Results from Streamlining of Custom Project Impact Evaluation

Jeremy J. Newberger, Principal Analyst, New England Electric System, Northborough, Massachusetts

ABSTRACT

Since 1995, New England Electric System has reviewed and streamlined its impact evaluation effort for custom (non-prescriptive) projects in an attempt to reduce costs while maintaining the quality of the results. Its streamlining plan was presented at the 1996 ACEEE Summer Study.

In this paper, results from the 1994 program year, which relied heavily on site-based metering, are compared to results from the 1995 and 1996 program years, which loosened the requirements for site-based metering and allowed alternative evaluation techniques. Comparisons are made between the evaluation budgets, realization rate results, and relative precision of the two populations.

The sources of the discrepancies between original savings estimates and final evaluated results are also analyzed and presented in terms of whether they could have been identified prior to program evaluation. The impact of the savings discrepancies on the utility's authorized incentive is estimated and compared to estimates of expenses for improvements in project quality "upstream" of evaluation, such as additional documentation, project management, or commissioning.

Based on these analyses, observations and conclusions are offered regarding future directions for the streamlining effort and project quality improvements. The potential trade-off between spending fewer resources on evaluation and potentially more on project quality are explored.

Introduction

Background

Customized energy efficiency measures implemented in the large commercial and industrial (C&I) sectors of the New England Electric System's (NEES) retail companies[1] account for a steadily increasing fraction of the gross C&I program energy savings achieved in each of the past five years. As the savings have increased, the utility has increasingly felt -- first internally and later with the restructuring process occurring throughout the states in which it operates -- the need to control evaluation costs. The utility reviewed and streamlined its custom project impact evaluation effort in an attempt to make it less costly while maintaining the quality of the results.

Based on an analysis of discrepancy classification and project type, guidelines were implemented beginning with evaluation of the 1995 program. The guidelines focus on reducing the use of short-term

[1]The New England Electric System (NEES) operates and offers DSM programs through four retail companies: Massachusetts and Nantucket Electric Companies in Massachusetts, The Narragansett Electric Company in Rhode Island, and Granite State Electric Company in New Hampshire. New England Power Service Co. (NEPSCo) is another subsidiary of NEES and handles many of the core tasks for all four retail companies, including impact evaluation of all DSM programs. While the work described in this paper was performed by NEPSCo, "NEES" will be used as the common reference.

metering for evaluating custom projects. The development of the discrepancy classification and guideline response are presented in depth in a prior study (Newberger, 1996). In summary, the primary reason was identified for the difference between the projected, or tracking, estimate of savings and the savings verified through the evaluation studies. The "discrepancy codes" are described in Table 1.

Table 1. Discrepancy Categorization

Code	Discrepancy Description
A	OK (no discrepancy).
B	Calculation errors.
C	Existing metering data used from application or utility billing.
D	Inoperable equipment.
E	Misapplication of technology.
F	Equipment changes from application, including count.
G	Adjustments to original analysis baseline assumptions and/or other methodological differences.
H	Changes in hours of operation or equipment loading.
I	Sub-optimal control of operating equipment.
J	Equipment performance different from predicted.

These discrepancy types may be further aggregated. Codes A, C, and H[2], represent negligible discrepancies or those beyond the control of everybody involved with the project. Any legitimate evaluation method is likely to come up with similar results. Codes G, I, and J, are fine discrepancies revealed only through a thorough analysis of the project which likely includes metering. Codes B, D, E, and F are gross discrepancies which should be apparent without any metering.

The Streamlining Plan

The plan which was implemented featured three parts:

---

[2] Category H, a change in loads or hours, may or may not be revealed through intensive metering. Loads discrepancies may only be apparent through metering, while significant hours changes may be apparent from discussion with facility operators. Therefore, an evaluation method without intensive metering may be appropriate for some of these sites as well. Furthermore, variations in loads and hours may occur randomly over the life of a project. For these projects, it would be appropriate to allow a band of tolerance around the original estimate of loads or hours.

(1) Development and implementation of evaluation guidelines to reduce the amount of on-site metering in evaluation studies
(2) Combining two years of studies for some end use groups
(3) A quality control initiative in response to some of the prevalent discrepancy types identified in the evaluation studies.

At the plan's inception, NEES projected that it would potentially reduce the number of intensively metered sites by about 40%, with a negligible impact on realization rate results. An evaluation cost reduction of 16% to 20% was projected. Further cost reduction was expected through combining two years of projects in some end-use studies.

Evaluation Guidelines

Guidelines were written for evaluators to consider using evaluation techniques other than site metering. These were incorporated into the evaluation scope of work, beginning with the impact evaluation of the 1995 program in 1996[3]. For each specific discrepancy type, a guideline was created in an attempt to facilitate evaluation without extensive site metering. The guidelines instructed evaluators:

- to cut short studies with major calculation errors (addresses code B);
- to steer evaluators toward use of hourly meter data, billing data, or post-installation metering data made available by others (code C);
- to cut short studies that may waste effort evaluating projects that don't work (codes D and E);
- to determine the impact of equipment changes before performing the evaluation (code F); and
- to assess the relative importance of hours or loads changes before metering (code H).

The Two Year Study Plan

NEES designed its 1996 evaluation studies to be used in reporting the results[4] for custom projects paid in both 1995 and 1996. Some of these two year studies were in end-uses whose savings are often cooling season dependent. Given the retail utilities' July filing deadlines, NEES had felt constrained to evaluate HVAC projects in the spring. The two year plan freed the evaluators to evaluate these projects in the proper season. Other end-use studies not planned for the second year were in areas where results had not shown much change from year to year or where participation was low. Significant cost savings were expected even though the evaluation period for some of the weather-dependent end-use groups would be somewhat protracted, with associated costs.

[3]The evaluation calendar lags one year behind the program year. Thus, projects implemented in 1995 were evaluated and reported on in 1996, and so on. Therefore, "1997 studies" and "1996 results" refer to the most recently completed evaluation studies of 1996 projects. This paper compares the results for 1995 and 1996 to the results for 1994.

[4]While there are differences in the programs implemented in each state, evaluation studies are usually designed to be representative and applicable for all states.

Quality Control (QC) Initiative

The QC initiative was in response to the discrepancy analysis which showed that many of the discrepancies in savings between the original values and estimated results were due to factors that could have been corrected "upstream" of evaluation such as calculation errors, poorly documented or unsupported assumptions, building simulation modeling errors in DOE-2 and Trace, and failure to update savings after installation when a project had changed measurably.

The following key recommendations were adopted at the end of 1996. Thus, their impacts will not begin to be evident until the 1997 program year is evaluated.

- An explicit link was created between the minimum requirements document (MRD)[5] and post-installation inspection: each MRD item must be checked off or initialed prior to rebate payment.
- Commissioning was made mandatory for projects with rebates of $100,000 or greater.
- Engineers were made aware of the importance of improving documentation of assumptions to support savings estimates in technical reports and of documenting post-report changes to a project. In addition, utility field Technical Representatives were given the option to reserve the right to sign-off on post-installation inspection, and change savings as necessary based on actual post-installation conditions.
- NEES commissioned a study to document standard default assumptions in commonly used building simulation models. The intent of this was to (1) give utility technical representatives better understanding of sensitive variables in these models and enable an improved review of savings estimates based on simulation and (2) alert the simulating engineers of the utility's increased awareness of these sensitivities and thus inspire them to better document their changes to building simulation models.

Results

Reduction in the Number of Intensively Metered Projects

A 40% reduction in the number of sites evaluated without metering was expected, based on an analysis of projects evaluated in 1995, virtually all of which involved extensive on-site metering. For the 1996 and 1997 evaluation studies, 19% incorporated no metering, while 32% involved spot metering of 5 hours or less, and 49% of the projects still employed metering installations for more than one day. The number of non-metered sites fell short of expectations, but this is partially compensated for by the number of spot metered sites. In place of extensive on-site metering, evaluators commonly used facility operating logs, EMS trend data, or meter data from utility records, supplemented by spot metering. This did not necessarily reduce the computational burden of the analysis performed by the evaluators, but did reduce the number of site visits by not requiring a follow-up visit to retrieve meters and, consequently, evaluation costs.

[5]A minimum requirements document (MRD) is created for most custom projects, with the exception of some end-uses, such as lighting. The MRD describes equipment and performance specifications for the project on which savings are contingent.

Evaluators frequently found post-installation metering had been provided with project data and in most cases found this data adequate for analysis. In some instances, evaluators also secured the cooperation of facility personnel to simulate loading of equipment during the spot measurement period. This was particularly employed in the evaluation of variable speed drive projects. NEES study managers worked with evaluators to leverage the use of available data and tailor the evaluation approach to the particulars of a given project. The extensive use of meters also was influenced by whether the evaluation contractor owned its own meters or leased them for the study, since the cost allocated to the study would be lower for owned meters than rentals.

Negligible Impact on Realization Rates or Relative Precision

The streamlining plan predicted a negligible impact on realization rate and precision results, reasoning that the guidelines were well suited for identifying discrepancies of large magnitudes such as those found in categories B, D, E, and F. Metering sites where large discrepancies had been identified prior to metering would not be likely to change results much.

Table 2 shows a comparison of realization rates for the 1995 studies compared to the 1996/1997 studies. A realization rate is the ratio of evaluated savings to predicted savings. Realization rates for energy are shown as they are the primary influence on a project's cost-effectiveness.

Table 2. Custom Evaluation Study Results

	Realization Rate	
End Use Group	**1996/1997**	**1995**
Compressed Air	1.21	1.22
Motors	n/a	n/a
Process Cooling	1.21	0.77
Process	0.88	0.96
Commercial Refrigeration	0.82	0.48
Industrial Refrigeration	0.85	0.86
VSDs	0.84	0.64
EMS	0.75	0.96
HVAC	0.88	0.82
Comprehensive Design	1.06	0.78
Chiller Initiative	0.77	n/a
Lighting	1.19	n/a
Lighting Controls	1.24	n/a
Overall	**0.95**	**0.84**
Overall relative precision	**8%**	**9%**

As seen in the table, the overall realization rate increased from 84% to 95% while relative precision improved by 1 percent (Wright, 1997). Among the 9 end use groups in which studies were performed in both study horizons, realization rates improved in 5 groups, decreased in 2 groups, and showed negligible change in 2 groups. Some of the variation from one year to the next is obviously influenced by the individual projects analyzed: one outlier performer in any one study can exert great influence on the realization rate. However, the results show that the introduction of the evaluation guidelines did not adversely affect these quantitative study results.

Evaluation Cost Reduction

The streamlining plan envisioned approximate savings of 16% to 20% from the reduced emphasis on metering. Table 3a shows the distribution of the studies performed in each year and Table 3b summarizes evaluation expenses for 1995 compared to 1996 and 1997.

Table 3a. Distribution of Evaluation Studies by End Use Group

End Use Group	1995 Study	1996 Study	1997 study
Compressed Air	x	x	
Motors			
Process Cooling	x	x	x
Process & Other	x	x	x
Commercial Refrigeration	x	x	
Industrial Refrigeration	x	x	
Variable Speed Drives	x	x	x
Energy Management Systems	x		x
HVAC	x	x	x
Comprehensive Design Approach	x	x	
Chiller Initiative		x	x
Lighting		x	
Lighting Controls			x

Table 3b. Comparison of Evaluation Costs

	1995	1996	1997	1996 and 1997	
				Total	Common with 1995
Number of end uses	8	10	7	12	8
Number of sites	88	116	41	157	141
Total evaluation expense	$645,961	$739,046	$278,093	$1,017,139	$744,871
Average cost per site	$7,340	$6,371	$6,783	$6,479	$5,283
Percent of 1995 $/site		87%	92%	88%	72%

Note: 1995 VSD study excluded from cost analysis because it was exclusively a phone survey.

The most dramatic impact is seen in the reduced spending in 1997 to evaluate the 1996 program. This is a function of the two year study plan which reduced the number of site surveys to 41 in 1997 from 116 in 1996.

Over the two year study horizon, the impact of the introduction of the evaluation guidelines may be seen in a comparison of average cost per site. The average cost per site for the 1996/1997 studies was 12 percent lower than the cost per site in 1995. For the eight end use groups in which studies were performed both in 1995 and 1996/97 study horizons, the cost per site declined by 28%. While the size and nature of the projects selected for evaluation influences the cost per site, much of the decrease is attributed to the introduction of evaluation guidelines. As an example, the Energy Management System (EMS) study performed in 1995 was the most expensive study in that year at a cost of $16,000 per site. The 1997 study of more complex EMS projects by the same evaluator cost $11,000 per site.

One of the ancillary benefits of the two year study plan was the flexibility it provided to include end use groups that had not been previously studied while still keeping overall costs down: the 1996/97 average was lower despite including a study of the Chiller Initiative program at a cost of $31,700 per site.

Discrepancy Distribution

Table 4 shows the distribution of discrepancy types for the two study periods. As seen in the table, the percent of projects evaluated with minor discrepancies increased, while those with gross discrepancies decreased. This trend is indicative of an improvement in quality. As quality improves, gross discrepancies may be caught before evaluation and their numbers will be reduced.

Table 4. Distribution of Applications by Discrepancy Type

	1995	1996/97
Minor/No Discrepancies (Codes A, C, H)	37%	48%
Fine Discrepancies (Codes G, I, J)	37%	37%
Gross Discrepancies (Codes B, D, E, F)	26%	15%

Current Efforts

End Use Group Aggregation

In developing the study plan for 1998, the innovation was to further aggregate the end use groups. The number of end use groups was reduced from 13 to 7, and a new group for Operations and Maintenance projects was introduced, bringing the total to 8. The aggregation serves to further reduce the number of site studies. For example, the Process Cooling, Commercial Refrigeration and Industrial Refrigeration groups were combined into one Process Refrigeration group. Instead of performing three separate studies with between 25 and 30 sites, one study of 10 sites is being performed.

Underlying this aggregation is a de-emphasis on the relative precision of any individual group in favor of emphasizing the precision of the study population as a whole. Based on results from previous studies, the relative precision for the 1998 studies is expected to be about 10% (Wright, 1998), which though worse than the 8% observed for 1996/97, is still very acceptable. Additionally, even with a reduced number of end-use studies, it is expected they will still yield valuable lessons to feedback into program design improvement.

Another metric which demonstrates the impact of the aggregation is that in 1998, sites comprising 33% of the savings predicted for all custom projects will be the subject of evaluation study[6]. This compares to 50% of the 1996/97 population savings, and 60% of the 1995 population savings.

Quality Control

Implementation of the quality control initiative and assessment of its impact are a longer term project. Since the QC recommendations were not adopted until the beginning of the 1997 program year, their impact will only begin to be seen in the evaluation studies taking place in 1998. Because of project lifecycles, some projects completed in 1997 may have begun in previous years and may not be affected by the QC guidelines. The simulation default study is just now being completed, so its impact may not be felt until projects completed in 1999.

Nevertheless, a comparison of evaluation costs to technical assistance costs and commissioning costs suggest that this effort is worth pursuing to a limited extent. Assuming that the relationship between savings and the utility's incentive is linear[7], the value of improving the realization rate can be analyzed. The overall realization rate for the 1996 program year was 95%. Improving this to 100% would gain the NEES Companies about $90,000 more in incentive, an increase of about 2% in the incentive earned for Commercial and Industrial programs.

The incremental cost of the QC improvements is likely to be less than the incentive bonus. The requirement to use the minimum requirements document will add no cost to project implementation, because the MRD was already being generated for each project. The requirement to improve technical documentation, including post-installation documentation, as well as simulation model documentation, will increase costs somewhat. Technical assistance costs in 1996 averaged about $5,000 per site, at an average cost of $70/hour. After a learning process is completed, documentation improvements may add an average of 5 hours to a project, or $350. In 1996, there were 200 technical assistance studies funded. On that basis, documentation improvement may cost about $70,000.

If the quality improvements lead to an increased reliability of savings estimates, it is possible that evaluation studies could be limited in some way, for example to recommend that some studies be performed with site inspection and engineering review only, and no metering. This would obviously reduce evaluation costs. Also, since evaluation costs average about $15 more per hour than technical assistance costs, the trade off of increased technical assistance costs for evaluation would also reduce overall costs.

The impact of mandatory commissioning on savings is complicated to assess. The primary reason for this policy was to focus limited commissioning resources on those projects where it would have the

[6]Studies at 55 sites in 5 end-uses are planned for 1998, compared to the 41 sites evaluated in 1997.

[7]This is a simplifying assumption. Some of the states have thresholds and other incentive features.

most value. In prior years commissioning would be provided upon request. The new policy dictates commissioning on larger projects, and on smaller projects only as feasible. The rebate threshold of $100,000 was chosen as a proxy for project complexity; there is some correlation between rebate amounts and savings as well. However, the new policy does not indicate an increase in commissioning funds: $195,000 was spent on commissioning in 1997 compared to $213,000 in 1996. In NEES' experience with evaluating the impact of commissioning on savings, commissioned projects had realization rates comparable to non-commissioned projects. These results indicate that commissioning may ensure that complex projects perform no differently than simpler projects. As with other quality initiatives, the mandatory commissioning initiative is seen as having positive customer benefits as well.

Conclusions

The results of the 1996 and 1997 evaluation studies suggest to NEES that it is on the right track in its efforts to reduce evaluation costs while maintaining the quality of evaluations. The use of guidelines for screening projects and identifying non-metering evaluation methodologies limited the use of intensive metering to 49% of the projects evaluated in 1996 and 1997. Costs per site decreased by 12% overall, and by 28% when compared to similar studies performed in 1995. The cost reduction did not adversely impact the evaluation results, as measured by realization rates, or their precision.

The NEES Companies continue to work to streamline its custom evaluation process, by aggregating end use groups to reduce the number of studies, and by introducing quality control measures. These continuing improvements are aimed at what remains the primary objective of this effort: to achieve cost reduction while maintaining the high quality of results.

Acknowledgments

The author acknowledges the contributions of Dottie Conant and David Jacobson at NEES for their constructive review of the annual evaluation plans as they have been developed, and to Roger Wright of RLW Analytics for his ongoing support in the development of the statistical basis for the study plans.

References

Newberger, J.J. 1996. "Streamlining Impact Evaluation for Customized DSM Projects," In *Proceedings of the 1996 ACEEE Summer Study on Energy Efficiency in Buildings*, 6:121-128. Washington, D.C.: American Council for an Energy-Efficient Economy.

Wright, R.L. 1997. "Analysis of 1996/97 Custom Evaluation Studies." Prepared for New England Electric System, July 21.

Wright, R.L. 1998. "Sample Design for the 1998 Custom Evaluation Studies." Prepared for New England Electric System, March 11.