Strategies for Improving the Accuracy of Industrial Program Savings Estimates and Increasing Overall Program Influence

Jennifer Fagan, Kris Bradley and Al Lutz, Oakland, CA

ABSTRACT

In recently completed evaluations of California's 2006-2008 Industrial energy efficiency programs, evaluated net savings ranged from one-fourth to less than one-half of program-claimed savings. This paper provides information on the root causes of these low verified savings levels, and presents a set of overarching recommendations aimed at improving the accuracy of savings claims and increasing the degree of program influence on rebated projects.

There is significant energy savings potential in the Industrial sector. The programs have capitalized on this by developing and implementing a large number of innovative energy savings projects. If program staff and evaluation staff can get on the same page with regard to savings estimation procedures and strategies for minimizing free ridership, this will greatly benefit Industrial sector programs in general. Developing reliable savings estimates for custom Industrial sector projects is not straightforward. These projects are custom in nature, meaning estimates are site-specific and there are no standard or deemed savings algorithms.

A key problem area that led to inaccurate claimed savings in the California Industrial programs concerned the assumed baseline conditions and gross savings estimation procedures in general. A second problem area was low program influence (i.e., high free ridership). These problem areas had been found in numerous Industrial program evaluations conducted since 1998. This paper will present a set of overarching recommendations and tools for improving the accuracy of claimed savings estimates and increasing the degree of program influence.

Introduction

There is a significant amount of cost-effective energy savings potential for the industrial sector, which accounts for nearly one-third of national energy use. Energy savings opportunities abound in outdated manufacturing buildings and machinery, industrial process boilers, and obsolete lighting systems. As noted in a recent ACEEE publication[1], improvements in energy efficiency are critically important in reducing overhead costs and helping businesses compete in a global economy.

In this environment, industrial businesses are highly motivated to make investments in their buildings and infrastructure to upgrade energy-using equipment and drive down energy use. Record numbers of businesses have participated in voluntary energy efficiency rebate programs in recent years, and continue to do so. This high level of interest in rebate programs presents industrial program managers with an interesting challenge – how to capitalize on this customer interest while continuing to exert a high level of influence over what gets installed. Since most programs are generally given credit only for those projects in which they had significant influence in the customer's purchase decision, they must be able to get out ahead of the decision making process in order to meaningfully contribute to the project concept, ideally at its conception. This is the crux of the challenge to increase program influence and reduce free ridership.

---

[1] American Council for an Energy Efficient Economy. 2011. Fact sheet, "America's Abundant Untapped Energy Efficiency Resource".

Another challenge for managers of voluntary custom rebate programs is the need to develop accurate estimates of project-level savings. Selection of the correct project baseline is a key consideration. The program must also take steps to ensure all projects comply with program rules for eligibility, since violations lead to zeroing out of energy savings.

This paper will present a set of recommendations and procedures aimed at improving the accuracy of claimed ex-ante savings estimates and increasing the degree of program influence. These recommendations are based on a set of recently-completed evaluations of industrial energy efficiency programs in California. Findings from a detailed analysis of nearly 150 projects indicate that the savings gap was largely due to a number of factors: improper baseline specification, lack of production adjustments, modest program influence on project decisions, and limited information on certain technologies with involved system interactions. The extent to which industry standards and common practices in an industry govern evolving baselines is highlighted, as is the importance of early and effective program influence. Useful life considerations, natural turnover, and the appropriateness of production level adjustments (particularly when enabled by newer industry-standard processes) are explored.

Definitions

Voluntary energy efficiency programs are typically only given credit for savings that directly result from program activities. These are considered **program induced** savings, the scope of which includes both gross and net impact components.

Gross savings are the change in energy use that are the direct result of a program without considering actions participants may have taken without the program. Gross savings are calculated using engineering assumptions and formulas, and are of a more technical nature. These are the savings that an energy end user, or program participant, most commonly associate with energy savings. The **net impact** component, in the form of a free ridership percentage or net-to-gross (NTG) ratio, takes into account the degree to which the program influenced the participant's decision to install energy efficiency technology through the program.

Study Approach

Gross Savings. Methods used to estimate gross savings ranged from end-use monitoring, to calibrated simulation models, calibrated engineering analysis, engineering review, and billing analysis. Factors that were considered in matching these approaches to different measures and programs included the size of the expected impact, the degree of site-by-site variation in per unit savings, the aggregate size of the measure's impact at the program and portfolio levels, the cost of applying the savings estimation method, the sampling size and associated sampling error, the reliability of the measured data, and the length of the evaluation and its timing relative to implementation (e.g., to assess whether billing analysis is feasible). Data collection and analysis techniques included application review, project documentation review, site visits, M&V activities, and on-site audits/interviews. For certain projects, metering/monitoring was conducted, and explanatory information was also obtained regarding pre and post implementation operating conditions. For complex measures where individual engineering algorithms were inadequate, and where direct measurement was either too expensive or otherwise not desirable as an exclusive method, calibrated engineering simulation models were used to estimate the load impacts.

For each sampled project, an ex-post savings estimate was developed based on actual as-found operating conditions. In developing this estimate, it was necessary to accurately determine the baseline condition, reflecting the pre-implementation operating conditions which would have prevailed in the absence of the program. Adjustments were included for production, occupancy, baseline, weather or other parameters that changed. Information was 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.

Net Savings. Net impacts were determined using a new method developed in California for nonresidential programs, which complies with the California Evaluation Framework and 2006-08 draft protocols. This approach differentiated projects and measures based on their complexity and importance to the portfolio, and assigned a NTG rigor level based on these considerations. More complex and important measures/projects were assigned a Standard or Enhanced Rigor Level and warranted a more comprehensive NTG inquiry involving "triangulation" of findings from participating customer and trade ally interviews, and integration of other market intelligence. Primary research was conducted to support this inquiry – with data collection efforts focused on participating customers, trade allies and equipment manufacturers. Net impact results were estimated using the Self-Report method, since other NTG methods are generally not feasible for large industrial custom projects. This method used a carefully designed questionnaire to tease out likely program influence in a manner which was designed to maximize customer recall of the circumstances surrounding their implementation decision, while minimizing bias in their responses. A detailed NTG scoring algorithm was then applied.

The most predominant data source was findings from telephone surveys, which consisted of both scripted questionnaires administered through a Computer-Aided Telephone Interview process and in-depth interviews, administered by professional interviewers, which are more fluid and personalized. Information from on-site interviews conducted by the gross impact team also aided the net to gross impact assessment. For certain equipment markets considered to be at least partially transformed, both primary and secondary research was done to assess the extent to which these measures were already standard practice.

Summary of Qualitative Impact-Related Findings and Related Recommendations

There are a number of specific findings that help to explain why the ex-post savings estimates were significantly below the ex-ante results. These related key findings are summarized below, along with recommendations for addressing them. The key problem areas were due to incorrect assumptions regarding baseline conditions, inaccurate gross savings estimation, and low program influence (i.e., high free ridership). A number of these findings, particularly those related to low program influence, had been found in numerous industrial customer program evaluations previously conducted. Each of these is discussed below.

1. Baseline Selection

Problem: There are Significant Problems with Baselines Used for Claimed Savings. The most common problem in the industrial programs is the pervasive use of pre-existing (often referred to as "in situ") equipment as the baseline for estimating program incentives and savings. In many cases, savings were calculated relative to an in situ baseline and then assumed to occur over the entire period of the effective useful life (EUL) of the new equipment. This assumption would only be justifiable in situations where the program induced an early replacement of equipment that would otherwise have had a very high probability of continuing in operation for a period equal to the EUL of the new equipment. Such cases are likely to be extremely rare in practice, yet they are the convention in the program claims. Instead, many of the projects were in fact replace-on-burnout or natural turnover events, or were early replacement events for a period of time (the existing equipment's remaining useful life) that was significantly less than the effective useful life of the measure. That is, the pre-existing equipment was either at the end (or near the end) of its physical life or at the end of its effective life because the customer decided to replace the equipment for reasons other than achieving energy savings (e.g., to improve product quality, respond to regulatory requirements, increase production, etc.).

Recommendation: Improve Baseline Specification. End the practice of using in situ baselines over the EUL of the measure as the baseline for estimating savings and paying incentives. Identify projects explicitly in program files as replace-on-burnout, natural turnover, or early replacement. For the replace-on-burnout and natural turnover cases, baselines should be based on the efficiency of alternative new equipment, not the existing in situ equipment. In the case of early replacement, provide evidence and documentation of the remaining useful life of the equipment replaced, the estimated time at which the equipment would have been replaced in the future, and the effect of the program in accelerating early replacement.

Problem: Inadequate Basis for Claimed Savings. In some cases there was inadequate engineering or physical basis for claiming savings or little or no reference to empirical information to substantiate the estimate of savings. Measurements were inconclusive for some of these measures and given the lack of empirical data on the basis for savings estimates, it was difficult both to accept the ex-ante claim and to develop an ex-post estimate of savings.

2. Enforcement of Program Rules

Problem: Inadequate Enforcement of Program and Policy Rules. There were a few cases where the project being evaluated did not qualify for the program because of violations of the program rules or the California Public Utilities Commission's (CPUC) energy efficiency policy rules.

Recommendation: Increase enforcement of program eligibility and policy rule requirements. Some of the evaluated projects were found to have violated program eligibility and policy rules. The CPUC should develop a process for reviewing projects for program eligibility prior to their being approved for a rebate.

3. Basis for Savings Assumptions

Problem: Unverified and Undocumented Assumptions Used as Inputs for the Savings Calculations for Many Applications. There were a number of cases where the assumptions for the program calculations were unverified and undocumented.

Recommendation: Put measures with inadequate empirical basis for savings estimates in the emerging technologies program until more reliable information is developed. Measures with highly uncertain savings in need of detailed research to establish validity, expected savings, and repeatable algorithms and measurement protocols should be included in the emerging technologies program until they are more widely understood.

4. Fuel Switching

Problem: Inadequate Declaration of Fuel Switching, Multi-Fuel Impacts, Distributed Generation. Several of the sampled sites involved multiple fuels, or fuel switching, but the savings claim and ex-ante analyses did not include these impacts.

Recommendation: Aggregate and Approve Fuel Switching and Distributed Generation-Related Projects in One or More Explicit Programs or Clearly Identified Program Elements. All multi-fuel project applications need to follow the three-prong test set forth in the Commission's Policy Manual as well as any other regulatory agency requirements.

Problem: Inadequate Consideration of Total System Energy Analysis. In some cases, energy usage was only analyzed for a portion of the system that was directly affected by the measure or project even though there were energy interactions with other systems that were also materially affected by the project.

Recommendation: Aggregate and Approve Fuel Switching and Distributed Generation-Related Projects in One or More Explicit Programs or Clearly Identified Program Elements. All multi-fuel project applications need to follow the three-prong test set forth in the Commission's Policy Manual as well as any other regulatory agency requirements.

5. Effects of Recession

Problem: Significant Effects of Recession. For several sampled projects, the facility had closed down, resulting in zero savings (since measures have to be operational according to program and policy rules). In other cases, production levels had been reduced, sometimes resulting in corresponding reductions in savings estimates.

6. Free Ridership

Problem: High Free Ridership – Limited Program Influence. One important finding was that the programs had a low to moderate percentage of claimed savings that are estimated to be program-induced. Program influence was relatively low for a number of different reasons. For example, program implementers frequently arrived late in the decision making process and offered incentives for projects that had already been decided upon. Program claims were made on a number of projects that customers initiated for non-energy savings reasons and for which no alternative was ever considered. At times, program incentives were offered for measures and technologies that were known to be industry standard practice (thus significantly increasing the odds of free ridership in any given application). Program attribution was also limited when incentives were offered for projects that were being implemented by end users in response to mandates from other regulatory agencies, for example, citations from air resource districts.

Recommendation: Increase the capability of the program to influence industrial efficiency improvements. To move these customers further along the efficiency spectrum takes time and advanced levels of technical expertise, often requiring expertise in specific industry production practices and options for improvement. This is a very difficult challenge in this sector. Development of the depth of technical expertise required to increase the net effects of the programs is a long term endeavor that requires both utility and regulatory support.

Recommendation: Influence and provide incremental energy efficiency options directly to end users at the earliest decision-making stages of major equipment or facility modifications. Program involvement at an early stage to identify large equipment and facility changes helps ensure efficiency opportunities are appropriately considered and maximize the chances of program influence. Utilization of sales or related tracking systems helps prevent projects from becoming lost opportunities.

Recommendation: Provide Continuity in Account Representative Assignments, Particularly for the Largest Customers. There were many instances where the utility account reps had been reassigned one or more times during the project lifecycle. In such cases, the likelihood of utility program influence is weakened because the assigned representative lacks the long-term relationship and continuity needed to provide a significant influence on the installed project. Utilities should seek to provide continuity in these account rep assignments, particularly for their largest customers.

Recommendation: Consider Using Early Project NTG and Baseline Screening Prior to the Incentive Being Approved for the Largest Projects and those with Significant Policy Issues. The evaluation team would review the baseline claim and conduct NTG interviews during the participant's project implementation and program participation decision process. The purpose of this screening would be to obtain critical information regarding program influence that could lead to the project being re-defined to increase efficiency levels and program influence or dropped for ratepayer-funded rebates if no influence is evident. This approach would also have the advantage of capturing critical information on program influence early in the decision making process, while the information is still fresh in the mind of the decisionmaker.

Recommendation: Carefully review the list of qualifying measures for each program and eliminate eligibility for those that are standard practice. Measures that are already extremely likely to be installed by the vast majority of the market should generally not qualify for incentives. Although identification of such measures can be difficult in practice in the industrial sector, a number of such measures can be identified through investigation of industry practices (e.g., interviews with manufacturers, distributors, retailers, and designers), analysis of sales data, and review of evaluation results. In determining which measures to retain and which to eliminate, a balance must be struck between reducing free ridership and avoiding significant lost opportunities.

Recommendation: Consider Limiting or Excluding Incentive Payments to Known Free Riders. One obvious and simple approach to reducing free ridership is for program administrators to simply exclude projects from the program that they believe have a high probability of being free riders. Administrators in several other jurisdictions have used this approach. In these cases, the administrator has the flexibility to determine total incentive amounts on a case-by-case basis, including zero incentives. Alternatively, or in conjunction with this type of approach, rules could be developed that exclude incentive payments for projects that are driven exclusively by non-energy factors that produce energy savings as a by-product.

Recommendation: Consider Incorporating a Payback Floor. The use of a payback floor (minimum payback level based on energy savings alone) can help to reduce free ridership by eliminating projects that have extremely quick paybacks and thus little need for ratepayer-funded incentives. With a payback floor, the program may also avoid incenting projects that are primarily being done for reasons other than energy savings (modernization, production efficiency, environmental compliance, etc.).

Recommendation: Set Incentive Levels to Maximize Net (Not Gross) Program Impacts. Free riders dilute the market impact of program dollars. Payback floors and increasing incentives with increasing payback levels are one approach. Another is to tie incentive levels to individual measures or types of measures that are known to have extremely high or low naturally occurring adoption levels.

Recommendation: Consider Tying Staff Performance to Independently Verified Net Results. Tying performance reviews of program staff to verified net savings as reported through an independent M&V or impact evaluation process is likely to increase project quality and the accuracy of initial savings estimates. Marketing staff, in particular, should have any financial incentives tied to savings that are independently verified.

Baseline Selection

Below are several principles that were used as guidance for determining the appropriate baseline to be used in calculating the gross savings for a project in this evaluation:

Code or market baselines were used for replace-on-burnout and 'natural turnover.' In situ baselines were only used for the portion of the remaining useful life (RUL) of the pre-existing equipment that was eliminated due to the program. Consideration was given to the specifics of the application with respect to the remaining life, if any, of the pre-existing equipment when selecting the baseline.

CPUC policy rules and IOU program eligibility rules governed the baseline. Careful review of utility and third-party program and CPUC policy rules were made and adjustments were applied to gross savings in some cases, while in others the adjustment was reflected in net savings. The adjustments were applied to gross when there was clear evidence from program or policy rules that savings claims could not be made nor rebates paid for the case in question. Program rules also came into play with respect to gross baseline requirements, e.g., specifying a given efficiency level or percentage above code.

Minimum production or energy service requirements govern the baseline. In some situations, a measure for which savings were claimed was determined to be the only acceptable equipment for an application. In such cases, the baseline was set at the minimum needed to meet the requirements. Care was taken to ensure that changes in production or energy service requirements were not merely preferences but were fundamentally required. An example would be an industrial process where only a variable-speed drive pumping system could meet the production requirements.

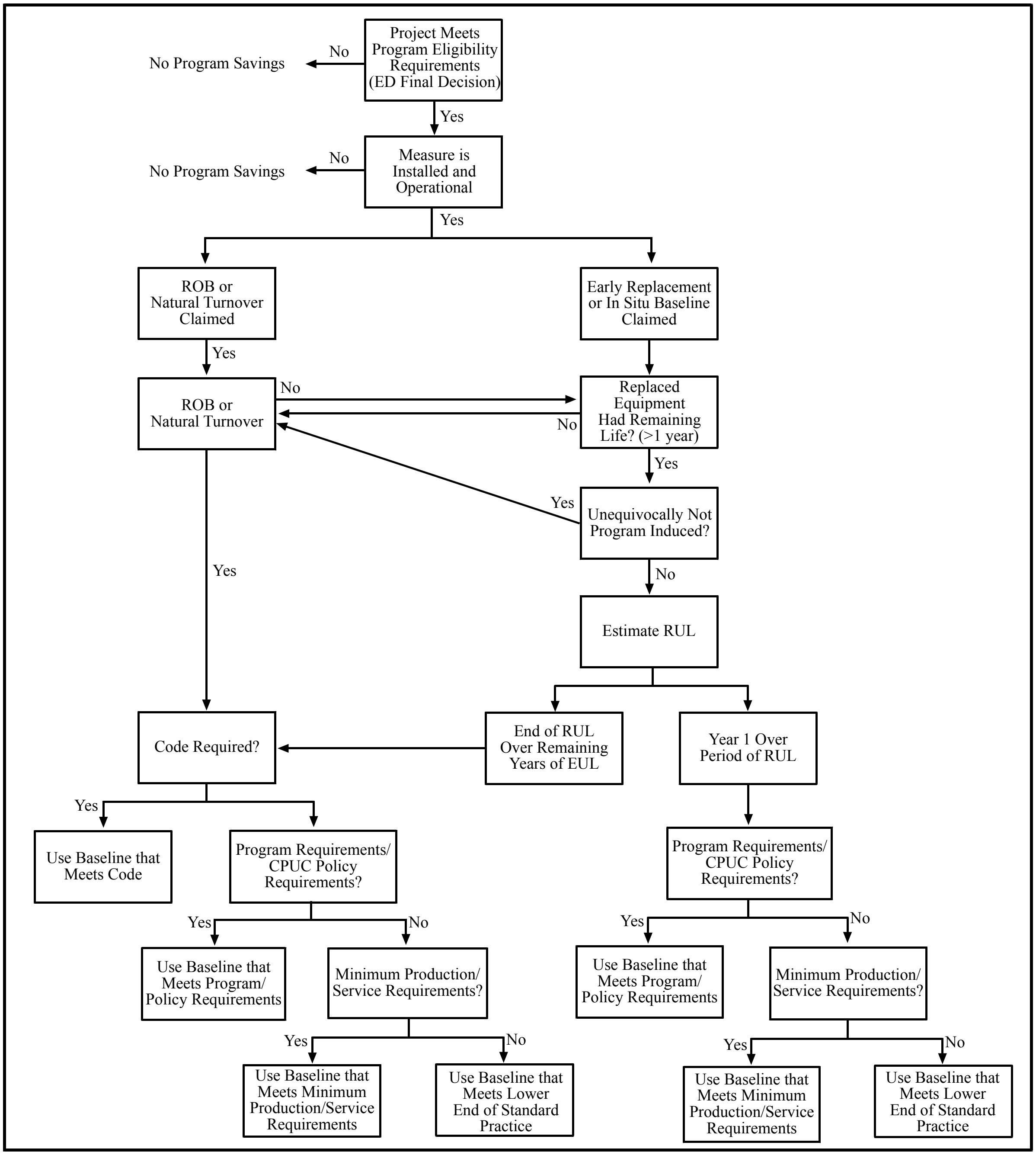
Evaluate early replacement RULs and program inducement. The gross engineering team determined whether there was evidence that early replacement actually occurred, and if there was, an estimate was made of the associated Remaining Useful Life (RUL). The net team determined whether the early replacement was program induced. If the early replacement was not program induced, the gross baseline was set based on the replace on burnout/natural turnover guidelines.

The decision tree below was used as guidance for determining the baseline for gross savings in all cases.

Baseline Guidance

Guidance for Determination of Baseline for Gross Savings

Take Most Efficient of All Applicable Cases



Summary of Other Recommendations

Below are several overarching recommendations from these evaluations that were aimed at improving other aspects of the evaluation process.

Recommendation: Clarify and enforce the definition of "industry standard practice". Industry standard practice should be used to set baselines for savings estimates and incentives (such that program savings estimates improve as reflected in improved evaluation gross and net realization rates). Program and policy references to "industry standard practice" should be more precisely defined with respect to program participation requirements, incentive level payments, gross versus net savings attribution, and energy efficiency goal attainment.

Recommendation: Be More Conservative in Estimating Savings. The programs should make more conservative assumptions for calculated (custom) savings projects in the industrial sector in the next program cycle until ex-post realization rates increase.

Recommendation: Use a Gross Realization Rate Adjustment in Savings Claims in Program Tracking Systems. Use of a realization rate adjustment in future program cycle ex-ante estimates of custom measure claims should be strongly considered until future evaluation results indicate higher gross realization rates.

Recommendation: Improve training of program implementation staff in several key areas. These areas are: proper baseline specification, enforcement of program and policy rules, reasonableness of claims, comprehensive facility systems analysis, and increasing program influence on end user's efficiency-related decisions.

Recommendation: Conduct analysis of customer incentives by customer and industry type. Conduct further research on the use of incentive caps. Customer incentive caps have been utilized in various forms for many years. During times of low budgets and low goals, caps were set low to spread incentives to a broad pool of participants. More recently, as goals and budgets have significantly increased, caps had increased greatly as well. We were not aware of any systematic study of the effect of the incentive caps. Similarly, research is needed to explore how much total incentive dollars have been distributed across or concentrated within certain customers to determine whether these patterns are aligned and supportive of efficiency policy goals.

Recommendation: More information is needed on industrial project costs, non-energy costs and benefits, net present value analysis, and associated participant cost-effectiveness analysis. There has been very little analysis conducted of the actual incremental costs of industrial energy efficiency projects. Further research is needed on industrial incremental measure costs in general. Increased financial analysis should be included in industrial project applications, especially for the projects with the largest incentives. Increased review of project financials inclusive of non-energy factors can also help to reduce free ridership.

Recommendation: Involve impact evaluators in large projects and a sample of projects on a real-time basis throughout the program cycle. The timing of evaluation processes should be accelerated. Moving the evaluation process forward in time to occur just after the project is installed helps to ensure the decisionmaker is still available, and that their memory of the basis for the project is still fresh. This can be accomplished through earlier contracting and implementation of the evaluation, combined with improved utility tracking and early reporting of installations (as well as projects in the pipeline), more frequent sampling and evaluation of projects throughout the program plan period.

Recommendation: Conduct a full complement of impact, process, and market evaluations. Large customer programs and markets are very dynamic and require regular assessment in order that they may be continuously improved by program managers and policymakers. Future evaluations need to consider more integration of process evaluation and market assessment to capture research economies and reduce customer and vendor interview burdens.

Recommendation: Conduct baseline research to establish standard industry practices for key measures in important industries. Significant research is needed to establish meaningful and defensible data, especially market share, for establishing industry standard practices for measures that are not completely site specific. Improved information on industry standard practices can then inform decisions about which measures to provide incentives for, which could in turn lead to reductions in free ridership.

Recommendation: Conduct a persistence study of industrial sector savings. Few studies of the persistence of program savings in the industrial sector have been conducted, particularly within the last decade. In some program years and cycles industrial production levels will be higher or lower depending on economic conditions. Some facilities that do close may stay closed while others may reopen and reutilize efficiency measures. Research is needed to measure the persistence of savings over time under a range of economic conditions.

Conclusions

The industrial sector has vast energy savings potential and is critically important to successfully achieving mandated energy savings goals. However, developing accurate estimates of energy savings from industrial projects is very challenging. This paper has highlighted the most important reasons for the significant gap between ex-ante and ex-post savings estimates, and has offered numerous recommendations for how to bridge this gap and improve program savings estimates going forward. Use of these recommended procedures will lead to improved estimates of ex-ante savings, increased program influence (and reduced free ridership), and will result in reducing the gap between claimed and evaluated results. These procedures will help to improve the overall quality of projects that are participating in the program, and support the general goal of continuous program improvement.

References

Itron, Inc. 2006-2008 Evaluation Report for PG&E Fabrication, Process and Manufacturing Contract Group. Prepared for California Public Utilities Commission. Oakland, California.

Itron, Inc. 2006-2008 Evaluation Report for Evaluation Report for the Southern California Industrial and Agricultural Contract Group. Prepared for California Public Utilities Commission. Oakland, California.

National Energy Efficiency Best Practices Study. 2005. Volume NR5 – Non-Residential Large Comprehensive Incentive Programs Best Practices Report. Itron, Inc. (formerly Quantum Consulting). Berkeley, California.

Itron, Inc. 2004-2005 Statewide Nonresidential Standard Performance Contract Program Measurement and Evaluation Study. Prepared for California investor-owned utilities. Oakland, California.