Savings You Can Count On: Can Reported Savings Be Used To Compare Program Performance Across States?

Scott Kessler and Douglas Mahone, TRC Energy Services
Luke Nickerman and Richard Aslin, Pacific Gas & Electric

ABSTRACT

In established industries, there are clear metrics that help management and investors determine whether a company is pursuing an effective commercial strategy. However, for energy efficiency, the primary output -- saved energy -- is not directly observed and therefore standardized comparisons have proved elusive. Variations in reported energy savings include: whether savings are reported as net or gross, if site or source savings are reported, how savings from projects spanning multiple calendar years or program cycles are accounted for, and if gas and electric costs are separated. Furthermore, savings estimates per measure vary across jurisdictions and even within jurisdictions if there are climate-specific values. This paper builds on previous work comparing values for common measures that showed savings can vary by more than two-fold even for significant and well understood measures such as CFLs. This paper takes the analysis one step further by comparing six PAs' (program administrators') portfolios and attempting to account for all of the issues mentioned above. The result is a discussion of the attempts to develop a methodology for adjusting savings to eliminate discrepancies and to enable comparison of program performance across states.

Introduction

This report describes a pilot investigation, initiated by Pacific Gas and Electric Company (PG&E), to explore the possibilities and challenges in comparing the reported results and the relative cost-effectiveness of energy efficiency portfolios administered by utilities and other entities across the country. As part of this analysis, portfolio savings were adjusted to account for differences in program savings reported by PAs. The results were reviewed to determine the efforts and adjustments needed to make accurate comparisons between PAs nationwide. This study attempted to answer the reasonable questions: "How are we doing relative to other PAs?" and "What factors might account for different levels of performance?" In making these comparisons it is necessary to understand exactly what the different reported savings values represent, as well as the assumptions and technical variations driving those reported values. This is one of the first efforts of its kind aimed at developing a repeatable methodology for adjusting PA portfolio savings based on differences in program and measure reporting.

An initial effort by PG&E to compare program performance and cost-effectiveness across multiple portfolios made use of data collected by the U.S. Energy Information Administration (EIA) to conduct a levelized cost analysis. Other previous works consulted include a study by the National Resources Defense Council (NRDC) and TRC Energy Services (formerly the Heschong Mahone Group) (Austin et al. 2012) which analyzed deemed savings and assumptions making up those deemed savings for four common measures, finding large discrepancies between savings values nationally. DOE also conducted a scoping study in 2011 that had similar

findings. In addition, E Source provides information on energy efficiency portfolios through their DSM Insights, a repository with tools for navigating state and utility reported data.

Building off of these initial efforts, PG&E asked the project team to conduct a more in-depth analysis of other energy efficiency program portfolios across the country. This effort was aimed at determining what steps would be needed to overcome reporting differences amongst PAs, and so to enable apples-to-apples comparisons of portfolio performance. Reported savings can vary due to many factors, among them a more/less conservative regulatory environment, different deemed savings for measures, climate differences or simply because some PAs with long-standing energy efficiency programs have already acquired much of the more easily obtained savings, or "low-hanging fruit". It is conceivable that, in such locations with well-established programs, energy efficiency measures are being installed in less used spaces, or are replacing existing equipment that is more efficient than other states' baselines. One can further hypothesize other possible explanations to account for differences in portfolio performance as well. The purpose of this pilot study was to identify and test these hypotheses and, hopefully, to develop a repeatable methodology for comparing PAs' performance.

Comparison Group Selection and Data Collection Tools

In an initial effort to make comparisons between various PAs, PG&E used information on 23 utility and non-utility entities based on information collected by the EIA. From this initial group, PG&E selected 10 representative PAs from across the country. These included a mix of utility and non-utility PA entities, selected to retain regional diversity, that serve electric, gas, and combined utility customers.

From this list of ten PAs, publicly available reports on their programs and portfolios were obtained and examined. PG&E and the project team decided that annual data for 2012 was most desirable, since it represented the most recent full calendar year for which savings data have been reported at the time of the study. For the majority of PAs, the report identification process consisted of a review of documents on their respective public utilities commission's websites. A smaller number of PAs had this information available on their own websites.

The PAs with the most similar and comparable reporting formats were selected for further analysis[1]. The five selected PAs were Baltimore Gas and Electric (BG&E), Focus on Energy (Wisconsin), Florida Power and Light (FP&L), MidAmerican (Iowa only), and NYSERDA.

Because of the disparity in content provided within the reported PA data it was necessary to develop consistent data collection tools for this pilot study. For example, while Focus on Energy's impact evaluation included clear explanations of each savings category along with the data tables, other PA annual reports did not include any explanation alongside their data tables. A questionnaire was developed to assist in comparing the reporting characteristics of each PA and to identify dissimilarities so they can be taken into account. The questionnaire included requests for additional information from each PA on a number of variables, including:

- Source of report
- Intended audience of the report

[1] Basing our selection on this criteria introduces a large amount of selection bias into this study. It is assumed that this comparison would prove more difficult to do among randomly chosen PAs as their reporting formats would not be as similar as those used by the selected group members.

- Time period covered
- Relevant multi-year program cycle for that time period
- Savings accounting method (i.e., does the PA claim savings when incentive funds are reserved or when paid?)
- Level of detail in reporting (i.e., by sector, program, measure, or other)
- Availability of demand reduction information
- Availability of natural gas savings information (if applicable)
- Combined or separate natural gas and electric funding
- Whether savings were gross, evaluated gross, or net
- Availability of net-to-gross ratios and line-loss factors
- Sources of deemed savings values
- Public availability of deemed savings calculation details and sources
- Whether reported savings were site (meter) or source (generation)

While analyzing each PA's reports, it became apparent that there were varying levels of narrative explanation provided with the data. Also, the data was provided in many different formats. During analysis of the available data from the five selected PAs, the team encountered significant differences in completeness, level of detail, supporting explanations, and reporting formats. A data collection form was created to ensure that the savings for each of the PAs are comparable. All data acquired from the various PAs were adjusted and standardized, to the extent possible, in each of the summary forms.

Primary Data Collection and Analysis

Program Level Analysis

This section details the process undertaken to make sure that variations in reporting were accounted for, including differences in net-to-gross ratios, line loss factors, accounting methodology, and program cycle dates. It was important to ensure consistency and so it was decided that the most useful data for this effort was gross savings for an average year of a whole program cycle, at the meter level, excluding savings from 1) previous program cycles, 2) solar water heating, and 3) solar photovoltaics (see subsection "TRM and Program Measure Labeling").

The first step in analyzing each PA's energy efficiency program portfolio was to break down the portfolio by program. All five selected PAs provided information at the program level so this was possible for each of them. Graphs of the electric savings by program for each PA's portfolio, by program, for electric savings are shown in Figure 1. Graphs of electric savings in 2012 by program. Residential lighting programs and commercial retrofit programs provided the majority of the electric savings and demand reduction. The majority of natural gas savings tended to come from whole home programs and non-residential programs.

Net-to-Gross ratio and measure adoption rate. Reporting variations caused difficulties even in this basic level analysis. For three of the PAs, it was clear how the savings were reported (gross, evaluated gross, or net). However for two PAs it was not clear whether the reported savings included adjustments for net-to-gross ratios or measure adoption rates (factors used to adjust from net to evaluated gross, and from evaluated gross to gross savings). Therefore, while general observations and conclusions drawn from looking at the programs composing each

portfolio will likely remain unchanged, it is possible that, in the case of programs with large or small net-to-gross ratios, their relative makeup may vary significantly depending on the type of savings. Net-to-gross ratios as low as 0.4 were encountered for programs with large amounts of free-ridership and as high as 1.6 for programs that anticipated a large amount of spillover.

Site and source savings. It was often difficult to determine whether savings that were being reported occurred at the site (meter) level or at the source (generator) level. While knowing site or source savings will usually not impact the percentage that each program contributes to the portfolio, this may have a significant impact on any analysis beyond that, or on any comparisons made using actual savings values.

Accounting methodology. For all PAs except two, it was unclear which accounting method was used for their reporting. Depending on their regulatory mandates, some PAs may count a project or measure's savings as soon as they are enrolled in a program and funding has been reserved for that project. In this case, savings are more truly aligned with marketing, recruitment, and enrollment costs. However, other PAs elect to count a project or measure's savings when the final incentive for the installed measure is paid. This aligns incentive costs with savings. The former is a variation of accrual accounting, while the latter is more akin to cash accounting. In addition, PAs report administrative and marketing costs differently, year-to-year. The year-to-year variations amongst the portfolios, and any conclusions drawn from comparing multiple portfolios, are impacted by which of these accounting systems (or any other) is in use.

Program cycles. Because of the issues arising from differing accounting systems, and because customer participation rates in the early years in a program cycle can differ from later years, PG&E and the project team determined that analyses should be completed on an average year's performance from the most recently completed program cycle, rather than on any individual year (e.g., 2012 as was initially assumed). This would reduce the influence of accounting method differences, as PAs are typically required to account for all costs within a multi-year program cycle. Additionally, this helps to reduce the chance that any changes in program results due to a ramp-up period, regulatory disruption, or changes to evaluation assumptions do not unduly influence the results attributed to a single year.

In an effort to address these differences and data issues, PG&E and the project team attempted to schedule conversations with each of the PAs under review. Only two phone interviews were completed due to PA time and resource constraints. These conversations proved helpful in resolving previously discovered issues, however it was difficult to get follow-up information resulting from those conversations. As a result, for this pilot study, many of the data issues were left unresolved and best assumptions as to what the data was reporting were applied. With more time to do the study, it likely would have been possible to overcome some of these difficulties.

Measure Level Analysis

In order to make adjustments at the measure level, it was important to understand the variations in measure level reporting and to modify reported measure savings so that they were more comparable across PA portfolios. While measure level reporting is often not available, it was publicly reported for three of the five selected PAs. For these, measure level reporting was

examined to determine which measures were contributing most significantly to their reported savings.

Measure categorization. To account for measures being labeled and reported differently for each PA, general measure categories were established. For example, one PA simply had the category "boilers" while another had "boiler controls", "boiler retrofits", and "boiler tune-ups". In this instance all three of the second PA's categories were combined into a "boiler" category. In order to get measure level reporting consistent across many PAs, it is necessary to aggregate measures to a much higher level than anticipated. If comparisons were being made among a smaller number of PAs, measure categories may be more aligned, allowing more detailed measure level information to be used.

Measure consistency. Another difficulty was tracking the terminology changes over the multiple years in a program cycle. In some instances, a PA began a program cycle using very general, non-specific measures like "lighting"; however, they progressed to more detailed and specific measure tracking in later years of their program cycle, such as "high intensity discharge lighting" and "lighting controls". To the extent possible, the more detailed reporting was aggregated to reproduce the more general categories. In a sense, the selected measure categories represent the lowest common denominator of measure reporting, in order to make the comparisons between PAs as broadly useful as possible.

Based on the reported data, the following measures were common enough to account for at least 5% of total electric savings, demand reduction (DR), or natural gas savings for at least one PA:

- Residential electric savings and DR: CFLs (compact fluorescent lamps)
- Residential natural gas savings: Attic Insulation, Boilers, Building Shell, Furnaces
- Commercial electric savings and DR[2]: HVAC, Lighting, Motors/VFDs, Refrigeration
- Commercial natural gas savings: Boilers, Energy Recovery, Furnaces, HVAC, Insulation

As with the program level analysis, it is necessary to know whether or not reported measure savings were gross, evaluated gross, or net. By cross-referencing measure level savings with program level savings, it was possible to make this determination for any portfolio that had distinct gross, evaluated gross, and net program savings (in this case, two of the three portfolios with measure level savings). It is also important to note whether net-to-gross ratios, line loss factors, and other assumptions were being applied at the program level or at the measure level, and to maintain consistency in this application across the analysis.

Technical resource manuals (TRM). Once the significant measures were identified and better understood, an attempt was made to adjust the measure savings estimates to compensate for systematic variations amongst the PAs. This was done by collecting the TRMs in effect for each service territory, and by identifying the variables used in the calculation for each measure's deemed savings. On the commercial side, there were three TRMs available for the PAs in the analysis, while on the residential side there were only two. Once the variables were known, an

[2] Savings were generally reported as residential and nonresidential making it impossible to identify industrial-only measures. When possible however, industrial measures were removed from this analysis due to their dependence on a site's specific processes.

average value could be calculated as the average of the available TRM values, and an adjustment could then be made for each PA, based on the variation between their assumed savings values and the average values. For example, the most basic formula for electric savings from commercial lighting measures is: kWh Savings = ΔWatts * Hours / 1000. The variables used in this deemed savings calculation are ΔWatts, the difference in watts between the efficient and baseline measures, and the hours of use. Those two variables, then, could be adjusted to reflect the TRM average values to arrive at a more consistent comparison of savings achieved by the different PAs for their commercial lighting programs.

Baseline and efficient measures: Many PAs provided multiple options on what to assume for baseline and efficient measures. While helpful for calculating the savings from one specific retrofit, this made making adjustments based on an assumed typical or representative retrofit more difficult. For example, a single TRM may have different savings values for a T-12 lamp with an electronic ballast or a T-12 lamp with a magnetic ballast. Similarly, not all PAs used a single lamp retrofit as their default, but instead use a two, three, or four lamp retrofit. Therefore, while the values selected in this effort provide a good estimation of the degree of variance between each PA's savings estimates, it is possible to get exact values that are different, depending on the baseline and efficient measures chosen for the analysis.

TRM and program measure labeling. Additional issues encountered during this phase included discrepancies between how TRMs identified measures and how PAs identified measures in their reporting. In all cases there were measures identified in the PA reporting that did not correlate with those listed in the TRM. For example, PA reports often provide deemed savings for categories such as "commercial lighting" whereas TRMs often go into greater detail, such as the exact lamp number and ballast type being replaced as well as whether or not it is an end of life replacement. So while an adjustment factor can be calculated for similar measures, it is unlikely that an adjustment factor can be calculated for exact measures as reported.

Based on the issues encountered while making the lighting adjustment, PG&E determined that this pilot effort would not attempt the more complicated adjustments for climate-sensitive measures, for schedule and resource reasons.

Other forms of adjustments are also possible, including calculating deemed savings based on measure quantities and savings claimed rather than based on TRM values. While a potentially more straight forward approach, this approach would require more granular data that typically isn't publically available.

Along with other possible adjustments, certain measures will likely be excluded from consideration, even in a follow-on effort. In this pilot analysis, certain measures such as industrial process, solar water heating, and solar photovoltaics were excluded as they were not available in all PA portfolios and could skew the comparison results. In future efforts, it may be useful to exclude all industrial and renewable measures for this same reason. Determining which measures are strictly or predominantly industrial requires a more detailed understanding of each portfolio.

The process for both program level analysis and measure level analysis yielded extremely valuable insights. These results are useful in determining which of the various attempted comparisons were most valuable.

Review of Findings

This section summarizes the findings from the analysis described above. There are a variety of ways that comparative portfolio performance data can be presented, each with advantages and disadvantages.

First-year Cost and Savings Comparison

The simplest form of comparison was the first-year cost over first-year savings, based purely on reported values. Though easy to perform, this analysis is flawed in a number of ways. While BG&E and Focus on Energy provided information on what type of savings are reported, and similar NYSERDA information was determined through an interview, it is still undetermined what type of savings are being reported for the remaining PAs in the analysis. NYSERDA's savings were initially reported as net savings, so they can be converted to evaluated gross savings by using reported net-to-gross ratios.

Both BG&E and Focus on Energy report gas savings, however neither reports gas program expenditures separately from electricity expenditures. Therefore their electric first-year cost over first-year savings is likely lower than the calculated value, while the value for their gas programs cannot be calculated at this time. These sorts of problems may be resolved in the future by working with the PAs to get unpublished information necessary to split expenditures.

Additionally, some programs such as FP&L's demand reduction programs were excluded as being outliers in the analysis of their portfolio, while some programs such as Focus on Energy's legacy programs (defined as savings for projects that were approved in a previous program year but completed in the current year, without using budget from a current program) were excluded because they would skew the cost-per-savings estimates. Differences in reporting practices by different PAs mean that it will often require additional digging to obtain the information needed to true-up those kinds of numbers.

This effort only looked at first year savings and the total reported costs associated with those savings as summarized in Table 1. As an additional refinement, cost-effectiveness can be recast as a benefit/cost ratio, where the energy savings are converted into dollar savings and then compared to program dollars (similar to the Total Resource Cost Test). That conversion requires the application of the levelized cost of energy over the life of the measures as well as avoided costs. While this is a reasonably straightforward calculation, there was insufficient time and information to do it for these PAs in this effort. In future work, this would provide some useful insight into the relative effectiveness of each portfolio.

Table 1. First-year costs and savings

Program Administrator	Savings (GWh)	Demand Reduction (MW)	Gas Savings (MDth)	Savings Type	Electric Program Expenditures ($M)	Gas Program Expenditures ($M)	$/ MWh	$/ Th	Program Years
BG&E	724	679	218	Evaluated Net	160	-	221		2009 – 2011
Focus on Energy	424	59	1,118	Evaluated Net	379	-	894		2011 - 2012
FP&L	191	90	-	Unknown	104	-	545		2012
MidAmerican	850	181	1,763	Unknown	264	83	310	4.71	2009 – 2012
NYSERDA	972	117	1,568	Evaluated Net	138	40	142	2.58	2010 – 2012

The two shaded columns in Table 1 reduce the total costs and savings to common metrics of dollars per unit energy consumption. As the numbers in Table 1 demonstrate, there are wide variations between the five portfolios included in this analysis. These should be considered as descriptive comparisons rather than definitive, because there was insufficient detail to allow for a balanced comparison using consistent values. More time and information than was available for this pilot study would be required to achieve that level of accuracy.

Program Composition Comparison

After completing the first year costs and savings comparison, the next attempt was to compare the portfolios by program composition. While this was a fairly straightforward process as well, the results did not provide much meaningful insight beyond which programs are the most significant in a PA's portfolio. The charts in **Error! Reference source not found.** display the electric savings by program in 2012 for all five of the comparison PAs. The labels only call out the biggest programs that account for 5% or more of the portfolio share. The labels reflect the terminology used by the respective PAs, so there is no consistency in program categories between portfolios.

Figure 1. Graphs of electric savings in 2012 by program.

Significant measures comparison

The third method of comparing portfolios was to identify the significant measures in each portfolio. This would test the hypothesis that by comparing savings of similar measures across portfolios, adjustment factors can be derived and applied to portfolio results, which would reflect a PA's tendency to report higher- or lower-than-average levels of savings. If the data supported

that hypothesis, then the savings levels could be normalized between portfolios, and so provide more meaningful comparisons. However, this adjustment proved difficult to carry out in practice.

Not all PAs provided measure level savings, and those that did were inconsistent in how those savings were reported.

Figure 2. , below, shows the most significant measures for each portfolio, and the percent of savings attributable to each. However, matching up measures by their reported values is often difficult to do between portfolios. For example, while some PAs list residential CFLs as one measure, BG&E reports different wattage types as individual measures, making direct comparisons between the portfolios difficult. Additionally, overly generalized measures, such as "commercial comprehensive measures", or measures that are labeled after their program ("Energy Solutions for Business") are similarly difficult to analyze.

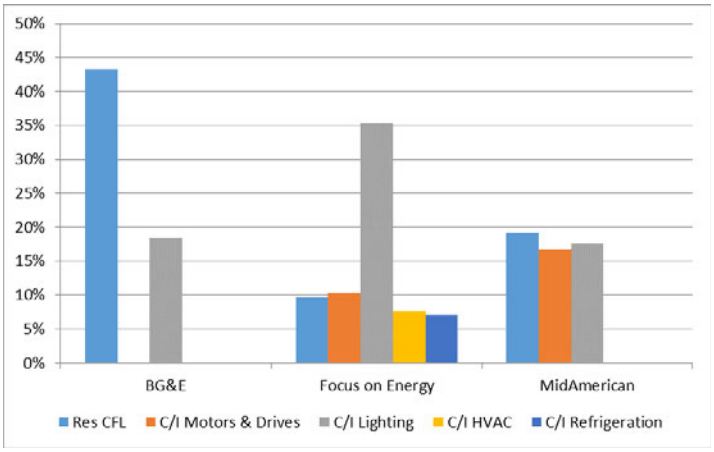

Figure 2. Savings by measure category by PA.

Adjusted significant measures comparison

The final method used to compare portfolios was to adjust the reported savings measure-by-measure, starting with the most significant measures (in terms of total savings). This was done in two ways. First, the measures were combined into more general measure categories, to make them more comparable between portfolios, and to reduce issues of categorization and labeling confusion. Secondly, the measures were adjusted based on the deemed values provided in their respective TRM. Due to time and budget constraints in this pilot study, this second adjustment was not completed for climate sensitive measures.

, below, provides the values used in adjusting residential CFLs. The formula commonly used for residential CFL savings is kWh savings = ΔWatts * Hours of use per year * In Service Rate / 1000.[3]

3 The In Service Rate (ISR) is a ratio that defines what amount of the CFLs sold or distributed are actually in use. The ISR is used to avoid counting savings for CFLs that do not end up getting installed.

Table 2. Deemed and adjusted residential CFL values

Program Administrator	ΔWatts	Hours of use per year	In Service Rate (ISR)	kWh Savings / CFL watts	Adjustment Factor
BG&E	CFL watts * 2.95	1088 (2.98/day)	0.88	2.82	0.86
NYSERDA	CFL watts * 2.53	1168 (3.2/day)	Not defined.	2.96	0.82
PG&E	CFL watts * 2.53	795.7 (2.18/day)	0.9	1.81	1.33
CFL Average Value	**CFL watts * 2.67**	**1017.2 (2.79/day)**	**0.89**	**2.42**	**N/A**

By calculating a ratio of kWh savings / ΔWatts, this provides a deemed savings value that is independent of CFL wattage. The average of this value for the three PAs for whom this analysis was possible delivered an average value. Dividing this average value by the individual PA values yielded an adjustment factor that modified residential CFL lighting savings to their savings, as if using the average values. In this case, NYSERDA's residential CFL lighting savings would be 0.82 times their reported amount if using average values in their TRM; i.e., NYSERDA values tend to be higher than average.

below shows how this adjustment is calculated. It accounts for differences in ΔWatts, in assumed hours of use, and in the In Service Rate (ISR).

These adjustment values could be used to modify portfolios where evaluated gross savings[4] from residential CFLs were available (in this case, BG&E). This illustrates how the adjustment factor derived from a major set of measures can be used as the adjustment for the entire portfolio savings. The assumption, of course, is that similar adjustments can be made to the overall portfolio, if the adjustment factor is derived from most significant measures in the portfolio. The results of this adjustment to the residential lighting savings and subsequent impact to the total portfolio savings are shown in Table 3 below. BG&E is the only PA that had measure level savings and a TRM for residential CFLs, therefore is the only PA that can be adjusted for this measure.

A similar analysis could be completed for other portfolios that have the same level of data, which would also serve to further refine the average values and adjustment factors. Due to the small sample size in this pilot study, it is difficult to draw general conclusions, but the initial results do indicate that PA savings can be impacted significantly should this adjustment be carried out for all measures.

Table 3. Adjusted residential CFL savings

Program Administrator	Reported Residential Lighting Savings (MWh)	Reported Portfolio Savings (MWh)	Adjustment Factor	Adjusted Residential Lighting Savings (MWh)	Adjusted Portfolio Savings (MWh)	% Difference Between Reported and Adjusted Portfolio Savings
BG&E	269,592	804,870	0.86	230,725	766,003	-5%

[4] In this instance, the difference between evaluated gross savings and reported gross savings is the application of the In Service Rate. Because the values for ISR are so similar, the same adjustment factors can be used for gross savings. However, if doing this for net savings, a new adjustment factor is needed to account for differences in the net-to-gross ratios for each PA.

A similar adjustment was made for retrofitting a T-12 light fixture with a High Performance T-8 lamp. Owing to the complexity of adjusting for multiple numbers of lamps in a fixture, the adjustment is more detailed and is not included here. The adjustment to the commercial lighting saving and subsequent impact on the two portfolios are provided in Table 4 below.

Table 4. Table of adjusted commercial lighting savings

Program Administrator	Reported Commercial Lighting Savings (MWh)	Reported Portfolio Savings (MWh)	Adjustment Factor	Adjusted Commercial Lighting Savings (MWh)	Adjusted Portfolio Savings (MWh)	% Difference Between Reported and Adjusted Portfolio Savings
Focus on Energy	304,812	865,066	1.19	362,822	923,075	7%
BG&E	149,130	804,870	0.76	113,854	769,594	-4%

While other adjustments are possible for other commercial lighting measures, it is unlikely that a deeper investigation would yield substantially different results. This is based on the assumption that conditions that cause one administrator to have more conservative values in their deemed savings, such as more efficient existing lighting, can be expected to equally impact the savings values for other types of commercial lighting. A T-12 to High Performance T-8 retrofit is considered the most representative type of commercial lighting project and should be the best indicator for how to adjust all commercial lighting savings. Of course, this assumption has not been rigorously tested due to time and data constraints.

This exercise of deriving adjustment factors for individual high impact measures and using those factors to adjust the overall portfolio reported savings indicated that it is possible to adjust PAs' overall savings to make comparisons more meaningful, if the necessary data is available. To do so for all significant measures, including climate dependent measures, however, will require more detailed data and time for analysis than was available for this pilot study.

If, in future analysis of this sort, adjusted portfolio savings are developed, then those results would cycle back to the first comparison of first-year costs over first-year savings described above. The result would be a more useful comparison of the cost-effectiveness of portfolios. PAs may also change what values they report, further complicating matters.

In addition to those numeric comparisons between portfolios, it will be necessary to supply detailed commentary to help explain the differences between portfolio results. For example, one might expect the cost-effectiveness of a mature portfolio to be lower than that of a relatively new portfolio, because there would be less "low hanging fruit" for the mature portfolio to "harvest". Or, as another example, it might be more costly to acquire air conditioner savings in a milder climate than in a hot climate, because the cooling loads and savings would be lower in the milder climate area. Lacking such explanatory information, the numeric comparisons cannot be fully understood.

Conclusions and Recommendations

This pilot study tested several hypotheses about the ability to develop a methodology to adjust reported savings data to enable meaningful, portfolio-level comparisons between PAs in different parts of the country. Due to the limits on available data and on the available time to clarify what those data were reporting, we could make no definitive conclusions about the best

method for making these comparisons. The process did, however, identify the key challenges, tested different approaches to overcoming them, and demonstrated the kinds of comparisons that can be useful for PAs and their senior management. The key issues are summarized here:

- Making adjustments to high impact measures is likely not feasible for an entire portfolio. The highly varied reporting practices of different PAs for both programs and measures would make it very difficult to complete a more comprehensive adjustment than found here without significant time and effort. However, these adjustments can be useful in comparing specific programs or measures, as demonstrated by the residential and commercial lighting adjustments.
- Exact adjustments are not feasible for most measures. The project team had to make a large number of assumptions just to achieve a reasonable adjustment factor for lighting measures, which are simpler than many other measures. Climate-dependent measure adjustments would be more complex. This analysis would require more effort than was possible for this pilot phase, but would likely to be important in generating meaningful comparisons amongst portfolios.
- Active involvement from the PAs, as the sources of comparison data, is necessary to do the analysis well. While some information on the portfolio savings was publicly available, much of it was not. This is likely to be more accessible for the PAs, using internally available data. In addition, given the amount of assumptions that must be made to interpret the programmatic and measure data, assistance from the authors of the subject reports is invaluable in terms of quickly understanding assumptions. Thus, to do comparisons well, requires the active assistance of data analysts at each participating PA.
- This study served its purpose of exploring potential data comparison methodologies, using publicly reported portfolio savings data. While rough adjustments are possible through the use of assumptions, exact comparisons are not possible at the current time without PA participation due to the fragmented and inconsistent manner of reporting at the portfolio, program, and measure level.

Based on the recommendations that arose from this pilot, future efforts should include:

- Development of work plans with the aim of eventual industry standardization in data reporting. While a collaborative project amongst the energy efficiency industry's many stakeholders, including PAs and regulators, would eventually be required, the task may be best initiated by a smaller group capable of addressing issues quickly, and so developing a framework for other entities to review. Efforts currently underway at the Department of Energy (SEE Action Network 2012) and the Consortium for Energy Efficiency are potential first steps toward this end.
- Combination of savings investigations with comparable cost investigations, as the two are necessary to ensure proper cost-effectiveness comparisons.
- Ensuring PA buy-in before investing significant analysis effort, to assist in the data collection process, as well as to enable better communication and participation.
- Encouragement to participating PAs to track all measures by quantity incentivized to enable simple comparisons between PAs, as well as to support comparisons between reported savings per unit and the TRM values.

- Addressing climate-specific measures and how their TRM evaluation parameters and savings should be adjusted.
- Development of consensus on standardized comparison reporting formats, so that PAs see the value of participation and support portfolio comparison efforts that can lead to more effective programs
- Investigation of alternative metrics for comparison and normalization, including number of measures implemented, market impact/penetration, square footage of the associated building stock, or economic output.

References

American Council for an Energy-Efficient Economy. *The 2012 State Energy Efficiency Scorecard.* Report Number E12C. Washington, D.C.: October 2012.

Austin, C., Goebes, M., Martinez, S., Miller, P., & Wei, J. *Tremendous Potential Benefits from Improving Energy Savings Estimates Across the Country.* Paper presented at the 2012 ACEEE Summer Study on Efficiency in Buildings. Asilomar, CA.: 2012.

State & Local Energy Efficiency (SEE) Action Network. U.S. Department of Energy and U.S. Environmental Protection Agency. *Scoping Study to Evaluate Feasibility of National Database for EM&V Documents and Measure Saving,* by The Cadmus Group. June 2011.