# To Do or Not to Do Redux: Triggering an Impact Evaluation?

*Erik Mellen, Northeast Utilities, Westwood, MA*
*Susan Haselhorst, ERS, North Andover, MA*

## ABSTRACT

Program administrators (PAs) of custom commercial and industrial (C&I) energy efficiency programs and their regulators inevitably must decide whether to conduct another round of an expensive impact evaluations. Two years ago at IEPEC 2013, the authors presented a novel approach ("Desk Review Assessment" or DRA) for developing objective criteria that would aid in deciding whether to proceed with an expensive full-scale evaluation. Although the DRA method had promising initial results, the test was not fully conclusive. This paper presents the second half of the DRA experiment and also includes a discussion of the California Program Practices Assessment (PPA), which is a recent initiative that has characteristics in common with the DRA.

This paper compares the desk review projections of change to realization rate (RR) stability for two different program years (Round One and Round Two) for a combination of four program administrators. The paper also discusses the California PPA.

## Introduction

Custom C&I program impact evaluations are particularly expensive, requiring on-site measurement and verification to confirm the applicant's savings estimates. The decision to repeat a study is influenced by available funds, program size, and the perceived stability of the program, and often just because an arbitrary period of time has passed since the last evaluation. The DRA method proposes a systematic desk review of a sample of projects to determine whether the engineering estimation process had changed sufficiently to warrant a full M&V impact evaluation. In concept, significant program changes, including changes to engineering savings estimation methods, should be the primary trigger for an impact evaluation because a stable program should produce stable realization rates. This desk review method tests a key element of program delivery – the measure savings estimation process.

In the method, a sample of projects files is selected for desk reviews using a structured rubric which produces a score of program performance in seven categories, or criteria. The performance of the test year can be compared to the score of previous years to measure program change. The inference is that if the present program is producing measurably different savings results from the benchmark, it is prudent to proceed with the full-scale impact evaluation. The incremental cost to complete an M&V impact assessment is about $10,000 per site, while a desk review of the same site is about an order of magnitude less expensive.

In 2013 at the IEPEC conference this team presented the desk review method for triggering an impact evaluation with results from one evaluation cycle. Although the findings were promising, they included only one confirmed data point; this round of testing provides four additional data points.

The experiment began in 2012 at a time when Massachusetts Custom Gas Working Group (Working Group) was faced with a dilemma: Should they move forward with a third consecutive impact evaluation with the hope of boosting the program RR, or postpone it to conserve resources, but potentially under-report savings?

The Working Group is responsible for the direction and execution of the evaluation of the natural gas CI programs. The Working Group is composed of the gas energy efficiency PAs of Columbia Gas, National

Grid Gas, Eversource Gas, Berkshire Gas, New England Gas, and Unitil, the evaluation contractors ERS and DNV GL, and the Massachusetts Energy Efficiency Advisory Council (EEAC) consultants. Programs are designed and evaluated jointly and statewide, although each PA individually administers the program, with unique processes for outreach, savings estimation oversight, and tracking. The PA's gas programs had been transformed from a small-budget, moderate technical review model in 2008 to a rapidly expanding program with more rigorously reviewed savings estimates thereafter. The program ambition had increased as well, expanding the portfolio to include a wide array of measures, such as high efficiency heating equipment, heating systems, heating controls, EMS, boiler combustion controls, building shell measures, and a variety of high efficiency gas industrial process equipment. As illustrated in Table 1, the program doubled in savings for 3 consecutive years statewide but has recently stabilized at about 8 million therms per year in natural gas savings. The table also notes the impact evaluation status of each program year and its role in the experiment.

**Table 1. Massachusetts Gas Energy Efficiency Program Accomplishments**

| All Program Administrators | 2009 | 2010 | 2011 | 2012[b] | 2013 |
|---|---|---|---|---|---|
| Number of participants | 339 | 335 | 369 | ≈500 | 623 |
| Total tracking savings (therms) | 1,978,536 | 4,427,361 | 7,915,793 | ≈7,760,000 | 8,345,907 |
| Statewide realization rate | 71%[a] | 68% | 82% | NA | 68% |
| Sample size | 43 | 48 | 16 | NA | 46 |
| Evaluation status | Full impact | Full impact | Partial | NA | Full impact |
| DRA status | Benchmark | | Round One test; Round Two benchmark | None | Round Two test |

[a] Controlled for outlier
[b] Custom share of total program-reported savings has been estimated.

The PA-sponsored evaluations in PY2009 and PY2010 were based on on-site M&V of a representative sample of participants. The RR each year was about 70% statewide. The past evaluations had concluded that administrative errors and factors that could have been identified in a more rigorous technical review contributed to variances in RRs. The PAs were taking steps to improve the technical review. However, since each PA independently administers a statewide common program, process improvements across PAs were not uniform. In considering the PY2011 evaluation plan, some of the PAs were convinced that significant improvements to the process had been made, while other PAs concluded that their process improvements were barely underway. It was not clear whether the PY2011 projects of the third year reflected enough improvement to warrant another impact evaluation, either statewide or for any particular PA.

## Try the Approach Once, Try It Again

Rather than embark on a full impact evaluation or postpone an evaluation entirely, the Working Group proposed a new approach. The group agreed to test, through a systematic review of a sample of PY2011 projects, whether the engineering estimation process had changed sufficiently to warrant one or more of the PAs proceeding to a full M&V impact evaluation. In concept, significant program changes, including changes to engineering methods, should be the primary trigger for an impact evaluation because a stable program should produce stable RRs. This proposed method offered a way to test a key element of program delivery – the measure savings estimation process.

In commissioning this task, the Working Group agreed to a process where a statistically selected sample of the test year (PY2011) sites would undergo desk reviews (the desk review sites) to characterize the current state of savings estimate quality. These results would be compared to similar reviews of sites that underwent M&V in the last two evaluations (the benchmark sites) to determine if there was a measurable improvement in the PY2011 methods.

The key ground rule was that objective criteria had to be determined prior to the completion and presentation of the PY2011 desk review results to avoid inadvertent tilting towards a preferred outcome. These were dubbed the decision criteria. It was also agreed that a decision whether to proceed to a full impact evaluation could be made independently for each of the three PAs with the largest savings (PA1, PA2, and PA3 for the purposes of this paper) and then statewide.

## Methodology

The method for implementing the framework is outlined in Figure 1 and described in some detail in the subsequent five sections.

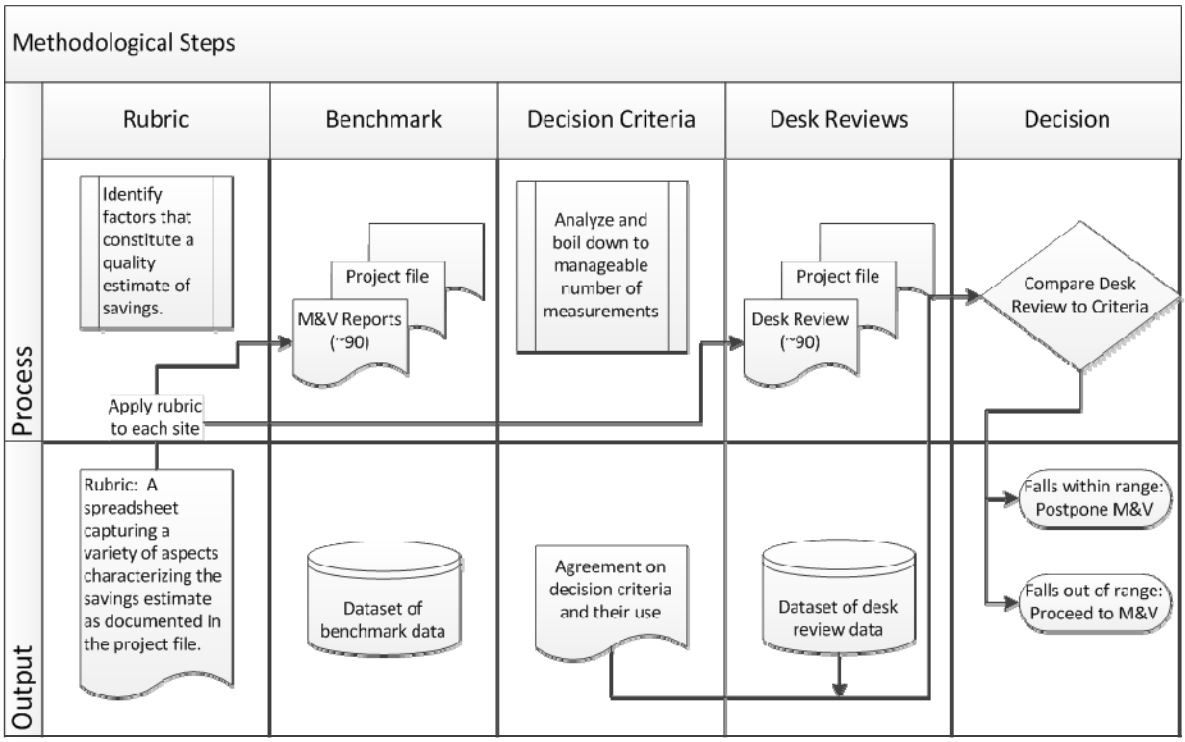


**Figure 1. Framework Methodology**

## Step 1: The Rubric

As a first step, the impact team created a rubric for assessing the quality of a project savings estimate. This rubric had to capture the judgments made by an engineer during a review of applicant savings and had to be based on the material available to a reviewer prior to the installation of the measure and stored in the project file. For example, the results of the review could not rely on post-installation gas usage, as that information would not be available for an administrator reviewing an estimate as part of an application approval process.

The team focused on characterizing aspects of the project that could be reviewed from the project file alone and, when done properly, would lead to a better estimation of savings. The characteristics that are considered important could be summarized as follows:

- Was the baseline correct for the measure?
- Was an appropriate savings methodology employed?
- Was there evidence that customer billing had been consulted in reviewing the savings estimate?
- Was the savings fraction (savings as a percentage of total pre-installation gas usage) reasonable?
- Were all the documents present in the file (application, invoice, savings estimation description, native spreadsheets, and models)?
- Were the savings reproducible?
- What was the quality of the overall savings estimate?

These questions were translated to a spreadsheet designed to capture the reviewer's responses systematically and consistently from site to site. Table 2 represents the rubric showing some of the more important fields, although altogether there were seventy parameters entered by the engineer for each site.

**Table 2. Excerpt of the Spreadsheet-Based Rubric**

| Item | Notes |
|---|---|
| **Customer and Measure ID** | |
| Site ID | Evaluation-assigned site ID |
| Measure ID | Evaluation-assigned measure ID |
| Customer type | Examples: retail, office, K-12, hospital, etc. |
| Tracking measure category | Measure type: boiler replacement, EMS, etc. |
| **Quantitative Savings Analysis** | |
| Tracking savings, therms | Per tracking data |
| Evaluator desk review savings, therms | For benchmark: actual evaluated savings. For desk reviews, evaluator estimates. |
| Evaluator pre-install weather normalized billed gas use | Best available weather normalized billed use. Used for calculating savings fractions. |
| Applicant savings fraction | Tracking savings / pre-install gas use |
| Evaluator savings fraction | Evaluator savings / pre-install gas use |
| **Documents Checklist** | |
| Document list: application, offer letter, TA study, calculations, invoice, inspection report | The engineer checks off each document type found in the project file. |
| **Tracking and Billing Review** | |
| Status of tracking and billing data in the file. | Checklist inventory and quality assessment |
| **Baseline Assessment** | |
| Baseline specified by the applicant | Indicates whether a retrofit or code baseline was used. The engineer also specified whether the baseline choice was clear, implied, or not clear. |
| Baseline determined by the evaluator | Evaluator judgment whether a retrofit or code baseline is appropriate. |
| **Assessment of Savings Estimation Methods and Quality** | |
| Building simulation | The engineer indicates which method was best suited to the measure and which method was used by the applicant. |
| Proprietary method | |
| 8760 or bin spreadsheet | |
| Factor driven, one-line calcs | |
| Estimation quality | Evaluator judgment of quality of estimate overall. |

Pick lists, predefined descriptors embedded in the spreadsheet, were defined for qualitative parameters to permit comparison across projects. For example, the "overall quality of the estimate" reflects the reviewer's overall judgment about the estimation method documented in the project file. A higher-quality savings estimate provides appropriate assumptions supported by site-specific information with transparent methods of calculation.

## Step 2: Creating the Benchmark Dataset

Once the rubric was designed, the engineering team went back to the site reports and project files selected and scored the earlier M&V sites from the PY2009 and PY2010 evaluations to create the benchmark for the Round One evaluation. In Round Two, the benchmark was expanded to include data from PY2011 as well as PY2009 and PY2010.

## Step 3: Defining the Decision Criteria

In the first round, the decision criteria had to be conceptualized and quantified in such a way that that the pass/fail test would be unambiguous once the desk review sites' results were in. In 2012 as this method was being developed, the Working Group agreed that the criteria had to be established before the desk review step to ensure objectivity. In addition, all the PAs and the EEAC consultants had to agree to the decision criteria even though some of the PAs hoped for opposite outcomes.

There were multiple options for how to proceed. How many criteria should there be, and what should they be? Should they be based on a simple site count or weighted in a manner reflecting the site's impact on program outcomes? Should individual criteria be weighted or each counted the same? How should non-numeric parameters, such as the quality of estimate, be translated into an objective score? How should the margin between passing and failing be defined?

Ultimately, the Working Group agreed to seven decision criteria, as shown in Table 3. Each criterion was presented as a percentage of total program tracking therms meeting the criterion. Thus, for example, a benchmark score of 75% for the "baseline is appropriate" criterion can be interpreted as indicating that the baseline was appropriate at benchmark sites representing 75% of the program therms. The criterion also specified the range of values (No Action Range) considered close enough to the benchmark to show that the process changes are insufficient to warrant an impact evaluation.

**Table 3. Statewide Decision Criteria and Benchmark Summary**

| Criterion | Benchmark Value | | Wtg |
|---|---|---|---|
| | 2011 | 2013 | |
| **Baseline is appropriate** – This criterion captures how often the applicant identified the correct baseline (retrofit or replacement at end of life). Inappropriate baselines were a major source of discrepancies in previous evaluations. | 75% | 78% | 40% |
| **Savings method was appropriate** – This criterion captures how often the applicant used an appropriate savings calculation method. | 47% | 54% | 10% |
| **Savings fraction** – This is the average program savings as a percentage of the average pre-installation bills. | 8.2% | 7.5% | 10% |

| Criterion | Benchmark Value | | Wtg |
|---|---|---|---|
| | 2011 | 2013 | |
| **Document inventory** – This criterion represents the frequency of certain documents observed in the project files. | 44% | 44% | 10% |
| **Evidence of bills in the file** – This criterion captures how often bills appeared in the project files since gas bills are so useful in estimating or benchmarking gas savings. | 35% | 40% | 10% |
| **Savings were reproducible** – This criterion indicates how often there was sufficient information for the reviewing engineer to reproduce the applicant savings. | 54% | 47% | 10% |
| **Quality of the estimate** – This is an overall assessment of the quality of the savings estimate. Table 3 specifies the five choices. | 67% | 69% | 10% |
| **Threshold standard** | 20% | | |

To finalize the criteria, the Working Group had to finalize the range of values for each criterion where no M&V would be required (No Action Range).

The degree of change in the criterion value considered significant enough to warrant proceeding to the on-site work (the "threshold standard") was 20%. This threshold is somewhat arbitrary. Finding that gas billing is factored into the savings analysis 20% more of the time, for example, shows an improvement in the estimation process, but it does not follow that savings will increase 20%. That being said, a 20% change in a criterion is likely to be large enough to rise above the noise in the results, indicating that more systematic changes have occurred and yet not so large as to preclude the identification of any improvements. The Working Group also agreed to weight the individual criteria, as shown in Table 5, into a single score. As an example, referencing the 2011 column in Table 3, assume that the first three rows exceed the benchmark by 30% (respectively, 98%, 61%, and 10.7%). The three criteria exceed the 20% threshold, and therefore each is considered an affirmed indicator of change and has a value equal to its weight (40%, 10%, and 10%). Assume that the remaining criteria vary from the benchmark by less than 20%; therefore, each is affirmed as indicating insufficient change and has a value of 0. The sum of the values is 0.6 (0.4+.1+.1), which the evidence weighs in favor of significant program change, and so an evaluation is warranted.

For Round Two, this process was much simpler since the benchmark was simply updated to include PY2011 results.

## Step 4: Desk Reviews of Test-Year Projects

Once the decision criteria had been defined and agreed upon, the engineering team commenced the desk reviews of a statistically selected sample of projects, applying the rubric to each. The sites were selected using an on-site M&V sample strategy. If the results proved a site M&V impact evaluation was warranted, the engineering team could quickly and efficiently convert the desk reviews to a site M&V plan. Samples of eighty-five and ninety-two sites were reviewed in Round One and Round Two, respectively.

## Step 5: Compare Desk Reviews to the Benchmark

As the final step, the desk review results from the test year are compared to the benchmark. If the test year strays from the benchmark by more than 20%, the deviation is considered substantive. Using the weighting, with a total score of 0.6 or greater, the PA desk reviews are considered eligible for reevaluation.

Before examining the actual results from Round One and Round Two of the DRA method, we will examine another interesting and related development in California with the PPA reviews.

## California Low Rigor Assessment

At the time of the publication of the first paper, there was an effort underway in California for C&I custom impact evaluations using a "low rigor assessment" (LRA). The LRA is quite similar to the process outlined in this paper and includes a rubric for scoring specific aspects of the project. The rubric is focused on project compliance with CPUC policy and ex ante review (EAR) guidance, conformance with program rules, use of best practices from industry M&V protocols, and baseline selection.

As instituted in 2010–2012 impact evaluation, every site in the sample received an LRA review. The purpose of the LRA was to systematically collect observations of the quality of the savings estimate from a sample of about five hundred sites. The LRA was designed as the first step in the M&V impact evaluation for a site. The LRA template-compiled data served as a data request form for the PAs when data was missing and identified other potentially contentious issues like baseline changes upfront. The LRA was also updated through the course of the site work, while new data was gathered from additional paperwork provided by the PA or from the site inspection. Data from each site LRA was compiled into a single database and analyzed to provide a numerically oriented summary of how well each PA prepared the custom savings estimates.

The study concluded that the LRA process was a positive contribution to the impact evaluation. First the LRA was helpful in the execution of the impact evaluation. The LRA imposed upon the site engineer a process for explicitly and consistently addressing key elements that should always be considered when conducting a site evaluation. The LRA also served as a template for identifying missing documentation and data that should have been in the project file. Because the LRA was comprehensive, all the missing data tended to be identified at once, leading to a single data request to the PA. Secondly, the aggregated scoring data was found to be useful to the PAs particularly in identifying weaknesses in their custom savings estimates. At least one of the PAs is now using the LRA form as a template for their review of custom measures.

## Results: Round One

The Round One (PY2011) criteria scores are presented in Table 4 for the state as a whole and also for the three largest PAs. Those cells that are shaded show where a criterion was out of the No Action Range, indicating that the savings estimation process had improved or regressed and an M&V impact evaluation was warranted. Each of the three largest PAs had its own unique benchmark, although only the statewide criteria are provided in the table. Criteria that remained within range are not shaded. Results are presented for the state and also for the three largest PAs. As noted previously, the Working Group had agreed that the results would be examined statewide and by each of the three largest PAs. The table concludes with the desk review conclusion of whether an impact evaluation was indicated ("Yes") or not and the M&V RR for the one PA where an impact evaluation was recommended and full site M&V impact evaluation was conducted.

**Table 4. Desk Review Results Compared to Decision Criteria in Round One, PY2011**

| Benchmark | Statewide Benchmark Value | State | PA1-Col | PA2-NG | PA3-NS |
|---|---|---|---|---|---|
| **Desk Review Results** | | | | | |
| *Desk review ratio* | *74%* | *70%* | *56%* | *75%* | *67%* |
| Baseline is appropriate | 75% of the time | 79% | 74% | 78% | 87% |
| Savings method was appropriate | 47% of the time | 61% | 85% | 47% | 72% |
| Savings fraction | 8.2% | 6.8% | 6.8% | 6.7% | 7.6% |
| Document inventory | 44% of docs found | 42% | 47% | 43% | 48% |
| Evidence of bills in the file | 35% of the time | 45% | 71% | 38% | 42% |
| Savings were reproducible | 54% of the time | 39% | 27% | 47% | 72 |
| Quality of the estimate | 67% good quality | 71% | 80% | 65% | 78% |
| Weighted score | | 0.3 | 0.4 | 0.2 | 0.9 |
| Recommend M&V | | No | No | No | Yes |
| M&V Results | | | | | |
| Prior RR | | 68% | 83% | 68% | 47% |
| New RR | | 82% | Not revised | Not revised | 84% |

These findings indicate that a significant change in practice was not indicated broadly enough to warrant another statewide impact evaluation (only three of seven criteria are out of range). However, when the results are examined according to PA, a different conclusion is reached for PA3. Both PA1 and PA3 did stray outside of the range more often than not. PA1 showed both improvements in three categories and an erosion in reproducibility. However, when the criteria are considered on a weighted basis, they indicate that only PA3 showed sufficient change to warrant another impact evaluation with all criteria indicating the same trend towards improvement.

These conclusions are aligned with the PA reports of process changes. PA3 reported that a significant and definitive change occurred in the late 2010 time frame. Prior to the change, the gas program manager conducted the savings estimate review; after that date, staff engineers were assigned the responsibility to review custom estimates of savings. The other PAs did not identify any such sharp change in practice.

Based on the evidence of this process and the confirmatory information from the PAs, the Working Group decided to proceed with an impact evaluation of PA3's program only. PA3 showed significant improvement in its RR, which resulted in a recalculation of the statewide RR, also improving it.

## Round Two DRA Results

Two years later, the Working Group had agreed to another round of impact evaluation based on on-site M&V of about fifty sites. The Working Group also agreed to precede the site work with desk reviews of

a larger sample (about one hundred sites) to test how well the method predicted a change in RR outcome. This work proceeded as noted above. The desk reviews scored projects prior to completing the on-site work.

Table 5 is similar to Table 4 summarizing the benchmark performance by the three largest PAs and statewide against their benchmarks but with additional notation indicating if the new criterion value should yield an improved RR (*I*), an eroded RR (*E,*) or no change (*N*). Overall, the scores indicated that compared to benchmark, the processes had improved almost universally. There was only one exception for a value with a significant difference: PA1's documentation criterion. However, only PA2 showed enough improvement in aggregate to warrant, in theory, another site-based M&V impact evaluation.

An informal poll was conducted with the three largest PAs prior to issuing the DRA results to determine if there had been any significant changes in the delivery of the program since the 2011 program year. While it was noted that there are ongoing improvements and staff changes, there was nothing that stood out in this period with the exception of National Grid's reclassifying of multifamily projects consisting of prescriptive measures only as prescriptive projects. However, while the PAs may not have implemented any abrupt change to the process, it did appear from the file reviews that there must have been an ongoing effort to improve the estimating process because the engineers observed files that were more often complete, with better algorithms, more native files, and other markers of an improved process.

**Table 5. Desk Review Results Compared to Decision Criteria in Round Two, PY2011**

| Benchmark | Statewide Benchmark PY2009–PY2011 | State | PA1-Col | PA2-NG | PA3-NS |
|---|---|---|---|---|---|
| **Desk Review Results** | | | | | |
| *Desk review ratio* | *72%* | *70% – E* | *78% – E* | *66%–E* | *67% – I* |
| Quality of the estimate | 69% | 81% – I | 80% – E | 91% – I | 77% – I |
| Appropriate methods | 54% | 74% – I | 65% – N | 84% – I | 73% – I |
| Savings was reproducible | 47% | 79% – I | 76% – I | 90% – I | 76% – I |
| Savings fraction | 7.5% | 6.0% – I | 5.7% –I | 6.7% – E | 5.6% – I |
| Bills included in folder | 40% | 89% – I | 93% – I | 93% – I | 86% – I |
| Baseline is appropriate | 78% | 85% – I | 92% – I | 99% – I | 75% – E |
| Document inventory | 44% | 47% – I | 22% – E | 73% – I | 58% – I |
| Weighted score | | 0.3 | 0.4 | 0.9 | 0.5 |
| Recommend M&V | | No | No | Yes | No |
| **Site M&V Results** | | | | | |
| Previous RR | | 82% | 83% | 68% | 84% |
| New Site M&V RR | | 68% | 67% | 64% | 71% |

## Round Two M&V Results

While the DRA results indicated that only one PA warranted an impact evaluation, it had been the intention of the Working Group to proceed with a full impact evaluation study all along. A sample of forty–six sites was drawn and full M&V work was completed. The RRs are shown in the last line of Table 5.

Much to the disappointment of the Working Group on multiple levels, the RR had not improved; RR improvement is always the goal of a program implementer and the DRA had predicted an improvement to the RR. In fact, the program RR had generally eroded since the last full impact evaluation for the three large PAs, which drove the lower statewide RR. PA2 may have expected an improved RR since the DRA had indicated that it met the test for a round of impact evaluations with improvements in almost every criterion, yet its RR remained the most stable. (Note, at the time of this paper, the PAs had not fully reviewed all the site reports and some change in RR may be expected, although the final RRs are should largely remain the same.)

What happened?

While the analysis is not complete at this time, there did appear to be two trends in projects where it appeared that the savings estimating process had improved, yet the RRs may have been poorer:

- Poorly performing ventilation controls. These measures appear to have particularly poor performance when examined with reliable billing analysis, even if the post-installation system appeared to be operating correctly. The implication is that the preexisting to code ventilation rates were much lower than had been assumed by the applicant.
- New construction and comprehensive design (CDA). There was an increase in new construction and CDA projects (from two in the previous sample to six in this round). The CDAs scored relatively well in the site reviews since the ex ante savings were supported with full TA studies and with building simulations. In a few cases, though, it appeared that the CDA analytic effort was focused on identifying electric savings with the natural gas savings accounted for as an afterthought and that a deeper examination of the models revealed flaws in the modeling of gas measures.

However, in the end, the reasons for discrepancies in RR were varied and there is no single reason why the RRs were lower than expected.

## Conclusions and Recommendations

The intent of the DRA as described in this paper is to provide a reproducible and systematic method for triggering a more expensive on-site M&V impact evaluation. A similar process, although with a different purpose, has been piloted in California, with more of a focus on systematizing aspects of the impact evaluation and to provide implementers with actionable feedback of where savings estimates are weak.

Neither effort has been able to show a strong ability to predict the likely change in RR outcomes either through the experimental design described in this paper or using statistical techniques that have attempted to correlate site criteria with site RR outcomes, at least as of yet. The impact evaluation team responsible for the California PPAs is convinced that with more sophisticated statistical analysis the PPA may become more predictive of RR outcomes. However, in the end, a good estimate may yield an RR much greater or much less than 1.0 due to changes that have occurred at a site, and a poor estimate may produce an RR of 1.0 just by luck.

The method was useful to the Working Group in Round One to resolve a fundamental disagreement between members of the group about next steps. Round Two has shown that the method was not successful at predicting changes in RR as had been hoped. The process, however, has merit as a tool for systematically collecting observations about the custom savings estimation process that can provide detailed and quantitative feedback to implementers. The incremental cost of completing the desk review is quite low

when implemented as part of the normal site M&V process and can lead to a higher quality and more complete site report, as it provides a ready-made M&V checklist.

Despite the limited success of either the Massachusetts or the California models in predicting RR, the tools used to do the assessments provided value as quality control and data collection mechanisms.

## References

Ehrlich, Charles, Al Lutz, Kris Bradley. *Using Structured Desk Reviews to Manage Risks Associated with Commercial and Industrial Custom Impact Projects.* Presented at AESP, February 2015.

Itron and DNV GL. *2010–12 WO033 Custom Impact Evaluation Final Report* submitted to CPUC.