Lessons Learned in Commercial and Industrial Non-Lighting Impact Evaluation

David I. Jacobson and Dorothy Conant, The NEES Companies

Impact evaluation of commercial and industrial non-lighting energy efficiency measures has proved to be much more complicated than evaluation of lighting measures. This paper, relating the experiences of a large New England utility in developing defensible impact evaluation results in a collaborative setting, will provide lessons and valuable insights for future impact evaluations. Topics to be discussed include:

- Use of evaluation techniques in program delivery;
- Generalizing end-use metering results;
- Applicability of billing analysis;
- Consistency of engineering estimates in customized site-specific evaluations;
- Measure retention versus savings persistence; and
- Importance of equipment loading measurements.

The paper will discuss each topic using specific real world examples where available.

INTRODUCTION

The New England Electric System (NEES) Companies have been promoting the installation of energy efficient equipment since 1987 through conservation programs offered by its retail affiliates in Massachusetts, Rhode Island and New Hampshire. To date, over 40,000 commercial and industrial customers have been served resulting in 250 MW and 900,000 annual MWh. The Company has performed comprehensive evaluations of all programs since 1990. These annual evaluations are required within six months of the year end. Overall, the Company has spent between $2 and $3 million per year on evaluation, an increasingly large portion for commercial and industrial non-lighting installations. This paper examines some of the general findings and lessons learned in the course of evaluating the more complex commercial and industrial non-lighting installations. The discussions are arranged by individual lessons and findings.

LESSON #1: EVALUATION TECHNIQUES CAN BE USED TO ENHANCE PROGRAMS DELIVERY

The expertise and equipment used for the evaluation of complex conservation measures should not be limited to documenting savings for regulators and program implementation staff. In addition to providing timely feedback on the performance of individual measures, the NEES Companies have used evaluation techniques in the day to day delivery of the programs. The three primary efforts in this area have been:

- Requiring installation of metering hardware at the time of participation for specific technologies;
- Providing customers, at no charge, the use of power metering equipment including basic data analysis services;
- Increased use of low cost current transducer and runtime loggers to pre-qualify measure savings;

End-Use Metering Requirements for Specific Technologies

Initial efforts to measure savings for one common measure, large thermal storage installations, were extremely expensive and time consuming. The first project involved measurement of the performance of a 2700 kW dynamic ice harvesting system on a 400,000 ft^2 commercial office building. Initial measurements illustrated numerous equipment performance problems as well as drastic equipment over sizing. The metering equipment was left in place for over two years and was used to help the building owner commission the system. The total cost for the metering and data analysis was approximately $90,000. This high cost for measurement of the performance of an individual installation was not acceptable. Given that the Company wanted to continue to promote this measure and better understand its performance, a lower cost method for monitoring these installations was sought.

Program planners and evaluators met with thermal storage equipment manufacturers to discuss equipment performance and how to provide easier and cheaper system performance measurements. Because of advances in microprocessor-based controls and data acquisition systems, the manufacturers were in a position to provide a complete set of diagnostic capabilities with each installation at no additional cost. The data points agreed upon included:

- Supply and return chilled water temperatures (°F);
- Ice storage outlet temperature (°F);
- Chilled water or ice brine flow rate (GPM);
- Ice storage system compressor power (kW); and
- Outdoor dry bulb temperature (°F).

The success of the effort has been slow but sure. To date, all manufacturers and installation contractors have met the basic requirements. Contractors installed the appropriate sensors and meters but not all provided the hardware or software for recording the data. Initially there were some difficulties in getting the different manufacturers to provide a means for recording the data in a usable format. Most of these issues have been resolved. For one equipment manufacturer, the utility agreed to pay a small fee for collecting and storing the data. This fee was required because of the proprietary nature of their data acquisition software, but was less than $200 per site per year. Company evaluation staff still spend time downloading and massaging data into a usable format and debugging various data acquisition systems, but the level of expenditures for evaluation of this technology has been greatly reduced. The total effort, including a consultant to help write the final report and make occasional site inspections, costs less than $1500 per system per year and only takes about one man-month per year of utility staff time for approximately 9 large thermal storage systems representing 3000 kW of peak demand savings.

Metering Equipment Loan Program

Another area in which we have tried to integrate evaluation techniques early in the delivery of programs has been through a metering equipment loan program. This program, officially begun in 1993, provides power monitoring equipment and basic data recording and analysis services to customers considering the installation of larger conservation measures. To date the program has served approximately 65 customers per year at an average cost of $900 per rental. The program has been successful in helping Company representatives demonstrate savings potential to skeptical customers as well as providing better documented estimates of savings for program tracking systems. In many cases, the equipment is simply used to establish a baseline of pre-retrofit consumption.

As an example of the program's success, a large paper mill rented a power monitor twice at a total cost of $1150. The data were used to quantify the savings for five separate customized industrial process conservation measures. The information was instrumental in convincing the customer to proceed with the installation of the measures which resulted in 1800 MWh of annual savings and rebates of $450,000. In addition, the metered data were used in developing savings estimates for regulatory reporting. The cost of the metering represented only 0.1 percent of the total measure costs.

Use of Low Cost Runtime Metering Equipment

The development of lower cost runtime metering equipment with user friendly data analysis software has provided additional opportunities for field implementation staff to use evaluation equipment in the delivery of conservation programs. With downloadable current and lighting loggers now approaching $150, their widespread use in qualifying runtime sensitive installations should increase. The NEES Companies have encouraged the use of loggers to demonstrate savings for measures such as air conditioning system retrofits and lighting controls. Loggers originally purchased for program evaluation studies have been made available to field staff for use at customer sites.

A recent example of the use of such equipment was for an energy management system at a large aerospace manufacturing facility. The facility manager claimed that the equipment to be controlled was left on during all unoccupied hours. A quick study using runtime loggers indicated that the equipment was already being controlled to some extent thereby reducing the savings potential for the installation from 2700 annual MWh to 1100 annual MWh. In this case both the customer and the utility got more accurate estimates of measure savings and bill reductions.

LESSON #2: GENERALIZING END-USE METERING RESULTS REQUIRES CAREFUL PREPARATION

Metered measurement of absolute savings at an individual or small number of sites is often of little value in determining program level impacts. Results are only valuable if they are easily generalizable to other sites where detailed metering has not been performed. For many measures this task is made difficult by engineering estimates that differ greatly from site to site, even within a similar technology grouping. Individual engineers often use different algorithms or simulation models to calculate savings. In order to develop a general realization rate from metered data, the methodologies used to produce engineering estimates should be consistent. This issue must be kept in mind when planning detailed end-use metering studies.

Often there is intense pressure to measure savings for new measures in a very short period of time. To meet these pressures, detailed measurements at one ''pilot'' site are often initiated with minimal planning as to how the results will be used. The studies usually involve longer term (>2 weeks) measurement of numerous variables such as temperature, flow rates, power, part loading etc. These studies can cost between $20,000 and $100,000. At the conclusion of the measurement period it is often discovered that what was thought to be a ''typical'' application of a specific technology turned out to be not so typical and not representative of the current application of that technology. Adding insult to injury, it is sometimes determined that some of the key findings regarding the causes of poor performance of the measure could have been determined from lower cost spot measurements. Newberger discusses how expensive metering can be avoided with careful planning and using evaluation strategies that can determine key savings variables through less expensive spot measurements (Newberger, 1996).

It is important to think ahead about how the results of a metering study will be extrapolated to multiple year's participants. Our experience has led us to develop some rough guidelines for conducting successful metering studies. The primary requirements include:

- Large enough sample to be representative of the variety of applications of the measure or technology in the population;
- The measures or group of measures to be metered should produce savings from similar process changes described by similar physical laws or equations, i.e. volume of air flow is a function of the cube of fan speed, etc.;
- Development of an analysis plan or engineering estimation methodology for the measure, i.e. a regression analysis of specific variables thought to impact savings or a simple model of savings, such as multiplying the change in demand and hours of use to determine kWh savings from a motor replacement; and
- Measurement and use of model variables which are easily collected from participants during the application process so results can be extrapolated to sites not metered.

A complete understanding of how the data will be analyzed and extrapolated using readily collected information from the entire population will help determine what variables should be metered in the first place.

One example of a successful end-use metering study at the NEES Companies is the measurement of savings resulting from the installation of variable speed drives on plastic injection molding machines. This was not a simple measure to analyze but there were similarities in the different installations and the basic physical mechanism for producing savings, in this case the reduction of pump speed during idle time. This measure has been extremely popular and has been installed on over 10,000 HP of injection molding machines through 1995. In 1993, the Company initiated a short-term metering study to estimate energy and demand savings (Englander et al, 1994). A model for estimating energy savings was developed and tested using the data from the study. The study included measurements for 12 machines at 7 sites. Because the high speed machinery goes through a very repetitive process some measurement periods were as short as one hour.

The final model, listed below, yielded percent savings as a function of the square root of percent idle time, motor HP, rated clamping force, and the ratio of clamping force to motor size:

$$S_f = -0.536 + 0.010H - 0.00114T + 0.0415(T/H) + 0.847I^{0.5}, \; R^2 = 0.95$$

where

S_f = fractional reduction in power
H = motor size, HP
T = rated clamping force, tons
I = fractional idle time

This percentage can then be used to estimate demand saving(S) using either pre-retrofit average power P_c:

$$S = S_f P_c$$

or, if post-retrofit power P_v data is available:

$$S = P_v \{S_f / (1 - S_f)\}$$

Pre-retrofit power was estimated as:

$$P_c = -0.567 + 0.080T - 34.1I + 79.9(H/T),\ R^2 = 0.91$$

The annual energy savings is then obtained by multiplying demand savings(S) by annual operating hours.

Because of the high R^2, the Company has relied on this one well planned study to calculate savings for new installations for several years. The key input variables, which are readily available for each installation, have been added to the application for rebate and are used to calculate savings for each installation. The regression equation has been built into the program tracking system.

LESSON #3: BILLING ANALYSIS HAS BECOME LESS IMPORTANT AS PROGRAMS MATURE

Billing analysis has been a valuable analysis technique for a number of specific technologies and conservation programs. The Company has had considerable success in using econometric billing analysis to evaluate small and large commercial lighting installations and an appliance removal program. For retrofit single and multifamily residential space heat retrofits, the Company has used weather normalized billing analysis.

As programs have evolved, the percent of savings that billing analysis can be used to evaluate has declined. In 1990, the share of Massachusetts Electric Company conservation program savings evaluated using billing analysis represented 76 percent of the annual MWh savings. In 1995, that figure fell to 35 percent. Since billing analysis can be considerably less expensive than metering and usually results in less intrusive customer contact and disruption, the Company has made considerable effort to increase its use. Our collaborative partners and their technical consultants who oversee our evaluation efforts have prodded us to consider billing analysis for more end-uses and programs. A brief examination of why this has not been possible may help guide future evaluation decisions.

One of the key requirements for using standard billing analysis techniques are distinct pre and post installation periods where conservation program related building changes can be isolated. This usually translates into a narrow, well defined installation time period when customers install a discreet set of measures all at once. A two to three year window of time is preferred—one year of pre-installation data, one year of post-installation data and a several month time window in between.

The NEES Companies have found that many larger commercial and industrial customers who install large numbers of measures in one or multiple end-use categories are not good candidates for standard billing analysis because they tend to spread the installations over a period of several years. Many customers are reluctant to make multiple changes at their facility simultaneously because of budget limitations or, more often, their lack of trust in the new energy efficient technology. They often like to experiment with a small portion of their facility, or portions of an industrial process, before committing to changes over their entire facility or processes. These customers who continuously install measures do not have clean pre and post installation billing periods. Thus, customers who are the most active program participants, the customers for whom you would most like to measure savings, may require more sophisticated modeling.

To date all of the commercial and industrial program billing analysis used to measure savings by the NEES Companies has been limited to lighting. The primary reason for this has been the limited number of non-lighting installations that met the time window requirements discussed above and had uniform methodologies for estimating savings. Having a common methodology or algorithm for estimating savings for the more complex non-lighting measures is a requirement if results are to be extrapolated to other participants not in the analysis. The sample sizes of measures with similar savings estimation methodologies are often too small to provide results with any statistical validity. This non-uniformity in engineering estimates often limits the applicability of billing analysis.

The primary limitation to the present and future use of billing analysis has been the change in emphasis of utility conservation toward more market driven programs. Market transformation programs are based on the assumption that customers were planning on making some major equipment change out before they decided to participate in the program, whether it be for a new facility, a major renovation, or the planned replacement of a specific piece of equipment. Because savings result from the installation of equipment that is more efficient than what customers originally planned to install, but did not, an appropriate pre-installation billing record does not exist. Without an actual pre-installation case to compare to, billing analysis for market driven programs has been limited for the most part to calibration of engineering simulation model post installation consumption predictions.

In summary, for billing analysis to be useful in helping assess technology and performance for commercial and industrial non-lighting applications we believe the analysis must:

- Consist of a sufficiently large sample of customers with the specific measure in question installed during the pre/post period under study (>30 sites);

- Have sufficient pre and post installation billing data;
- Have savings estimates calculated using a similar methodology; and
- Include only retrofit applications, where a true pre-retrofit case exists.

LESSON #4: CUSTOMIZED SITE-SPECIFIC EVALUATION TECHNIQUES HAVE PROBLEMS TOO

Given the limitations of the most common evaluation techniques discussed above, evaluators must look beyond standard billing analysis and metering techniques for many of the more complex commercial and industrial non-lighting measures. The NEES Companies have developed a number of hybrid techniques specific to certain technologies, end-uses and program delivery mechanisms. Most involve on-site assessments of a sample of facilities. Results from the sample are extrapolated to the sites not visited. The primary data are collected from spot measurements of key input assumptions, short term metering, and a review of simulation inputs. Key lessons learned in developing these techniques include:

- Limiting conflicts between different engineering estimation methods;
- Building on existing savings estimation methodology; and
- Importance of consistency of estimates when extrapolating results.

When trying to use hybrid evaluation techniques, especially those that do not involve billing analysis or long term end-use or whole building metering, there is a tendency to hire another ''independent'' engineer to review the original engineer's estimates. This second engineer may use a completely different simulation tool or algorithm to calculate savings. Unless this second engineer uses real after-the-fact measure performance—such as equipment usage and maintenance logs, and spot measurement of key performance parameters,—this second analysis may just be viewed as a second opinion and not actual measurement of savings. Impact evaluation should not be a contest between professional engineers or about whose simulation algorithm is more accurate. Instead of trying to replicate the original savings from scratch, the engineer hired to do the after-the-fact evaluation should, whereever possible, use a similar calculation methodology, enhanced only by actual in-service performance data. The second engineer should build upon the original engineer's work, not try to prove him wrong. In actuality, there will be instances where gross errors are made in the original estimates that do require a complete new analysis. There will be many instances where the original savings algorithm is too simple, incomplete or simply not appropriate. In these cases the evaluation should provide an ''independent'' assessment of savings and not be constrained to just add new data to a bad model. Fortunately, our experience is these cases are the exception not the rule.

As mentioned above, in order to extrapolate sample results to a population, the engineering methods used to estimate savings for the sample need to be consistent with those used for the rest of the population. This is particularly important when feeding evaluation results back into program tracking estimates. A specific example illustrating potential problems involved a refrigeration measure where savings were based on a vendor's computer simulation. The first year's evaluation of that measure showed that the vendor had overestimated savings during certain portions of the refrigeration cycle. Evaluation staff met with the vendor to correct the assumptions in the vendor's model. During the next year of program implementation, the vendor's model for calculating savings for the measure was updated. The measures installed during the next two years included some using the older model and some using the updated model. Because of this combination, the results from the previous year's study could not be extrapolated directly to the newer installations. It is important to understand how the estimates for the current population were developed before applying results from a study of previous year's participants.

LESSON #5: LONG TERM PERSISTENCE OF SAVINGS MAY ALREADY BE ACCOUNTED FOR IN OTHER EVALUATION STUDIES

The NEES Companies have conducted a number of longer term persistence studies of both lighting and non-lighting measures. Results of these studies have provided some valuable lessons which should be considered when planning these types of studies. The primary lessons relate to definitions of persistence and distinguishing the results of short term evaluation from longer term assessments of persistence.

Persistence can be defined at two distinct levels. The first involves measure retention; is the measure still in place a predetermined number of years after installation. Measure retention can easily be determined from short on-site surveys or telephone interviews. The second more challenging level involves determining actual savings retention. This level requires the evaluator not only to assess if the measure is still in place, but if it still operating as originally intended and producing the predicted savings. This second level is much more difficult to assess. For simple measures like commercial lighting fixture replacement, measure retention is usually a good proxy of savings retention—provided the pre installation fixture wattage is deemed reliable and the hours of use are determined to be relatively stable. For more complex non-lighting measures, measure retention does not always represent savings persistence. A study designed to measure savings retention for these complex measures must mimic the methodologies used during the original assessment of savings. This can be complex and costly and has rarely been carried out. Often some of the assumptions that went into the original calculation of savings cannot be evaluated. Clear study goals must be established and the limitations of the methodology understood.

In cases where there is an attempt to calculate savings retention, care must be taken to distinguish longer term saving degradation from short-term measure performance. For one measure, liquid refrigerant pumps, the NEES Companies did both a persistence study of one to two year old installations and an evaluation of the savings from the current year's installations. These refrigerant pumps are used to boost the pressure of refrigerant leaving the condenser in order to allow the systems to operate at lower system head pressures. Both studies used original estimates of savings from the vendors as baselines.

Results of each study showed considerably less savings than originally predicted. The kWh realization rate for the current installations was 41% while the realization rate of the older sites was 59%. Assuming the persistence realization rate captured longer term savings increases or degradation, and the realization rate from the current sites covered short term performance issues, one would be tempted to use the product of the results of the two studies 0.59 and 0.41 to discount the savings from sites not studied. Upon careful examination of the details of the two studies, one would conclude that this would not be the appropriate use of the two studies. Both studies uncovered many of the same factors for poor measure performance which were for the most part unrelated to the age of the installation. In most cases the factors for poor performance had to do with baseline conditions of the pre-retrofit refrigeration system, similar errors in the original calculation methodology and overestimates of annual operating hours. The lesson here is that care must be taken not to double count the effect of savings degradation (or enhancement) determined from evaluation of current years sites with those effects determined from longer term persistence studies. A savings baseline for sites studied for longer term persistence would help isolate savings longer term measure performance issues.

LESSON #6: FAILURE TO ACCURATELY ACCOUNT FOR PART LOADING CAN LEAD TO OVERESTIMATION OF SAVINGS

Estimates of savings for many conservation measures are highly dependent on equipment sizing and loading factors. Estimates of savings for HVAC and drivepower applications have been particularly vulnerable to lower than expected estimates of equipment loading. Examples of evaluations where this has been a factor at NEES include drivepower and thermal storage.

In two different studies of motor loading, results indicated the average loading factor was actually around 60 percent instead of the 75 percent originally assumed in program planning estimates (Savage Engineering 1993, 1994). In addition to those specific studies, evaluation of various VSD installations indicated that some drives were set to constant settings of 50 percent of full speed.

In the course of measuring cooling load for thermal storage sites, it was determined that peak cooling loads were on average overestimated by over 80 percent (Norford, 1995). Over sizing of cooling equipment is common practice in the HVAC industry and utility efficiency programs have had a difficult time changing those design practices. Realistic estimates of over sizing need to be factored into original estimates of savings.

Besides the implications for accurate assessment of post-installation program savings, equipment sizing also has implications for program design. Instead of assuming the existing equipment is properly sized, engineers should work harder to downsize the recommended retrofit or replacement equipment to take advantage of higher loading factors. Why install a 100 HP motor with a VSD set at a constant 50 percent speed when a much smaller motor will get the same job done.

CONCLUSIONS

Evaluation of more complex commercial and industrial non-lighting measures is considerably more challenging than evaluation of commercial and industrial lighting retrofits. Because of the change in emphasis toward market driven programs and a more diverse measure mix, traditional billing analysis cannot be relied upon as much as in the past. End-use metering, thought to be the most likely alternative, is very expensive and often produces results which cannot be applied to the population. Consequently, hybrid techniques which involve on-site visits and short term measurements are being used more and more. Regardless of the specific techniques used, evaluators must take into account the consistency of the original engineering estimates of savings across the population, equipment loading and the proper application of persistence studies.

REFERENCES

Englander, S.E. and Remley, C.H., ''Measured Performance of Variable Speed Drives on Injection Molding Machines,'' ACEEE 1994 Summer Study on Energy Efficiency in Buildings, August 1994.

Newberger, J.J., ''Streamlining Impact Evaluation for Customized DSM Projects,'' ACEEE 1996 Summer Study on Energy Efficiency in Buildings, August 1996.

Norford, 1995, L.K., ''Evaluation of Avoided Electricity Demand in 1994 at Nine Thermal Storage Sites,'' Appendix IV-4, 1994 DSM Performance Measurement Report, Massachusetts Electric Company, July 1995.

Savage Engineering, ''Motor Performance Study: Site Surveys and Measurements of Loading and Operating Hours for 1993 Impact Evaluation,'' Appendix IV-5, 1993 DSM Performance Measurement Report, Massachusetts Electric Company, June 1994.

Savage Engineering, ''Motor Performance Study,'' Appendix G, 1992 DSM Performance Measurement Report, Massachusetts Electric Company, June 1993.