

Measure by Measure: the Real Reasons for Gaps in Claimed and Evaluated Savings

Al Lutz and Vishy Tirumalashetty, Itron, Inc.

ABSTRACT

Energy efficiency programs can deliver on large energy efficiency savings goals, but when they don't, the reasons for the gap needs to be examined. Many projects of various types were evaluated through California's Investor Owned Utility (IOU) sponsored energy efficiency programs for program years 2006 through 2008. Analysis of these individual projects, mainly involving lighting and Heating, Ventilation and Air Conditioning (HVAC) systems, revealed that the technical or gross savings gap at a specific site was due to a combination of factors at each site, including improper assumptions about operating hours, operating load, or the applicable baseline. This paper highlights the primary factor in the savings gap for different types of measures, concentrating on the gross savings claims (as opposed to the net-to-gross or net-of-free-ridership claims). The primary factor affecting energy savings for projects in various programs is identified, and the results are aggregated by end use and measure type. Identifying the importance of these factors can lead evaluators and implementers in better focusing on and documenting the parameters surrounding those factors for a particular measure. This topic is especially relevant as utility energy efficiency goals change and as utilities nationwide are being rewarded for their efficiency efforts – and penalized when they fail.

Importance of Energy Evaluation in the Buildings Sector

The buildings sector accounts for a large portion of total annual national energy use (EIA 2010). Energy use reductions in this sector have been the target of energy savings efforts and program offerings by governments and utilities. The evaluation of any program is important in ascertaining the achieved results as compared to the claimed results, and the evaluation of energy programs targeted at the buildings sector is no exception. The diversity of programs across different subsectors and technologies, along with multiple project drivers and market actors, make these evaluation efforts particularly important and instructive in determining the important factors to concentrate on when claiming – and verifying – energy savings. This attention is particularly useful in ensuring the ex-ante claimed and ex-post verified energy savings are accurate and defensible. A reliable estimate of the energy savings from an equipment-retrofit project is important for utility companies, performance contractors, and building owners (Yalcintas 2008).

This paper focuses on the commercial and institutional portion of the buildings sector. The sector is diverse and comprises a large percentage of total building energy use. Commercial and institutional buildings are generally larger than buildings comprising the residential sector (with the exception of some multi-family housing complexes) and entail greater energy use per building.

Energy impact evaluation reports authorized by the California Public Utility Commission (CPUC) Energy Division (ED) for the 2006-2008 evaluation cycle were reviewed for this study. We analyzed the impact evaluation reports for four contract groups; these are publicly available

at www.calmac.org. Further, we reviewed a large number of site specific energy project impact evaluations that formed the basis for these reports. This review enabled a deeper assessment and quantification of the frequency for the possible reasons for the savings gap that exists between claimed and evaluated savings. The results and findings of the energy evaluations of energy efficiency programs in the program years 2006 – 2008 conducted for the CPUC (ADM et al. 2010; Itron et al. 2010; SBW Consulting et al. 2010; Summit Blue et al. 2010) are examined to assess the critical areas of importance for improvement in energy savings estimation and verification.

When we consider the program induced energy savings, it is important to realize that there are both gross impact and net impact components. These components lead to gross realization rates, net to gross ratios, and net realization rates combining those factors. This paper concentrates on the gross energy savings and gross realization rate.

The gross impact component of energy savings refers to the savings technically possible through the implementation of the measure. These are the savings that an energy end user or program participant most commonly associates with energy savings. The gross realization rate is calculated by dividing the ex-post gross savings by the claimed ex-ante savings. Gross realization rates were calculated and are shown in this paper for different measure types based on site specific energy project realization rates. The measure type realization rate was calculated as a simple average from the site specific project gross realization rates collected from the site specific project impact evaluations that form the basis of the impact evaluation reports reviewed for this paper (ADM et al. 2010; Itron et al. 2010; SBW Consulting et al. 2010; Summit Blue et al. 2010)

Research Performed to Determine Factors for the Energy Savings Gap

Gross impacts are typically determined for larger commercial energy projects through application review, project documentation review, site visits, measurement and verification (M&V) activities, and on-site interviews. Metering/monitoring may be accompanied by data collection of pre-retrofit and post-retrofit operating conditions. The evaluators determine an ex-post savings figure, reflecting the gross energy savings found based on actual ‘as found’ operating conditions. The evaluators attempt to ascertain the most accurate baseline condition to determine actual grid-level impacts over the life of the installed measure, reflecting the pre-retrofit operating conditions which would have prevailed in the absence of the program. In some cases, a different or more accurate methodology for energy savings may be used, perhaps involving the knowledge of post installation energy use or other parameters not available during project implementation and ex-ante savings estimation. Adjustments are included for occupancy, baseline, weather and other parameters that have changed from the pre-retrofit conditions to the post-retrofit conditions. Information may be collected on the remaining useful life (RUL) of the equipment and the normal replacement schedule for equipment, in order to differentiate retrofits where the program induced early equipment replacement verses a normal replacement at the end of equipment life. In the case of early equipment replacement, dual baselines are appropriate. The baseline until the end of the useful life of the equipment is the existing in-situ equipment; the baseline after the end of the remaining life is the code required or standard practice replacement equipment. More details and a description of dual baseline are provided in a recent decision for the CPUC addressing this and other topics (ALJ 2011).

Site specific project reports for four contract groups in the CPUC funded energy evaluations for the 2006-2008 cycle were analyzed, *across contract groups*, in order to compile a relatively large number of projects in each measure category; this was done to enable a deeper assessment of the possible reasons for the savings gap that existed between claimed and evaluated savings. The evaluated contract groups included 129 projects. Program descriptions from abstracts on the CALMAC website and the number of projects per program are summarized as follows:

- **Commercial Facilities [67 projects] (ADM et al. 2010)**
The Commercial Facilities Contract Group evaluated two market-sector focused incentive programs. This report presents the energy impact results for the Pacific Gas and Electric (PG&E) High Tech and Large Commercial customer segments. The programs claimed savings through equipment retrofit, retrocommissioning and new construction activities with particular emphasis on providing design assistance for installing custom energy efficiency measures.
- **Major Commercial [43 projects] (SBW Consulting et al. 2010)**
The Major Commercial contract group includes four measure categories (Custom Lighting, Custom HVAC, Custom Other and Audit) within commercial, industrial, and agricultural programs that were implemented by Southern California Edison (SCE), Southern California Gas (SCG) and San Diego Gas and Electric (SDGE).
- **Government Partnerships [13 projects] (Summit Blue et al. 2010)**
This report presents the evaluation results for the energy efficiency projects and programs in the governmental sector. This report summarizes the impact evaluation activities related to following subset of partnerships: the University of California, California State University (UC/CSU) Partnership Program, the California Community Colleges (CCC) Partnership Program, and the Palm Desert Partnership Program. The measures evaluated in this program include: early retirement of residential air conditioners, refrigerant charge and airflow (RCA) adjustment for both residential and commercial refrigerators.
- **PG&E Fabrication, Process and Manufacturing [6 projects] (Itron et al. 2010)**
This report presents the evaluation results for the energy efficiency projects and programs within the scope of the California Public Utilities Commission's (CPUC) Pacific Gas and Electric Fabrication, Process, and Manufacturing (PG&E Fab) evaluation contract group. PG&E carried out a very large and extensive implementation effort in the industrial sector. There are a number of specific findings that help to explain why the evaluated ex-post savings estimates are significantly below the ex-ante estimates. These include the following: errors in baseline determination; inadequate basis for claims of how some measures save energy; inadequate enforcement of program and policy rules; inadequate consideration of total system energy analysis; effect of the economic recession; and high free ridership. The evidence also indicates that program claims were made on some projects that customers initiated for non-energy savings reasons and for which there was no alternative ever considered.

The evaluation reports and site-specific impact reports from these programs were studied to provide the realization rates by project and to reveal the primary reasons for the difference between claimed and verified energy savings. From this review, projects from the four programs

were segmented into end use and the major type of measure. These measure categories and a brief description are included in Table 1.

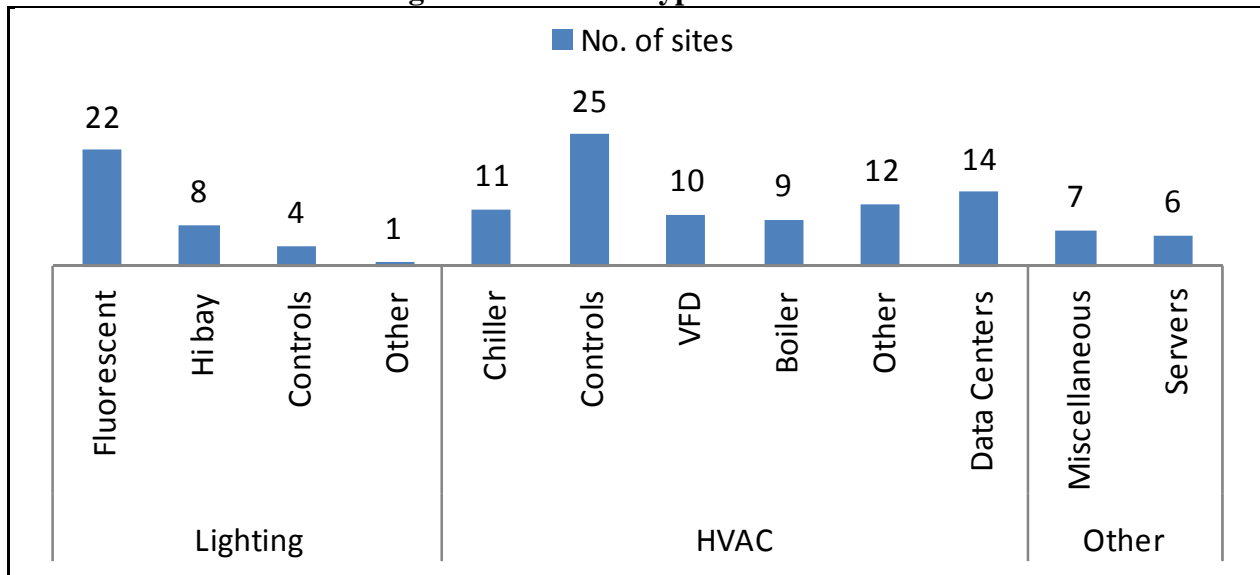
Table 1. End Use, Measure Category, and Measure Description

End Use	Measure Category	Description
Lighting	Fluorescent	Linear Fluorescent Replacements (Low Bay)
	High Bay Fixtures	High Bay Fixtures – HID and Fluorescent
	Controls	Occupancy Sensors / Daylighting Control Systems
	Other	Other Fixtures – LED's
HVAC	Chiller	Chiller Replacements / Major Modifications
	Controls	Simple Dedicated Controls and EMS/BAS Systems
	VFD	Air Side and Water Side Variable Speed Drives
	Boiler	Boiler Replacements / Major Modifications
	Other	Dx units, Economizers, Other Equipment
	Data Centers	Data Center HVAC System Measures
Other	Miscellaneous	Domestic Water Booster Pumps, Hot Water Systems, Pool Covers
	Servers	Replacement and Control of Servers

Source: ADM et al. 2010; Itron et al. 2010; SBW Consulting et al. 2010; Summit Blue et al. 2010

As discussed above, site reports were reviewed from the Commercial Facilities, Major Commercial, Government Partnerships and PGE Fabrication programs. Only measures broadly applicable to commercial buildings were included. Any measures that were identified to be industrial-focused were not included in the analysis. A total of 129 projects were selected from these programs for analyzing the reason for savings gap between the claimed and the evaluated savings. Figure 1 below shows the measure type distribution based on the review of the 129 projects selected from these programs.

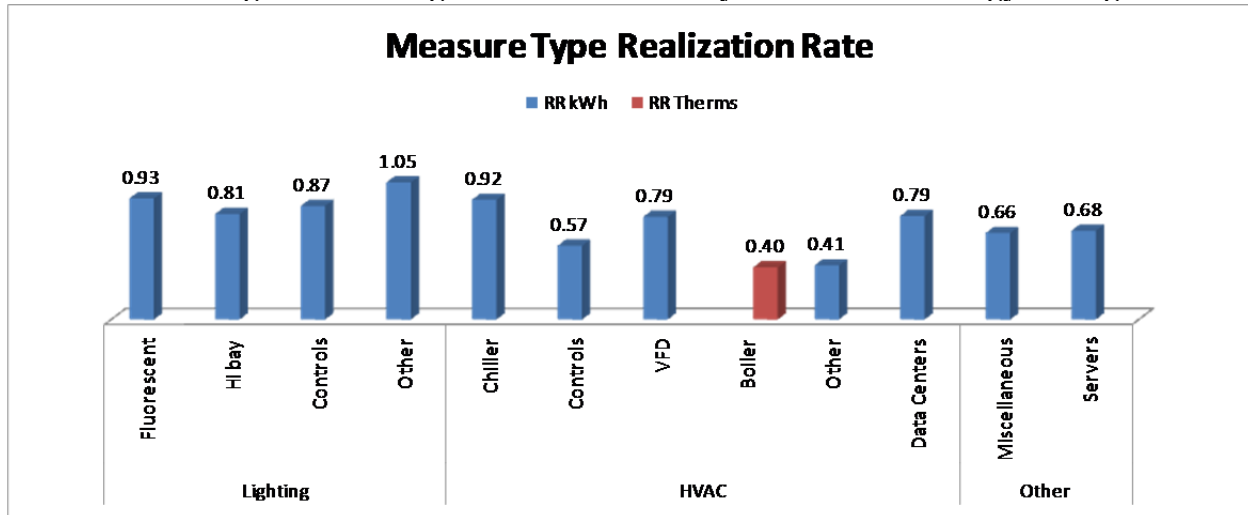
Figure 1. Measure Type Distribution



Source: ADM et al. 2010; Itron et al. 2010; SBW Consulting et al. 2010; Summit Blue et al., 2010

We calculated the average gross realization rates for each of these measure categories (types) based on site specific (project) gross realization rate obtained from the four evaluation programs for each measure type. The measure type gross realization rate was calculated as a simple average from the site specific realization rates collected from the site specific impact evaluations in the four program reports. Figure 2 displays the average energy savings realization rates calculated for each of these measures. Note that for the boiler measure, savings realization rate reported is for gas savings.

Figure 2. Average Realization Rate by Measure for Energy Savings



Source: ADM et al. 2010; Itron et al. 2010; SBW Consulting et al. 2010; Summit Blue et al., 2010

For each of the measure types identified in Table 1, the site specific project impact evaluations for the projects in that measure type were analyzed. A primary reason for a gap in claimed and evaluated in energy savings was noted for each project. These reasons and a brief description of each reason along with the frequency of their occurrence are provided in Table 2.

Table 2. Description and Frequency of the Primary Reasons for Savings Gaps

Reason	Description	Frequency of Primary Reason for Gap
Ex ante calculations (RR < 100%)	Ex ante calculations revised based on available information	12%
Ex ante input values (RR < 100%)	Estimates revised based on ex-post values (excluding operating hours)	16%
Baseline selection	Baseline selection modified based on code or standard practice	17%
Operating hours	Operating hours revised based on more recent ex-post values	14%
Lack of functioning equipment	Lack of functioning or operable equipment or systems; equipment or system modifications not used	6%
Ex-ante input values (RR > 100%)	Realization rate greater than 100% due primarily to change in input values	23%
Ex-ante calculations (RR > 100%)	Realization rate greater than 100% due primarily to change in savings calculation methodology	12%

Source: ADM et al. 2010; Itron et al. 2010; SBW Consulting et al. 2010; Summit Blue et al. 2010

Summarizing Measure Specific Reasons in the Gap between the Claimed and Evaluated Savings

The following sections describe in more detail several key areas that contributed to the gap between gross impact claimed and evaluated savings, i.e., the difference between the gross savings claimed in the ex-ante case and the gross savings found attributable to the energy efficiency programs in the ex-post case. Other authors have previously highlighted improper use of input values, simplified calculation methods, and improper equipment installation as causes of a gap in gross realization rates (Nadel & Keating 1991). In this section, we expand on these reasons and examine the frequency of their occurrence.

Ex-ante Values Different than Ex-post Values (including Operating Hours)

One of the most significant reasons for differences between ex-ante and ex-post estimates of savings was a change in equipment operating parameters, schedules and profiles. The ex-ante calculations sometimes utilize overestimated or underestimated critical parameters such as operating hours or kW reductions; in such cases the ex-post savings were significantly different than the ex-ante claimed savings. Such instances were found at 53% of the sites examined (68 of 129 sites). Of these 68 sites, a change in operating hours caused a lower realization rate in 18 cases. Changes in ex ante values caused an increase in realization rates (i.e., realization rates over 100%) at 30 of the 129 sites examined, or about 23%. Also, changes in ex-ante values resulted in a decrease in realization rate at 20 of the 129 sites examined, or about 16%. One practitioner indicates that the main challenge in determining the energy savings from an equipment retrofit is properly identifying the data (before and after a retrofit takes place) that is to be used in the analysis (Yalcintas 2008).

Improper Baseline Specification

Another significant reason for differences between ex ante and ex post estimates of savings was improper baseline specification. Often times, existing or in situ systems are used as the baseline, although in many instances this will not be an appropriate baseline claim. The correct baseline for energy evaluation purposes would be the operating system that would have been installed without the program. This is typically governed by codes, industry standards and common practices.

Consider a case where standard practice has changed the normal baseline for an energy end use. Commercial lighting has moved from standard electromagnetic ballasts to electronic ballasts and from T12 lamps to T8 lamps, with higher efficiency T5 lamps being a step above the industry standard or standard practice. For a new installation, the more efficient T8 lighting system would typically be an appropriate baseline. For a retrofit application, the T8 system would also constitute the baseline if the existing system was replaced at the end of its useful life or on burnout (replace on burnout). If existing systems have functional remaining useful life (RUL) and the retrofit was an early system replacement, the verified savings could be based on the existing equipment for the RUL period.

A total of 22 of the 129 cases examined – about 17% - showed the baseline as a cause for claimed savings being greater than verified savings. This percentage may understate the

importance of this factor, however, as the lack of proper baseline specifications can lead to a complete disqualification of claimed energy savings.

Lack of Functioning Equipment or Systems

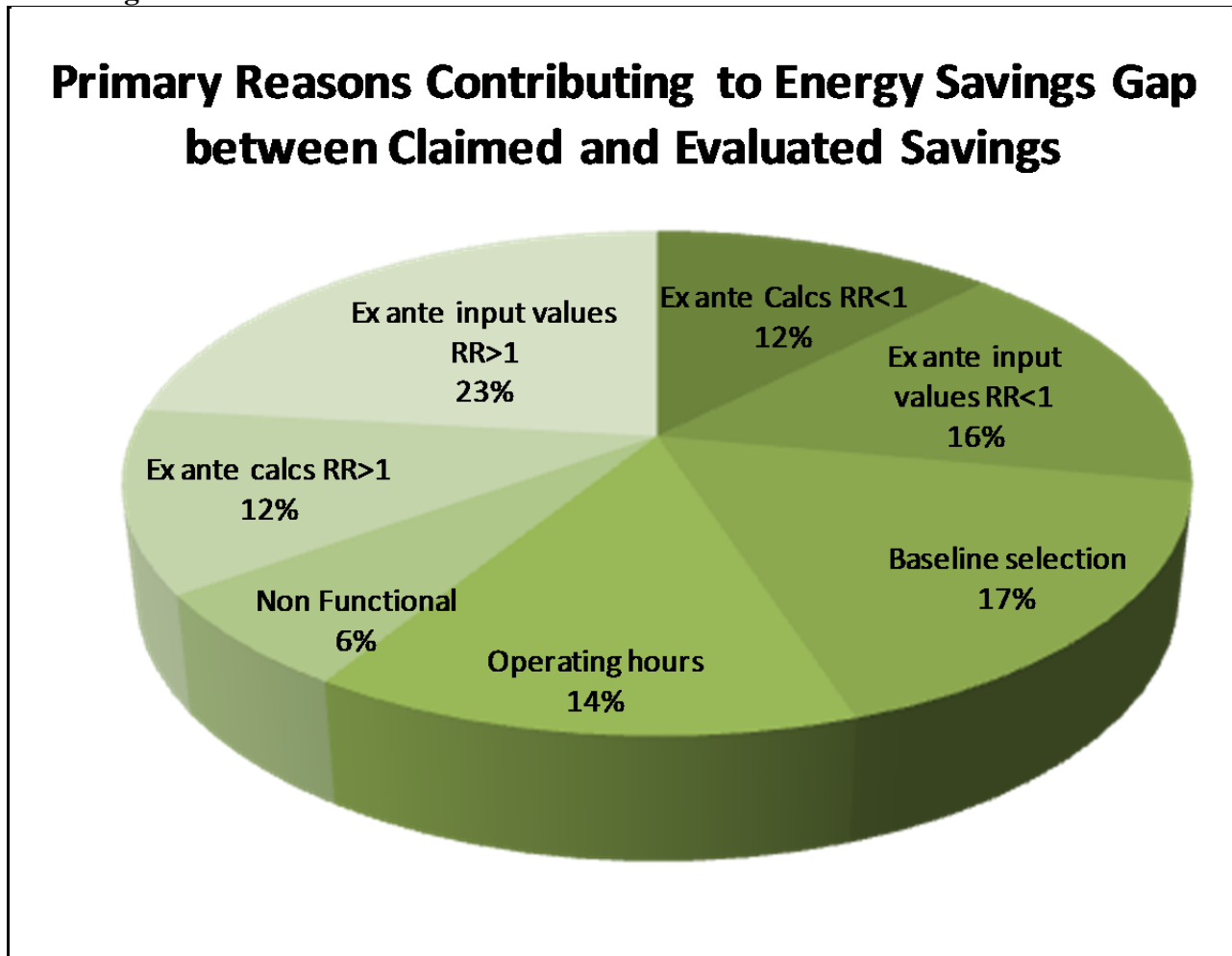
There were found to be instances where the retrofit failed to work as intended and the existing system condition was reverted to immediately after the retrofit. In one case, a high tech manufacturer operating a commercial building did not wish to risk any negative ramifications – and reverted to Heating, Ventilation and Air-Conditioning (HVAC) system operation and chiller variable speed pumping as had existed prior to the EMS control system installation. The recurrence of equipment or systems that were not installed, were removed, or did not function as intended highlights the importance of verification efforts. A total of 8 of the 129 cases examined – about 6% - showed the lack of functioning systems to be a primary cause for claimed savings being lower than verified savings. This percentage again may understate the importance of this factor, however, since the lack of functional equipment and systems can lead to a complete disqualification of claimed energy savings.

Change in Ex-ante Calculations

Interestingly, a change from the ex-ante calculations employed in determining the claimed energy savings resulted in an almost equal number of projects that showed an increase and a decrease in realization rates. Of the 129 projects evaluated, the verified savings at 15 sites increased as a result of different calculation method, while savings at 16 sites decreased as a result of this change. Overall, 24% of the examined projects were affected. Authors have demonstrated that the type of methods used for conducting M&V on installed measures varies, with partially measured retrofit isolation producing realization rates closer to unity than other methods (Kaufman & Palmer 2009, 19).

Figure 3 summarizes the percentage for each primary reason to the savings gap between the claimed savings and the evaluated savings discussed above. The largest primary reason for the savings gap was changed input values and operating hours (53%), while next in importance was baseline selection (17%).

Figure 3. Reasons for the Variation on Realization Rates across All Measures



Source: ADM et al. 2010; Itron et al. 2010; SBW Consulting et al. 2010; Summit Blue et al., 2010

Measure-Specific Realization Rates and Reasons for Differences between Claimed and Evaluated Savings

The various end use categories and measure types along with the average realization rates and the primary reasons for any savings gaps are shown in Table 3. The realization rates are averages across sites only and not weighted by energy savings; these rates are meant to highlight the relevance of reasons to measure types. This parameter helps to pinpoint measure categories that estimators may want to focus on to achieve higher realization rates and better estimates. The prevalence of each primary reason is instructive across end uses and measure categories.

Table 3. Prevalence of Reasons and Realization Rates across Measure Types

End Use	Measure Type	No. of Sites	RR kWh	RR kW	RR Therms	Ex- ante calcs RR<1	Ex- ante input values RR<1	Baseline selection	Operating hours	Non Functional	RR>1 Ex- ante calcs	RR>1 Ex- ante input values
Lighting	Fluorescent	22	0.93	0.77	-	2	1	1	5	0	0	13
	High Bay	8	0.81	0.89	-	0	0	0	5	0	0	3
	Controls	4	0.87	0.57	-	0	0	0	2	1	0	1
	Other	1	1.05	0.93	-	0	0	0	1	0	0	0
HVAC	Chiller	11	0.92	0.67	-	1	2	4	1	0	2	1
	Controls	25	0.57	1.44	0.80	5	1	1	3	1	6	8
	VFD	10	0.79	0.85	-	3	1	1	0	2	3	0
	Boiler	9			0.40	1	1	3	0	1	3	0
	Other	12	0.41	6.52	-	1	6	4	0	1	0	0
	Data Centers	14	0.79	0.47	1.32	2	5	5	0	1	1	0
Other	Misc	7	0.66	1.03	0.44	1	0	2	1	1	0	2
	Servers	6	0.68	0.63	-	0	3	1	0	0	0	2
Total		129				16	20	22	18	8	15	30

Source: ADM et al. 2010; Itron et al. 2010; SBW Consulting et al. 2010; Summit Blue et al. 2010

Lighting measures had a high realization rate across measure types. The analysis indicates that operating hours were a large cause of the realization rate gap. Interestingly, input values caused a large number of projects to have realization rates higher than 100%. Thus, estimators may chose to focus on hours and other input values to obtain better estimates of savings. On the other hand, there were not many cases where an inappropriate calculation method (the reason listed as ‘ex-ante calculations’ was identified as a primary reason for any savings gap, and no changes – or only minor enhancements - in that area may be appropriate.

The HVAC end use is instructive. Chillers showed a high realization rate for kWh and a somewhat lower rate for kW. In this instance, the analysis highlights baseline as an important consideration. HVAC measures in data centers also highlight the importance of baseline, as well as the importance of input values. From the above summary, attention to the ex-ante (and ex-post) calculation method may be an important area to concentrate on for HVAC controls and VSDs in HVAC applications.

For non-HVAC data server modifications and replacements, it may be instructive to concentrate on input values as the best way to accurately estimate savings.

Suggestions for Reducing the Gap between Claimed and Evaluated Energy Savings

Significant energy efficiency goals (in terms of the physical units of energy saved) for energy efficiency exist overall and for many programs in California’s energy efficiency portfolio. The targets for commercial energy efficiency programs are high. These goals, in the authors’ opinion, may at times cause a high estimation of energy savings, which more closely relates to a maximum value verses a realistic expected value. How do we reduce the size of this discrepancy between claimed savings and evaluated savings? Focusing on gross energy savings

estimation, there are a number of questions the participants and implementers can ask as savings estimates are prepared.

- Is there an alternate more realistic baseline than the baseline presently selected?
- What changes are expected in the near future? Are input values realistic for the present conditions and into the future, in light of any technical or behavioral changes that can be expected?
- Are all interactions - HVAC and other interactions – accounted for?
- Is the technical basis for each savings claim documented and believable?
- Are there other calculation methods that give better estimates of kW savings in peak periods, kWh savings and gas savings?

The following summary list of recommendations describes baseline and other issues that may assist in closing the gap between claimed and evaluated savings, resulting in the evaluated ex-post results being closer to the claimed ex-ante savings. These recommendations apply to various market actors, including utilities, participants, implementers, regulators, and evaluators.

- Improve baseline specifications by explicitly identifying the project and identifying the remaining use life of existing equipment
- Consider using early project baseline screening for larger projects and those with policy issues such as fuel switching, self generation, and greenhouse gas impacts
- Carefully reviewing qualifying measures for each program and eliminate eligibility for those that are standard practice
- Making more conservative assumptions, including values used in savings calculations, based on present and expected operation
- Placing measures with an inadequate empirical basis for savings estimates in the emerging technologies program
- Incorporating greater levels of real-time measurement and pre- and post-installation measurement based verification for the most uncertain measures
- Using a true up period to confirm operation after savings are claimed may be appropriate if it complies with policy requirements related to timely filing of claimed – and amended – results

Relevance of the Research and Future Directions

The commercial energy savings potential is a critical component for the utilities to achieve targeted energy savings goals. Commercial energy savings measures do, however, face gaps in the evaluated versus the claimed energy savings. This paper highlights the reasons for the gap between the ex-ante and ex-post savings and breaks it down by measure, providing recommendations that can bridge this gap and improve program claimed savings. This paper should guide the energy efficiency community in better understanding the factors that need to be addressed to ensure that the accuracy of the utilities' projections for achieved energy savings is improved.

This assessment may help in identifying improvements to the design, optimization, and implementation of energy efficiency programs, particularly in the commercial sector.

Several areas where further work may be fruitful are indicated. It may be beneficial to expand the examined projects to get a better representation by measure. The weighted realization rates and the range of realization rates, segmented by measure, will help utilities determine the importance of that measure to their savings goals. Finally, a very detailed analysis into all relevant reasons for the gap in evaluated and claimed savings, and not just the primary reasons, with a weighting on each, would help all stakeholders understand the complexities inherent in determining energy savings and the numerous and varied reasons why differences can exist.

Attention to the factors discussed in this paper can reduce the gap between ex-ante and ex-post gross savings estimates. Use of these findings can lead to energy efficiency projects incorporating gross ex ante estimates and savings claims that are highly accurate and that result in relatively high overall gross realization rates, thereby closing the gap between claimed and evaluated results. This will allow energy efficiency programs to truly meet their goals and to continue their offerings in the future with greater confidence in the promised energy savings.

References

- ADM, Innovologie LLC, C. J. Brown Energy, P.C., Claridge, David Ph.D., Marketing Excellence. 2010. *Commercial Facilities Contract Group 2006-2008 Direct Impact Evaluation Report*. Study ID: CPU0016.01. www.calmac.org. Sacramento, Calif.: ADM Associates, Inc.
- ALJ Decision 11-07-030 July 14, 2011. 2011. Third Decision Addressing Petition for Modification of Decision 09-09-047. http://docs.cpuc.ca.gov/published/FINAL_DECISION/139858.htm
- [EIA] Energy Information Agency, August, 2010. *Annual Energy Review 2009; Report No. DOE/EIA-0384*. U.S. Energy Information Agency, www.eia.gov/aer. Washington, D.C.
- Itron, Inc., Energy and Resource Solutions, Energy Metrics, Mad Dash, Inc., Michaels Engineering, PWP, Inc., Katin Engineering. 2009. *2006-2008 Evaluation Report for PG&E Fabrication, Process and Manufacturing Contract Group*. CPU0018.01. www.calmac.org. Oakland, CALIF.: Itron, Inc.
- Kaufman, N. & Palmer, K. 2010. “Energy Efficiency Program Evaluation: Opportunities for Learning and Inputs to Incentive Mechanisms”. Report RFF DP10-16. Washington, D.C.: Resources for the Future.
- Lutz, A., Bradley, K. & Tirumalashetty, V. 2011. “Would the True Industrial Energy Efficiency Savings Please Stand Up?” *In Proceedings of the ACEEE Summer Study for Energy Efficiency in Industry*. Niagara Falls, NY. American Council for an Energy-Efficient Economy.
- Nadel, S.M. and Keating, K.M. 1991. “Engineering Estimates vs. Impact Evaluation Results: How Do They Compare and Why?” Research Report U915. Washington, D.C.: American Council for an Energy-Efficient Economy.

SBW Consulting, Inc., Itron, Inc., Kema Inc., KDVR, Inc., Jai J. Mitchell Analytics, Modest Systems, Digital Energy, ASW Engineering. 2010. *Major Commercial Contract Group Final Impact Evaluation Report 2006-2008 Program Years*. Study ID: CPU0021.01. www.calmac.org. Bellevue, WA: SBW Consulting, Inc.

Summit Blue, ECONorthwest, PA Consulting Group, Robert Thomas Brown Co., ADM, SBW Consulting, Inc. *Government Partnerships Programs Direct Impact Evaluation Report*. 2010. Study ID: CPU0023.01. www.calmac.org. Walnut Creek, CALIF.: Summit Blue Consulting.

Yalcintas, M. 2008. "Energy-savings predictions for building-equipment retrofits". *Energy and Buildings* 40 (12): 2111-2120.