Report by **SBW CONSULTING, INC.**

Report No. **1302**

FINAL REPORT

IMPACT EVALUATION OF THE EXISTING BUILDINGS PROGRAM

PROGRAM YEARS 2010-2011

Submitted to
ENERGY TRUST OF OREGON
851 SW SIXTH AVE, SUITE 1200
PORTLAND, OREGON 97204

Submitted by
SBW CONSULTING, INC.
2820 Northup Way, Suite 230
Bellevue, WA 98004

In association with
CADMUS
720 SW Washington St., Suite 400
Portland, OR 97205

October 31, 2013

sbw
ENERGY + WATER + EFFICIENCY

TABLE OF CONTENTS

TABLES

EXECUTIVE SUMMARY

Introduction

Energy Trust of Oregon (Energy Trust) began operating the existing commercial building retrofit program in early 2003. The program provides technical assistance and financial incentives to commercial and institutional customers who install qualifying efficiency improvements that save electricity or natural gas. It is open to all commercial customers that pay the public purpose charge. The program is market-driven and builds on existing market relationships, which is consistent with best practices among resource acquisition and market transformation efforts. It is administered by a third party Program Management Contractor for the Energy Trust.

SBW Consulting, Inc. (SBW) conducted this evaluation to assess the gross impacts of the commercial retrofit program for the 2010-2011 program years. It included a review of the expected savings estimates prepared by the program and a re-estimation of realized annual savings for a representative sample of program participants. The realized savings were based on on-site data collection and a re-application of the program algorithms under as-built and operated conditions. Results from the sampled cases were extrapolated to the entire program population for each program year.

Objectives

Specific objectives include the following:

- **Verify installations**. Confirm through field inspection that the sampled measures were installed and operational. Also verify the hours of operation and other conditions that affected the expected savings from the program-installed measures;
- **Review program measure-specific energy savings methods**. Critique energy savings algorithms used by the program to calculate savings for custom measures;
- **Calculate project-specific gross savings**. Calculate project-specific gross savings (kWh and therms) impacts for a sample of projects for the 2010 and 2011 program years using best practice evaluation methods. Also calculate project-specific realization rates for kWh and therms. Calculate project-specific gross demand savings for each sampled project; however, the kW savings estimate is of secondary concern.
- **Calculate program level gross savings**. Estimate an overall gross energy savings realization rate for the program and estimate program level gross savings (kWh and therms) for the 2010-2011 program years.
- **Recommend program improvements**. Recommend changes in the gross savings calculation methods or other program processes that will enhance future realization rates and program cost effectiveness.

Program Accomplishments

Table E-1 summarizes the expected kWh and therm savings, as claimed by the Energy Trust Existing Buildings Program in 2010-2011. In total, the program incentivized nearly 34,000

measures at almost 6,600 sites during the two-year cycle, accounting for an expected 242 million kWh and 4.1 million therms in annual energy savings. The table also compares the 2010-2011 claims with those from the previous cycle (2006-2007 and 2008-2009), showing the significant increase in measures completed and savings achieved in the most recent cycle.

Table E-1: Expected Savings for 2010-11 Existing Buildings Program

Program year	Number of sites	Number of measures	Expected program kWh savings	Expected program therm savings
Previous evaluation cycle				
2006	1,611	3,446	31,326,511	985,727
2007	1,463	3,667	26,531,894	526,998
Total	**3,074**	**7,113**	**57,858,405**	**1,512,725**
2008	1,506	3,839	42,397,819	1,180,882
2009	1,960	7,781	74,503,452	1,083,537
Total	**3,466**	**11,620**	**116,901,271**	**2,264,419**
Current evaluation cycle				
2010	2,810	12,605	104,674,358	1,832,026
2011	3,773	21,110	136,791,255	2,246,637
Total	**6,583**	**33,715**	**241,465,613**	**4,078,663**

Methodology

This evaluation employed standard energy program impact evaluation methods to provide the best available estimate of the total program energy impacts.

Kickoff Meeting - This task included the review of program data and the collection of information necessary to finalize the evaluation work plan. Information was collected through a kickoff meeting and a series of discussions with the Energy Trust relevant to the study methodology.

Sample Design - We analyzed the Energy Trust program database to determine the distribution of sites by savings in each program year and then implemented the sample based upon sampling decisions made by the Energy Trust. Two separate samples were selected for the 2010 and 2011 program years. For each year, one sample was selected from the list of sites that had non-zero electric savings and the other was selected from those sites with non-zero gas savings. A total of 120 cases were selected across the four selected samples.

Site Data Collection - This task included the collection of measure-specific information needed to support the analysis of gross realized savings from the program, as specified in the workplan. This task also included the determination of analysis methods and recruitment of the sampled sites. A site visit was performed for all sampled measures to collect measure performance data.

We developed a project evaluation database that was used to assemble and perform quality control checks on all data needed to conduct the program-level impact evaluation. The database was created in Excel and was formulated to be consistent with all Energy Trust data formatting and content requirements.

Project-level Gross Savings Analysis - This task included the analysis of gross realized savings for the sampled measures, using the measure-specific data collection and analysis methods developed in Task 3. Gross realization rates were computed for each measure.

Program-level Gross Savings Analysis - Program level gross results for the two-year period were estimated by extrapolating the gross savings (kWh and therms) from each measure, using methods specified in the workplan.

Study Findings

A separate engineering analysis of gross savings (kWh and therms) was performed for each measure designated for inspection using the measure-specific data that was collected during the site visits. Realization rates were computed for each sampled measure using the realized kWh/therm savings results and the annual savings estimated by the program as follows:

Realization Rate = realized annual energy savings / expected annual energy savings

The expected savings values were taken from the Energy Trust tracking database. If the realized energy savings equaled the expected energy savings, then the realization rate equaled unity (1.0).

Figure E-1 below summarizes the realization rates for the 310 sampled and inspected measures, in the form of a distribution graph sorted from lowest to highest realization rates for electric and gas measures, as well as all measures combined. Measure realization rates ranged from a low of zero to a high of nearly two. This figure reveals that over half of these measures had realization rates at or very near one. About 8 percent of the measures yielded little or no savings; three percent had no savings, while another 5 percent realized less than half of their expected savings. These low savers were a mixture of gas and electric measures. Low savers represented 8 percent of the electric measures and 10 percent of the gas measures. Of the 9 measures that had no savings, the primary reasons were that the measures had not been installed (or were installed incorrectly) in the first place or were removed after installation, or that the facility was vacant at the time of inspection. Table E-2 shows these nine measures along with measures that were incorrectly installed. A small number of electric measures (3 percent) had savings 50% or more than expected. Unusually high realization rates were typically caused by overly conservative estimates of expected savings by the program.

Table E-2: Measures by Installation Status

Installation Status	Number of Measures
Not Installed	7
Installed Incorrectly	14
Removed	1
Vacant	1

The results from the individual sample points were extrapolated to two-year (2010-11) program totals. Program extrapolations were made for gross realized savings in both two fuel groups (electric and gas) for each program year. For the 2010 program year, the program-level realization rates were estimated to be 1.07 and 0.86 for kWh and therms, respectively. For the 2011 program year, the realization rates were estimated to be 0.91 and 1.01 for kWh and therms, respectively. The 2-year realization rates were 0.98 for kWh and 0.94 for therms. These represent a significant improvement from the previous evaluation cycle. One measure involved a vacant building in which the equipment had been removed. The measure belonged to the therm 2010 sample frame. If the measure had still been operational, the realization rate for this frame would have changed from 0.86 to 0.88.

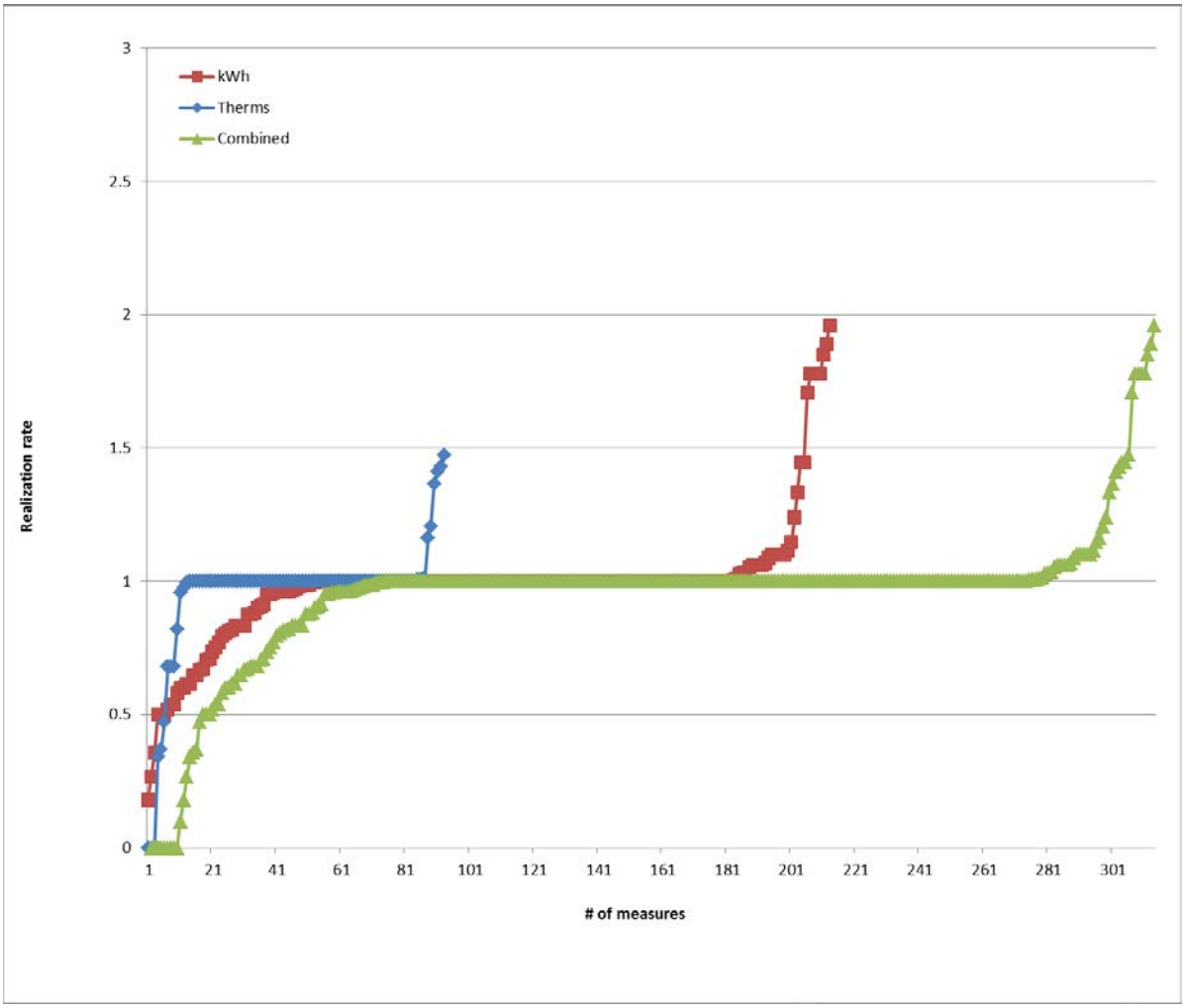

Figure E-1: Distribution of Measure Realization Rates.

Conclusions

From the results of the impact evaluation of gross realized savings for the 2010-11 program years, the following key conclusions were drawn.

1. **Installation Verification** – The evaluation verified through field inspection that 97 percent of the sampled and assessed measures were either fully or partly installed and operational. When they occurred, partial installations and variations from the expected measure operation (differing schedules, set points, etc.) accounted for many of the differences between expected and evaluated savings. Oftentimes, the timing of when discrepancies occurred (eg. when a particular set point had been changed, etc.) and reasons for the discrepancies we encountered (eg. why did a particular set point vary from the original design, etc.) were unclear.
2. **Gross savings realized** – Significant gross energy savings were found for both fuel types in each program year. Unweighted measure-level results show a range in realization rates from zero to nearly two. Domain-level results for the electric and gas fuel types indicate that realized savings were less than expected savings for both fuel types in both program years. The electric realization rate was estimated to be 0.99 across the two program years. The therm realization rate was estimated to be 0.94 for the two-year period. These results represent a significant improvement from the previous evaluation cycle. Table E-3 shows the confidence level and relative precision for each sample frame.

Table E-3: Sample Frame Confidence Level and Relative Precision

Sample Frame	Confidence Level	Relative precision
kWh2010	90%	11%
kWh2011	90%	8%
Therms2010	90%	9%
Therms2011	90%	3%

3. **Major reasons for differences between the expected and realized savings –** The realized and expected gross savings differed for a variety of reasons. In most cases the differences tended to reduce savings. The most common reasons for reduced savings were lower operating hours and “Other”. The “Other” category included situations such as measures not actually being installed, vacant buildings, inappropriate algorithms, or questionable algorithm inputs.
4. **Algorithm review –** In many instances, custom algorithms could not be reviewed due to a lack of documentation. When they could be reviewed, the custom algorithms were generally found to be reasonable. Custom algorithm types included eQUEST models, TRACE 700 models, and MS Excel spreadsheet calculation models. Typically, these custom algorithms were applied to complex HVAC measures such as controls or central plant equipment

replacement measures. Some general recommendations were made for reviewed algorithms where improvements were appropriate.

5. **Project documentation –** Major improvements were made by the Project Management Contractor, since the last evaluation cycle, in providing the evaluation with complete and accurate documentation of the sampled projects. However, additional improvements can still be made for future evaluations. These include providing electronic copies of all calculations, and ensuring that provided documentation versions matches the reported savings values.

Recommendations

Findings from this evaluation led to the following recommendations:

1. **Feedback to Energy Trust savings estimates**. Energy Trust estimates of expected savings have the disadvantage of having to predict the future performance of a measure before it is installed. The evaluation has the advantage of estimating realized savings for a measure under as-built and operated conditions. Because of these different perspectives, differences between expected and realized savings are unavoidable. The evaluation was able to gauge measure conditions at a single point in time, but further in-depth study of measures in the evaluation sample, where significant differences in the estimates were found, can improve the ability of Energy Trust to predict savings and/or ensure good performance for future measures. It is recommended that Energy Trust carefully study these cases, such as by revisiting sites and speaking with customers and vendors to understand better the reasons why certain measures performed poorly. Information from such customer follow-up might lead to improved procedures for inspection, quality control, and training, which in turn may increase realization rates in future evaluations.

2. **Project documentation**. The value and cost-effectiveness of the evaluation was very dependent upon access to accurate and complete program documentation for each sampled project. Documentation should be provided to the evaluator in sufficient detail for an independent third party to understand expected measure performance. It should include the information and tools necessary to understand the algorithm that was used to calculate the expected savings, in an electronic format necessary to reproduce the savings estimate.

   Adequate documentation was provided for most sampled projects. The completeness and accuracy of the documentation was an improvement from the previous evaluation cycle. However, there were a number of cases where sufficient documentation was not provided. This was especially true for non-lighting projects where the expected savings were calculated with an hourly simulation or proprietary software. Compensating for this lack of documentation significantly increased the cost of the evaluation. It is recommended that Energy Trust improve the completeness of the project documentation in future program years so that the evaluations can be supplied with the information and tools necessary to cost-effectively complete the evaluation scope specified by Energy Trust. This includes items such as measure performance specifications, documentation of assumptions made and calibration methods used during the analysis of expected savings, and backup information related to the calculations made.

All spreadsheet calculation tools should be provided in Excel format (not pdf) so that the equations can be understood and the savings estimates can easily be reproduced. When hourly simulations tools are used, documentation should include the final as-built energy modeling files, in electronic format, that are needed to reproduce the expected savings analysis. It is also recommended that the use of proprietary software be minimized.

The most expeditious way to provide program files to future evaluators would be to store all pertinent files for a given site in a well-documented electronic folder, which could be passed on in its entirety should that site be sampled in the evaluation.

When the folder is first received from the ATAC, it should be thoroughly checked to ensure that all necessary material is included and that the versions of the supplied documentation match exactly the reported savings values in the database. This check should be completed for every site.

3. **Measure interactive effects**. For interior lighting measures, the expected savings did not capture interactions with the HVAC system, when they were relevant. It is recommended that the expected savings methodologies be upgraded to capture interactive effects where they are significant.
4. **Program Communications**. Some customers did complain about the number of times they had previously been asked to allow sites visits as part of other aspects of program implementation. It is recommended that the program communicate clearly that evaluations are happening (in some cases multiple evaluations at the same site) and required to improve the programs as well as support ongoing funding.

MEMO

Date: January 9, 2014
To: Board of Directors
From: Phil Degens, Evaluation Manager
Spencer Moersfelder, Existing Buildings Program Manager
Subject: Staff Response to the 2010 and 2011 Existing Buildings Impact Evaluation Report

The 2010-2011 impact evaluation report shows that the Existing Buildings program improved its realization rates for both electric and gas savings in the years 2010 and 2011 relative to previous evaluations. At the same time, the program significantly increased the level of savings and the number of customers served. The program also did a good job of providing these services to a wide array of commercial buildings. The evaluators found the program was effectively implemented as site visits found that in almost all cases the project measures had been installed.

The program also made strides towards gathering and storing project data consistently and electronically. All of the files were successfully transferred electronically to the evaluators via Energy Trust's secure file transfer protocol (SFTP) site and required minimal Energy Trust resources to carry out. There was still an issue with obtaining complete simulation models for a few projects. However, this issue was much less prevalent than in earlier years and the program now requires that all simulation models be provided to the Program Management Contractor (PMC) before an incentive is paid.

It is worth noting that the program changed its PMC at the beginning of 2013. The program achievements that are presented in the report are due to the successful program implementation of the prior PMC. It is anticipated that the current PMC will integrate appropriate learnings from this report so as to continue the program's successful implementation.

Additionally, Energy Trust evaluation staff plans on evaluating savings for one program year (in lieu of two or more years) at a time. This will ensure that that the program receives more frequent and faster delivery of evaluation results and this timely information and feedback will in turn improve program delivery.

1. INTRODUCTION

Energy Trust of Oregon (Energy Trust) began operating the existing commercial building retrofit program in early 2003. The program provides technical assistance and financial incentives to commercial and institutional customers who install qualifying efficiency improvements that save electricity or natural gas. It is open to all commercial customers that pay the public purpose charge. The program is market-driven and builds on existing market relationships, which is consistent with best practices among resource acquisition and market transformation efforts. It is administered by a third party Program Management Contractor for the Energy Trust.

SBW Consulting, Inc. (SBW) conducted this evaluation to assess the gross impacts of the commercial retrofit program for the 2010-2011 program years. It included a review of the expected savings estimates prepared by the program and a re-estimation of realized annual savings for a representative sample of program participants. The realized savings were based on on-site data collection and a re-application of the program algorithms under as-built and operated conditions. Results from the sampled cases were extrapolated to the entire program population for each program year.

1.1. Goals and Objectives

The overall goal of this impact evaluation was to quantify the magnitude of energy savings (annual kWh and therms) captured by the commercial retrofit program through an assessment of a representative sample of efficiency projects that were implemented during the 2010 and 2011 program years.

Specific objectives include the following:

- **Verify installations**. Confirm through field inspection that the sampled measures were installed and operational. Also verify the hours of operation and other conditions that affected the expected savings from the program-installed measures;
- **Review program measure-specific energy savings methods**. Critique energy savings algorithms used by the program to calculate savings for custom measures;
- **Calculate project-specific gross savings**. Calculate project-specific gross savings (kWh and therms) impacts for a sample of projects for the 2010 and 2011 program years using best practice evaluation methods. Also calculate project-specific realization rates for kWh and therms. Calculate project-specific gross demand savings for each sampled project; however, the kW savings estimate is of secondary concern.
- **Calculate program level gross savings**. Estimate an overall gross energy savings realization rate for the program and estimate program level gross savings (kWh and therms) for the 2010-2011 program years.
- **Recommend program improvements**. Recommend changes in the gross savings calculation methods or other program processes that will enhance future realization rates and program cost effectiveness.

2. PROGRAM ACCOMPLISHMENTS

Table 1 summarizes the expected kWh and therm savings, as claimed by the Energy Trust Commercial Existing Buildings Program in 2010-2011. In total, the program incentivized nearly 34,000 measures at almost 6,600 sites during the two-year cycle, accounting for an expected 242 million kWh and 4.1 million therms in annual energy savings. The table also compares the 2010-2011 claims with those from the previous cycle (2006-2007 and 2008-2009), showing the significant increase in measures completed and savings achieved in the most recent cycle.

Table 1: Expected Savings for 2010-11 Commercial Existing Buildings Program

Program year	Number of sites	Number of measures	Expected program kWh savings	Expected program therm savings
Previous *evaluation* cycle				
2006	1,611	3,446	31,326,511	985,727
2007	1,463	3,667	26,531,894	526,998
Total	**3,074**	**7,113**	**57,858,405**	**1,512,725**
2008	1,506	3,839	42,397,819	1,180,882
2009	1,960	7,781	74,503,452	1,083,537
Total	**3,466**	**11,620**	**116,901,271**	**2,264,419**
Current evaluation cycle				
2010	2,810	12,605	104,674,358	1,832,026
2011	3,773	21,110	136,791,255	2,246,637
Total	**6,583**	**33,715**	**241,465,613**	**4,078,663**

Table 2 shows the breakdown of measures, by type, installed by the 2010 and 2011 programs separately and combined. Definitions of these measure types can be found in Appendix A. Efficient lighting measures were by far the most common measure type across the 2-year period. It accounted for more than seventy percent of the measure count and half of the claimed electric savings for both program years. Across the 2-year period, HVAC was the second most commonly installed measure type, accounting for a relatively small portion (11 percent) of the electric savings but the greatest portion (47 percent) of the gas savings. Custom Controls measures were less commonly implemented (1 percent of measure count) but accounted for 21 percent of the two-year claimed therm savings and 12 percent of the claimed electric savings. In general, Custom measures were few in number, but accounted for significant portions of savings.

Table 2: Expected Savings by Measure Type

2010

	Measures		Expected program kWh		Expected program therms	
Measure Type	**Count**	**Percent**	**Savings**	**Percent**	**Savings**	**Percent**
Custom Controls	155	1%	10,375,020	11%	202,938	11%
Custom Gas	27	0%	-	0%	20,072	1%
Custom Other	678	6%	13,095,831	13%	512,289	28%
Food Service	783	6%	1,380,932	1%	35,677	2%
HVAC	760	6%	9,979,724	10%	710,452	39%
Insulation	285	2%	2,416,345	2%	214,752	12%
Lighting	8,985	73%	49,256,847	51%	-	0%
Motors	80	1%	6,318,851	7%	4,662	0%
Water Heating	520	4%	4,311,363	4%	107,521	6%
Totals	**12,273**	**100%**	**97,134,913**	**100%**	**1,808,364**	**100%**

"Motors" category includes VFDs.

2011

	Measures		Expected program kWh		Expected program therms	
Measure Type	**Count**	**Percent**	**Savings**	**Percent**	**Savings**	**Percent**
Custom Controls	233	1%	17,216,967	13%	644,440	29%
Custom Gas	8	0%	23,691	0%	54,073	2%
Custom Other	573	3%	6,263,740	5%	112,452	5%
Food Service	454	2%	1,519,522	1%	26,380	1%
HVAC	1,822	9%	15,514,475	12%	1,167,878	52%
Insulation	201	1%	1,346,378	1%	118,610	5%
Lighting	14,547	75%	75,421,718	58%	-	0%
Motors	121	1%	1,970,797	2%	-	0%
Water Heating	1,425	7%	11,472,192	9%	105,931	5%
Totals	**19,384**	**100%**	**130,749,480**	**100%**	**2,229,765**	**100%**

"Motors" category includes VFDs.

COMBINED						
	Measures		Expected program kWh		Expected program therms	
Measure Type	Count	Percent	Savings	Percent	Savings	Percent
Custom Controls	388	1%	27,591,987	12%	847,378	21%
Custom Gas	35	0%	23,691	0%	74,145	2%
Custom Other	1,251	4%	19,359,571	8%	624,741	15%
Food Service	1,237	4%	2,900,454	1%	62,057	2%
HVAC	2,582	8%	25,494,199	11%	1,878,331	47%
Insulation	486	2%	3,762,723	2%	333,362	8%
Lighting	23,532	74%	124,678,565	55%	-	0%
Motors	201	1%	8,289,648	4%	4,662	0%
Water Heating	1,945	6%	15,783,555	7%	213,453	5%
Totals	**31,657**	**100%**	**227,884,393**	**100%**	**4,038,128**	**100%**

"Motors" category includes VFDs.

Table 3 provides a breakdown of sites by site building type, in a similar manner. The table shows that offices are the most common building type in the 2-year participant population and account for the largest fraction of both the electric and gas savings. Grocery, restaurant, warehouse and retail are other common building classifications. The grocery, warehouse, retail and “other” facilities represent a significant fraction of the electric savings. Colleges/universities, schools, retail and hospitals are major contributors to the gas savings.

Table 3: Expected Savings by Building Type

2010						
	Sites		Expected program kWh		Expected program therms	
Building Type	Count	Percent	Savings	Percent	Savings	Percent
Assembly	51	1.9%	2,322,864	2.3%	61,134	3.4%
Auto Services	85	3.1%	1,723,333	1.7%	17,472	1.0%
Automotive Repair/Sales	-	0.0%	-	0.0%	-	0.0%
Church	42	1.5%	353,851	0.3%	10,314	0.6%
College/University	33	1.2%	5,897,332	5.8%	301,529	16.6%
Data Center	11	0.4%	2,839,715	2.8%	-	0.0%
Grocery	277	10.2%	11,382,954	11.3%	11,526	0.6%
Gym/Athletic Club	25	0.9%	1,496,831	1.5%	27,548	1.5%
Hi Rise Residential	5	0.2%	17,446	0.0%	-	0.0%
Hospital	25	0.9%	2,834,251	2.8%	133,405	7.4%
Institution/Government	119	4.4%	3,000,287	3.0%	130,087	7.2%
Laundry/Dry Cleaners	36	1.3%	127,509	0.1%	29,402	1.6%
Lodging/Hotel/Motel	94	3.4%	1,582,465	1.6%	86,017	4.7%
Manufacturing	4	0.1%	79,228	0.1%	19,834	1.1%
Office	420	15.4%	17,163,745	17.0%	296,206	16.3%
Other	121	4.4%	10,888,157	10.8%	105,799	5.8%
Other Health	71	2.6%	1,363,281	1.3%	48,900	2.7%
Restaurant	400	14.7%	1,681,087	1.7%	62,488	3.4%
Retail	268	9.8%	12,785,653	12.6%	61,017	3.4%
Schools K-12	178	6.5%	3,528,677	3.5%	301,806	16.7%
Warehouse	180	6.6%	8,215,436	8.1%	34,330	1.9%
Unassigned	282	10.3%	11,840,007	11.7%	73,691	4.1%
Totals	**2,727**	**100.0%**	**101,124,109**	**100.0%**	**1,812,506**	**100.0%**

2011						
	Sites		Expected program kWh		Expected program therms	
Building Type	Count	Percent	Savings	Percent	Savings	Percent
Assembly	66	1.8%	4,205,283	3.2%	44,598	2.1%
Auto Services	131	3.6%	4,592,305	3.5%	28,931	1.4%
Automotive Repair/Sales	-	0.0%	-	0.0%	-	0.0%
Church	7	0.2%	38,724	0.0%	9,739	0.5%
College/University	48	1.3%	4,495,133	3.5%	188,383	8.8%
Data Center	9	0.2%	1,479,562	1.1%	11,693	0.5%
Grocery	315	8.7%	9,202,952	7.1%	10,877	0.5%
Gym/Athletic Club	29	0.8%	957,607	0.7%	67,328	3.1%
Hi Rise Residential	1	0.0%	-	0.0%	310	0.0%
Hospital	17	0.5%	2,583,043	2.0%	159,476	7.4%
Institution/Government	25	0.7%	541,861	0.4%	59,114	2.8%
Laundry/Dry Cleaners	52	1.4%	524,767	0.4%	6,398	0.3%
Lodging/Hotel/Motel	121	3.4%	2,286,887	1.8%	65,089	3.0%
Manufacturing	1	0.0%	5,464	0.0%	269	0.0%
Office	661	18.3%	25,901,014	20.0%	637,396	29.8%
Other	146	4.0%	9,878,732	7.6%	125,457	5.9%
Other Health	81	2.2%	1,207,157	0.9%	22,250	1.0%
Restaurant	312	8.6%	1,666,079	1.3%	83,107	3.9%
Retail	614	17.0%	19,783,200	15.3%	254,533	11.9%
Schools K-12	156	4.3%	4,014,951	3.1%	213,655	10.0%
Warehouse	302	8.4%	14,996,251	11.6%	59,410	2.8%
Unassigned	514	14.2%	21,355,576	16.5%	92,749	4.3%
Totals	**3,608**	**100.0%**	**129,716,548**	**100.0%**	**2,140,763**	**100.0%**

COMBINED						
	Sites		Expected program kWh		Expected program therms	
Building Type	Count	Percent	Savings	Percent	Savings	Percent
Assembly	117	1.8%	6,528,147	2.8%	105,732	2.7%
Auto Services	216	3.4%	6,315,638	2.7%	46,403	1.2%
Automotive Repair/Sales	-	0.0%	-	0.0%	-	0.0%
Church	49	0.8%	392,575	0.2%	20,053	0.5%
College/University	81	1.3%	10,392,465	4.5%	489,912	12.4%
Data Center	20	0.3%	4,319,277	1.9%	11,693	0.3%
Grocery	592	9.3%	20,585,906	8.9%	22,403	0.6%
Gym/Athletic Club	54	0.9%	2,454,438	1.1%	94,876	2.4%
Hi Rise Residential	6	0.1%	17,446	0.0%	310	0.0%
Hospital	42	0.7%	5,417,294	2.3%	292,881	7.4%
Institution/Government	144	2.3%	3,542,148	1.5%	189,201	4.8%
Laundry/Dry Cleaners	88	1.4%	652,276	0.3%	35,800	0.9%
Lodging/Hotel/Motel	215	3.4%	3,869,352	1.7%	151,106	3.8%
Manufacturing	5	0.1%	84,692	0.0%	20,103	0.5%
Office	1,081	17.1%	43,064,759	18.7%	933,603	23.6%
Other	267	4.2%	20,766,889	9.0%	231,256	5.8%
Other Health	152	2.4%	2,570,438	1.1%	71,150	1.8%
Restaurant	712	11.2%	3,347,166	1.4%	145,595	3.7%
Retail	882	13.9%	32,568,853	14.1%	315,550	8.0%
Schools K-12	334	5.3%	7,543,628	3.3%	515,461	13.0%
Warehouse	482	7.6%	23,211,687	10.1%	93,740	2.4%
Unassigned	796	12.6%	33,195,583	14.4%	166,441	4.2%
Totals	**6,335**	**100.0%**	**230,840,657**	**100.0%**	**3,953,269**	**100.0%**

3. METHODOLOGY

This section describes the methodology that was used by the evaluation team to determine the gross energy impacts of the Energy Trust commercial retrofit program for program years 2010-2011. This effort employed standard energy program evaluation methods to provide the best available estimate of the total program energy impacts.

Kickoff Meeting

To begin the evaluation, SBW received and reviewed an electronic copy of the program tracking database. The evaluation team then met with the Energy Trust evaluation and program staff to discuss a variety of topics relevant to the work plan and evaluation methodology. SBW incorporated agreements reached at these meetings into the work plan.

Sample Design

SBW analyzed the Energy Trust program database to determine the distribution of sites by savings in each program year and then implemented the sample based upon sampling decisions made by the Energy Trust during the kickoff meeting. SBW established and documented a fixed sample size for the 2010 and 2011 program years of 120 cases. Further details of the sample design can be found in Section 4.

Site Data Collection

The purpose of this task was to collect all of the information needed to support the analysis of gross savings for the sampled measures. This includes traditional capital measures as well as the tune-up and O&M measures introduced in the 2010-2011 program cycle. For this evaluation, special emphasis was placed on the data collection and analysis of controls measures, since there is more uncertainty about the performance of this measure type.

SBW developed a project evaluation database that was used to assemble and perform quality control checks on all data needed to conduct the program-level impact evaluation. The database was created in Excel and was formulated to be consistent with all Energy Trust data formatting and content requirements. The database included data accumulated from standardized site workbooks (implemented in Excel) that the evaluation staff used to assemble the data needed for gross impact analysis and to prepare the analysis results. At the end of the evaluation, the entire database was provided to Energy Trust as an electronic record of work performed.

Data collection was accomplished through the following steps.

- **Copy Project Files**. Energy Trust provided scanned copies of the project files for each selected project. When available for custom measures, Energy Trust also provided electronic versions of the savings calculation spreadsheets, simulation files and other information that supported the program savings estimate.
- **Sample Recruitment**. Energy Trust sent a letter of introduction to the contact for each of the sampled sites prior to recruitment. This letter helped to establish for these customers

the importance of this work and their role in performing the work. It also introduced SBW as the evaluation contractor.

The appropriate site contact was called to recruit the customer for the evaluation. The call confirmed that the customer was able to provide a person that was knowledgeable about the location of the measure(s) in the sampled project and provide access to the measure(s). If successful, SBW also determined the most appropriate contact for the site visit. If the customer refused to participate, a knowledgeable person could not be confirmed, or reasonable access to the measure was not possible, the sampled measure was replaced with another site from the same stratum in the sample.

- **Review Project Files**. Each site that was successfully recruited was assigned to a lead analyst, who in most cases was responsible for both data collection and analysis of gross savings. The lead analyst reviewed all relevant information in the project file and extracted data that was important to the evaluation. In general this included the performance specifications of each measure (baseline and as-built) and operating hours for the equipment that comprised the sampled measures.

  For each sampled prescriptive measure, SBW examined the program application form and determined the measure performance information that was input to the program software that estimated expected savings. For each sampled custom measure, SBW reviewed the savings analysis that was performed by the program and determined the input parameters that were required to re-estimate savings using this algorithm. Based on the findings from this review, the analyst selected an appropriate measure-specific data collection methodology. For cases where documentation of the program algorithm could not be provided, SBW worked with the Energy Trust to establish an appropriate method for estimating realized savings and selected an appropriate measure-specific data collection methodology.

- **Schedule Site Visit**. The appropriate site contact was called to schedule a site visit and obtain other information needed to adequately plan for the site visit. SBW field staff worked with the site contact to be sure that tenants or other occupants were notified about the site visit so that the field staff could get appropriate access to all of the installed measures.

- **Data collection**. On-site data collection involved the collection of the measure performance information (baseline and as-built) that was needed to re-estimate savings using the prescriptive or custom algorithms. Data was collected through direct observation by field staff or interviews with the site contact or other site personnel. The information was documented for use in rerunning the prescriptive software or re-estimating gross savings using the custom algorithm, or other agreed-upon methods for cases where the program algorithm could not be applied. If the evaluation values were significantly different than the program values, then reasons for the differences were noted in the documentation.

- **Data collection - short term metering**. Short term metering and the analysis of EMS trend data collected by the customer were performed by the SBW team on an as needed basis to

support the analysis of savings for a select set of measures. Metering was applied to measures with the highest degree of uncertainty in the expected savings.

- **Data collection – implementation of previous recommendations**. SBW determined whether recommendations made in previous impact evaluations of this program to improve the program evaluability were implemented. As part of data collection, SBW examined the project files to determine if all or some of the evaluability recommendations from the previous evaluations were implemented.

Project-level Gross Savings Analysis

The purpose of this task was to analyze gross realized savings for each measure in the sampled projects. This was accomplished through the following steps.

- **Review program algorithms.** This was limited to custom measures with customized engineering calculations. It did not include a review of the underlying algorithms behind prescriptive deemed savings values or standardized calculation software (e.g., standard lighting spreadsheet). SBW documented the results of each review, with particular attention paid to recommendations for improvements (if any) to the algorithms that would increase the accuracy of the gross savings estimates.
- **Re-estimation of measure-level gross savings with program algorithm**. For prescriptive measures, gross realized savings were estimated by applying the program algorithm. For custom measures, the program algorithm was applied for cases where it was supplied by Energy Trust. For custom measures where the program algorithm could not be applied, an alternative algorithm was applied, as agreed to with Energy Trust. The baseline and as-built inputs to the algorithms came from the field data collected under Task 3 above. To the extent possible, reasons for differences between the program and evaluation savings estimates were documented. Measure-level savings were summed to project totals. Average demand (kW) savings were also computed for each electric measure.
- **Treatment of O&M measures.** O&M measures were introduced to the existing commercial building program in the 2010-2011 program cycle. Since they were not previously evaluated for this program, the methods used to estimate evaluation savings were carefully considered on a case-by-case basis for each O&M measure in the final sample.

Program-level Gross Savings Analysis

The purpose of this task was to extrapolate gross savings from the sampled projects to the program level for program years 2010 and 2011. This was accomplished through the following steps.

- A project-specific savings realization rate was computed for each sampled project. This was the ratio of realized savings for the project to the expected savings in the program database. This could be more or less than 1, depending on whether more or fewer savings than the program originally estimated were found.
- An appropriate sample weight was assigned to each project. The large savers selected with certainty were assigned a sample weight of 1. Other projects, selected at random,

represented more than one project in the population and thus were assigned a weight greater than 1.

- The aggregate realization rates, by sampling stratum, were applied to all other projects in the respective stratum. In this manner, sample results were extrapolated to the populations for each program year, yielding evaluation estimates of total realized savings by year.
- A program-level savings realization rate was computed for the 2010 and 2011 samples by dividing the realized program savings by the Energy Trust expected program savings estimate.

4. SAMPLE DESIGN

Two separate samples were selected for the 2010 and 2011 program years (four samples total). For each year, one sample was selected from the list of applicable sites that had non-zero electric savings and the other was selected from those sites with non-zero gas savings. Some sites had both gas and electric savings so some sites were selected in both samples.

The Energy Trust excluded icemaker and steam trap measures from this evaluation because they had been studied in other evaluations. These and other excluded measures accounted for a negligible portion of the 2-year electric savings and less than two percent of the gas savings. The excluded measures are summarized in Table 4 below.

Table 4: Measures Excluded from Evaluation

2010						
	Measures		Expected program kWh		Expected program therms	
Measure Type	Count	Percent	Savings	Percent	Savings	Percent
Icemakers	63	0.5%	64213	0.1%	0	0.0%
Steam traps	39	0.3%	0	0.0%	51,685	2.8%
Measures w/no savings	825	6.5%	0	0.0%	0	0.0%
Totals	**927**	**7.4%**	**64,213**	**0.1%**	**51,685**	**2.8%**

2011						
	Measures		Expected program kWh		Expected program therms	
Measure Type	Count	Percent	Savings	Percent	Savings	Percent
Icemakers	44	0.3%	40584	0.0%	0	0.0%
Steam traps	6	0.0%	0	0.0%	17,696	1.0%
Measures w/no savings	2163	17.2%	0	0.0%	0	0.0%
Totals	**2,213**	**17.6%**	**40,584**	**0.0%**	**17,696**	**1.0%**

	Combined					
	Measures		Expected program kWh		Expected program therms	
Measure Type	Count	Percent	Savings	Percent	Savings	Percent
Icemakers	107	0.3%	104,797	0.0%	0	0
Steam traps	45	0.1%	0	0.0%	69,381	1.7%
Measures w/no savings	2988	8.9%	0	0.0%	0	0.0%
Totals	**3,140**	**9.3%**	**104,797**	**0.0%**	**69,381**	**1.7%**

After removing the excluded measures catalogued above, we divided the remaining population into the four sample frames by year and fuel type. A stratification design was developed for all four samples. It is summarized in the Table 5 below.

Table 5: Recruited Sample

Sample frame (program year + fuel saved)	Strata	Strata lower bound (Site savings kWh/yr. or Therms/yr.)	Number of sites	Sample % of savings	Sampled cases	Sampling percentage	Estimated sampling error at 90% confidence level
2010 kWh	Certainty	714,319	15	21%	15	100%	
	5	303,348	31	17%	3	10%	
	4	139,624	90	20%	3	3%	
	3	57,964	164	17%	3	2%	
	2	19,807	380	14%	3	1%	
	1	2,855	987	10%	3	0%	
	Excluded	0	303	0.50%	0	0%	
			1,970	**100%**	**30**	**2%**	**10%**
2011 kWh	Certainty	635,445	19	20%	19	100%	
	5	287,551	52	19%	4	8%	
	4	121,578	103	17%	3	3%	
	3	51,270	241	17%	3	1%	
	2	17,792	531	15%	4	1%	
	1	2,640	1,406	11%	5	0.4%	
	Excluded	0	524	0.8%	0	0%	
			2,876	**100%**	**38**	**1%**	**9%**
2010 therms	Certainty	51,568	5	32%	5	100%	
	4	20,973	12	23%	5	42%	
	3	7,078	26	19%	5	19%	
	2	2,016	67	15%	4	6%	

Sample frame (program year + fuel saved)	Strata	Strata lower bound (Site savings kWh/yr. or Therms/yr.)	Number of sites	Sample % of savings	Sampled cases	Sampling percentage	Estimated sampling error at 90% confidence level
	1	108	282	11%	5	2%	
	Excluded	0	182	1.0%	0	0%	
			574	**100%**	**24**	**4%**	**10%**
2011 therms	Certainty	30,841	10	31%	10	100%	
	4	13,140	19	18%	3	16%	
	3	5,008	45	17%	3	7%	
	2	1,669	136	17%	4	3%	
	1	171	578	17%	7	1%	
	Excluded	0	114	0.5%	0	0%	
			902	**100%**	**27**	**3%**	**10%**
TOTALS			**6,322**		**119**	**2%**	

The table lists the characteristics of the strata in each design, including the number of sites in the population, the total savings, and the sample quota. The sample included 68 cases with electric measures and 51 cases with gas measures, for a total of 119 sampled cases. These cases were included in 111 sites. Eight sites included both sampled electric and gas measures. It is noted that the final site count was one site less than the 120 sites specified in the work plan. This slight reduction was due to the fact that one of the replacement sites (see section 5.1 sample disposition) included therm measures for a site that was previously selected as a kWh site in the original sample.

Many sites included measures with small expected savings. Small savings were observed for many measure types. To improve the efficiency of the evaluation, a large number of measures were treated as “Pass-thru.” These measures were assigned a realization rate of 1, and were not analyzed in detail. Table 6 shows the percentage of expected savings these measures represent for each sample domain. Table 6: Pass-Thru Measure Percentage of Expected Savings

Sampling Domain	% Energy Trust Savings of PassThru Measures
2010kWh	5%
2011kWh	9%
2010therm	3%
2011therm	7%

All four sample designs included a certainty stratum (shown in Table 5 as stratum 9). All sites in this stratum were selected for evaluation. In each design, these strata and these sites account for at least 20 percent of the total savings for each population. In addition, all designs had an excluded stratum (stratum 8). These contain sites with very small savings. Excluding them from

the sample improved the precision of the final results. Also shown in the table is the expected precision for an estimate of total savings from the sample design.

The fraction of the cases selected in each stratum determined how the final results were weighted. The weight assigned to a case in each stratum was the inverse of the case's probability of selection. The probability that a particular case was selected equals the number of cases selected from the stratum divided by the number of total cases in the stratum. For the certainty stratum the weight was 1. Each sample was selected by randomly ordering sites within each stratum and then selecting in random order the first set of sites that equal the strata quota as shown in the table above.

Table 6 shows the distribution of the program population and sampled sites across the building types treated by the program. The table shows that, for the combined two-year period, office, warehouse and retail were the building types most frequently included in the sample. Twenty-five offices and twenty retail sites were included in the sample. Warehouses were also selected quite often with 10 sites in the sample.

Table 7: Sample by Building Type

	2010			2011			Combined		
Building Type	Number of sites	Sampled sites	Percent sampled	Number of sites	Sampled sites	Percent sampled	Number of sites	Sampled sites	Percent sampled
Assembly	43	1	2.3%	58	1	1.7%	101	2	2.0%
Auto Services	71	-	0.0%	126	1	0.8%	197	1	0.5%
Automotive Repair/Sales	-	-		-	-		-	-	0.0%
Church	31	-	0.0%	4	-	0.0%	35	-	0.0%
College/University	26	4	15.4%	45	2	4.4%	71	6	8.5%
Data Center	10	1	10.0%	7	1	14.3%	17	2	11.8%
Funeral/Cremation	-	-		2	-		2	-	
Grocery	268	2	0.7%	312	3	1.0%	580	5	0.9%
Gym/Athletic Club	22	1	4.5%	20	1	5.0%	42	2	4.8%
Hi Rise Residential	1	-	0.0%	1	-	0.0%	2	-	0.0%
Hospital	12	3	25.0%	14	3	21.4%	26	6	23.1%
Infrastructure	1	-		2	1		3	1	
Institution/Government	63	1	1.6%	12	2	16.7%	75	3	4.0%
Laundry/Dry Cleaners	23	2	8.7%	51	-	0.0%	74	2	2.7%
Lodging/Hotel/Motel	87	3	3.4%	111	2	1.8%	198	5	2.5%
Multifamily Residential	5	-		2	-		7	-	
Manufacturing	3	-	0.0%	1	-	0.0%	4	-	0.0%
Office	361	10	2.8%	601	15	2.5%	962	25	2.6%
Other	109	2	1.8%	130	2	1.5%	239	4	1.7%
Other Health	55	3	5.5%	74	-	0.0%	129	3	2.3%
Parking structure/Garage	14	1		21	-		35	1	
Religious/Spiritual	30	-		82	2		112	2	
Restaurant	373	1	0.3%	291	1	0.3%	664	2	0.3%
Retail	263	6	2.3%	596	14	2.3%	859	20	2.3%

	2010			2011			Combined		
Building Type	Number of sites	Sampled sites	Percent sampled	Number of sites	Sampled sites	Percent sampled	Number of sites	Sampled sites	Percent sampled
Retirement/Assisted Facilities	18	-		28	3		46	3	
Schools K-12	162	5	3.1%	133	1	0.8%	295	6	2.0%
Warehouse	168	5	3.0%	298	5	1.7%	466	10	2.1%
Unassigned	282	-	0.0%	506	-	0.0%	788	-	0.0%
Totals	**2,501**	**51**	**2.0%**	**3,528**	**60**	**1.7%**	**6,029**	**111**	**1.8%**

Does not count measures excluded from the evaluation

Some sites were selected in both years.

The program database classified these measures according to standardized measure types, as shown in Table 8 below. The majority of sampled measures were lighting (fixtures and controls), reflecting the preponderance of this measure in the portfolio. HVAC, motors, and custom measures were also well-represented in the sample.

Table 8: Sample by Measure Type

	2010			2011			Combined		
Measure Type	Number of measures	Measures at sampled sites	Percent sampled	Number of measures	Measures at sampled sites	Percent sampled	Number of measures	Measures at sampled sites	Percent sampled
Custom Controls	155	13	8%	233	24	10%	388	37	10%
Custom Gas	27	-	0%	8	2	25%	35	2	6%
Custom Other	237	23	10%	149	9	6%	386	32	8%
Food Service	720	2	0%	410	0	0%	1,130	2	0%
HVAC	721	34	5%	1,816	119	7%	2,537	153	6%
Insulation	285	2	1%	201	3	1%	486	5	1%
Lighting	8,909	233	3%	14,494	362	2%	23,403	595	3%
Motors	80	11	14%	121	8	7%	201	19	9%
Other	-	-		1	0	0%	1	-	0%
Water Heating	520	3	1%	1,425	5	0%	1,945	8	0%
Totals	**11,654**	**321**	**3%**	**18,858**	**532**	**3%**	**30,512**	**853**	**3%**

Includes "Pass-thru" measures

5. FINDINGS

This section presents and discusses the energy impact results for measures implemented by the commercial retrofit program during the 2010 and 2011 program years. It begins with a discussion of the sample disposition, then presents and analyzes evaluated kWh and therm savings and realization rates for the measures in the sample. Next, the section briefly discusses the results from the review of measure savings algorithms. The section ends with a presentation and discussion of evaluated savings results at the program level.

5.1. Sample Disposition

The evaluation team made considerable efforts to recruit each site, and complete data collection and analysis for each sampled measure. In the end, all but two sites in the primary sample were successfully recruited. Two sampled sites were replaced when the customers refused to participate despite multiple attempts to recruit them. In general the recruited customers had a positive attitude towards participation, although some of them did complain about the number of times they had previously been asked to allow sites visits as part of other aspects of program implementation.

5.2. Measure Realization Rates

A separate engineering analysis of gross savings (kWh and therms) was performed for each measure designated for inspection using the measure-specific data that was collected during the site visits. The savings were computed using the data collection and analysis methods described in Section 3 above. Realized average kW savings were computed by dividing the annual kWh savings by the number of annual hours that the measure was in operation.

At each sampled site, the evaluation analysts focused on the significant measures (310 measures across all sampled sites). Ones accounting for a negligible portion of the overall program savings were "passed through," and assigned evaluation savings equal to expected savings. Of the electric measures, 34 percent (219 of 639) were inspected, and the rest were passed through. These latter measures were overwhelmingly (nearly 80 percent) small lighting measures. Of the gas measures, 43 percent (91 of 214) were inspected, and the rest were passed through. The pass-through measures were a variety of measure types.

Realization rates were computed for each sampled measure using the realized kWh/therm savings results and the annual savings estimated by the program as follows:

Realization Rate = realized annual energy savings / expected annual energy savings

The expected savings values were taken from the Energy Trust tracking database. If the realized energy savings equaled the expected energy savings, then the realization rate equaled unity (1.0). kW realization rates were not computed because expected kW savings were not included in the Energy Trust tracking database.

Figure 2 below summarizes the realization rates for the 310 sampled and inspected measures, in the form of a distribution graph sorted from lowest to highest realization rates for electric

and gas measures, as well as all measures combined. Measure realization rates ranged from a low of zero to a high of nearly two. This figure reveals that over half of these measures had realization rates at or very near one. About 8 percent of the measures yielded little or no savings; three percent had no savings, while another 5 percent realized less than half of their expected savings. These low savers were a mixture of gas and electric measures. Low savers represented 8 percent of the electric measures and 10 percent of the gas measures. A small number of electric measures (3 percent) had savings 50% or more than expected. Of the 9 measures that had no savings, the primary reasons were that the measures had not been installed (or were installed incorrectly) in the first place or were removed after installation, or that the facility was vacant at the time of inspection[1]. Unusually high realization rates were typically caused by overly conservative estimates of expected savings by the program. Details of the realized annual gross savings and realization rate computed for each sampled measure can be found in Appendix C – Measure-level Evaluation Results.

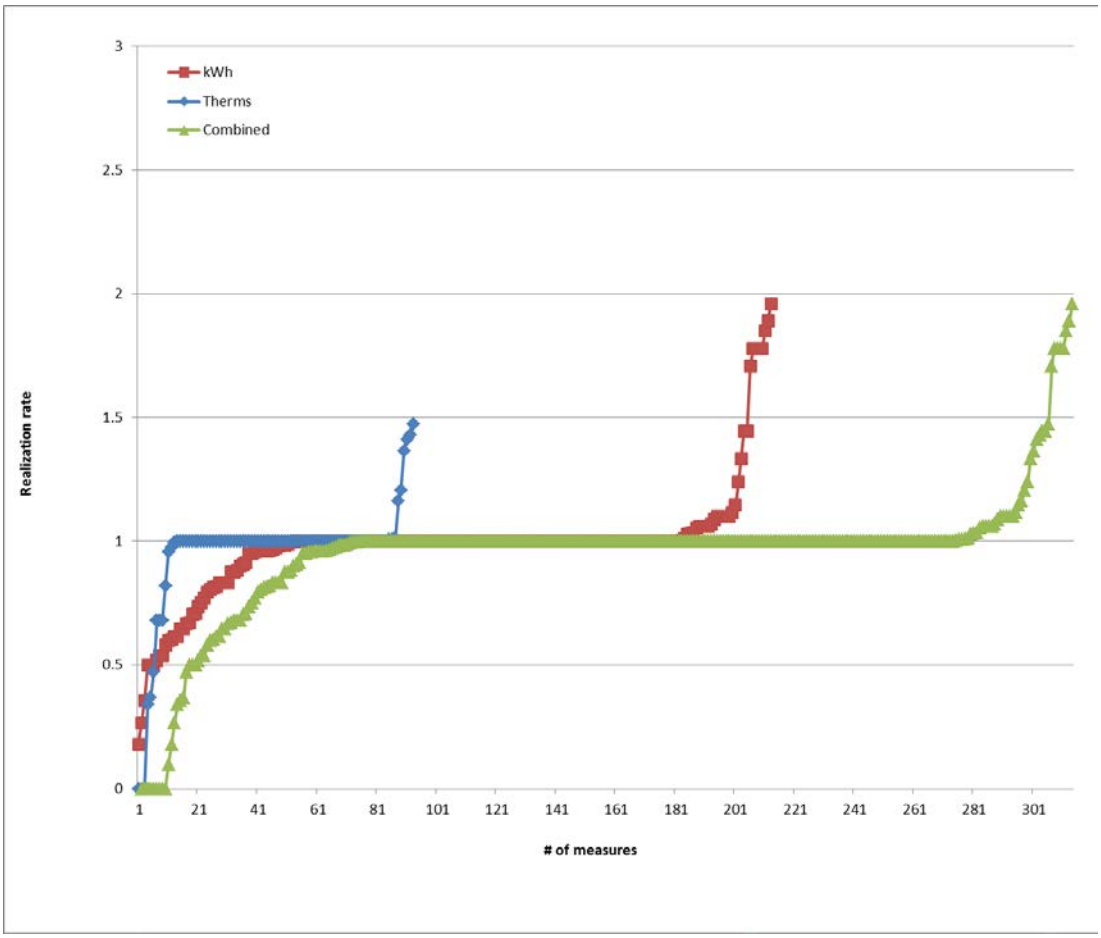

Figure 2: Distribution of Measure Realization Rates.

[1] Assessing the true impact of vacancies is challenging, however, as vacant buildings may be reoccupied, yielding savings again, though these savings could be affected by future remodels.

5.2.1. Measure Type Results

Table 9 shows the aggregated realization rates by measure type for the inspected sample. Note that these results are unweighted, and thus do not reflect portfolio-level results. They are provided here solely to provide insights into whether groups of measures are performing better, worse, or the same as expected. This table also provides corresponding realization rates from the most recent past evaluation, though any comparisons should take into account the fact that possible methodological differences between the two evaluations could explain some of the differences present.

Measure types with lower-than-average realization rates include: Custom Controls (0.79 electric), HVAC (0.82 electric), water heating (0.60 gas) and Food Service (0.0 gas). The zero Food Service savings was due to measure equipment that had been reclaimed after the business was closed. We also compared all lighting measures in aggregate against all non-lighting measures, and found that non-lighting measures did not perform as well as lighting measures.

Table 9: Realization Rates by Measure Type

			2010-11 Realization rates*		2008-09 Realization rates (for comparison)	
Measure Type	Number of evaluated electric measures**	Number of evaluated gas measures**	kWh	therms	kWh	therms
Custom Controls	17	10	0.79	1.00	0.79	0.56
Custom Gas	-	2		1.00		
Custom Other	15	9	0.90	0.93	0.99	0.89
Food Service	-	1		0.00	0.93	0.83
HVAC	59	62	0.82	0.89	0.73	1.07
Insulation	-	4		1.00		1.00
Lighting	120	-	0.93		0.94	
Motors	8	-	1.02		0.99	0.37
Water Heating	-	3		0.60		0.97
Subtotal - Lighting	120	-	0.93		0.91	
Subtotal - Non Lighting	99	91	0.84	0.93	0.89	0.92
Totals	**219**	**91**	**0.86**	**0.93**	**0.90**	**0.92**

* Presented for information only. Measure-type results are not statistically valid

**Does not Include "Pass-thru" measures

Table 9 provides similar aggregated results by building type for the 2010-11 program years. Again, these results are unweighted, and thus do not reflect portfolio-level results. They are provided here solely for insights into building classifications that are performing better, worse, or the same as expected.

Building types with lower-than-average realization rates include: Data Centers (0.75 electric), Hospital (0.63 electric, 0.76 gas), Infrastructure (0.41 electric), and Other Health (0.69 electric). A zero realization rate (i.e., no savings realized) was computed for the Laundry/Dry Cleaners and Restaurant classifications. Both of these building types included one site. The laundry site was found to be vacant and the equipment removed. The measure equipment at the restaurant site had been destroyed and was no longer operational.

Higher-than-average realization rates were computed for the Lodging/Hotel/Motel (1.17 gas) and Retirement/Assisted Facilities (1.41 gas) building classifications. Additional information about the differences between expected and evaluated savings for the sampled cases is provided in section 5.2.2 below.

Table 10: Realization Rates by Building Type

	kWh				Therms			
Building Type	**Number of evaluated measures**	**Expected program savings**	**Realized savings (evaluated)***	**Realization rate***	**Number of evaluated measures**	**Expected program savings**	**Realized savings (evaluated)***	**Realization rate***
Assembly	7	1,077,947	1,077,947	1.00	2	1,680	1,680	1.00
Auto Services	4	236,078	235,460	1.00	-	-	-	
College/University	9	3,762,430	3,700,454	0.98	5	410,047	407,683	0.99
Data Center	3	2,693,231	2,026,292	0.75	-	-	-	
Grocery	16	936,528	951,037	1.02	-	-	-	
Gym/Athletic Club	5	598,141	570,548	0.95	3	30,471	30,471	1.00
Hospital	7	3,711,011	2,320,580	0.63	4	234,325	156,013	0.67
Infrastructure	2	2,590,835	1,062,152	0.41	-	-	-	
Institution/Government	-	-	-		6	92,638	79,592	0.86
Laundry/Dry Cleaners	1	3,966	3,966	1.00	1	7,078	-	0.00
Lodging/Hotel/Motel	1	20,000	20,000	1.00	4	83,110	97,369	1.17
Office	62	10,887,602	9,889,224	0.91	8	277,736	268,171	0.97
Other	8	3,059,970	2,866,495	0.94	7	114,083	114,083	1.00
Other Health	10	31,379	21,569	0.69	1	190	190	1.00
Parking structure/Garage	1	1,165,037	947,693	0.81	-	-	-	
Religious/Spiritual	2	235,273	225,269	0.96	2	30,797	25,255	0.82
Restaurant	-	-	-		2	1,395	-	0.00
Retail	34	5,857,260	5,342,393	0.91	2	14,490	14,676	1.01
Retirement/Assisted Facilities	8	51,187	45,037	0.88	1	5,743	8,110	1.41
Schools K-12	2	1,471,554	1,471,554	1.00	14	179,428	179,428	1.00
Warehouse	37	3,001,740	3,016,876	1.01	-	672	672	1.00
Totals	**219**	**41,391,169**	**35,794,546**		**62**	**1,483,883**	**1,383,392**	

* Presented for information only. Measure-type results are not statistically valid.
Does not Include "Pass-thru" measures

5.2.2. Reasons for High/Low Values

One objective of the evaluation was to determine reasons actual savings for measures varied significantly from expected savings, resulting in particularly high or low realization rates. This information will assist Energy Trust in formulating ways to improve program calculation methods and processes to enhance future realization rates and program cost effectiveness. During the evaluation analysis of realized savings, we noted key reasons for differences between the realized and expected energy savings. These reasons were categorized and documented in the individual site workbooks. The six categories we applied, along with illustrative examples, are listed below:

- **Count**: the number of measure-affected devices actually installed and operational was much higher/lower than expected, e.g., the program application documented 100 lighting fixtures being installed, but only 80 were found, since 20 had been removed as part of a remodel.
- **Control settings**: the controls for a measure-affected device are programmed differently than expected, e.g., a setback thermostat programmed to maintain a 70°F unoccupied temperature, rather than 65°F.
- **Equipment size**: the size of the measure-affected device (e.g., fixture wattage, chiller tonnage, nameplate capacity) is significantly different from the expected value.
- **Equipment efficiency**: the efficiency of the measure-affected device (e.g., a chiller) is significantly different from the expected value.
- **Operating hours**: the measure-affected equipment runs for a longer or shorter period than expected.
- **Other**: other reasons not captured in the previous categories.

In cases where more than one reason existed for differences, we also noted which reason was the primary one. The reasons we uncovered for the sampled measures are summarized in Table 11. Of the 310 measures we inspected, we identified 134 reasons for differences. Some measures had more than one reason, which could increase or reduce savings. Overwhelmingly, these differences tended to reduce savings. The most common reasons for reduced savings were lower operating hours and “Other”. The “Other” category included situations such as measures not actually being installed, vacant buildings[2], inappropriate algorithms, or questionable algorithm inputs. Appendix B provides more detailed measure-level information on the reasons for differences between expected and evaluated savings for cases where the realization rate was less than 0.80 and greater than 1.2.

[2] The evaluation identified one vacant site, which was part of the 2010 kWh sample. It was vacant because the customer went out of business. We did not inspect the measures because of the vacancy, so we do not know if the measure was successfully installed and achieving the expected savings. According to the customer, the affected equipment had been installed and operating but was subsequently removed.

In many cases, reasons for differences were evenly distributed between increased and decreased savings. In the case of equipment count, equipment size, and "other" however, the majority of instances resulted in reduced savings.

Table 11: Frequency of Reasons for High/Low Realization Rates

Sample/ Reason for difference		Primary reason	Secondary reason	Increased savings	Reduced savings
N		**112**	**22**	**50**	**84**
N as % of all inspected measures		**36%**	**7%**	**16%**	**27%**
Reasons	Count	14	5	5	14
	Control settings	15	5	9	11
	Equipment size	4	3	1	6
	Equipment efficiency	6	2	4	4
	Operating hours	44	4	22	26
	Other	29	3	9	23

Does not include "Pass-thru" measures

Nine measures were found to be not installed, which resulted in zero savings (six kWh measures at three sites, and three gas measures at three sites):

- Of the six kWh measures, four were lighting measures representing 3% of the non-installed kWh savings. Two were controls measures representing 97% of the non-installed kWh savings. All together, the non-installed kWh measures represented 1% of the 2010 kWh ex-ante savings and 2% of the 2011 kWh ex-ante savings.
- Of the three gas measures, one was an improperly installed CO2 control, one was a gas fryer than had been destroyed, and one was a laundry facility that had gone out of business in which the equipment had been removed. The latter two were beyond the control of the program. These non-installed therm measures represented 3% of the 2010 therm ex-ante savings and 1% of the 2011 therm ex-ante savings.

5.3. Review of Program Algorithms

The evaluation included a review of program algorithms for custom measures with customized engineering calculations. Excluded from this review were algorithms behind prescriptive deemed savings values or standardized calculation software. We documented the results of each custom algorithm review, and developed, when possible, recommendations for improvements that would increase the accuracy of the savings estimates.

The custom measures review relied on a wide range of techniques to estimate savings, including bin analysis, hourly simulation (eQUEST/Trace 700), utility billing analysis, and proprietary vendor programs. In some cases, the program documentation we were able to obtain was insufficient to fully analyze the underlying algorithms in these analyses. We have,

however, listed below some overarching recommendations regarding how custom savings are calculated.

- Fully document assumptions / inputs into spreadsheet and bin analyses.
- Ensure all necessary files for recreating outputs to hourly simulation models are preserved for full transparency.
- Ensure that files and outputs to hourly simulation models match the reported database savings.

Program algorithms for lighting measures were not formally included as part of this review. Nonetheless, evaluation analysts found that in general, the equations for calculating lighting savings were reasonable. Two possible refinements would be to: (1) include a utilization factor to account for burnt-out lamps, and (2) incorporate factors to account for heating and cooling system interactions[3]. Since lighting heat-cool interactions were not part of the program algorithms, they were not included in the evaluation calculations. As a result, the evaluated lighting savings presented in this report are slightly overstated, though Energy Trust may apply downstream corrections that mitigate or eliminate this effect.

5.4. Annual Program Energy Savings

The results from the individual sample points were extrapolated to two-year (2010-11) program totals using the methods described in Section 3. Program extrapolations were made for gross realized savings in both fuel groups (electric and gas) for each program year. The results from the extrapolation within each of the four sample frames are shown in Table 11. For the 2010 program year, these tables show program-level realization rates of 1.07 and 0.86 for kWh and therms, respectively. For the 2011 program year, the realization rates were estimated to be 0.91 and 1.01 for kWh and therms, respectively. . One measure involved a vacant building in which the equipment had been removed. The measure belonged to the therm 2010 sample frame. If the measure had still been operational, the realization rate for this frame would have changed from 0.86 to 0.88.

[3] These factors might be generic, based on typical expected performance for a given building type and/or climatic region. The Regional Technical Forum (RTF) has developed enhanced sets of these factors.

Table 12: Evaluated Savings by Sampling Domain

Sample frame (program year + fuel saved)	Strata	Strata lower bound (Site kWh/yr. or Therms/yr.)	Number of sites	Sampled cases	Strata weight	Sample expected savings	Sample evaluated savings	Weighted expected savings	Weighted evaluated savings	Realization rate
2010 kWh	9	714,319	15	15	1.0	17,705,291	16,508,851	17,705,291	16,508,851	
	5	303,348	31	3	10.3	1,146,610	1,078,964	11,848,303	11,149,295	
	4	139,624	90	3	30.0	510,699	658,856	15,320,970	19,765,680	
	3	57,964	164	3	54.7	315,892	413,231	17,268,763	22,589,961	
	2	19,807	380	3	126.7	90,752	80,145	11,495,253	10,151,700	
	1	2,855	987	3	329.0	29,169	27,228	9,596,601	8,958,012	
	8	0	303	0		-	-	0	0	
			1,970	**30**				**83,235,181**	**89,123,499**	**1.07**
2011 kWh	9	635,445	19	19	1.0	22,193,893	18,120,606.00	22,193,893	18,120,606	
	5	287,551	52	4	13.0	1,651,686	1,170,848.00	21,471,918	15,221,024	
	4	121,578	103	3	34.3	752,154	741,532.00	25,823,954	25,459,265	
	3	51,270	241	3	80.3	243,961	237,803.00	19,598,200	19,103,508	
	2	17,792	531	4	132.8	92,902	84,186.00	12,332,741	11,175,692	
	1	2,640	1,406	5	281.2	51,033	57,126.00	14,350,480	16,063,831	
	8	0	524	0		-	-	0	0	
			2,876	**38**				**115,771,185**	**105,143,926**	**0.91**
2010 therms	9	51,568	5	5	1.0	545,385	490,008.00	545,385	490,008	
	4	20,973	12	5	2.4	176,804	143,670.72	424,330	344,810	
	3	7,078	26	5	5.2	54,736	49,221.00	284,627	255,949	
	2	2,016	67	4	16.8	14,046	14,046.00	235,271	235,271	
	1	108	282	5	56.4	2,336	1,200.28	131,750	67,696	
	8	0	182	0		-	-	0	0	
			574	**24**				**1,621,363**	**1,393,733**	**0.86**

Sample frame (program year + fuel saved)	Strata	Strata lower bound (Site kWh/yr. or Therms/yr.)	Number of sites	Sampled cases	Strata weight	Sample expected savings	Sample evaluated savings	Weighted expected savings	Weighted evaluated savings	Realization rate
2011 therms	9	30,841	10	10	1.0	660,984	653,813.43	660,984	653,813	
	4	13,140	19	3	6.3	69,311	68,554.00			
	3	5,008	45	3	15.0	23,817	26,184.00	357,255	392,760	
	2	1,669	136	4	34.0	13,132	13,620.00	446,488	463,080	
	1	171	578	7	82.6	2,993	2,736.00	247,136	225,915	
	8	0	114	0		-	-	0	0	
			902	**27**				**1,711,863**	**1,735,569**	**1.01**

Combining the sample frames yields the annual kWh and therm savings shown in Table 12. This table also compares the 2-year realization rates from this evaluation to the equivalent rates from the previous evaluation of this program. The comparison shows higher kWh and therm realization rates in 2010/11 when compared to 2008/09 for both fuel types. The kWh realization rates increased from 0.90 to 0.98 and the therm realization rates increased from 0.81 to 0.94.

Table 13: Evaluated Savings by Program Year

		Expected program savings		Realized savings (evaluated)		Realization rates	
Program year	Number of sites	kWh	therms	kWh	therms	kWh	therms
Previous evaluation cycle							
2008	1,170	42,105,793	862,294	41,887,080	746,564	0.99	0.87
2009	1,590	74,426,951	941,618	63,537,310	705,644	0.85	0.75
Total	**2,760**	**116,532,744**	**1,803,912**	**105,424,391**	**1,452,208**	**0.90**	**0.81**
Current evaluation cycle							
2010	2,544	85,813,714	1,729,547	91,884,445	1,486,729	1.07	0.86
2011	3,778	108,759,845	2,118,681	98,776,194	2,148,020	0.91	1.01
Total	**6,322**	**194,573,559**	**3,848,228**	**190,660,639**	**3,634,749**	**0.98**	**0.94**

6. CONCLUSIONS AND RECOMMENDATIONS

From the results of the impact evaluation of gross realized savings for the 2010-11 program years, the following key conclusions were drawn.

1. **Installation Verification** – The evaluation verified through field inspection that 97 percent of the sampled and assessed measures were either fully or partly installed and operational. When they occurred, partial installations and variations from the expected measure operation (differing schedules, set points, etc.) accounted for many of the differences between expected and evaluated savings. Oftentimes, the timing of when discrepancies occurred (e.g. when a particular set point had been changed, etc.) and reasons for the discrepancies we encountered (e.g. why did a particular set point vary from the original design, etc.) were unclear.

2. **Gross savings realized** – Significant gross energy savings were found for both fuel types in each program year. Unweighted measure-level results show a range in realization rates from zero to nearly two. Domain-level results for the electric and gas fuel types indicate that realized savings were less than expected savings for both fuel types in both program years. The electric realization rate was estimated to be 0.99 across the two program years. The therm realization rate was estimated to be 0.94 for the two-year period. These results represent a significant improvement from the previous evaluation cycle. Table 14 shows the confidence level and relative precision for each sample frame.

Table 14: Sample Frame Confidence Level and Relative Precision

Sample Frame	Confidence Level	Relative precision
kWh2010	90%	11%
kWh2011	90%	8%
Therms2010	90%	9%
Therms2011	90%	3%

3. **Major reasons for differences between the expected and realized savings –** The realized and expected gross savings differed for a variety of reasons. In most cases the differences tended to reduce savings. The most common reasons for reduced savings were lower operating hours and "Other". The "Other" category included situations such as measures not actually being installed, vacant buildings, inappropriate algorithms, or questionable algorithm inputs.

4. **Algorithm review –** In many instances, custom algorithms could not be reviewed due to a lack of documentation. When they could be reviewed, the custom algorithms were generally found to be reasonable. Custom algorithm types included eQUEST models, TRACE 700 models, and MS Excel spreadsheet calculation models. Typically, these custom algorithms were applied to complex HVAC measures such as controls or central plant equipment replacement measures. Some general recommendations were made for reviewed algorithms where improvements were appropriate.
5. **Project documentation –** Major improvements were made by the Project Management Contractor, since the last evaluation cycle, in providing the evaluation with complete and accurate documentation of the sampled projects. However, additional improvements can still be made for future evaluations. These include providing electronic copies of all calculations, and ensuring that provided documentation versions match the reported savings values.

Findings from this evaluation led to the following recommendations:

1. **Feedback to Energy Trust savings estimates**. The Energy Trust estimates of expected savings predicted the future performance of a measure before it is installed. The evaluation estimated realized savings for a measure under as-built and operated conditions. Because of these different perspectives, differences between expected and realized savings are unavoidable. The evaluation was able to gauge measure conditions at a single point in time, but further in-depth study of measures in the evaluation sample, where significant differences in the estimates were found, can improve the ability of Energy Trust to predict savings and/or ensure good performance for future measures. It is recommended that Energy Trust carefully study these cases, such as by revisiting sites and speaking with customers and vendors to understand better the reasons why certain measures performed poorly. Information from such customer follow-up might lead to improved procedures for inspection, quality control, and training, which in turn may increase realization rates in future evaluations.

   Areas to focus on may include the following underperforming measure categories:

   - Hospitals: This included three sites. One site was a large zero saver in which the controls measures were not installed. The controls firm had gone through staff reductions and the work was never completed, even after multiple requests from the customer. The post inspection did not reveal this. The second case involved a fan night setback routine that was not programmed. Both of these cases involved controls measures that were not implemented and that were not caught during post-inspections. The third site did not have major issues.
   - Data Centers: All three sites in this category had reduced realization rates. In two cases, controls were the primary issue. In the third case, the installed equipment count was too high (a unit that had been installed as a backup was assumed to be operating).

- Infrastructure: One site was included in the sample for this category. A lower realization rate resulted from an incorrectly calculated baseline.

Issues with controls measures appear to be more prevalent. Controls measures can be difficult to evaluate because many times there is nothing physical to inspect. The control routines and set points within the system must be inspected. This requires familiarity of the control system either by the evaluator or a knowledgeable site contact. Controls measures also tend to be operated outside of design parameters more so than other measure types because they can be easily changed by operators, and the changes are not visible but can have a significant effect on savings, usually in the negative direction.

Another theme that has appeared has been claiming savings for equipment that was installed as a backup (as described in the data center site above). A new piece of equipment, for instance a chiller, will be installed to handle the load for the facility. A second identical unit will also be installed to provide backup in case of failure of the primary unit. Both do not operate at once, but oftentimes we see that savings are claimed for both. A thorough knowledge of the control sequence is necessary to ensure that the savings are handled correctly.

In two cases where zero savings were found, the evaluators determined that the reasons were outside the control of the Energy Trust. In one case, a company went out of business and the measure equipment was removed. In a second case, measure equipment at a site was destroyed.

2. **Project documentation**. The value and cost-effectiveness of the evaluation was very dependent upon access to accurate and complete program documentation for each sampled project. Documentation should be provided to the evaluator in sufficient detail for an independent third party to be able to understand expected measure performance. It should include the information and tools necessary to understand the algorithm that was used to calculate the expected savings, in an electronic format necessary to reproduce the savings estimate.

   Adequate documentation was provided for most sampled projects. The completeness and accuracy of the documentation was an improvement from the previous evaluation cycle. However, there were a number of cases where sufficient documentation was not provided. This was especially true for non-lighting projects where the expected savings were calculated with an hourly simulation or proprietary software. Compensating for this lack of documentation significantly increased the cost of the evaluation. It is recommended that the Energy Trust improve the completeness of the project documentation in future program years so that the evaluator can be supplied with the information and tools necessary to cost-effectively complete the evaluation scope specified by Energy Trust. This includes items such as measure performance specifications, documentation of assumptions made and calibration methods used during the analysis of expected savings, and backup information related to the calculations made.

   All spreadsheet calculation tools should be provided in Excel format (not pdf) so that the equations can be understood and the savings estimates can easily be reproduced. When hourly simulations tools are used, documentation should include the final as-built energy modeling files, in

electronic format, that are needed to reproduce the expected savings analysis. It is also recommended that the use of proprietary software be minimized.

The most expeditious way to provide program files to future evaluators would be to store all pertinent files for a given site in a well-documented electronic folder, which could be passed on in its entirety should that site be sampled in the evaluation.

When the folder is first received from the ATAC, it should be thoroughly checked to ensure that all necessary material is included and that the versions of the supplied documentation match exactly the reported savings values in the database. This check should be completed for every site.

3. **Measure interactive effects**. For interior lighting measures, the expected savings did not capture interactions with the HVAC system, when they were relevant. It is recommended that the expected savings methodologies be upgraded to capture interactive effects where they are significant.
4. **Program Communications**. Some customers did complain about the number of times they had previously been asked to allow sites visits as part of other aspects of program implementation. It is recommended that the program communicate clearly that evaluations are happening (in some cases multiple evaluations at the same site) and required to improve the programs as well as support ongoing funding.