF DID NOT REPLACE (R5 ≠ 1), SKIP TO C1] What type of lighting did you use to replace the incentivized TLEDs? [ACCEPT MULTIPLE RESPONSES]

1 New/different TLEDs
2 LEDs
3 Linear fluorescent
4 Compact fluorescent
5 Incandescent
6 Halogen
97 Other [SPECIFY]
98 Don't know
99 Refused

Controls (All Types)

Next I would like to ask a few questions about the lighting controls you use for the for the TLED fixtures in your facility. I'd like to talk about ALL the controls you have associated with TLEDs in your facility, not just the incentivized TLEDs.

C1. I'd like to read you a short list of control types. Can you tell me which of these you use to control the TLED fixtures in your facility? If you're not familiar with something I mention, it's no problem. [READ LIST]

1 On/off switch
2 Dimmer switch
3 Photosensor or photocell
4 Occupancy sensor
5 Daylighting controls
6 Energy Management System
7 Timer
97 Something else? [SPECIFY]
98 [DO NOT READ] Don't know
99 [DO NOT READ] Refused

C2. [IF ONLY CONTROL IS ON/OFF SWITCH (C1 = 1 ONLY), SKIP TO C5] Did you receive incentives for these TLED controls through <Program>?

1 Yes
2 No
98 Don't know
99 Refused

C2A. [IF YES, D/K, or Refused, SKIP TO C3] Why did you not receive an incentive?
[RECORD VERBATIM RESPONSE]

C3. Did you install the TLED controls before, at the same time as, or after you installed the incentivized TLEDs?

1 Before
2 Same time
3 After
98 Don't know
99 Refused

C4. On a scale of 1 to 5 where 1 means "not at all satisfied" and 5 means "very satisfied," how satisfied are you with the performance of your controls for the TLEDs?

1 1 – not at all satisfied
2 2
3 3
4 4
5 5 – very satisfied
98 Don't know
99 Refused

C4A. [IF SATISFACTION IS NOT LOW (IF C4 ≠ 1, 2, OR 3), SKIP TO C5] Why do you say that? [IF NECESSARY: What could be improved?]
[RECORD VERBATIM RESPONSE]

C5.

1 Yes
2 No
98 Don't know
99 Refused

C5A. [IF NO PLANS TO INSTALL CONTROLS (IF C5 ≠ 1), SKIP TO D1] What type(s) of controls do you plan to install [within the next 12 months to control your TLEDs]?

1 On/off switch
2 Dimmer switch
3 Photosensor or photocell
4 Occupancy sensor
5 Daylighting controls
6 Energy Management System
7 Timer
97 Other [SPECIFY]
98 Don't know
99 Refused

Decision-Making

D1. Did you have any role in the decision to install the incentivized TLEDs?

1 Yes
2 No
98 Don't know
99 Refused

[WHETHER OR NOT ANY ROLE IN DECISION, SKIP TO D2]
[IF INSTALLED MULTIPLE TYPES OF TLEDS (TYPE A, B, C), SKIP TO D3A]

D2. Did you consider (or do you think the decision maker considered) any other types of TLEDs in addition to the incentivized TLEDs?

1 Yes
2 No
98 Don't know
99 Refused

D2A. [IF DID NOT CONSIDER OTHERS (IF D1 ≠ 1), SKIP TO D3A] What type or types did you consider? [ACCEPT MULTIPLE RESPONSES]

1 Type A
2 Type B
3 Type C
97 Other [SPECIFY]
98 Don't know
99 Refused

D3A. [IF <Measuredesc> ≠ TYPE A, SKIP TO D3B] What motivated you to select **Type A** TLEDs? [ACCEPT MULTIPLE RESPONSES]

1 Recommendation from electrician or contractor/electrician or contractor
2 Low first cost
3 Low lifetime/lifecycle cost
4 Ease to install

5 Building limitations/requirements
6 Compatibility with existing light fixtures
7 Compatibility with existing lighting controls
97 Other [SPECIFY]
98 Don't know
99 Refused

D3B. [IF <Measuredesc> ≠ TYPE B, SKIP TO D3C] What motivated you to select **Type B** TLEDs? [ACCEPT MULTIPLE RESPONSES]

1 Recommendation from electrician or contractor/electrician or contractor
2 Low first cost
3 Low lifetime/lifecycle cost
4 Ease to install
5 Building limitations/requirements
6 Compatibility with existing light fixtures
7 Compatibility with existing lighting controls
97 Other [SPECIFY]
98 Don't know
99 Refused

Article IV. **D3C.** [IF <Measuredesc> ≠ TYPE C, SKIP TO W1] What motivated you to select **Type C** TLEDs? [ACCEPT MULTIPLE RESPONSES]

1 Recommendation from electrician or contractor/electrician or contractor
2 Low first cost
3 Low lifetime/lifecycle cost
4 Ease to install
5 Building limitations/requirements
6 Compatibility with existing light fixtures
7 Compatibility with existing lighting controls
97 Other [SPECIFY]
98 Don't know
99 Refused

D4. Did you consider LED fixtures (for instance, flat panels or troffer retrofits) instead of the incentivized TLEDs?

1 Yes
2 No
98 Don't know
99 Refused

[IF DID NOT CONSIDER LED FIXTURES (IF D4 ≠ 1), SKIP TO W1]

D4A. What motivated you to select TLEDs instead of LED fixtures? [ACCEPT MULTIPLE RESPONSES]

1 Recommendation from electrician or contractor/electrician or contractor
2 Low first cost
3 Low lifetime/lifecycle cost
4 Ease to install
5 Building limitations/requirements
6 Compatibility with existing light fixtures
7 Compatibility with existing lighting controls
96 We did install both TLEDs and LED fixtures
97 Other [SPECIFY]
98 Don't know

99 Refused

Wrap-Up

W1. On a scale of 1 to 5 where 1 means "not at all likely" and 5 means "very likely," how likely would you be to recommend TLEDs like the ones you received incentives for to another business owner?

1 1 – not at all likely
2 2
3 3
4 4
5 5 – very likely
98 Don't know
99 Refused

W1A. [IF LIKELIHOOD IS NOT LOW (IF W1 ≠ 1, 2, OR 3), SKIP TO W2] Why do you say that? [IF NECESSARY: Why would you be unlikely to recommend TLEDs to another business owner?]
[RECORD VERBATIM RESPONSE]

Those are all the question I have for you today. Thank you for your time.

APPENDIX C. SEM ANALYSIS METHODOLOGY

This appendix presents the methods used in this evaluation to develop gross savings, followed by a discussion of the results.

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 MT&R 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 MT&Rs, 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 C-1 shows the model types used for the evaluation models developed. Twelve (12) models also utilize a non-weather independent variable.

Table C-1: Selected evaluation model types

Fuel	Temperature Response Model Type	Model Count
Electric	Constant	1
Electric	CDD Only	4
Electric	CDD & HDD, Single Reference Temperature	12
Electric	CDD & HDD, Dual Reference Temperature	1
Electric	HDD Only	4
Electric	Subtotal	22
Gas	HDD Only	28
All	**Total**	**50**

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 C-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 C-2: SEM program year assignment, standard cohort schedule

Month	Cohort 1	Cohort 2	Cohort 3	Cohort 4	Cohort 5	Cohort 6	Cohorts 7 & 8	Cohort 9
Kick Off Date	Nov-11	Jan-13	Oct-13	Jan-14	Oct-14	Jan-15	Oct-15	Oct-16
Oct-11	BeforeSEM	BeforeSEM	BeforeSEM	BeforeSEM	BeforeSEM	BeforeSEM	BeforeSEM	BeforeSEM
Nov-11	BeforeSEM	BeforeSEM	BeforeSEM	BeforeSEM	BeforeSEM	BeforeSEM	BeforeSEM	BeforeSEM
Dec-11	PY12	BeforeSEM	BeforeSEM	BeforeSEM	BeforeSEM	BeforeSEM	BeforeSEM	BeforeSEM
Jan-12	PY12	BeforeSEM	BeforeSEM	BeforeSEM	BeforeSEM	BeforeSEM	BeforeSEM	BeforeSEM
Feb-12	PY12	BeforeSEM	BeforeSEM	BeforeSEM	BeforeSEM	BeforeSEM	BeforeSEM	BeforeSEM
Mar-12	PY12	BeforeSEM	BeforeSEM	BeforeSEM	BeforeSEM	BeforeSEM	BeforeSEM	BeforeSEM
Apr-12	PY12	BeforeSEM	BeforeSEM	BeforeSEM	BeforeSEM	BeforeSEM	BeforeSEM	BeforeSEM
May-12	PY12	BeforeSEM	BeforeSEM	BeforeSEM	BeforeSEM	BeforeSEM	BeforeSEM	BeforeSEM
Jun-12	PY12	BeforeSEM	BeforeSEM	BeforeSEM	BeforeSEM	BeforeSEM	BeforeSEM	BeforeSEM
Jul-12	PY12	BeforeSEM	BeforeSEM	BeforeSEM	BeforeSEM	BeforeSEM	BeforeSEM	BeforeSEM
Aug-12	PY12	BeforeSEM	BeforeSEM	BeforeSEM	BeforeSEM	BeforeSEM	BeforeSEM	BeforeSEM
Sep-12	PY12	BeforeSEM	BeforeSEM	BeforeSEM	BeforeSEM	BeforeSEM	BeforeSEM	BeforeSEM
Oct-12	PY12	BeforeSEM	BeforeSEM	BeforeSEM	BeforeSEM	BeforeSEM	BeforeSEM	BeforeSEM
Nov-12	PY12	BeforeSEM	BeforeSEM	BeforeSEM	BeforeSEM	BeforeSEM	BeforeSEM	BeforeSEM
Dec-12	PY13	BeforeSEM	BeforeSEM	BeforeSEM	BeforeSEM	BeforeSEM	BeforeSEM	BeforeSEM
Jan-13	PY13	BeforeSEM	BeforeSEM	BeforeSEM	BeforeSEM	BeforeSEM	BeforeSEM	BeforeSEM
Feb-13	PY13	PY13	BeforeSEM	BeforeSEM	BeforeSEM	BeforeSEM	BeforeSEM	BeforeSEM
Mar-13	PY13	PY13	BeforeSEM	BeforeSEM	BeforeSEM	BeforeSEM	BeforeSEM	BeforeSEM
Apr-13	PY13	PY13	BeforeSEM	BeforeSEM	BeforeSEM	BeforeSEM	BeforeSEM	BeforeSEM
May-13	PY13	PY13	BeforeSEM	BeforeSEM	BeforeSEM	BeforeSEM	BeforeSEM	BeforeSEM
Jun-13	PY13	PY13	BeforeSEM	BeforeSEM	BeforeSEM	BeforeSEM	BeforeSEM	BeforeSEM
Jul-13	PY13	PY13	BeforeSEM	BeforeSEM	BeforeSEM	BeforeSEM	BeforeSEM	BeforeSEM
Aug-13	PY13	PY13	BeforeSEM	BeforeSEM	BeforeSEM	BeforeSEM	BeforeSEM	BeforeSEM
Sep-13	PY13	PY13	BeforeSEM	BeforeSEM	BeforeSEM	BeforeSEM	BeforeSEM	BeforeSEM
Oct-13	PY13	PY13	BeforeSEM	BeforeSEM	BeforeSEM	BeforeSEM	BeforeSEM	BeforeSEM
Nov-13	PY14	PY13	PY14	BeforeSEM	BeforeSEM	BeforeSEM	BeforeSEM	BeforeSEM
Dec-13	PY14	PY13	PY14	BeforeSEM	BeforeSEM	BeforeSEM	BeforeSEM	BeforeSEM
Jan-14	PY14	PY13	PY14	BeforeSEM	BeforeSEM	BeforeSEM	BeforeSEM	BeforeSEM
Feb-14	PY14	PY14	PY14	PY14	BeforeSEM	BeforeSEM	BeforeSEM	BeforeSEM
Mar-14	PY14	PY14	PY14	PY14	BeforeSEM	BeforeSEM	BeforeSEM	BeforeSEM
Apr-14	PY14	PY14	PY14	PY14	BeforeSEM	BeforeSEM	BeforeSEM	BeforeSEM
May-14	PY14	PY14	PY14	PY14	BeforeSEM	BeforeSEM	BeforeSEM	BeforeSEM
Jun-14	PY14	PY14	PY14	PY14	BeforeSEM	BeforeSEM	BeforeSEM	BeforeSEM
Jul-14	PY14	PY14	PY14	PY14	BeforeSEM	BeforeSEM	BeforeSEM	BeforeSEM
Aug-14	PY14	PY14	PY14	PY14	BeforeSEM	BeforeSEM	BeforeSEM	BeforeSEM
Sep-14	PY14	PY14	PY14	PY14	BeforeSEM	BeforeSEM	BeforeSEM	BeforeSEM
Oct-14	PY14	PY14	PY14	PY14	BeforeSEM	BeforeSEM	BeforeSEM	BeforeSEM
Nov-14	PY15	PY15	PY15	PY14	PY15	BeforeSEM	BeforeSEM	BeforeSEM
Dec-14	PY15	PY15	PY15	PY14	PY15	BeforeSEM	BeforeSEM	BeforeSEM
Jan-15	PY15	PY15	PY15	PY14	PY15	BeforeSEM	BeforeSEM	BeforeSEM
Feb-15	PY15	PY15	PY15	PY15	PY15	PY15	BeforeSEM	BeforeSEM
Mar-15	PY15	PY15	PY15	PY15	PY15	PY15	BeforeSEM	BeforeSEM
Apr-15	PY15	PY15	PY15	PY15	PY15	PY15	BeforeSEM	BeforeSEM
May-15	PY15	PY15	PY15	PY15	PY15	PY15	BeforeSEM	BeforeSEM
Jun-15	PY15	PY15	PY15	PY15	PY15	PY15	BeforeSEM	BeforeSEM
Jul-15	PY15	PY15	PY15	PY15	PY15	PY15	BeforeSEM	BeforeSEM
Aug-15	PY15	PY15	PY15	PY15	PY15	PY15	BeforeSEM	BeforeSEM
Sep-15	PY15	PY15	PY15	PY15	PY15	PY15	BeforeSEM	BeforeSEM
Oct-15	PY15	PY15	PY15	PY15	PY15	PY15	BeforeSEM	BeforeSEM
Nov-15	PY16	PY16	PY16	PY16	PY16	PY15	PY16	BeforeSEM
Dec-15	PY16	PY16	PY16	PY16	PY16	PY15	PY16	BeforeSEM
Jan-16	PY16	PY16	PY16	PY16	PY16	PY15	PY16	BeforeSEM
Feb-16	PY16	PY16	PY16	PY16	PY16	PY16	PY16	BeforeSEM
Mar-16	PY16	PY16	PY16	PY16	PY16	PY16	PY16	BeforeSEM
Apr-16	PY16	PY16	PY16	PY16	PY16	PY16	PY16	BeforeSEM
May-16	PY16	PY16	PY16	PY16	PY16	PY16	PY16	BeforeSEM
Jun-16	PY16	PY16	PY16	PY16	PY16	PY16	PY16	BeforeSEM
Jul-16	PY16	PY16	PY16	PY16	PY16	PY16	PY16	BeforeSEM
Aug-16	PY16	PY16	PY16	PY16	PY16	PY16	PY16	BeforeSEM
Sep-16	PY16	PY16	PY16	PY16	PY16	PY16	PY16	BeforeSEM

Oct-16	PY16	PY16	PY16	PY16	PY16	PY16	PY16	BeforeSEM
Nov-16	PY17	PY17	PY17	PY17	PY17	PY17	PY17	PY17
Dec-16	PY17	PY17	PY17	PY17	PY17	PY17	PY17	PY17
Jan-17	PY17	PY17	PY17	PY17	PY17	PY17	PY17	PY17
Feb-17	PY17	PY17	PY17	PY17	PY17	PY17	PY17	PY17
Mar-17	PY17	PY17	PY17	PY17	PY17	PY17	PY17	PY17
Apr-17	PY17	PY17	PY17	PY17	PY17	PY17	PY17	PY17
May-17	PY17	PY17	PY17	PY17	PY17	PY17	PY17	PY17
Jun-17	PY17	PY17	PY17	PY17	PY17	PY17	PY17	PY17
Jul-17	PY17	PY17	PY17	PY17	PY17	PY17	PY17	PY17
Aug-17	PY17	PY17	PY17	PY17	PY17	PY17	PY17	PY17
Sep-17	PY17	PY17	PY17	PY17	PY17	PY17	PY17	PY17
Oct-17	PY17	PY17	PY17	PY17	PY17	PY17	PY17	PY17

APPENDIX D. CUSTOM MEASURE RESULTS

This appendix provides summaries of the custom track evaluation results by measure category. Realization rates (RR) shown are calculated as the sum of evaluated savings divided by the sum of reported savings. The first table is for PY2017 only and the second table combines the PY2015, PY2016, and PY2017 evaluation results. DNV GL provided measure-specific results to Energy Trust separately.

Table D-1: Custom track evaluation results by measure Category, PY2017 Only

Custom Evaluation Category & Measure Description	# Evaluated Electric Measures	# Evaluated Gas Measures	Electric RR (%)	Gas RR (%)
Controls	**14**	**9**	**70%**	**82%**
Custom Building Controls	13	9	70%	82%
Custom EMS	1		68%	
HVAC	**7**	**4**	**100%**	**106%**
Custom HVAC	2		75%	
Custom VFDs	3	1	82%	100%
Custom VFD Pump	1		100%	
Custom Demand Control Ventilation	1	1	616%	181%
Custom Gas		2		103%
Motors	**3**		**89%**	
Custom VFDs	2		90%	
Custom Motors	1		88%	
Other	**9**	**4**	**102%**	**138%**
Custom Other	8	4	101%	138%
Custom Economizers	1		104%	
Process Cooling	**2**		**96%**	
Custom Chillers	2		96%	
Process Heating	**1**	**2**	**0%**	**32%**
Custom Heat Recovery	1	1	0%	8%
Custom Boiler		1		105%
Grand Total	**36**	**19**	**87%**	**86%**

Table D-2: Custom track evaluation results by measure category, PY2015, PY2016, & PY2017

Custom Evaluation Category & Measure Description	# Evaluated Electric Measures	# Evaluated Gas Measures	Electric RR (%)	Gas RR (%)
Controls	**50**	**31**	**83%**	**83%**
Custom Building Controls	46	31	83%	83%
Custom EMS	2		72%	
EMS for BPTaC	2		100%	
HVAC	**36**	**15**	**101%**	**97%**
Custom HVAC	19	7	86%	87%
Custom Chillers	8	1	131%	109%
Custom VFDs	3	1	82%	100%
Custom Demand Control Ventilation	2	2	74%	233%
Custom Economizers	2		83%	
Custom VFD Pump	1		100%	
Custom Boiler	1	2	100%	98%
Custom Gas		2		103%
Motors	**18**	**1**	**64%**	**1,201%**
Custom VFDs	14	1	65%	1,201%
Custom VFD Pump	3		57%	
Custom Motors	1		88%	
Other	**16**	**9**	**80%**	**133%**
Custom Other	15	9	78%	133%
Custom Economizers	1		104%	
Process Cooling	**2**		**96%**	
Custom Chillers	2		96%	
Process Heating	**1**	**2**	**0%**	**32%**
Custom Heat Recovery	1	1	0%	8%
Custom Boiler		1		105%
Grand Total	**123**	**58**	**90%**	**89%**

APPENDIX E. 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	Gas Fryers	
Track: Measure Type	Standard: Food Equipment	
Measure Code(s)	GASFRY	
This measure covers the installation of new ENERGY STAR-compliant gas fryers. Fryers account for about 34% of gas savings. Key evaluation parameters include number of vats, vat capacity-size (Large or Standard), cooking energy efficiency, business hours, equipment operating hours, and pounds of food cooked per day. DNV GL evaluated 42 fryer measures from 2015-17. The savings weighted realization rate across these 42 measures is 77%. Energy Trust should consider adjusted measure assumption to align with this result.	**RR: Avg.** (Min-Max)	**52%** (0% - 197%)
	Sample Target	22
	Survey Completes	10

Measure Information

Program Data Review: 377 unique measure lines were reported over the program year 2017. These lines accounted for 0% of electricity savings and 34% 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 22 sites for verification, 10 interviews were complete, 3 of these were confirmed to be out-of-business, 2 refused, and 7 were no-response.

- Telephone interviews were completed for data collection
- Verifying tracking savings with MAD, all sites were Standard size vat
- Pounds of food cooked per day are based on customer response, or an adjusted-default value based on the opeating hours if they are unable to estimate amount of food cooked. Customer reported pounds of food cooked are mostly higher than default, but one is lower due to not being a restaurant.

Evaluation Recommendations

Use the ENERGY STAR calculator for traceability and transparency.

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

Add a low-use incentive category/measure to address non-restaurant sites, their use-hours are much different than restaurant sites (much less, so less potential for savings)

Adjustment to evaluation plan

None

DNV GL Measure Description	Cooler Doors (Add doors to open cases)	
Track: Measure Type	**Standard: Refrigeration**	
Measure Code(s)	COOLDRETFITG, COOLDRETFITE	
This measure covers the retrofit of open refrigerated cases to include doors or the replacement of open cases with closed door cases. The measure has electric & gas (interactive HVAC) savings. Key evaluation parameters include number of doors, linear feet of casing, retrofit type, and HVAC heating and cooling type.	**RR: Avg.** (Min-Max)	**100% (elec.), 100% (gas).**
	Sample Target	6
	Survey Completes	2

Measure Information

Program Data Review: 43 unique measure lines were reported over the program year 2017. These lines accounted for 29% of electricity savings and 18% of gas savings reported for the standard track.

Program Delivery: Standard

Evaluation Summary

The evaluation sample included 6 sites for verification. One telephone interview was completed, and one other site was verified via site-visit as a part of strip curtain evaluation site visits.

Both sites are retrofit projects, 1 site claimed both gas and electric savings, 1 site only claimed electric savings. The electric-only site is also evaluated to have gas savings due to confirming gas heating system on site.

Claimed savings matched MAD values.

Phone interview site contact had little technical knowledge, and was unable to verify type of fuel used for heating/cooling, only able to verify quantiy/linear feet of doors.

Evaluation Recommendations

Improve identification of HVAC heating system type

Differentiate between gas and electric MAD savings, there should be two records, one for electric only, and one for gas & electric.

Best way to truly verify installed lengths is with onsite inspection

Adjustment to evaluation plan

None

DNV GL Measure Description	Strip Curtains	
Track: Measure Type	Standard: Refrigeration	
Measure Code(s)	STRIPCURTLTWFS, STRIPCURTMTWCS	
This measure covers the retrofit/replacement of strip curtains for walk-in freezers. Key evaluation parameters include square-feet verification, replacement/add-on verification, replacement practices, and average daily operating hours.	**RR: Avg.** (Min-Max)	**68%** (0% - 100%)
	Sample Target	14
	Survey Completes	12

Measure Information

Program Data Review: 252 unique measure lines were reported over the program year 2017. These lines accounted for 40% of electricity savings and 0% of gas savings reported for the standard track.

Program Delivery: Standard

Evaluation Summary

The evaluation sample included 7 sites for verification (each site/location claimed 2 measure lines). 6 sites were visited (of those visited, 1 site is closed); the 7th (P00001265951) was not visited due to travel costs.

Site visits were completed for data collection.

- Installation rate are adjusted for partial savings.
- Savings achieved are less than expected due to modification of the curtains, such as removal of curtains, draping of curtains, and business closed.
- Partial savings are estimated based on the percent of each curtain observed to be in use during site visit.
- Curtains at the closed site did not achieve savings.

Evaluation Recommendations

Include the RTF calculator name with version used to calculate the prescriptive savings. If multiple sources are used, then provide the sources used for each savings estimate.

Adjustment to evaluation plan

None

DNV GL Measure Description	Boiler	
Track: Measure Type	Standard: Boiler	
Measure Code(s)	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, steam end-use load and operational parameters.	**RR: Avg.** (Min-Max)	**74%** (50% - 161%)
	Sample Target	14
	Survey Completes	11

Measure Information

Program Data Review: 46 unique measure lines were reported over the program year 2017. These lines accounted for 0% of electricity savings and 22% of gas savings reported for the standard track.

Program Delivery: Standard

Evaluation Summary

The original evaluation sample included 14 measure lines across 10 sites. The final sample included 11 boilers at 8 sites. The evaluation developed two estimates of savings 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 is not specifically a backup.

2. The evaluation adjusted savings for boilers found to not operate in the 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.

Site	Estimated Operating Efficiency	Operating Efficiency Scaling	Multiple Boiler Scaling	Evaluated UES, Therms	MAD UES, Therms
Site 1: boiler 1	89%	67%	75%	1.10	2.19
Site 2: boiler 1	95%	104%	75%	2.21	2.85
Site 2: boiler 2	95%	104%	75%	2.21	2.85
Site 3: boiler 1	94%	100%	75%	2.14	2.85
Site 3: boiler 2	94%	100%	75%	2.14	2.85
Site 4: boiler 1	94%	100%	100%	2.85	2.85
Site 5: boiler 1	89%	67%	75%	1.44	2.85
Site 6: boiler 1	95%	107%	75%	2.28	2.85
Site 6: boiler 2	95%	107%	75%	2.28	2.85
Site 7: boiler 1	98%	123%	100%	3.51	2.85
Site 8: boiler 1	89%	67%	75%	1.44	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-using 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. For the sites where we had adequate billing data, we encountered the following scenarios:

- The boiler is only used for space heating and the boiler is 100% of the gas bill. In this case, we can use the regression savings approach.
- The boiler is used to heat an indoor swimming pool. Other gas loads include DHW and space heating. The model identified an optimum temperature threshold for HDDs of 84°F. This is much higher than typical for other building types. This is probably due to the heating pool indirectly providing space heating and a high indoor air temperature setpoint. We were able to identify the base load (DHW) but were not able to isolate the boiler load from other space heating equipment. Therefore, we were not able to apply the regression-based approach.
- The boiler is used for space heating and heating an indoor swimming pool. The only other load on the gas bills is the DHW. For this site, we were able to isolate the boiler load and use the regression-based approach.

Using the isolated boiler load from the post period, we can calculate 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. The table below provides a summary of the results for the sites that used this approach.

Site	Site 1	Site 2	Site 4	Site 8
Adjusted R2	0.98	0.99	0.98	0.93
HDD Temperature Threshold, F	62	67	70	60
Constant	29.0	450.5	1,626.9	3,325.9
HDD Coefficient	10.1	7.0	5.0	5.2
Non-Weather Dependent, therms	348	5,406	19,523	26,607
Weather Dependent, therms	34,116	33,016	28,288	16,919
Percent of Meter Weather Dependent	99.0%	85.9%	59.2%	38.9%
Boiler Load, therms	34,116	36,501	47,811	41,350

The high adjusted R^2 shows that the billing data fits this model. The boiler at site 1 is used for space heating. Of the four sites where we were able to use the regression-based savings estimate, this is the only site that is used for the same end-use assumed in the MAD savings estimate. The other sites have other end uses including space conditioning, reheat, and pool heating. The boiler at Site 2 is for space heating and reheating the discharge air in air handling units. The model shows the reheat load (and to a lesser extent domestic hot water) as non-weather dependent load. Site 4 and Site 8 are recreational facilities with indoor swimming pools. These sites have a significant percentage of non-weather dependent gas usage. The boiler installed at site 3 is used for pool heating only. Other gas-using equipment at this site provides space heating and domestic hot water. The boiler at Site 8 is used for both pool and space conditioning.

The Typical Meteorological Year (TMY) data was compared to the actual weather conditions over the billing period. These were typically within 3 to 4%. Therefore, it was not necessary to normalize results to TMY. The following table shows the approach and savings results for each site.

Site	Evaluation Approach	Evaluated Savings Therms	Reported Total therms	Realization Rate
Site 1: boiler 1	Regression	4,222	4,380	96%
Site 2: boiler 1	Regression	3,422	5,700	60%
Site 2: boiler 2	Regression	3,422	5,700	60%
Site 3: boiler 1	Adjusted MAD	6,413	8,550	75%
Site 3: boiler 2	Adjusted MAD	6,413	8,550	75%
Site 4: boiler 1	Regression	2,295	1,425	161%
Site 5: boiler 1	Adjusted MAD	8,621	17,100	50%
Site 6: boiler 1	Adjusted MAD	4,555	5,697	80%
Site 6: boiler 2	Adjusted MAD	4,555	5,697	80%
Site 7: boiler 1	Adjusted MAD	1,402	1,137	123%
Site 8: boiler 1	Regression	5,117	5,700	90%
Total	-	**50,437**	**69,636**	**72%**

Weather Sensitivity and MAD Assumptions

The evaluation team conducted a weather regression analysis on gas meter billing data for boiler sites from 2015/1016 and 2017 program cycles. The goals of this analysis are:

- to assess the weather-sensitivity of gas usage
- to compare the MAD-based boiler consumption to the total meter consumption

We received billing data for a total of 58 meters across 43 sites. A total of 18 meters were removed from the analysis due to insufficient data or the likelihood that the boiler was not included on the gas meter. In these cases, some sites had multiple meters or the MAD-based boiler consumption was greater than 250% of the meter consumption. The 35 meters that had reasonable regression results are summarized in the following table:

Average of Meters	Weather Sensitive % of Meter	Claimed Boiler Consumption % of Meter
Straight Average	75.6%	75.8%
Meter Consumption Weighted Average	74.5%	60.9%

The weather-sensitive percentage of metered gas and the claimed boiler consumption percentage of metered gas are similar. This suggests that the estimated boiler load in the MAD savings are reasonable. However, a number of sites sampled in the evaluation were found to have loads other than space heating.

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. It is recemended that the boiler load end-uses be identified and alternative savings estimates be developed.
- Some boilers were found to have conditions that made it unlikely that they typically operate in condesing mode. This reduces the operational efficiency as well as savings. It may be beneficial to have sites demonstrate that boilers will operate in condesing mode based on loading and estimated setpoints. As an alternative, the savings estimate could account for a percentage of boilers that will not operate in condesing mode.
- 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.

DNV GL Measure Description	Wall Insulation, Roof/Attic Insulation	
Track: Measure Type	**Standard: Wall Insulation, Ceiling Insulation**	
Measure Code(s)	INSWALLG, INSROOFG, INSROOFER	
Insulation is offered for wall insulation, roof insulation or attic 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, building vintage, roof/attic, existing insulation verification, and space heating/space cooling verification.	**RR: Avg.** (Min-Max)	**Wall** **100%(gas)** (100% - 100%, gas) **Roof/Attic 100%(elec.)** **9.7%(gas)** (100% - 100%, elec.) (0% - 100%, gas)
	Sample Design	Wall: 1 (gas) Roof/Attic: 4 (3 gas, 1 electric)
	Survey Completes	Wall: 1 (gas) Roof/Attic: 4 (3 gas, 1 electric)

Measure Information

Program Data Review: 52 unique measure lines were reported. These lines accounted for 2% of electricity savings and 7% of gas savings reported for the standard track.

Program Delivery: Standard

Evaluation Summary

- The evaluation sample did not include site visits for verification, 5 interviews were complete, none were non-response.
- Telephone interviews were completed for data collection.
- Measure Information
  - Low roof/attic gas RR due to zeroed-out savings for 2 of 3 projects. These projects included major remodel-rehab triggering code requirements: roof replacements, reported by customer. These projects did not exceed code minimum R value. These two projects made up the majority of the tracking savings for the 3 sampled projects.
  - One site was not a standard building type: ice arena. MAD savings consider this space type; savings would be greater for air conditioning. However, this was a site with zeroed out RR due to major remodel-rehab.
  - 3 of 5 projects had 100% RR, heating system and R-value match measure definition.

Evaluation Recommendations

- Ensure that projects are true retrofits, not complete remodel-rehab

Adjustment to evaluation plan

None

DNV GL Measure Description	Smart (Motion Sensor) Power Strips		
Track: Measure Type	Standard: Power Strip		
Measure Code(s)	OCCPLUGSTRIPSI, OCCPLUGSTRIP[10], LOADPLUGSTRIP		
The smart strip was a TrickleStar 183SS-US-8XX Motion Sensor Power Strip. Key evaluation parameters were the installation and operation of the power strip plus the use of the motion sensing feature of the equipment.		**RR: Avg.** (Min-Max)	**38% (25%-100%**
		Sample Target	4
		Survey Completes	4

Measure Information

Program Data Review: 2,925 units were claimed across 228 measure lines for 355,170 kWh in savings in the Standard track. An additional 88 units were claimed across 73 measure lines for 10,736 kWh through the Direct Install track. Measure quantities of 1 through 14 per project were most common, 56 measures lines had quantities ranging from 15 to 80. Five measure lines claimed 90 or more units, which were mostly schools and offices.

Program Delivery: These devices were purchased and self-installed by the customers

Evaluation Summary

The evaluation focused on projects in which the customer purchased the power strips. The evaluation sample excluded power strips given away to customers. The sample included 6 measure lines across 4 projects for verification. Telephone interviews were completed for data collection.

- Site-level realization rates were either 25% or 100%. Sampled sites consist of 6 schools: 5 schools that were a part of one school district and 1 school that was from another school district. Interviews were conducted to each of the school district representative, not individual schools.
- The 5-school site had used a quarter of what they have ordered, because the customer ordered as many as possible, and deploys them as needed (25% savings).
- The 1-school site uses all of the plug strips and its motion sensing feature as claimed (100% savings)
- The sample only consists of schools and did not cover other building types

Evaluation Recommendations

- Measure still shows savings potential, but installation rate is at risk based on site procurement practices (schools in this case)
- Consider implementing controls on program policies (certain restrictions, etc.) to mitigate the risk that a customer purchases units but does not install them.
- If distribution continues and savings are claimed:
  - Do not make customers sign an agreement about how and where they will use the device. Signing an agreement does not guarantee the smart strips will be installed, nor installed correctly.
  - Do not incentivize large quantities without an installation verification plan in place.

Adjustment to evaluation plan

None

[10] Only the OCCPLUGSTRIP measure code was sampled in this evaluation, the other two codes LOADPLUGSTRIP & OCCPLUGSTRIPSI were not sampled. This evaluation focused on projects where customers purchased significant quantities of power strips. The PY2015-16 evaluation researched the impact of power strips given away to customers.

Evaluated Standard measures, project-specific results. (Low frequency measures)

Measure type	Measure Description	Measure Code(s)	RR	Evaluated Sites	Findings	Recommendations
Controls	Anti-sweat Heater Controls	REFANTISWTL, REFANTISWTM	Elec: 100%	1 out of 5 sampled 0 out of 4 backup-promoted	Reported equipment verified and installed. Savings adopted into the MAD are reasonable, since they were derived from RTF.	None
Motor	ECM Motors	BEMOTREF, ECMWIU	Elec: 100%	1 out of 1 sampled 0 out of 2 backup-promoted	ECMs retrofit to walk-in coolers. The sampled site is a supermarket. Savings are evaluated with MAD document's prescriptive savings, depending on whether the measure is new install/retrofit, as well as reach-in/walk-in cooler and freezers. The savings method is from RTF, which uses reasonable inputs and assumptions.	None
Food Equipment	Vent Hood - Gas Heat	VENTHOODE, VENTHOODG	Gas: 402% Elec: 273%	1 out of 2 sampled, 1 out of 1 backup promoted	New control system for dedicated commercial kitchen exhaust hood and make-up air system. It works with VFD to modulate fan speed. Verified the system is installed and operating. High RR due to customer claiming the lower HP tier savings, and sites had higher operating hour than MAD assumption (14 hours/day, 6 days/week). Evaluated savings are based on applying hours-of-use and HP adjustment to MAD savings values.	Provide documentation on the calculator mentioned in the MAD.
Food Equipment	Electric Hot Food Holding Cabinets	HOTCAB, HOTFOODCABHALF	Elec: 5%	1 out of 1 sampled	Evaluated site is half-size. Hot food cabinets are only used partially during business hours. Low realization rate due to site being event planning and not full-service restaurant.	Use the Energy Star calculator with actual performance specifications and operating hours.
Icemaker	Ice Machines	ICEIMHSMT1, ICESCUSMT2, ICERCULGT1, ICEIMHLGT1	Elec: 66%	0 out of 2 sampled, 1 out of 3 backup-promoted	Evaluated site is Ice Making Head Tier 1. Lower RR due to ice machine ice harvest rate being lower than default assumption in Energy Star calculator.	Use the Energy Star calculator with actual performance specifications and operating hours.
Food Equipment	Gas Griddles	GASGRID	Gas: 69%	1 out of 1 sampled	One site only. Lower RR due to slightly lower hours of operation and pounds of food cooked per day than the EnergyStar calculator default.	Use the Energy Star calculator with actual performance specifications and operating hours.
Food Equipment	Dishwashers	UCHITEMPELE, STDRUPHITEMPELE	Elec: 5%	0 out of 1 sampled, 1 out of 2 backup-promoted	One site only (undercounter, high temp – electric water heater). Very low RR due to site being a school, therefore reduced days of operation, also racks washed per day is much lower than EnergyStar calculator assumption.	Use the Energy Star calculator with actual performance specifications and operating hours.

About 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 business. We provide classification and technical assurance along with software and independent expert advisory services to the maritime, oil and gas, and energy industries. We also provide certification services to customers across a wide range of industries. Operating in more than 100 countries, our 16,000 professionals are dedicated to helping our customers make the world safer, smarter and greener.