ELECTRIC POWER RESEARCH INSTITUTE

Guidebook for Energy Efficiency Program Evaluation, Measurement and Verification

1016083

Guidebook for Energy Efficiency Program Evaluation, Measurement and Verification

1016083

Technical Update, December 2008

EPRI Project Manager

O. Siddiqui

ELECTRIC POWER RESEARCH INSTITUTE

3420 Hillview Avenue, Palo Alto, California 94304-1338 ▪ PO Box 10412, Palo Alto, California 94303-0813 ▪ USA
800.313.3774 ▪ 650.855.2121 ▪ askepri@epri.com ▪ www.epri.com

ORGANIZATION THAT PREPARED THIS DOCUMENT

The Cadmus Group, Inc.

This is an EPRI Technical Update report. A Technical Update report is intended as an informal report of continuing research, a meeting, or a topical study. It is not a final EPRI technical report.

NOTE

For further information about EPRI, call the EPRI Customer Assistance Center at 800.313.3774 or e-mail askepri@epri.com.

CITATIONS

This document was prepared by

The Cadmus Group, Inc.
Energy Services Division
720 SW Washington St., Suite 400
Portland, OR 97205

Principal Investigators
M. S. Khawaja
C. Mulholland
J. Thayer
K. Smith

This document describes research sponsored by the Electric Power Research Institute (EPRI).

This publication is a corporate document that should be cited in the literature in the following manner:

Guidebook for Energy Efficiency Program Evaluation Measurement & Verification. EPRI, Palo Alto, CA; 2008, 1016083.

PRODUCT DESCRIPTION

Energy efficiency program evaluation is increasingly important as utilities implement programs to meet regulatory requirements, such as energy efficiency portfolio standards. While utilities need internal staff to oversee evaluation activities, most evaluations are actually conducted by outside consultants. Thus, utility staff require a sufficient understanding of the evaluation process to plan program evaluation activities as well as to manage internal stakeholders and evaluation contractors. This guide is intended to help prepare utility staff to accomplish these tasks.

Results and Findings

Evaluation is best conducted through a theory-based approach that articulates program goals as well as activities it will generate to accomplish these goals; describes causal relationships between program activities and intended results; and explains why program activities are designed to lead to intended outcomes. Programs are best designed with their "evaluability" in mind (the practicality of collecting data necessary to measure key metrics of performance). In this regard, designing evaluation explicitly into programs prior to their launch helps ensure the eventual availability of verifiable and accessible data about program performance.

To rigorously calculate savings attributable to an energy efficiency program, key external factors beyond the control of the program or consumer need to be accounted for, not just the simple difference in measured energy consumption pre- and post-program. Such factors can include corrections for variances in weather conditions, occupancy levels and hours, and production levels in industrial facilities. These adjustments are made to normalize pre- and post-program energy use to the same set of conditions such that program impacts are neither under- or overstated.

An effective evaluation technique is a quasi-experimental design that measures program impacts through observed differences in electricity consumption or other metrics between treatment and control groups pre- and post-program. A properly constructed quasi-experimental design also is the best approach for estimating net savings, which accounts for the phenomena of free riders, spillover, and the rebound effect.

Challenges and Objectives

EPRI commissioned this guide to provide utility staff with a clear and easily accessible explanation of the principles inherent in energy efficiency program evaluation, as well as the information to develop an evaluation plan and oversee the program evaluation process.

Applications, Values, and Use

The guide is targeted at middle-level utility management who are responsible for designing and implementing energy efficiency program evaluations. Utility staff who use this guide will be able to better develop evaluation plans and parameters and to provide more effective oversight of all evaluation activities. Better management results in a better evaluation and, thus, more reliable energy savings estimates that stand the scrutiny of regulatory oversight.

EPRI Perspective

There are many guidelines, manuals, and other documents available today that focus on energy efficiency program evaluations. Among these are the National Action Plan for Energy

Efficiency's *Model Energy Efficiency Program Impact Evaluation Guide* and the American Public Power Association's *Market Research and Evaluation for Energy Efficiency Professionals.* What distinguishes this guide from these and others is its focus specifically toward large utilities with slightly more complex evaluation needs. The guide is written for utility staff generally familiar with energy efficiency but new to the discipline of program evaluation and is intended to help make readers conversant with evaluation concepts, nomenclature, and practices.

Approach
This guide was developed by seasoned practitioners with decades of experience in conducting energy efficiency program evaluation activities for utilities. The project featured a detailed review of the most current literature on energy efficiency program evaluation. This review was augmented by interviews with a number of evaluation experts from across the country.

Keywords
Evaluation
Impact evaluation
Market assessment
Process evaluations
Baseline data
Cost-benefit test
Free rider
Spillover
Measurement and verification
Evaluability
Rebound
Persistence
Grid-tied appliances
AMI
Co-benefits
Net-to-gross
NTG
Market evaluation

CONTENTS

1
INTRODUCTION

About the Guide

Under its 2007-2008 Energy Efficiency Initiative, EPRI published a guidebook[1] that provided an overview of the energy efficiency planning process, including: defining strategic objectives, establishing baseline forecasts, estimating the potential for energy and demand savings, designing and implementing programs, and conducting evaluation, measurement, and verification (abbreviated as EM&V, or referred to simply as “evaluation”). That guidebook drew upon fundamental principles and practices of demand-side management, as documented in seminal EPRI studies of the 1980s and 1990s and honed over decades of utility experience.

This document is an extension of that preceding guidebook focused on the component of energy efficiency deemed by EPRI members as most warranting greater attention – *evaluation*.

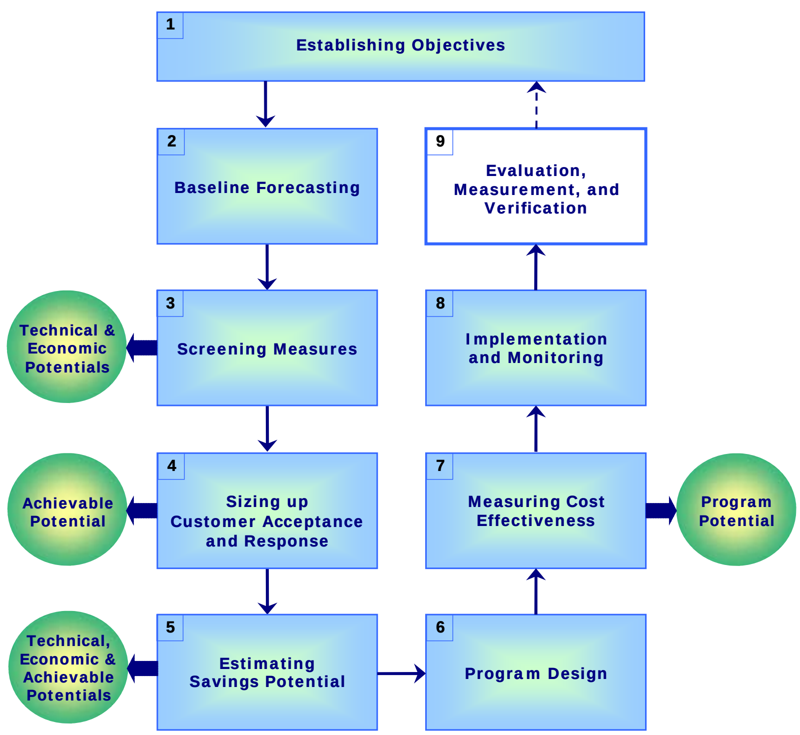

Figure 1-1
Evaluation, Measurement & Verification in the Context of Energy Efficiency Planning Process

---

[1] *Energy Efficiency Planning Guidebook: Energy Efficiency Initiative*. EPRI, Palo Alto, CA: 2008. 1016273.

This guide provides the following elements required to successfully manage energy efficiency program evaluations[2]:

- Principles of energy efficiency program evaluation
- Discussion of various types of evaluation
- Explanations of critical dynamics affecting the interaction of program evaluation elements
- Review of best practices employed to support a variety of energy efficiency programs
- Glossary of commonly used terms

It provides general guidance for conducting various aspects of energy efficiency program evaluation, with an emphasis on how to design and implement impact evaluations to estimate energy and demand savings.

This guide primarily addresses evaluations of resource acquisition programs, and also reviews how evaluation processes, methodologies and best practices can be applied to process evaluations, market transformation programs and market evaluations. Because many utilities operate a portfolio of diverse measures to encourage efficient energy use across all segments of their marketplace, the guide discusses evaluation activities appropriate for all of them.

Why this Guide?

This guide is designed to serve as a resource to help utility staff understand the most effective means of designing, implementing and managing evaluations of energy efficiency programs or full portfolios. It provides a framework that utilities and organizations can use to help design programs and evaluations to reflect their specific situations.

Many utilities today face mandates to conduct efficiency programs and, more importantly, to demonstrate measurable results that pass the scrutiny of external review. Some states reward those utilities with successful programs and penalize those whose programs have not met their projected energy and demand savings goals. Whether driven by incentives or penalties from outside entities such as regulators or state legislatures, or by capacity constraints, utilities are under pressure to show that their programs mitigate demand growth and yield energy savings. Moreover, most utilities are challenged to meet ever increasing energy savings targets with fewer staff, many of whom are not experienced in energy efficiency program evaluation. This guide has been developed to address the evaluation needs of utilities facing these circumstances.

What this Guide is Not

This guide is **not** a step-by-step approach to evaluating energy efficiency programs or portfolios. Such an approach would be prescriptive, and would not allow utilities the flexibility they need to

[2] EPRI has also produced the following companion document, focusing on evaluation of demand response programs: *Demand Response Evaluation, Measurement, and Verification: A Synthesis of Evolving Protocols and Practices*. EPRI, Palo Alto, CA: 2008, 1018194.

design appropriate evaluations for their individual programs and portfolios. There is no single solution that fits all markets, nor will the ever-changing nature of our markets support an unchanging or static solution. This guide presents the current array of “best practices” available to address a great variety of differing programs. The reader is encouraged to use this guide as a framework to assemble the optimal combination of program components, practices and metrics that reflect the needs of their unique energy marketplace.

Using this Guide

The guide is written particularly to benefit utility professionals whose responsibilities include managing and evaluating energy efficiency programs. It contributes to establishing a common understanding of the challenges and opportunities in working with energy efficiency programs and hopefully serves to align expectations and goals among all of those who work with energy efficiency programs within a utility.

Each chapter addresses a discrete component of the overall evaluation planning and implementation process, from setting evaluation goals, and designing an evaluation through data collection and analysis and producing a final report.

Chapter 3 discusses what data resources are needed to establish a solid foundation of performance metrics upon which the planned evaluation can be based. It also discusses the timing of evaluation planning, and the emerging trend of linking evaluation planning with energy efficiency program and portfolio design, so that the requirements for effective “evaluability” are embedded in the program. Chapter 4 considers the basics of an impact evaluation. The following chapter explores Net Savings and how to calculate the net-to-gross conversion. Chapter 6 considers the issue of non-energy benefits, and chapter 7 examines the way technology is changing the way we measure and meter programs. Finally chapters 8 and 9 cover cost effectiveness and process evaluation.

Throughout, this guide focuses on the overall management of these evaluation processes, as well as the effective communication of results. Throughout all chapters, the guide seeks to define a standard evaluation planning and implementation process. It covers several standard approaches to calculating savings, and provides useful advice for many important evaluation issues. As mentioned elsewhere, it also provides useful definitions to common terminology, and provides links to other useful resources. Utility program examples are interspersed throughout the guide to illustrate the overall evaluation process.

2

WHY PERFORM EVALUATION?

Why Do Utilities Offer Energy Efficiency Programs?

Let's begin with why utilities engage in energy efficiency promotion as a part of their business model. Why would any business engage in activities that effectively induce their customers to reduce purchases of their product? There are four central reasons for utilities to adopt energy efficiency as an integral component of their business model:

- Consumer pressures, public opinion pressing for energy efficiency
- Regulatory pressures forcing utilities to reduce production of electricity from non-renewable (i.e. fossil-based) resources
- National energy management policies are imminent and likely to increase costs of carbon-dependent energy
- Energy efficiency can have significant impacts on reducing the need for capacity expansions as well transmission and distribution upgrades. This deferment of capital expenditures can be very beneficial to utilities. Freeing up capacity allows utilities to encourage targeted economic development, and to concentrate on supporting a synergistic mix of residential, commercial and industrial energy usage that produces the highest benefits on an environmental, economic and societal basis.

Energy Efficiency is More Cost-effective than Supply-Side Options

Energy efficiency generally represents a utility's most cost-effective means to address growing demand for electricity demand while reducing reliance on carbon-emitting energy sources. Several recent studies, for example, have shown that:

- Energy efficiency is 67 percent cheaper than new supplies of energy.[3]
- The cost to run energy efficiency programs is, on average, 2¢ to 3¢ per avoided kWh, versus five times that amount for electricity generated from new nuclear, natural gas, and coal-fired power plants.[4]

Ratepayer-funded energy efficiency programs require a greater degree of accountability to substantiate the prudence of their investment than does the construction of power plants. Energy

---

[3] *www.neep.org/files/NEEP_Achievable_Potential_Presentation_UPDATED.ppt*

[4] "Why We Never Need to Build Another Polluting Power Plant". Joseph Romm, 7/28/08, www.salon.com

efficiency programs are typically funded by mandatory state ratepayer contributions that are strategically dedicated to underwrite programs to transform energy markets.

As is the case for any major investment, utilities must substantiate to regulatory bodies that the benefits that accrue from investment in energy efficiency programs outweigh their costs. Regulatory oversight emphasizes several key factors for energy efficiency programs:

- Process transparency
- 3rd party certification of program results
- Clear objectives, consistent with public goals, that can be measured
- Established criteria for success and failure
- Adoption of so-called industry "best practices"
- Uniform procedures to review program performance

The bottom line is that the use of ratepayer funds, along with strict regulatory oversight, means that programs must be evaluated on a periodic basis to assess whether, and to what extent, they are meeting the objectives of the utility and its regulators.

Evaluation Determines the Effectiveness of Energy Efficiency Programs

Evaluation validates or disproves energy efficiency program effectiveness, and provides the evidence needed to allow utilities to recover program costs and foregone revenues. Program evaluations serve as periodic assessments that determine what progress efficiency programs have made toward their stated goals. Evaluations also can be used to determine and/or adjust goals to conform to revised performance estimates, or to evaluate how effectively a program is being implemented. Finally, evaluations underscore the value of promoting greater energy efficiency as the lowest cost approach to reducing energy use without a commensurate decline in energy dependent services or standard of living.

A good evaluation will inform all key stakeholders – including utility program staff, external implementers, utility management, and regulators about the effectiveness of the deployed programs. It also can inform the public and direct customers about the utility's efforts to increase energy efficiency.

In addition to providing feedback about a program's performance, evaluation provides valuable information about what is or is not working well within a specific program or portfolio. That information can then be used to revise existing programs or redesign future programs so that they will be more effective at reaching the targeted audiences and producing the desired energy savings. The types of improvements to which evaluation can contribute include:

- Adapting programs to changing market conditions
- Resetting regulatory performance metrics
- Implementing "in-stream" improvements in program delivery strategies
- Improving program design and administrative procedures
- Adjusting incentives levels

- Adjusting service packages used to stimulate desired activity
- Discovering program overlap or areas not currently addressed
- Discovering and exploiting inter-program synergies, and the design of program portfolios that achieve broader efficiencies

In summary, evaluations provide the following benefits:

1. Quantify the effects of a program to determine its impact, and consequently whether it has met its intended goals
2. Help utility staff understand whether the logic model behind the program design functioned as expected, whether program implementation was effective, and whether it could be improved to transform the market more effectively
3. Determine whether programs are meeting their regulatory requirements

Perhaps Professor John Kenneth Galbraith stated it best: “Things that are measured tend to improve.”

Evaluation as Part of Program Planning and Implementation

Evaluation provides invaluable feedback on program performance. Traditionally planning for evaluation began after programs had been implemented, and evaluation had no interaction with program planning or design prior to initial program implementation. Today, that sequence is shifting. The sections below discuss both the old and new evaluation paradigms.

Old Paradigm

In the past, it was common practice to launch energy efficiency programs without attempting to gather any input from the prospective evaluator. Indeed, the evaluator wasn’t chosen until well after the program was up and running, as shown in Figure 2-1. In these cases the evaluator was brought in at the end of the project life cycle, in much the same way as auditors are deployed at the culmination of a project or fiscal cycle. While the separation of implementers and auditors makes good fiduciary sense in the realm of financial accounting, it creates challenges and barriers to effective energy efficiency program evaluation.

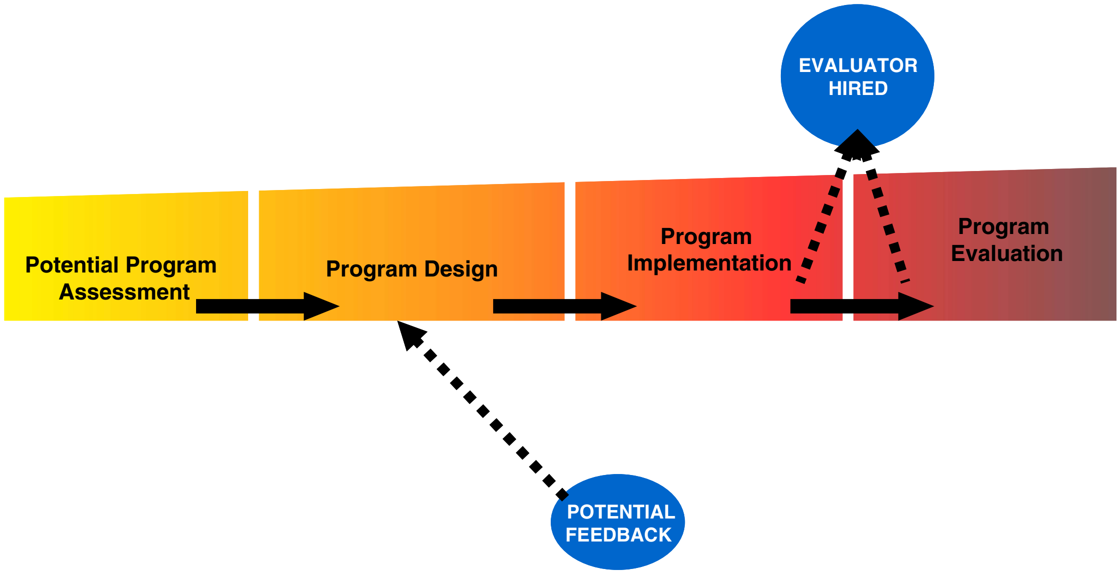

Figure 2-1
Project Life Cycle – Old Paradigm

The result of this disjointed process was that the evaluator often had to conduct evaluations with incomplete, mismatched or inaccurate data. Even when full data were available, the timing of the evaluators' findings came at such a late stage in the program cycle that there was little opportunity to improve program performance.

Over time, it became clear that early coordination between implementer and evaluator could provide timely feedback that could be immediately integrated into program implementation, and would produce more robust evaluation findings on the basis of the improved data provided by the implementers.

There may be some distinctions between the implementer's goals and those of the evaluator. The former is intent upon maximizing the energy savings each program delivers, while the latter is concerned about the veracity and defensibility of the methodologies and estimates they produce. The implementer may err on the side of higher savings numbers, but both the utility and evaluator are ultimately concerned with savings that can be verified.

New Paradigm

The potential for program evaluations to provide early and dynamic support for ongoing program refinement argues strongly for the evaluator to become involved near program inception, rather than after the program has been implemented. As shown in Figure 2-2, the paradigm that has emerged over the last few years places the evaluation, measurement and verification functions earlier in the process. It recognizes that evaluation produces valuable insights that can positively affect energy efficiency program outcomes over the program's life cycle. Evaluation can provide course correction advice in a manner that actually allows for course correction.

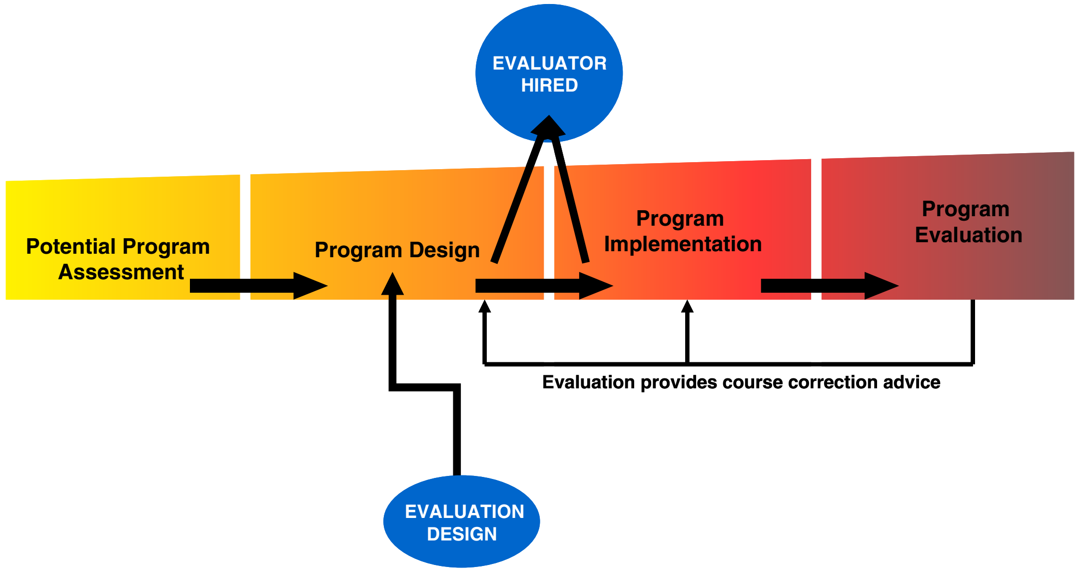

Figure 2-2
Life-Cycle Process – New Paradigm

Types of Evaluation Studies

There are three main types of evaluation studies: impact evaluation, process evaluation, and market assessment. Each measures a different component of program performance.

Impact Evaluation

Impact evaluation measures a program's quantitative results (e.g., the amount of energy saved). Impact evaluation can also assesses a program's indirect and derivative effects, including avoided greenhouse gas (GHG) emissions, improved health benefits, increased job creation, better energy security, more efficient transmission/distribution benefits, and water savings. Finally, impact evaluations determine overall program and portfolio cost-effectiveness.

Process Evaluation

Process evaluation examines the effectiveness of program delivery and implementation (e.g. what worked and what did not work throughout the overall program process). Process evaluations try to describe the program from the perspectives of all those involved, ranging from utility staff to implementers, to trade allies, and to customers. Process evaluations assess program designs to discover inefficiencies, bottlenecks, and misalignments as well as seeking out missed opportunities within the scope of the program.

Market Effects Evaluation

Market effects studies are designed to expose the underlying transformations attributable to programs that are taking place in the various markets within a utility's service territory. For example, a market effects study would capture the extent to which energy-efficient appliances have become more available within a utility's service territory since a program was offered.

Measuring What Doesn't Exist

Evaluating energy efficiency programs poses a fundamental challenge: the goal is to measure the absence of something, or the relative shape of a non-existent asset.

Energy or demand savings cannot be directly measured since "savings" denotes the absence of energy or demand. Therefore the measurement must be accomplished by comparing the "before" and "after" conditions to extrapolate the effective change that has occurred. Until recently it was common to try to adjust energy efficiency savings data to net out all other factors that might influence consumer behavior other than the targeted energy efficiency program. Because non-utility programs such as ENERGY STAR® are so prevalent in national markets, it has become nearly impossible to net out their influence, so the focus today is on understanding customer behavior before and after a program has been implemented.

The following chapter focuses on the steps to be taken as you prepare for an evaluation.

3
ORGANIZING FOR AN EVALUATION

This chapter is intended to provide utilities with guidelines on how to prepare for evaluating energy efficiency programs. It is particularly germane to utilities that are relatively new to implementing and evaluating energy efficiency programs, or are revisiting the discipline after a period of hiatus.

Objectives of the Evaluation

An evaluation's first and most important objective is to provide guidance for program implementation. Evaluations need to provide required information in real time, or as close as possible, so course corrections, when needed, can be made in a timely fashion.

To develop evaluations that will achieve this goal, various stakeholders' needs should be compiled and addressed. Program implementers may have different needs than program managers or regulators. A potential list of applicable program impact objectives includes:

1. Estimate gross and net kWh savings
2. Estimate gross and net kW savings
3. Determine impacts on system peak, summer peak, or winter peak
4. Calculate cost-effectiveness
5. Assess customer satisfaction
6. Estimate greenhouse gas emission impacts

Program developers need to decide which of these objectives are most important and to which programs or portfolio of programs they apply.

Understanding exactly how programs achieve their results allows a utility to adjust its program management to increase "traction" and deliver greater results. Measurement and verification will also reveal changes in areas only tangentially related to energy efficiency, allowing for positive operational improvements across a wide swath of utility activities.

By linking results to program implementation (and to subsequent fine-tuning, based on evaluation results), utilities can reduce their capital costs as the perception of uncertainty and risk diminishes.

By providing clarity around the costs and benefits of investments in energy efficiency, the utility also helps to educate regulatory stakeholders about the nature and function of effective energy-efficiency programs. This education will do much to ease the dialog between regulators and the regulated, thus encouraging a more collaborative approach. Similarly, evaluations can often reach out to elements of the end-user base to elicit their views, which offers another avenue for communicating the utility's engagement in publicly endorsed efforts to reduce climate change.

Finally, accounting for stakeholders' objectives and continuing to "bring them along" will greatly increase their buy-in of both the programs and how they are measured. This is absolutely critical. Adoption of a "no surprise" approach to evaluation is the best approach to minimizing conflict.

Program Theory-Based Evaluation

One helpful approach to planning and communicating an evaluation's intent is to explain the program's theory. A program theory provides a logical progression of steps the program takes and explains how these will deliver the intended results. Notably, the program theory should:

- Present program goals as well as activities it will generate to accomplish these goals
- Describe the causal relationships between program activities and intended results
- Explain why program activities are designed to lead to the intended outcomes

Developing the Program Theory and Logic Model

In recent years, this recommended step in the evaluation planning process has been increasingly adopted as a way to organize the most important researchable issues and plan evaluation outcomes.

The most common elements of a logic model contain:

- **Program activities**. These should include services offered, incentive mix (e.g., rebate, financing, free services), and marketing.
- **Resources to support program activities**. These include program costs to the utility (e.g., marketing, administration, incentives, and evaluation budget).
- **Program outputs**. These would include the number of participants and the number of technologies or measures installed.
- **Expected program outcomes**. These should include program objectives resulting from program activities, and results such as energy/peak demand goals or market effects for market transformation programs. These outcomes should be categorized into short, medium, and long-term outcomes.

Evaluability Assessment

Evaluations depend upon an assortment of both primary and secondary data used to probe, qualify, and extrapolate a realistic picture of how programs operate within the sometimes chaotic context of an ever-changing consumer behavior, a volatile environment, and the marketplace's capricious behavior.

"Evaluability" is a new term introduced to describe the probability evaluation data will be available as evaluations are being conducted. This way of reviewing data considers the following considerations:

- Is there a way to track the data?

- Is there a way to track non-participants?
- Are measures' specific locations being tracked?
- Are program assumptions, such as times of operation, being tracked?
- Are the delivered energy savings being tracked?
- Does the recording device used include both the outcome and the activity?
- Are savings assumptions being documented?
- Are sources of savings assumptions being documented?
- Does the database record the "as-found" values for parameters used to estimate ex ante savings?
- Does baseline monitoring need to take place?
- Can one of the impact evaluation methods specified in this guide be used?
- Are there code compliance or program overlap issues for savings estimation?

Data Tracking System

The choice of data tracking system is one of the early items to consider for program implementation and evaluation, especially as the utility chooses an implementation approach. Some utilities outsource program implementation to a single contracting firm, while some opt to engage multiple contractors. Most contractors that offer energy efficiency program implementation services employ their own data tracking system. Therefore, in instances where implementation is outsourced to a single contractor, that contractor's data tracking system is typically used as the repository for all program data. When, on the other hand, a utility chooses to outsource program implementation to several contractors, it is recommended that the utility require all contractors to input all program data into a single data tracking system of the utility's own specifications. This is preferable to using multiple tracking systems (i.e. each contractor employing its own system), although multiple tracking systems are still more typical in the field. A single data tracking system forces consistency and allows access to all programs in one location. If the tracking system is well developed, it can provide several stakeholders with the ability to access all data in one place and aggregate it in any way desired (e.g., all residential, all commercial, all lighting, all northern territory, etc). Most current data tracking systems have "dashboards" that can be customized to a utility's needs.

As mentioned above, many program implementation contractors have their own proprietary tracking systems that they make available to their clients. Some utilities choose to use such systems, while others prefer to develop their own. While EPRI does not recommend any particular system, it strongly advises utilities to determine what kind of system best meets their needs to ensure that the data needed for evaluation can be properly tracked.

One final point that utilities should consider when assessing a data tracking system is whether the potential system can interface with the utility's existing customer information system. Matching customer data with program participation information not only makes evaluation easier, but it also allows utilities to better target customers for overall program participation.

Reviewing Data Requirements

Reviewing data requirements is a critical step for impact evaluations as it requires that data be organized and accessible. If this is not the case, the evaluation could take much longer and be much more costly to execute.

Billing Data

Customer billing data are extremely important and valuable because they describe customer behavior in some detail and are available for all customers, whether program participants or not. Customer billing data should include the following fields:

- Account number
- Premise number
- Customer name
- Business or key contact for C&I customers
- Customer address
- Customer rate code
- Monthly kWh consumption
- Monthly billing demand (where applicable)
- Customer revenue codes
- SIC (NAICS) for C&I customers
- Date the bill was read (and if it was estimated)

Engineering Estimates

The nature of technical data collected can vary greatly, depending on the program used because different programs necessitate collection of different types of technical data. The technical requirements for a residential audit program, for example, include:

- The recommendation the auditor makes to the customer
- Demand savings
- Energy savings
- Measure cost
- Payback period
- Audit date
- Report date

As another example, the technical requirements for a business lighting program may include:

- Installation location
- Equipment type, quantity, model, and wattages (lamps and ballasts)
- Equipment being replaced: type, quantity, model, and wattages
- Hours of use of lighting equipment
- Use of lighting controls

Load Profiles

Depending on evaluation objectives, specifically if estimates of demand impacts are desired, load profiles will be needed. Load profiles show consumption of either an entire structure (building, house, etc.) or consumption of a specific end-use over a period of time (e.g. one day to one year).

To prepare for evaluation, a utility's existing load profiles should be obtained, if available. For those unavailable or not covering the period of interest, secondary sources should be investigated. Often, other utilities have load profiles they are willing to share, if appropriate. Alternatively, engineering simulation models may be used to create fairly accurate proxies of hourly load profiles for buildings or end-uses.

Program Data

Whether the utility chooses one tracking system or several, serious thought should go into deciding the data to be tracked. This choice is mostly dependent on evaluation objectives. Each objective should be mapped to one or two metrics, which then need to be mapped to one or more data inputs. For example, an evaluation objective may be a measurement of energy savings for residential lighting. This requires mapping to measurement of pre- and post- wattage and hours of use. The data tracking system should include these inputs in the program data collection forms.

There is an old adage that a question properly posed is a question half answered. By the same token, an effective evaluation relies in great part upon data derived from the program itself. Thus, it is necessary to consider at the outset the program data related to the verifiable results of initiatives launched under the aegis of the program being evaluated. Examples of questions to guide fundamental data requirements for evaluation include:

- How many CFLs were installed?
- How many refrigerators were hauled away?
- How many homes met the required standards set forth by the program?

As a basis, verifiable achievements are required to begin a performance review. At this point, it is probably opportune to reiterate the importance of designing the evaluation as programs are launched; so evaluators can be certain the programs produce verifiable and accessible data about performance. Fundamental to this process is the availability of good primary data about program impacts.

External Data

Aside from programmatic data, whether provided from program results or from the utility's partners, a solid evaluation must rationalize raw data to reflect the impact of external events. Thus, an integral part of the analysis requires the identification and securing of data pertinent to external forces shown to influence the results.

Perhaps the most obvious of these elements are the "elements" themselves. For example, weather affects consumers' behavior regarding heating and cooling, making it necessary to factor these externalities out of results to provide a view of what would have occurred under "normal" circumstances. Similarly, demographic composition of the population under consideration will affect their overall behavior patterns. People at opposite ends of the economic spectrum are unlikely to react similarly to all stimuli, whether these are delivered as changes in prices or as widely broadcast social marketing messages. These variances, though generally recognized, need to be quantified statistically to account for the effects of demographic diversity.

Data Quality Assurance

Another important concern is the completeness and usability of data; incomplete data may be impossible to use.

To assure complete data are collected, it is recommended that all data fields be organized in an identical format, i.e. all fields such as names, addresses, phone numbers, etc., should be formatted uniformly. This allows for smooth data manipulation, but also requires the discipline of interviewers to enter information in a consistent manner.

Additionally, it is important that all data fields are completed to the extent possible. While it may not always be possible or necessary to have complete records for every data series, every effort should be made to populate the most essential data fields.

Define the Eligible Market

As evaluations typically focus on customers eligible for the programs being evaluated, it is important to ascertain, to the extent possible, the size of that population. This allows the evaluator to determine participation rates and to size samples appropriately.

Population size is often estimated at the start of a program, so it may be necessary to re-estimate that population if the character of the utility's customer base has changed significantly in the interim.

If planning estimates for population size are not available, it may be necessary to refer to secondary sources, such as U.S. Census data or to extrapolate from similar groups previously identified in nearby and roughly identical service territories.

Participants and Non-participants

One of the methods for normalizing “raw” results from energy efficiency programs is to establish a control group against whose behavior the target audience can be assessed. Since we do not live in a vacuum that allows analysis to be conducted without undue “outside” influence during the course of the program implementation and evaluation, development of this “baseline” response is crucial to understanding how the program impacts its target audience, even as it reacts to ongoing changes. Without the external data a non-participating sample group provides, it would be impossible to show the true variance between expected behavior, program-affected behavior, and desired behavior.

Evaluation Costs

State regulatory agencies often define the percentage of an energy efficiency program’s budget that may be allocated to evaluation. At the lower end of the funding spectrum, some jurisdictions allocate around 2 to 3 percent of estimated savings to cover evaluation costs. Such meager funding typically yields evaluation results with a higher degree of uncertainty and lack of program-specific detail with which implementers and regulators can effectively review and manage the energy efficiency initiative.

Recently, many states have increased their funding levels for evaluation. New York, for instance, recently raised its funding level from 2 percent to 5 percent[5] and justified the increase based on the following assumptions:

1. Greater detail and accuracy in the evaluation should validate a much higher estimate of savings, allowing utilities to establish significantly higher recovery of “lost revenues.”
2. Reduced uncertainty in estimation of savings more effectively validates the role of energy efficiency in reducing the need for transmission and generation upgrades or new plant investments. The avoided cost these savings enable can be significant and go a long way to justify public investment in energy efficiency initiatives, both for regulators and the public.
3. The increased accuracy of evaluation results permits more reliable demand projections that aid in planning future energy needs.
4. At the operational level, regulators recognize that more accurate and detailed performance data allow implementers and utilities to improve the performance of energy-efficiency programs more quickly and in a more targeted manner. These “in-stream” adjustments assure programs achieve the highest return (in savings) on the given investment of ratepayer set-aside funding.

[5] Bill Saxonis. “New York State’s Energy Efficiency Portfolio Standard”. Office of Energy Efficiency and Environment, New York State Department of Public Service. Presented at NARUC Workshop on Energy Efficiency Program Performance. Washington, D.C. February 16, 2008. (http://www.narucmeetings.org/Presentations/Saxonis.ppt#398)

At the more generously funded end of the spectrum, some states have pegged the percentage of savings dedicated to the production of evaluation results as high as 8 percent.

This higher funding level permits these jurisdictions to examine many of the key assumptions affecting energy savings estimation and redefine them in ways that more closely reflect their regional experience, while validating relationships that have been surmised but not adequately documented to meet the rigor required by regulatory bodies that accept or reject those redefined or emerging standards.

Often, allocating a fixed percent of program budgets for evaluation can lead to difficulties. For example, in the year programs are launched, budgets are not very high as participation rates are not expected to be high. However, evaluation costs are usually proportionally higher in the first year, when evaluation plans are constructed, survey instruments created, tracking data bases designed, and so on. Yet, the evaluation budget, defined as a percent of program budget, is likely to be lowest in the first year.

Reducing Evaluation Costs

The most effective way to manage evaluation costs without impairing the essential value of the resulting evaluation report is to apply a “bottoms-up” budget prioritization. Two perspectives should be considered in this prioritization effort.

First, consider which programs deliver the most estimated savings and which are the most cost-effective. Based on gross savings estimates and raw program data (prior to being normalized to exclude external influences), a capable manager should be able to determine which programs deliver the most savings, while simultaneously validating that these programs do not have inordinately high implementation costs.

Given the importance of these very productive programs, it is reasonable to prioritize the evaluation work associated with them. This “bottoms-up” prioritization and budget allocation ensures documentary support is made available for documents of the energy efficiency programs that contribute the highest estimated savings. This ensures the greatest cost recovery for the utility and provides the most convincing validation of the overall efficacy of energy efficiency programs for regulators and the public.

Second, the utility manager should consider the inherent costs of various types of programs. Depending on the intent of the energy efficiency program, it is clear that some programs are much more costly to evaluate than others.

In general, programs that attempt to suppress overall energy use are easier to evaluate because gaining information on total energy consumption of users is easier to obtain than information about *when* they use energy. The latter information is necessary if the measure’s purpose is to reduce or shift demand away from periods of peak usage.

Securing information about programs that aim to reduce peak energy demand (expressed as kW) requires knowledge about customers’ total energy consumption and when they are drawing that power. This secondary variable requires sophisticated data gathering and cannot be gleaned from analyzing customers’ utility bills. Depending on the nature of the target, securing this added dimension of customer usage may require the installation of meters. This adds considerable complexity and cost to data collection activity and makes programs aimed at energy demand shifting more expensive to evaluate.

By a similar token, an examination of data collection methods proposed as part of any program will be a good measure of the expense associated with that program. Retail store intercepts, stocking analyses, or examination of building permits at a governmental office are all time consuming and require significant set-up time. By contrast, billing analysis and a review of program records are the easiest tasks to accomplish. In short, evaluating programs that measure demand will be more costly than those measuring energy uses. Beyond that generality, it behooves utility managers to consider estimated costs and time requirements for collection methods proposed to ascertain the overall allocation of program evaluation costs.

Evaluation Schedule

Increasingly, jurisdictions are initiating evaluation planning nearly contemporaneously with design and deployment of actual energy efficiency programs. In some instances, evaluators are introduced into the process even as program implementation procedures are being established with the implementation contractors. In this way, procedures and data sources can be jointly agreed upon so that evaluation efforts can use primary data resulting from the program without having to re-query the participants. This coordination of data gathering and real-time sharing of program performance data greatly improves the evaluation's quality, reduces data acquisition costs, and results in more effective feedback, which can improve the program even as it is being deployed.

Evaluations are rarely conducted for purely internal reasons. Typically, they are used as part of a public process to review the results of a utility's energy efficiency programs. These public processes are scheduled along a predictable time frame, allowing utility staff, implementers, and evaluators to forecast when program results will need to be validated.

These public processes will, to a large extent, dictate the schedule to which any evaluation should conform. The experienced evaluator will identify critical review dates and work backwards to establish a time line sufficient to allow the program to generate sufficient results to be evaluated, then allow the evaluation to take place, and, finally, allow staff and program participants to review the results before recommending actions to regulatory bodies charged with overseeing the cost-effective investment of public funds. Table 3-1 provides a typical evaluation preparation timeline, with the frame of reference that "time = zero" is the point of issuance of a request for proposal (RFP) for an evaluation contractor.

Table 3-1
Evaluation Preparation Timeline

Item	Milestone	Month of Occurrence
1	Energy-efficiency program initiation *(typically phased approach for portfolio program)*	- 3 to 0
2	Appointment of program implementer	- 1
3	Issuance for evaluation RFP Deadline for Evaluation response	0 to 1
4	Appointment of Evaluator	2
5	Kick-off meeting	2.5
6	Delivery of Logic model and evaluation process	3.5
7	Review of Technical Resource Manual *(defines processes & protocols)*	3.5
8	Define QA/QC procedures	3.5
9	1st Annual Review **or dependent on utility's program intervals*	12*

4
IMPACT EVALUATION BASICS

Introduction

Impact evaluation assesses the changes in energy use that can be attributed to a particular intervention, such as the installation of energy efficient equipment, participation in a demand response program, or increase of awareness.

The challenge of impact evaluation is that it seeks to measure “what did not happen.” In other words, this type of evaluation assesses impact as the difference between what actually happened (easily measured) and what would have happened had the program not existed (significantly more difficult to measure).

Energy Savings

“Energy savings” refers to the reduction in energy use after treatment or installation of an energy efficient device. Mathematically, energy savings are the difference between energy use post treatment and a *baseline* of energy use. A baseline attempts to define what *energy use would have been had the treatment not been applied*. Often the baseline is simply the pre treatment energy use. Enhancements on this simple approach include adjustments of the pre treatment use to reflect post treatment conditions such as weather, production levels, hours of use, square footage, number of occupants, etc.

Energy savings is computed as the product of reduction in gross kW draw of a particular end use and its associated hours of operation (or full load hours). For example, if a lighting program changes a 100 watt incandescent light bulb for a 23 watt compact fluorescent light bulb, the reduction in gross kW is computed as:

$$\Delta kW = \frac{(100-23)}{1{,}000}$$

If the incandescent light bulb was used approximately 3 hours a day (1,095 hours annually) and the CFL usage is expected to remain the same with the CFL in place, energy savings can be computed as:

$$\Delta kWh = \frac{(100-23)}{1{,}000} * 1{,}095 \approx 84$$

Gross Savings Estimation

Since the absence of energy use is an elusive substance to measure, the common approach is to compare the before and after states to determine what has changed. This element of change should reflect the elimination of some portion of the prior energy use after implementation of a program. Thus, the following general equation applies for energy savings and demand:

$$\text{energy savings} = (\text{baseline energy use}) - (\text{reporting period energy use}) \pm (\text{adjustments})$$

Definition of Terms

The important terms for this formula are as follows.

Adjustments

To rigorously calculate the savings attributable to an energy efficiency program, key external factors beyond the control of the program or consumer need to be accounted for beyond the simple difference in measured energy consumption pre- and post- program. Such factors can include corrections for variances in:

- Weather conditions— particularly for programs involving heating, ventilation, or air-conditioning (HVAC) systems in buildings;
- Building occupancy levels and hours
- Production levels in industrial facilities

These "adjustments" are made to align energy use in the pre- and post- program time periods to the same set of conditions in order to neither understate nor overstate the impact of the program.

Demand Savings

"Demand savings" are reductions in the rate of use of electric power measured in kilowatts (kW). Typical incandescent light bulbs, for example, demand between 25 to 100 watts of electricity. Replacing a 100 watt incandescent light bulb with a 23 watt compact fluorescent lamp (CFL) reduces the energy use by 77 watts during the time the light bulb is in use.

Co-Benefits

"Co-benefits" refer to ancillary benefits derived from energy savings, typically due to the reduction in power generated from carbon-based fuels, including: avoided greenhouse gas (GHG) emissions; improved health benefits; and better energy security. These co-benefits can extend into more efficient transmission/distribution, water savings, and macro-economic benefits such as job creation. For energy efficiency programs targeted to low income households, energy efficiency often also leads to better payment behavior and reduction of bad debt.

Persistence

"Persistence" refers to the changes in the savings level of a measure over time. The performance of an energy saving measure may degrade over its life due to many factors including consumer behavior, removal, burn out, and declining efficacy as the equipment ages. Maintenance practices can also impact persistence.

Net Savings

"Net savings" refers to the total change in load that is attributable to an energy efficiency program. This change in load may include, implicitly or explicitly, the effects of free drivers, free riders, energy efficiency standards, changes in the level of energy service, and other causes of change in energy consumption or demand. Of the "observed" savings, the amount that can be attributed directly to programmatic effort is net savings.

Uncertainty

"Uncertainty" in the evaluation process arises from the difficulty in measuring the *absence* of energy use. Evaluations, therefore seek to determine "expected values" based on methodologies that compare "after" and "baseline" states. Of course, these estimates are only as good as the methodologies and sample characteristics used to generate them, and the degree of confidence provided by the methodologies is expressed in terms "confidence" and "precision." Thus, an expected value might be characterized as having a confidence/precision quotient of 90/10, with 90 percent describing the probability that the true savings value are expected to occur with the ±10 percent margin of error. In the example above, if the estimated 84 kWh was derived from a sample designed to produce 90/10, one would make the following statement: average energy savings was estimated at 84 kWh, and we are 90 percent confident that the *true value* (which is unobserved as we would not have included all the CFLs in the program in the study) is between approximately 76 and 92 kWh.

Uncertainty can also arise from other factors such as equipment accuracy and non response bias from surveys.

Reducing uncertainty can be achieved at a cost. Thus, the practical concern about uncertainty hinges upon finding the right degree of precision given the importance of the data under consideration. Successful evaluations try to be **cost-effective** by balancing risk, uncertainty and cost considerations.

Goals

Typically, impact evaluations attempt to answer some or all of the following questions:

- How much energy was saved? How much is attributable to the program (net savings)?
- How much was demand (kW) was reduced? How much was attributable to the program? When does this demand reduction occur (e.g., hour of day, season)?
- How long are program benefits expected to last?
- What were the co benefits of the program?

- Were these reductions achieved in a cost effective manner? From which perspective (i.e., participant, utility, society, ratepayers)?

Evaluation Methods

As mentioned above, at the heart of impact evaluation is measurement of gross and net energy and demand reductions. Savings may occur through a wide variety of program-induced actions, such as weatherizing a home, educating customers, installing energy-efficient air conditioners, enrolling customers into time-differentiated pricing programs, etc.

To estimate program impacts, an analyst can use the following data sources:

1. Secondary data – results from another similar program with some minor adjustments for local conditions and installation rates
2. Deemed savings
3. Engineering models
4. Statistical models
5. Metered data

Next, we explain the use of each type of data for estimating energy and demand impacts.

Secondary Data

A significant body of knowledge exists from evaluation of energy efficiency programs over the last three decades. Some programs are fairly straightforward and may not require direct evaluation (e.g., residential lighting). Results from other jurisdictions may be applied to your program, though some adjustments will likely be required. Such adjustment may include installation rates of your program, your weather conditions, daylight hours in your region, etc. There may be instances where one can use some secondary data from another program, and collect primary data for one's own program on most critical or uncertain inputs. For example, one may use hours of use from a large scale metering effort in a nearby utility, but still conduct customer surveys to obtain estimates of installations, removals, free ridership, etc.

Leveraging secondary data can greatly reduce evaluation costs, and should always be explored.

Deemed Savings

"Deemed savings" are pre-determined energy savings values for most common energy efficiency measures. Sometimes deemed savings values are approved by regulators, not just as planning estimates but also for purposes of conducting impact evaluations. In these cases, the evaluation effort may consist of periodic review of these estimates along with frequent comparison to commonly used values from sources such as the Database for Energy Efficient Resources (DEER) in California, or the Regional Technical Forum (RTF) estimates in the Northwest. States that use deemed savings databases currently include Vermont, Texas, New York (NYSERDA), Wisconsin, and Minnesota (which is using a customized version of the RTF estimates). Most states have developed their own list of deemed saving values, calculated to take their respective climates into account. The DEER database and the RTF both incorporate significant climate variations because they cover such broad geographic areas and can assign savings values based on the zip code in which a measure is installed.

Deemed savings are computed using assumed typical installations. The actual installations will need to be compared to the assumed typical cases and adjustments made as necessary.

Engineering Models

"Engineering models" include a family of approaches ranging from simple definitional calculations to complex building simulations. These methods vary over a wide range from simple algorithms like the CFL savings equation shown in the energy savings section above to fairly complex simulation tools such as DOE-2. Simple engineering calculations are best suited for cases where the energy savings can be computed through the use of predictable inputs such as change in wattage and hours of use for lighting measures. Simple models may also be appropriate for other low-cost, "back-of-the-envelope" estimates.

When the energy relationships are more complex or the required level of detail is more rigorous, the preferred analysis method is Calibrated Simulation as outlined in International Performance Measurement & Verification Protocols, Option D. There are several types of simulation tools available as shown in Table 4-1.

Table 4-1
Simulation Tools

Energy Simulation Models	
Commercial	**Residential**
eQUEST	Energy-10
EnergyPro	REMRATE
EnergyPlus	EnergyGauge USA
TRACE	ESPRE
BLAST	
EZSIM	

With these simulation tools, the user develops a theoretical model of a commercial building or a residence by entering the necessary data to describe the building characteristics, such as building area, envelope elements HVAC systems and operation, occupancy schedules, and weather data. A sample input template is shown in Table 4-2 for a restaurant and a single family home.

Table 4-2
Critical Inputs for Most Simulation Models

ConstructionType	**Existing**	**Existing**
Building Type	***Restaurant***	***Single Family***
General Information		
Exterior Wall Construction	2x6 -24" o.c. Wood with Brick Exterior Finish, Medium Abs.	Stucco, Standard 2*4 Wood Framing, Insulation
Roof Construction	Standard Wood Frame, Built Up Roof	Shingle Roof, insulation, Dark Colored
# of Floors	1	2
Floor to Ceiling Height	20	10
Total Floor Area [sqft]	2247	1921
Roof Area [sqft]	2247	960.5
Envelope		
Window U-factor	U=0.65	U=0.67
Window to Wall Area	15%	20%
Wall Insulation (R Value)	R-15	R-11
Roof Insulation (R Value)	R-19	R-25
Floor Insulation (R Value)	R-11	R-15
Lighting Density [W/sqft]	1.75	1.52
Occupancy Schedule		
Occupancy Schedule WkDay	9am-9pm (Customer Operating Hours)	5pm-9am (Only Occupancy)
Occupancy Schedule WkEnd	9-9 Sat 11-7 Sun (Customer Operating Hours)	24 Hours (Only Occupancy)
% occupancy Wkday	38%	NA
% occupancy WkEnd	33%	NA
Annual Operating Hours	4171	NA
Water Heater		
Water Heater Capacity (gal)	80	50
Water Heating Fuel Type	Gas	Gas
Water Heater Energy Factor	0.62	0.62
Supply Temperature	135	125
HVAC		
Heating and Cooling Type	***Rooftop Package DX Unit***	***Central AC w/ Gas Furnace***
Heating Efficiency	2.7 COP	7.7 AFUE
Percent Of Building Heated	100%	100%
Cooling Efficiency	9.2 EER	SEER 14
Percent Of Building Cooled	100%	100%
Set Points		
Heating Daytime Set Point [°F]	70	68
Heat. Setback/Setup Set Point [°F]	67	64
Cooling Daytime Set Point [°F]	72	72
Cool. Setback/Setup Set Point [°F]	75	75

With the theoretical model created, a preliminary simulation is performed for comparison to the billing history. If the predicted energy consumption is within ±10 percent of the billing data, the model is considered to be valid for impact evaluation. If not, a calibration process is necessary. This involves adjusting assumed input values until the model is within a reasonable level (e.g., within ±10 percent).

Once the simulation model is calibrated to the actual baseline billing history, the user can "install" the specific measure data into the model. The resulting post-installation consumption is then compared to the baseline data to determine the net energy savings for the energy efficient measure. A multiple measure evaluation can be simulated as individual measures in a specific order to determine unique measure savings, or as a combination of measures in order to determine the interactive effect of various different energy efficiency measures.

In addition to calculating total building energy consumption, most simulation tools provide specific end-use consumption results such as space heating and cooling, lighting, domestic hot water, and miscellaneous equipment. An example eQUEST (DOE-2) output converted to a pie chart for end-use percentages of total usage is shown in Figure 4-1.

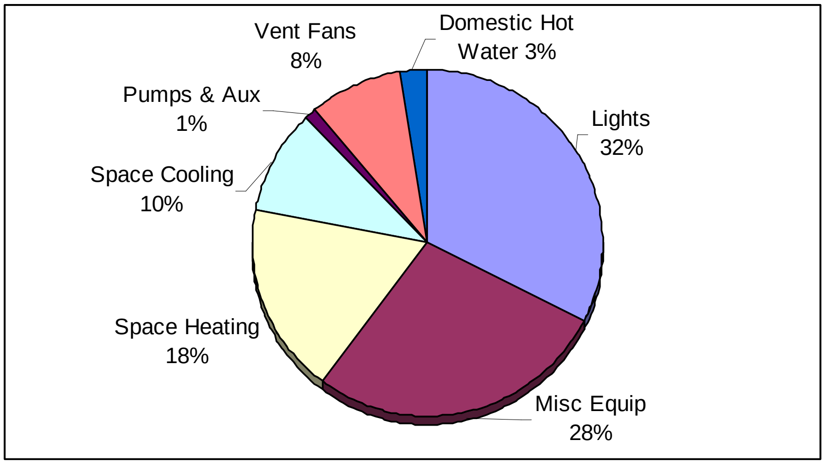

Figure 4-1
End Use Level Summary of a Building Simulation Model in eQuest[6]

Simulation tools such as eQuest, Energy-10, and EnergyPro have the ability to generate hourly data for a year (8,760 load profile) by end-use. These results give the user the ability to predict the typical peak time interval (peak kW demand) for a building and can be used to calculate the average and/or peak demand savings associated with specific measure analyses.

End-Use Metering

"End-Use metering" is used where energy and demand impacts may be affected by external variables such as weather conditions, seasonal variation, and variable or unpredictable operating schedules. Candidate examples would be heating/cooling equipment replacement, variable speed controllers, energy management control systems, and automated lighting controls.

The primary step is to develop a detailed plan identifying the end-use equipment to be metered, the parameters and data involved, the length of time required for data collection, and the metering equipment required to capture the necessary data for impact analysis. A description of the baseline (pre-retrofit) and the anticipated post-retrofit equipment and operating conditions, as well as any assumptions made for unknown conditions should also be included.

Metering equipment ranges from simple stand-alone devices recording on and off states of equipment operation to sophisticated devices recording multiple time-series parameters of equipment performance characteristics. Important factors involved in selection of metering

[6] Modeling based on small office building in Seattle with DX cooling package and electric resistance heating.

devices are the number and type of data elements, metering time period and duration, location of equipment to be metered, data retrieval considerations, and the required accuracy.

The duration of metering is always a concern due to cost and timeline considerations. As a general rule, if the equipment does not have seasonal pattern in its usage, short time periods are sufficient. For example, many evaluations of commercial lighting use tend to be conducted with lighting loggers left in place for approximately one month. If the equipment has a strong seasonal pattern, then metering needs to include the season of interest. For example, programs focused on reducing summer peak demand through installation of efficient HVAC equipment, need to include summer peak in the metering period.

Figures 4-2 and 4-3 show commercial lighting and residential air conditioning load profiles over the course of one year (i.e. 8760 hours). The figures clearly show that some load profiles are very seasonal and some are not.

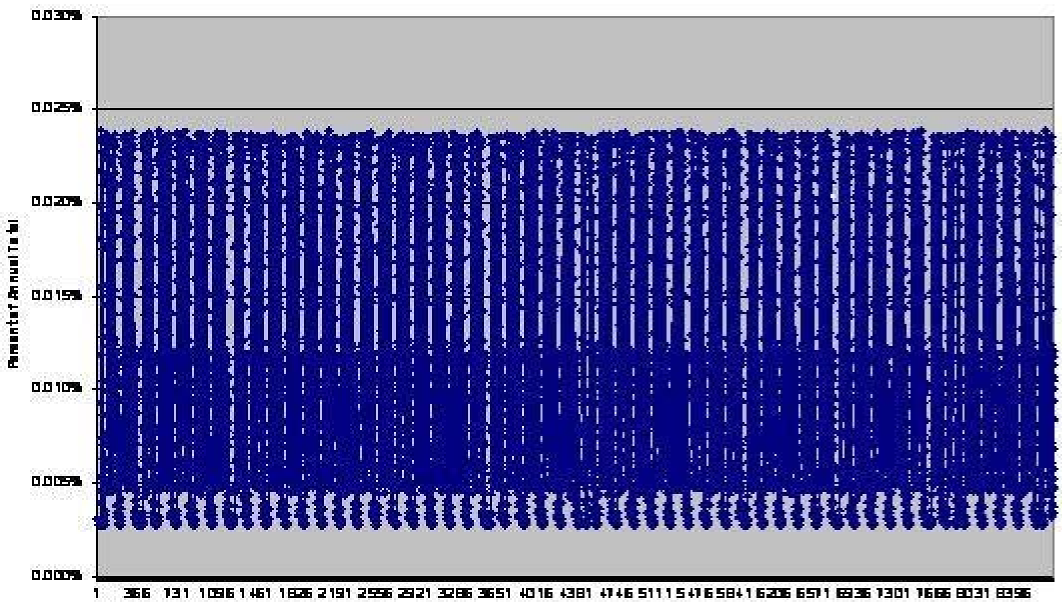

Figure 4-2
Commercial Lighting Illustrative 8760 Load Profile

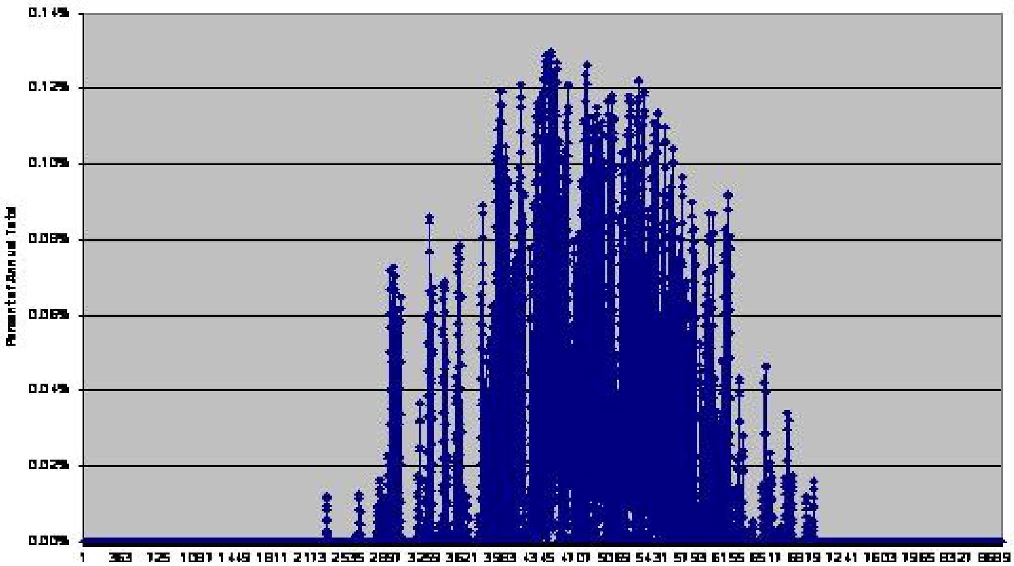

Figure 4-3
Residential Air Conditioning Illustrative 8760 Load Profile

It may not always be possible to meter the baseline energy usage, in which case an assumed baseline needs to be created. Baseline equipment manufacturers' data may be used to estimate operational characteristics along with the post-retrofit usage profile to establish a proxy for equipment runtime. Appropriate engineering algorithms would then need to be developed to determine the energy and demand savings impact.

Careful sample design is essential with end-use metering, as the time and cost involved in data collection per metered piece of equipment can be prohibitive. There are nationally recognized guidelines, such as the IPMVP and ASHRAE Guideline 14-2002, which detail metering requirements by retrofit measure category and the related analysis uncertainties. Both documents are cited in the Reference section. Neither of these documents provides explicit sample design instructions, but rather provides guidelines for when, and when not, to meter.

Statistical Analysis

Statistical models attempt to derive estimates of demand and energy savings by analyzing variances in consumption data, typically though monthly bills or end use metered data. Through this analysis of variance, statistical models attribute observed changes to various explanatory variables such as weather, occupancy, or participation in an energy efficiency program. The following types of impact evaluation models are used in estimating energy efficiency program impacts:

- Simple Pre/Post (difference of means)
- Simple Regression (accounting for weather)
- Quasi Experimental Design
- Detailed Regression

- Combined Engineering-Statistical Models

Simple Pre/Post (Difference of Means)

A simple pre/post difference of means is the simplest statistical model. Though almost never used, it is included here as starting point for more complex and realistic models. Estimated savings are simply the difference between average annual consumption in the pre- and post-program periods. At minimum, the two averages should represent the same time period (e.g., both have to cover total year and not have different number of bills).

This is not a very useful approach as it does not take into account any potential differences between the two time periods and assumes that *all* change is due to program participation.

Simple example, average annual actual pre = 10,000 kWh, average annual actual post = 8,000 kWh, simple difference: Program savings = 2,000 kWh.

Simple Regression (Accounting for Weather)

A significant improvement over the simple difference of means is to start to account for other potential causes of difference in annual consumption, most notably the weather. In order to achieve this goal, you need to use regression models.

Regression models (also called causal models) attempt to *explain* the variations in the dependent variable (Y) through the use of explanatory variables (Xs). Assume we have a regression model with the Y variable being energy use and one explanatory variable being weather. Regression models may "observe" that energy consumption is higher in certain months of the year. It attempts to explain these variations by the given explanatory variable X. Where positive correlations are observed, such as high temperature with high energy use (i.e. driven by air conditioning load), the model assigns a positive sign to the relationship (i.e., both X increase so does Y and X decreases do does Y). The model then attempts to quantify the exact form of the relationship, i.e., change in energy consumption (in kWhs) per degree change in temperature. Simple regression models have one X and multiple regression models have several Xs.

The model often used takes the following form:

$$Daily\ EnergyUse = \alpha + \beta_1\ HDD + \beta_2\ CDD$$

In this formula, α represents a constant value and β_1 and β_2 represent coefficients to heating degree days (HDD) and cooling degree days (CDD). Using this formula, energy consumption data from customer bills are converted to average daily values to eliminate differences due to different meter read periods (e.g., some billing periods may covers 25 days whereas others may cover 35). Variations in daily energy use are explained through heating and cooling degree days.[7] Once this model is estimated, long run average CDD and HDD are entered and the result is called normalized annual consumption, i.e., normalized for weather.

For example, you start by running a regression model for both the pre and post periods. Consider the following illustrative data:

- Pre kWh = 4,000 + 2 HDD + 0.5 CDD
- Post kWh = 3,700 + 1.5 HDD + 0.4 CDD
- Average 10 year weather shows
- Average HDD = 2,500
- Average CDD = 1,000
- Normalized Annual Consumption = NAC
- Pre NAC = 4,000 + 2 (2,500) + 0.5 (1,000) = 9,500 kWh
- Post NAC = 3,700 + 1.5 (2,500) + 0.4 (1,000) = 7,975 kWh
- Normalized savings = 1,525 kWh (Not 2,000 in the example above)

The NAC values show what the consumption would have been had the weather been normal. They are unlikely to match the actual consumption from metered data.

Quasi Experimental Design

In the section above, the analyst has "removed" the impact of a major cause of the difference in consumption from the pre to the post period. However, the analyst still assumes that the remaining difference of 1,525 kWh is caused by the program. This may or may not be true.

In a pure experimental design (usually conducted in a laboratory), the researcher can control for all important inputs and allow one of them to change at a time and observe its impact. In social sciences, laboratory conditions do not exit. Researchers rely on "quasi experimental design" methods instead. Under these conditions, one group of subjects undergoes one kind of experiment (e.g., students are taken through one learning approach) and another group is not. The first group is called the treatment group and the second is called the control group. Both groups are tested pre- and post- experiment. The differences between the pre and the post test results are compared and that difference is compared (difference of difference). The result is a net impact of the experiment. In other words, the observed difference in the treated group is what occurred to the treated group. And the observed difference in the control group measures what would have happened absent the treatment. The difference between the two changes is the *net* impact.

---

[7] CDD=max{temp-65, 0} and HDD=max {65-temp, 0}, 65 degrees is called reference temperature. Other reference temperature values can be used.

Ideally the comparison group is selected randomly (in which case it is called a control group), and it is isolated from any interventions similar to the original intervention, in order to generate an experimental design and to avoid selection bias. Also, ideally, the comparison group is similar to the treated group in all aspects except the actual treatment.

In energy efficiency programs, the comparison group is randomly selected from non participating customers. Similarity is achieved by selecting from similar customers by geography, size, type, energy use levels, etc.

For example, assume similar group of homes, on average, had a change in average normalized annual consumption of 500 kWh:

- Program gross impact is 1,525 kWh
- Program net impact is 1,025 kWh

In other words, had the program not existed, the treated homes would have also witnessed a reduction in use of 500 kWh (for all other reason other than weather). The reduction in consumption may have taken place due to economic conditions, change in rates, etc.

Table 4-3 shows an example of a program in Indiana in which changes in normalized annual consumption were compared between treated and untreated homes. As the table shows, the non-participants were selected to be similar to the participants in terms of pre-program normalized energy use (12,953 kWh for non-participants compared to 12,414 for participants). In the post period they no longer appear similar; participants' use dropped by 1,086 kWh, while non-participants' use increased by 462 kWh.

Table 4-3
Example of Quasi Experimental Design8

Client Group	Preprogram Normalized Energy Consumption (kWh)	Postprogram Normalized Energy Consumption (kWh)	Change in Consumption (kWh)
Nonparticipants	12,953	13,415	462
Participants	12,414	11,328	(1,086)
Net program Impact			(1,548)
Net impact in percentage terms			12.5%

The quasi experimental design approach assumes that, had the treated homes not been treated, they also would have witnessed an increase in use on the order of 462 kWh. In other words, the

[8] Based on program evaluation work conducted by The Cadmus Group, Inc.

comparison group shows what the treated homes would have done had they not been treated. The net impact is 1,548 kWh.

- Treated homes decreased consumption by 1,086.
- Had they been left alone, they would have increased consumption by 462.
- The net impact is 1,548 kWh.

Detailed Regression Models

More detailed regression models may also be used to provide additional information. Explanatory variables may include occupancy, square footage, production levels, details of installed measures, etc. For example, if the explanatory variables included those indicating presence or absence of individual measures, the estimated regression coefficients would provide information on the impact of these measures. In other words, the total impact may be decomposed to impacts due to specific measures. A model used often in energy efficiency program evaluations is called conditional savings analysis (CSA) and takes the following form:

$$ADC_{it} = \alpha + \beta ECM_i + \lambda_1 HDD_{it} + \lambda_2 CDD_{it} + \varepsilon_{it}$$

Where,

ADC_{it} is the average daily consumption during the pre- and post-program periods

ECM_i is a binary variable (0 or 1) indicating the presence or absence of the specific energy conservation measures at site i. The associated beta represents the savings associated with the specific measure.

HDD_{it}, is average daily heating degree days based on facility location

CDD_{it} is the average daily cooling degree days based on facility location

ε_{it} is the error term

Table 4-4 shows the result of CSA analysis for a utility program in the Midwest. In this program, air sealing measures did not save any energy; insulation measures saved 26.89 kWh per month on average, etc. Overall, average participant in the program saved 426 kWh.

Table 4-4
CSA Model Summary for One Program Example

Measure	Savings (kWh/Month)	Proportion of Units With Measure Installed
Air sealing	0	27.8%
Insulation	-26.89	19.4%
Lighting	-27.26	99.5%
Water	-8.24	33.9%
Other	0	35.5%
Common	-24.65	1.49%
Weighted Average x 12		426

Combined Engineering-Statistical Models

In these models, engineering-based estimates are used as explanatory variables in regression analysis of whole customer loads. Parameters verify the reasonableness of the engineering estimates or identify systematic biases. These coefficients are often called realization rates.

A very commonly used model in energy efficiency program evaluation is called statistically adjusted engineering estimate model (SAE). Using energy consumption during the post-installation period as the dependent variable and weather and engineering estimates of savings as the primary independent variables, this approach involves estimating a regression model with the following specification:

$$ADC_{it} = \alpha + \beta EE_i + \lambda_1 HDD_{it} + \lambda_2 CDD_{it} + \varepsilon_{it}$$

Where,

ADC_{it} is the average daily consumption during the pre- and post-program periods

EE_i is the initial engineering estimate of savings; the associated beta represents the savings realization rate

HDD_{it}, is the average daily heating degree days based on facility location

CDD_{it} is the average daily cooling degree days based on facility location

ε_{it} is the error term

Table 4-5 shows an example of running SAE for the same utility in the Midwest mentioned above for the same program. Similar to the CSA model, the air sealing measures had a realization rate of 0 percent. In other words, of the initial estimate of about 456 kWh annual savings, 0 percent was realized. Of the 493 kWh expected from insulation, 57.4 percent were realized. Overall, 840 kWh were expected and 418 kWh realized, on average.

Table 4-5
SAE Model Summary

Measure / Common Area	Average Tracking System Savings per Unit	Realization Rate	Proportion of Units with Measure Installed	Net Savings per Unit From SAE Model
Air sealing	456	0	27.8%	0
Insulation	493	0.574	19.4%	55
Lighting	405	0.808	99.5%	325
Water	585	0.170	33.9%	33
Other	45	0	35.5%	-18
Common	1343	0.984	1.49%	20
Weighted Average	840	49%	NA	418

Upstream Programs

Many utilities are resorting to upstream programs targeted at equipment installers, retailers, distributors, and manufacturers to achieve energy savings and transform their own markets. This type of program poses a new challenge to evaluators. Traditional energy efficiency programs have involved rebates given to specific participants, each with a specific account number, billing history, physical address, and other unique identifiers.

The ultimate goal of upstream programs is to achieve energy savings through transforming the market of energy-consuming equipment that customers purchase.

Progress toward this end can be segmented into short-term, intermediate-term, and long-term goals. Short-term goals include increasing market awareness of energy efficient products and reducing energy use. Intermediate-term goals include increased consumer awareness of the benefits of energy-efficient products and increased promotion, which would lead to the long-term goals of sustained demand and increased market share for these products.

Estimating the impacts due to market transformation programs is an inherently difficult task, particularly when program do not offer direct incentives to end-use customers. In fact, such programs may be “invisible” to the end-use customer.

To measure the program’s effect on the market, the evaluation must first establish a baseline. As shown in Figure 4-4, this baseline represents the market share of the measures in absence of the program, and includes any naturally occurring adoption. From this baseline, the impact of the program can be categorized either as directly attributable to the program or as market effects.

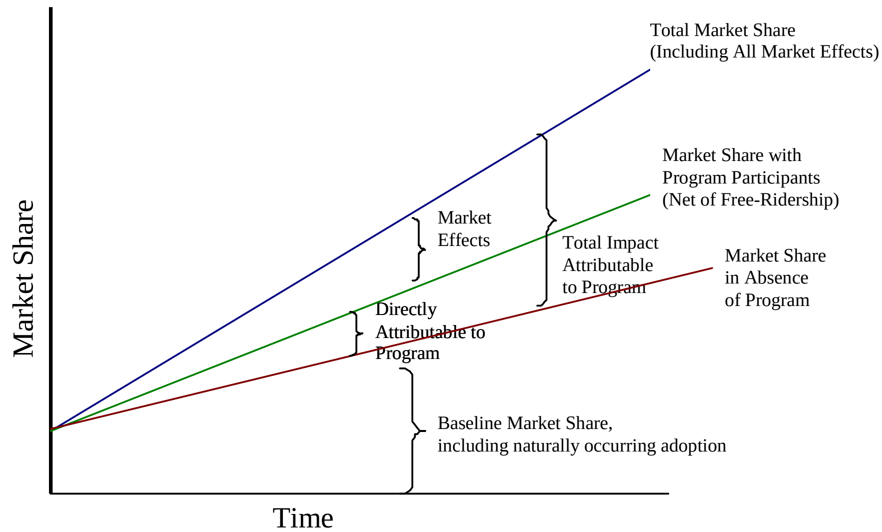

Figure 4-4
Upstream Results

Based on this logic, a number of researchable issues associated with program outputs can be formulated. Table 4-6 provides a brief overview the activities that can be employed to evaluate these programs.

Table 4-6
Upstream Lighting Evaluation

Action	Details
Consumer Surveys	Assesses market share of specific products, as well as other progress indicators and issues related to program delivery
Participating and Nonparticipating Retailer Surveys	Investigates changes in market share, net-to-gross ratio, and program delivery and satisfaction
Shelf Stocking study	Measures product availability and changes in market share over time
Stakeholder Interviews	Provides insight into program design and delivery
Analysis of DOE/EPA data	Analysis of DOE and EPA retailer partner sales data to calibrate market share findings and determine program attribution
Analysis of pricing data	Understand incremental cost of energy efficient products

Note on Line Losses

As electricity is transmitted from power plants and distributed to the final end-user, some power is dissipated through I^2R losses, or "line loss". This amount is usually on the order of 8% to 10%. Therefore, a kWh saved at the end-user level induces a reduction of more than 1 kWh at the generation level. The equation used is:

$$kWh\,Saved\ at\ Plant = \frac{kWh\,Saved\ at\ End - Use}{(1 - lineloss)}$$

When comparing the resource equivalency of energy efficiency measures to supply-side alternatives, this conversion should be applied to ensure the distinction between generated- and delivered- electricity. This is particularly important for calculations of program cost-effectiveness measured in terms of avoided cost of generation.

5
NET-TO-GROSS

Net Savings

Programs may induce the purchase of such equipment as CFLs, energy-efficient refrigerators, or variable-speed motor drives. Program tracking data would account for the numbers of efficient devices sold, installed, or given-away as a result of the program, and "gross savings" would be calculated based on this accounting using the techniques described in the previous chapter.

However, a CFL removed or not installed in the first place does not save energy. Similarly, an energy-efficient refrigerator estimated to have a useful life of 10 years that breaks down or is otherwise replaced after 4 years will not deliver projected savings in future years. Likewise, a variable-speed drive applied to a constant load motor application does not save energy.

Relying on estimates of gross savings, based on tracked metrics, to make system decisions is problematic because actual delivered savings, referred to as "net savings" are subject to variances in application, usage, and behavior in the field.

Components of Net Savings

Net savings are derived by adjusting gross energy savings estimates to account for a variety of circumstances that can substantially change the realized savings. These include:

- Installation rates
- Failure rates
- Baseline assumptions
- Leakage

This adjustment of gross energy savings estimates for the above-mentioned factors produces an adjusted gross energy savings estimate, which is then typically corrected to account for the impacts of:

- Free riders
- Spillover
- Rebound or take-back effects

Installation Rates

Often, measures are purchased and rebated but never installed or not installed right away. For example, a customer may purchase a 10-pack of CFLs but only install a couple. The rest sit on a shelf until a light bulb burns out. These CFLs do not produce savings sitting on a shelf. This factor is also referred to as the "storage" issue. The way to establish the extent of storage is through customer surveys and site visits.

Failure or Removal Rates

Quantifying failure rates—or how often energy-efficient devices or measures are removed—can be expensive to determine. Such studies need to occur over a long period to allow time to elapse before devices or measures actually fail or are removed. Customer surveys and site visits may provide some reliable estimates if sample sizes are large enough.

Baseline Assumptions

Baseline assumptions are very critical for estimating program impacts. The previous chapter discusses baseline issues at length. Baseline describes what would have happened absent the program intervention. As the program is implemented, it is crucial that the appropriate data are collected. For example, in an air conditioning program, the program implementer must record in detail the equipment that is being replaced, including working condition, nameplate information, and, if possible, age. In most cases, the existing equipment represents the baseline. However, if the equipment were at the end of its useful life, the appropriate baseline is what code would have required.

Leakage

An example of leakage is a CFL purchased in one utility service territory that benefits from subsidized pricing that is then transported outside the territory to impact energy savings elsewhere. Data captured at the point of sales would suggest a greater local impact, but "leakage" diverts these measures to another location. While the impact is negative from the perspective of the region affected by the measure, the impact is positive in the region the measures are finally installed. Leakage is part of a general class of effects referred to as "secondary effects," and can be addressed early in a program's implementation if retailers are required to ask for zip codes when selling an item that has been "bought down" though an upstream program or is receiving an instant rebate.

Free Riders

"Free riders" are program participants who would have installed the measures on their own without the program. Since they were likely to make the desired change without inducement, the argument suggests the program was irrelevant to new behavior and their action cannot be attributed to the program being evaluated. Different levels of free ridership (such as deferred free riders and partial free riders) introduce further complexity into this key factor differentiating gross savings from net savings.

Various methods are available to account for free riders. One of the better approaches uses quasi-experimental design (see Chapter 4). Most evaluations use self-reports through surveys. Direct elicitation in the form "would you have installed x without the program incentive" tend to have exaggerated "yes" responses. The customer surveyed is likely to provide the answer they believe the surveyor is seeking. In this case, the question is equivalent to asking: "Would you have done the right thing on your own?"

More sophisticated survey designs ask the question in a number of different ways to check for consistent responses.

Spill Over

"Spillover" refers to actions taken outside the program that are directly attributable to program participation. Several types of spillover have been defined, but the most common are:

- Participant spillover, which are attributable actions taken by a program participant, typically at the same site (but which may be taken by the participant at another site); and
- Non-participant spillover, which refers to actions taken by non-participants due to the program, which may result from the increased availability of efficient products, the training of participating trade allies, etc.

Examples of spillover include

- Adoption of program measures without an incentive.
- Consumers act upon the programs' influence because of changes in the availability of energy-consuming equipment in the marketplace.
- Changes brought about by more efficient practices employed by architects and engineers, which ultimately force consumer behavior into desired patterns.
- Changes in the behavior of non-participants resulting from utilities' direct marketing or changes in stocking practices.

The spillover effect of energy efficiency programs serves as an additional impact that can be added to the program's valid results, in contrast to the impact of free riders, which reduce net savings attributable to the program.

Non-participants who adopt the desired behavior are often referred to as "free drivers"; they contribute to the net savings attributable to the program, in contrast to the reduction in net savings attributable to "free riders."

Rebound

Rebound is a term used to describe changes in consumer behavior resulting from the installation of energy efficiency measures that diminish expected energy savings associated with the original installation. An example of such behavior is the more frequent use of air conditioning equipment occasioned by the reduction in cost resulting from the installation of a more energy-efficient model. Increasing the warmth of a home due to the installation of more insulation is another common occurrence. This is also referred to as “snapback” or “take-back” and is material in measuring net savings for residential programs.

Approaches to Measurement of Free Riders, Spillover, and Rebound

As the effects of free riders, and the impact of spillover and the rebound effect are hard to determine, a variety of approaches are used to estimate these effects. The best approach to estimating net savings is the use of properly constructed quasi-experimental design. When this is not possible due to difficulty of identifying similar non-participants, other options are available. The various approaches to calculating so-called net-to-gross conversion are listed below.

Self Reporting

The most common approach to estimating impacts of free ridership, spillover, and the rebound effect is through direct elicitation from participants. Most often, however, survey respondents may give answers they think are expected or “appropriate” instead of what actually happened. Hence, the survey design is critical to producing reliable estimates.

Enhanced Self-Reporting Surveys

This approach involves conducting surveys, but augmenting this information by also analyzing external documents, program tracking data, and conducting interviews to validate survey findings.

Qualitative Choice

Some of the above factors can be estimated indirectly through “qualitative choice models.” These models are fairly complex and sensitive to inputs and model construction, and only fairly experienced evaluators should build these models. Though similar to the regression models described above, they mainly differ in that the dependent variable is discrete (0, 1), indicating participation or non-participation in a program offering. The explanatory variables include characteristics of the respondents as well as their attitude toward energy efficiency and other environmental factors. The explanatory variables may also include the financial incentives offered by the program. The models then estimate the probability a single household would have participated in the program at different incentive levels. When the incentive level is set to $0, the estimated probability provides an estimate of free ridership.

Stipulated Net-to-Gross Ratios

Sometimes, the determination of how net savings are extracted from gross program results can often fall under the purview of the regulatory body, and thus be defined at either the utility or regulatory level. In such cases, typically those determining the method of calculation use historical ratios and apply them to the gross savings estimates without recourse to validation by survey, interview, or statistical means. A number of factors, such as size of target market or technology maturity, will influence which of these approaches the evaluators want to use.

Net-to-Gross (NTG) ratios may be specified for a particular evaluation cycle, but it is more common for regulatory bodies to stipulate the methodology to be used to calculate them than to stipulate the actual ratios. For example, the California Public Utilities Commission (CPUC) developed a battery of survey questions to be used to collect the information that will then be used to calculate NTG for the various programs and technologies covered by the 2006-2008 program evaluations in California. The Energy Trust of Oregon is also developing methodology to determine NTG ratios for its programs.

NTG usually is calculated at the technology level, but in some cases, such as that of a customized industrial program, may be calculated at the program level. Every utility should be familiar with the approaches taken in the jurisdiction(s) it serves.

Size and Scale of Target Market

Both survey approaches can be used to determine the difference between gross and net savings of any target group, regardless of size. However, to use quasi-experimental design, it is necessary to have a significant target market size and statistical pool to draw from. Finally, the use of stipulated net-to-gross values relies on sufficient historical data to ensure stipulated values are truly representative.

Example

A residential lighting program offers an 18 watt CFL to replace a 100 watt incandescent light bulb. The rebate level is $5. The program replaces 1,000 bulbs. Change in watts is equal to 82. Hours of use estimated at 1,200 hours annually. Estimated savings can be obtained as follows (division by 1,000 converts watts to kilowatts):

- $\text{kWh saved} = \dfrac{82\ \text{W} * 1{,}200\ \text{hrs}}{1{,}000} = 98.4\ \text{kWh}$
- $98.4\ \text{kWh} * 1{,}000\ \text{bulbs} = 98{,}400\ \text{kWh}$

However, customer surveys reveal the following facts about the 1,000 CFLs introduced into the market:

- 12% were put in storage
- 3% failed or were removed
- 24% replaced 75W (Δ Watt = 57)
- 12% replaced 60W (Δ Watt = 42)

- 15% were installed in closets, where the annual useful life is actually closer to 300 hours
- 75 people purchased the CFLs but did not apply for rebate
- 50 people indicated that they would have purchased the CFLs without the rebate, but took the rebate anyway

Adjusted Gross Savings

Estimating the gross energy savings associated with replacing an incandescent bulb with a CFL is straightforward. As shown in the formula below, gross energy savings are a function of the difference in wattage and annual hours of use. The planning scenario of a 100 W incandescent being replaced by an 18 W CFL operated 1,200 hours a year yields a pre-CFL energy savings of 98.4 kWh.

$$\text{Gross Savings} = \frac{\text{Watt}_{pre} - \text{Watt}_{post} * \text{Annual Hours of Use}}{1{,}000} = \frac{(100 - 18) * 1{,}200}{1{,}000} = 98.4\ kWh$$

Assuming 1,000 CFLs are installed through the program, the total gross energy savings generated by the program is therefore 98,400 kWh.

However, an adjustment to gross savings is necessary to account for deviations between planning assumptions regarding pre-program wattage and hours of use and actual observed pre-program wattages and hours of use. This step is referred to as the baseline adjustment to gross savings.

As mentioned earlier, 24% of the program CFLs replaced 75 Watt incandescent bulbs and 12% replaced 60 Watt incandescent bulbs, not 100 Watt bulbs as assumed during program planning. In addition, 15% of program CFLs were installed in closets with 300 hours of annual use (not the assumed 1,200 hours of use). In each of these cases, the savings in the planning stage were overstated for those specific installations. The data for pre-program wattages are presented in Table 5-1 below.

Table 5-1
Illustrative Distribution of CFL Replacements of Incandescent Bulbs

Baseline Wattage	Percent of Program
60	12%
75	24%
100	64%
Total	100%

To determine the actual average pre-program wattage, a weighted average is then calculated using the data above. For this program, the weighted average pre-program wattage is 89.2 Watts, which is 10.8 Watts lower than the planning estimate of 100 W. A similar calculation can be used to determine the actual average hours of operation. The weighted average of 300 hours (15%) and 1,200 hours (85%) is 1,065 hours.

Replacing the planning estimates with the actual values determined for pre-program wattage and hours of operation and recalculating the initial gross savings calculation yields a baseline adjusted per-CFL value of 75.8 kWh. The total program baseline adjusted gross savings is therefore 75,828 kWh, or 77.1% of the unadjusted planning estimate for total gross program savings.

$$GrossSavings = \frac{(Watt_{pre} - Watt_{post}) * Annual\ Hours\ of\ Use}{1{,}000} = \frac{(89.2 - 18) * 1{,}065}{1{,}000} = 75.8\ kWh / CFL$$

A second adjustment to gross savings is also needed to account for the fact that not all of the CFLs sold through the program were installed (i.e. were stored for later use) or operable (failed prematurely). Of the 1,000 CFLs in the program, 12% were not installed (left in storage) and 3% had failed (burned out or broken). In other words, only 85% of CFL sold through the program are currently installed and generating energy savings.

$$(1 - StorageRate - Failure\ Rate) = (1 - 0.12 - 0.03) = 0.85$$

Applying this installation factor to the baseline adjusted total gross energy savings further reduces total program gross savings estimate to 64,454 kWh (75,828 kWh * 0.85 installation factor). The final total program gross savings estimate represents 65.5% of the original program planning estimate.

Net-to-Gross Savings

Finally, two final adjustments are needed to determine the net energy savings of the program.

Any CFLs that would have been installed regardless of program efforts are considered free-riders. Simply put, program resources were used to generate savings that would have taken place without program intervention. As a result, the savings associated with these CFLs is removed when calculating net program savings. In this case fifty of the program CFLs would have been installed completely independent of the program. This represents a free-ridership rate 5% (50/1,000).

Conversely, 75 additional CFLs not part of the program were installed by participants without rebates as a result of the program's impact on participants' perceptions of CFLs. This represents a spillover rate of 7.5%. Since program resources were not used to generate the savings associated with these CFLs, but were responsible for their installation, their savings can be added when calculating net program savings.

The cumulative impact of free-ridership and spillover is captured by the net-to-gross ratio (NTGR). In this example, the program's NTGR is 102.5% (1.00 + -5% +7.5%). Applying the NTGR to the final baseline adjusted gross savings yields a net program energy savings of 66,065 kWh. This represents 67.1% of the program's original planning estimate.

Table 5-2
Summary of Net-to-Gross Calculations

Calculation	Value
Per CFL Energy Savings (Planning Estimate)	98.4 kWh
Total Gross Program Savings (Planning Estimate)	98,400 kWh
Per CFL Energy Savings (Baseline Adjusted)	75.8 kWh
Total Gross Program Savings (Baseline Adjusted)	75,828 kWh
Program Installation Rate	85%
Total Gross Program Savings (Baseline and Installation Adjusted)	64,454 kWh
Program Free-ridership	5%
Program Spillover	7.5%
Program Net-to-Gross Ratio	102.5%
Total Program Net Energy Savings	66,065 kWh

As mentioned earlier, careful survey design is critical to producing reliable estimates. Survey instruments should be crafted to elicit all of the information needed to make the adjustments from gross to net savings illustrated in the CFL program example in this chapter. In addition, evaluation plans should allow for adequate time for such surveys to be developed, disseminated and processed. To ensure a sufficient survey response rate, utilities may consider online web-based surveys to make the process more convenient for customers, or tie the completion of such surveys to some form of incentive.

6
NON-ENERGY BENEFITS

In the jargon of energy efficiency, "non-energy benefits" refers to the ancillary benefits derived from the savings of energy. Primarily, these arise due to the reduction in power generation from carbon-emitting fuels, avoided greenhouse gas (GHG) emissions, improved health benefits, increased job creation, better energy security, reduced transmission/distribution losses, and water savings.

Environmental Impacts

Energy savings have direct impacts upon the environment. These occur primarily due to reduction of amount of fuel used at the generation level. The environmental impacts occur in a variety of ways. Primarily, they affect the region by reducing the amount of carbon-related fuel that is expended, thus avoiding the emissions associated with the "saved" energy.

Energy efficiency programs cause utilities to "avoid" generation of the power that would have been needed absent the programs. They can also defer transmission and distribution upgrades and capacity expansion. Avoided costs are dependent on the mix of generation employed by a given utility. Base load is most often met with large coal plants that operate nearly continuously. Other plants tend to be "load following" and primarily use natural gas. This is obviously not the case with all utilities. Some states include other fuel types in this mix, such as nuclear for base load.

Externalities are by-products of activities that affect the well-being of people or the environment, where those impacts are not reflected in market prices. Several states have made attempts at including these factors in the energy efficiency cost-effectiveness calculus. Some states have created a conservation "advantage" by increasing all avoided cost annually by a fixed percentage (usually 10 percent). Others have attempted to place a value on carbon reductions in their cost-benefit tests. Not doing anything attributes zero value on environmental benefits.

In estimating environmental impacts of energy efficiency programs, the transmission and distribution line losses need to be accounted for. Thus a kWh saved at the end use level means more than 1 kWh is avoided at the generation level.

The exact amount and mix of reduced GHG emissions depends on when the energy savings take place. At different times of the day, depending on the electric load, different fuels are used to meet customers' demand. Utility models show the type of fuel being "dispatched" at each hour of the day. Most energy efficiency programs reduce energy use on the margin and impact "load following" generation plants. Base load plants are less likely to be impacted by energy efficiency programs.[9]

Economic Impacts

Energy efficiency programs may also have significant impacts on the economy, such as added value and job creation. Employment can increase directly, due to program expenditures and staffing, and indirectly because program participants have additional disposable income as a result of lower energy bills. In the residential sector, this may mean that participants purchase more goods and services while in the non-residential sector participants are able to devote funds that used to be spent on energy to hiring, business infrastructure, etc. These increased expenditures lead to indirect effects through spending "multipliers," which identify how a dollar spent in one market segment affects other segments.

The best approach to estimating these impacts is the use of input/output models. These models use inter-industry multipliers to estimate the direct and indirect employment and economic impacts of an event. For example, increased spending on high-efficiency windows will exert an indirect beneficial impact on the glass manufacturing industry. These multipliers are available for all regions across the United States and can be purchased for nominal fees.

IMPLAN® is one of the most commonly used input/output models. It is a data and software package designed by the Minnesota IMPLAN Group, Inc. IMPLAN® uses classic input-output analysis in combination with regional data and multipliers, and can measure the effect of changes or events upon a defined economic region. It supports user-defined modeling to forecast economic effects resulting from proposed measures in the region.

It is important to note that while an energy efficiency program may have positive net economic effects, individual events in the delivery of the program may have negative effects. Take for example, a home weatherization program which provides free weatherization for low income households. There are four distinct events that affect the economy during the implementation of this program:

- **Taxpayers contribute money to fund the program.** This is a negative influence on the economy because it is taking disposable income from households and businesses that may otherwise have been spent on goods and services.
- **Tax funds are used to fund the program.** This is a positive economic event, as program funds are used hire staff and purchase weatherization materials.

[9] The World Resource Institute "Guidelines for Quantifying GHG Reductions from Grid-Connected Electricity Projects," provides an excellent explanation of this approach.

- **Households experience increased disposable income as a result of lower energy bills.** This is a positive economic event. Over the life of the measure, money that was devoted to energy costs can be redirected to purchase goods and services.
- **Utilities lose revenues due to decreased sales.** This is a negative economic event, since lost revenues will have a negative effect on utilities and may cause an increase in energy rates

For example, the Ohio Home Weatherization Assistance Program was credited with creating 403 new jobs and adding $17.7 million to the state's economy.[10]

Table 6-1
Economic Impact of Weatherization Program in Ohio

	Value Added (2003 $ Million)				Employment
	Personal Income	Property Income	Indirect Business Taxes	Total	
Program Spending	16.8	6.7	1.3	24.7	449
Reduced Household Expenditures	-10.6	-6.9	-1.9	-19.4	-360
Fuel Provider Lost Revenue	-3.8	-3.9	-1.2	-9.0	-79
Increased Household Expenditures	12.0	7.3	2.0	21.4	393
Total	**14.4**	**3.1**	**0.2**	**17.7**	**403**

Payment Impacts

This type of co-benefit is most likely to occur among low income populations. As their homes are weatherized and their energy bills reduced, their ability to make payments increases. This leads to reductions in bad debt, terminations, forced mobility, and collection costs. These benefits can be quite significant to the utility, participants and society at large.

[10] Ohio Office of Energy Efficiency. "Ohio Home Weatherization Assistance Program Impact Evaluation." July 6, 2006. Developed by Quantec.

7
USING TECHNOLOGY TO ENHANCE EVALUATION

One could question the cost-effectiveness of the common scenario where a utility operates a meter to charge a customer, the customer operates computerized controls and energy management systems, and an evaluator arrives after a measure is installed to deploy separate metering equipment. Fortunately, as the cost of measurement technology is decreasing, its power is increasing. As integration between meters and controls improves, some of these inefficiencies can be eliminated or at least, reduced. For direct evaluation work, meters are becoming increasingly more powerful and their costs are dropping.

This chapter is not an exhaustive review of all enabling technology for energy efficiency evaluation, but rather gives some examples of technology you can harness in evaluation, and shows where this same technology can actually help increase the success of programs.

Utility Scale Metering

Advanced Metering Infrastructure (AMI) is being considered in many jurisdictions as a way to modernize the electricity grid and enable reductions in peak demand and electricity consumption. AMI entails replacing traditional electro-mechanical meters that are manually read with solid-state electronic meters that can be automatically read. Sometimes referred to as "smart metering" it allows for remote measuring of energy usage for all customers, enables two-way communication, and can even be tied in to measure gas consumption as well.

These AMI devices are part of an evolving strategy being considered by many utilities to implement a new "smart grid" that can reduce the need to provide additional power to cover peak demand conditions. AMI uses technology to capture and transmit energy use information on an hourly, and sometimes sub-hourly, basis to a central collection point. In contrast, standard meters provide a daily electricity usage total and a cumulative monthly billed amount that is determined by multiplying total usage by a fixed, non-time sensitive price. Advanced metering systems currently account for about 6 percent of all meters in the United States.

Figure 7-1
Installation of an Advanced Meter

The benefits to utilities and consumers from adopting advanced metering technology are relevant to the operation of demand response programs that attempt to shave or shift peak load demand, based on signals from the utility to curtail local energy use as a result of increased demand on the system.

AMI technology can enable utilities to monitor and manage power usage in response to real-time information and relay detailed use data and price signals to customers. Utilities can use the data from advanced systems to more accurately forecast load, reduce the need to purchase additional power, lessen energy supply and demand imbalances, reduce energy waste, and improve system reliability. Utilities can also shift peak energy use, reducing their costs and the need for new power plants and transmission lines.

AMI bridges the arenas of customer billing, demand management and control, and evaluation. By effectively using the interval metering aspects of AMI, evaluation can be enhanced, and in some cases the costs of evaluation metering can be reduced. For example, interval metering can provide data on the effectiveness of demand control at small intervals such as 15 minutes. Where two-way transmission allows for control signals, the technology both implements and evaluates. Real time evaluation enables adjustment of a program in a way that after the fact M&V approaches cannot.

Customer-Side Metering

There are several technologies that allow customers to view displays of their home power usage moment by moment, giving them real time feed back and potentially influencing their behavior.[11] Several utilities and energy efficiency entities have conducted pilot projects using direct feedback technology. For example, National Grid and NSTAR in Massachusetts, Hydro One in Canada, and the Energy Trust of Oregon, all conducted pilots that deployed PowerCost Monitors from Blue Line Innovations.

Figure 7-2
Blue Line Innovations PowerCost Monitor for Customer-Side Metering Display

[11] More information on the effect of electricity use feedback can be obtained from the following EPRI document: *Residential Electricity Use Feedback: A Research Synthesis and Economic Framework*. EPRI, Palo Alto, CA December 2008. 1016844.

The technology to meter electrical use circuit by circuit is relatively simple and available today. In the future utilities could insert circuit level metering into new homes, allowing the customer, the utility, and the evaluator to obtain more detailed data than socket level metering. A technology available in Europe but not yet in North America is the Plogg, a simple plug-through meter that sends data to a nearby PC wirelessly. Evaluators could deploy these types of devices to collect sub-metering data on energy consumption specific to the end-uses targeted by a particular energy efficiency program.

Figure 7-3
Plogg Customer-Side Metering Technology

Customers that deploy energy management systems (EMS) in a sense already have an M&V system built into their building. These systems collect temperatures, pressures, and equipment position. In some cases they also collect energy consumption data. Many EMS systems are currently not programmed or enabled to correctly capture and record data, a feature called trending. Correcting this can provide built-in measurement and verification. So-called "continuous commissioning" services are now being offered in which EMS outputs are continuously compared and anomalies are detected as they develop. Data logged and compared by these systems can be a powerful tool in both extending evaluations power and decreasing its cost. Makers of EMS systems have been improving their interoperability and subscribing to open standards like BACnet to improve how these systems work in buildings. The ability of modern EMS to operate and report over the web means that evaluation can occur continuously. This is not only convenient, it also allows program managers to catch problems in programs and adjust procedures throughout the program, greatly increasing the probability of positive results and decreasing risk.

Increasingly, self monitoring is being built into systems and appliances. For example, Microsoft's Windows 7 operating system tapped to replace Windows Vista is being designed with sophisticated features that allow users to manage the power use of networked computers and track actual power use. The server version of Window 7 is planned to allow users to set the desired power use of a set of servers, a feature that potentially enables enhanced measure implementation, demand reduction, and evaluation.

Residential and commercial HVAC systems are being designed to self diagnose critical parameters like airflow through their system, high static pressure that can indicate a dirty filter, and whether they are correctly charged with refrigerant. At present, HVAC program managers employ verification firms that follow contractors verifying through measurement the quality of installations, and evaluators who follow the verifiers and evaluate the larger program. On board diagnostics may replace one of these M&V levels or at least will decrease the cost of measurement. Another development in the HVAC field is the development of gauges used by contractors that guide them to a correct installation then send the data to a verifier's or utility's server. By collecting data directly from the technology, "gaming' is minimized, data collection is streamlined, and verification is built in. While some additional level of evaluation is still required it is greatly streamlined.

Figure 7-4
HVAC Diagnostic Tool

Grid–Tied Appliances

There is a great potential for peak demand and energy savings from devices that can respond to utility signals to reduce or cease operation of an appliance (e.g. water heater or electric dryer), and allow residents to remotely control energy-consuming appliances via the Internet (e.g. to lower the heating set point temperature when on vacation). Energy consuming devices at customer premises that utilities are able to directly control and monitor (through direct load control programs) for event-driven demand reduction or ancillary services are referred to as "grid-tied" appliances.

The approach to evaluating the demand and energy savings of grid-tied appliances is not substantially different than for standard appliances. There are two key questions. The first is whether the technology has been deployed and implemented effectively by the utility. That is, whether the utility is managing the technology and program(s) cost effectively and maximizing the benefit. Or, in the case of home owners, if and how they are using the web enabled control to save energy.

The second question is whether the appliances are performing as expected after receiving commands. A standard metering approach such as recording the energy consumed by an appliance at 15-minute intervals, and then comparing that data to the commands sent by the utility will verify correct operation of an appliance.

If the homeowner has the prerogative to override utility commands then collecting data on the time and number of overrides will also be important. With some smart appliances on-site metering may be replaced with data reported by the appliances themselves.

Evaluation Metering (M&V)

In recent years, metering technology for evaluation has grown better able to deliver enhanced features in more applications, and has become cheaper and simpler to operate, allowing basic meters to be deployed by less specialized, less expensive staff.

Figure 7-5
3-phase True Power Meter

Just a few years ago, a meter capable of measuring and logging true power cost several thousands of dollars, was too large to fit within an electrical panel and was difficult to use. Now for less than $1,000 an evaluator can purchase a meter that is durable, easy to use, includes sophisticated power quality measurement features, and can fit within electrical panels. Integrated cell phone cards not only allow collection of data, but easy to use interfaces port data to the web for use by evaluators, program managers, and customers. Nearly real time data allows adjustment of measures and evaluation protocols increasing effectiveness and decreasing cost.

Figure 7-6
Web-Tied Meters and Software

In addition to increased capability, meter venders also offer simple, inexpensive meters that are easy for non-engineers to deploy, thereby decreasing the cost of deployment. For example, data loggers that meter temperature, relative humidity, and light levels can be purchased for roughly $100. They can be pre-configured by higher level evaluation staff and installed by basic-level staff that simply affix them to a wall and enable them with a push of a button. These meters, although simple, can also capture sophisticated data including air handler operation. Similar meters are also available that can log changes in position, and electrical field. Other inexpensive, simple meters can log changes in voltage and current. These simple meters offer another cost advantage. They can be removed by the customer and returned in a pre-addressed box, eliminating the cost of a trip to remove the meter.

Figure 7-7
Simple meters for current, magnetic field, lighting, and run time

One of the simplest meters now available is the Kill-a-Watt, a plug load meter selling for less than $30. It can meter Volt-Amps, Watts, and power factor, and can calculate kWh for as long as the meter is plugged in. While it does not collect time series data these simple meters are important in two ways: (1) their simplicity and affordability mean that customers can now meter the power use of their appliances; and (2) evaluation of appliance program can be performed cheaply and simply. Account mangers can help their customers directly meter safely and easily. For spot metering these meters give a quick and easy snapshot for appliance meters. While these meters are not particularly accurate for very low readings, they perform very well for readings above 5W.

Figure 7-8
Kill-a-Watt Meter

In summary new developments in evaluation meters allow collecting data remotely through cell phone and Ethernet technologies over the web, and provide powerful capability at ever decreasing cost. These meters help shift after-the-fact evaluation to real time evaluation allowing ongoing program adjustment. This is the most powerful cost advantage of these metering advances.

Simple, easy-to-deploy meters allow placement meters at more sites increasing statistical power at reasonable cost. These simple meters also shift metering to non-technical staff and even to customers, changing how measures are implemented and tracked.

Smart Grid Technology

Development of Smart Grid technology offers the promise of metering affecting not only EM&V, but also the installation of efficiency measures in the home or in larger sites.

Recent development of "gateway devices" that could serve as a hub for home- or building- area energy networks are offering an alternative to the capabilities of discrete metering installations. These networked devices, such as the next generation of power distribution panels can provide information on electricity consumption at the circuit level. Applying Smart Grid technology to measure and verify consumption at the end-use level can greatly enhance the EM&V of end-use specific programs.

Regulatory progress towards establishing an interoperability framework for Smart Grid devices and systems is proceeding, with the National Institute for Standards and Technology (NIST) expected to publish a report on establishing standards and protocols for Smart Grid devices by the end of 2008. This process will culminate with recommendations on standards that could be adopted by the Federal Electric Regulatory Commission (FERC).

Currently the framework of standards and practices pertaining to emerging Smart Grid technologies includes:

- Building-to-Grid (B2G)
- Industrial-to-Grid (I2G)
- Home-to-Grid (H2G)
- Transmission and Distribution (T&D)

Currently utilities trying to implement Smart Grid technologies face a diverse and conflicting array of regulations affecting these new technologies, which is a serious barrier to early adoption of Smart Grid and Advanced Metering technologies. This patchwork of regulation is a target of the Energy Independence and Security Act of 2007 (EISA) whose mandate is to define a national Smart Grid framework that will lead to accelerated adoption. This will be a particularly important contribution to the deployment of EE programs in the industrial sector due to energy intensity and concentration of these customers.

A recent EPRI report articulated the potential energy savings impact of several applications enabled by a Smart Grid, including enhanced EM&V capability for energy efficiency programs.[12] The following excerpts from that EPRI report are provided here to provide further

[12] *The Green Grid: Energy Savings & Carbon Emissions Reductions Enabled by a Smart Grid*. EPRI, Palo Alto, CA: 2008. 1016905.

explanation of the causal link between a Smart Grid, enhanced EM&V capability, and increased energy savings.

> The advanced metering and communications infrastructure of a Smart Grid will support the measurement, storage, and retrieval of data to readily verify energy savings and demand reductions. This relative ease of M&V, in turn, will give utility planners more confidence to incorporate energy efficiency and demand response into integrated resource plans and facilitate greater utility use and reliance upon demand-side resources. Through its enhanced M&V capabilities, a Smart Grid will support emerging business models that increase the cost-effectiveness of energy efficiency programs for utilities. This, in-turn, has the potential to accelerate deployment of energy efficiency programs beyond current levels.
>
> A Smart Grid with advanced metering infrastructure could provide interval metering used to approximate or simulate sub-metering, i.e., measurement of energy consumption at the end-use level. Consider the example of a utility program offering a rebate for an energy-efficient refrigerator. Today, the energy savings associated with that appliance would likely be deemed or calculated ex ante for evaluation purposes. Typically, such ex ante energy savings assumptions are discounted due to their inherent uncertainty, since they cannot be measured at the end use level. A Smart Grid infrastructure, however, could allow the utility to sub-meter the energy consumption of refrigeration for households that participate in the refrigerator rebate program. By being able to measure and compare refrigeration consumption of program participants ex ante and ex post, the utility can reduce the uncertainty of, and more precisely gauge, the program's energy savings impact. Assuming that the program proves cost-effective in yielding energy savings, the utility would be encouraged to expand the program and pursue it more aggressively, resulting in incremental energy savings attributable to a Smart Grid.
>
> Insofar as more precise and reliable M&V, as enabled by a Smart Grid, reduces the uncertainty of program impact and encourages greater program investment – as previously postulated – its impact can contribute towards bridging the gap between realistic achievable and maximum achievable potential. Moreover, a Smart Grid infrastructure would automate aspects of M&V and thereby reduce M&V costs. Since M&V costs are typically included in a utility's administrative costs to implement energy efficiency programs, this automation capability would render programs more cost-effective. The link between enhanced M&V and greater realization of energy efficiency is reinforced under the assumption that new business models may emerge that allow utilities to recover costs for energy efficiency in a manner competitive with cost recovery for generation resources, thereby providing greater incentives for verifiable energy efficiency.[13]

Conclusion

Advances in metering technology blur the line between implementation, verification and evaluation. Rather than a serial approach of installing a measure, verifying that it is installed correctly, then late evaluating it, meters can be agents of implementation themselves and allow continuous verification and evaluation, thereby cutting overall program costs.

---

[13] *The Green Grid: Energy Savings & Carbon Emissions Reductions Enabled by a Smart Grid*. EPRI, Palo Alto, CA: 2008. 1016905.

8
COST-EFFECTIVENESS

Assessing the cost effectiveness of energy efficiency programs is a critical step in deciding which programs to offer. Five economic tests are widely used to conduct this type of analysis. These tests are often referred to as the "standard practice" tests because they are based on the California Standard Practice Manual.[14] As described below, these tests answer the questions of whether a program is cost-effective from different points of view.

This chapter describes each of the five "standard practice" tests and the inputs for each. Five different tests were developed because energy efficiency programs impact five distinct stakeholder groups differently. Therefore, the perspective of each test is as important as its results in evaluating energy efficiency programs.

Test Perspectives

- **Participant test:** "Are participants better off as a result of the program?"
- **Total resource cost (TRC) test:** "Does the total resource cost go down?"
- **Ratepayer impact (RIM):** "Will rates increase as a result of the program?"
- **Utility cost test:** "Are revenue requirements lower as a result of the program?"
- **Societal test:** "Will societal costs go down as a result of the program?"

Since the dawn of energy efficiency programs, policy makers have argued about the appropriateness of the various tests. After three decades, the issue is still being argued. There is no universal acceptance of one test as being the one upon which policy decisions should be made. If the goal is provision of energy services at the lowest possible cost, then which test will promote the selection of an optimal mix of resource acquisitions? The critical question here is "cost to whom"?

At the extreme ends of the debate are the advocates of the RIM test against the advocates of the TRC and Societal tests. The RIM test examines the impact of the program on rates while the TRC and Societal tests examine whether a unit of energy efficiency is less or more expensive than a unit of operation or addition of a power plant.

---

[14] California Standard Practice Manual: Economic Analysis of Demand Side Programs and Projects, July 2002. http://drrc.lbl.gov/pubs/CA-SPManual-7-02.pdf

The RIM advocates see energy efficiency through the eyes of customers that do not participate in the programs and, as a result, end up paying higher rates.[15] The TRC and Societal test advocates see energy efficiency through a public policy or societal lens.

Other issues that are often debated include choices of:

- Discount rate to use for the various tests;
- Avoided costs; and
- Value of externalities (most importantly, the value of carbon emissions)

While a full discussion of each of these topics is beyond the scope of this guide, we offer the following high level guidance on these issues and discuss the industry standard practice, where applicable.

Discounting

Discounting "brings future values to the present." One dollar today does not equal one dollar tomorrow or a year from now. Even in an inflation free economy, temporal differences in value exist. One dollar can be invested today to produce more than one dollar a year from now.

Discounting brings all future benefits and costs (in dollars) to the present to allow for comparison to initial investment costs. Such comparison is conducted in the form of Net Present Value benefits ("NPV"), the difference between present value of benefits and present value of costs, and Benefit/Cost ratios, ratios of the present value of benefits to the present value of costs. Programs or projects with positive NPV values or B/C costs greater than or equal to 1 are generally regarded as economically favorable endeavors.

This process facilitates comparisons that allow optimal choices between investment options based on financial return. Many economists, however, do not agree that discounting makes sense for public goods investments (e.g., parks, clean air, or conservation), since, in essence, the benefits that accrue to future generations are regarded as less valuable than benefits that accrue today. In fact, many argue that benefits to future generation should have higher value than those accruing in the present. This argument is not entirely based on moral grounds. A pure economic argument is that as resources dwindle and emissions increase, the value of future resources increase and the value of one fewer ton of carbon in future should also increase. This argument, at its extreme, calls for negative discount rate.

---

[15] In the opinion of the RIM advocates, this cross subsidy is a matter of equity and needs to be considered in assessing selection of programs. It should, however, be pointed out that rate structure of utilities is fraught cross class subsidies.

In practice, different tests use different discount rates. Utility cost test and RIM are most likely to use a discount rate based on utility cost of capital. Societal and TRC are likely to use a "societal rate" often based on a 30-year Treasury Bill rate (typically 3% to 5%).

Avoided Cost

Energy efficiency program cause utilities to "avoid" generation of the power that would be needed absent the programs. They also defer transmission and distribution upgrades and capacity expansion. For some utilities, what is "avoided" is wholesale purchased of power.

Avoided costs are a function of a utility's generation mix. Base load is most often met with large coal plants that operate nearly continuously. Other plants tend to be "load following," and are most commonly natural gas fired units. This is not the case with all utilities; some states include other fuel types in this mix, such as nuclear energy. The result is that carbon-based supplies become the supply of choice unless energy efficiency is less expensive.

Value of Externalities

Externalities are by-products of activities that affect the well-being of people or the environment, where those impacts are not reflected in market prices. Several states have made attempts at including these factors in the cost-effectiveness calculus. The societal test can include these benefits which may include a variety of components such as health benefits, income and employment generation, improved payment behavior and reduction in bad debt, and, most importantly, reduction in carbon footprint.

Some states have created a conservation "advantage" by increasing all avoided cost annually by a fixed percentage (10% usually). Others have attempted to more explicitly place a value on carbon reductions for inclusion in cost benefit tests.

As shown in Table 8-1, the inclusion of specific benefits and costs varies by test.

Table 8-1
Benefits and Costs in the Five Cost-Effectiveness Tests (Historical Tests)

Test	Benefits	Costs
Participant Test	Participant bill reductions Incentive payments to participants Tax credits	Program costs paid by participant
Total Resource Cost (TRC) Test	Utility avoided energy costs Utility avoided capacity costs Utility avoided distribution costs Utility avoided transmission costs Tax credits	Program costs paid by utility Program costs paid by participant
Ratepayer Impact (RIM) Test	Utility avoided energy costs Utility avoided capacity costs Utility avoided distribution costs Utility avoided transmission costs Fees paid to utility by participant	Program costs paid by utility Revenue losses
Utility Cost Test	Utility avoided energy costs Utility avoided capacity costs Utility avoided distribution costs Utility avoided transmission costs Fees paid to utility by participant	Program costs paid by utility
Societal Cost Test	Utility avoided energy costs Utility avoided capacity costs Utility avoided distribution costs Utility avoided transmission costs Externality benefits	Program costs paid by utility Program costs paid by participant (investment, operating) Externality costs

Test Results

For each of the tests, the discounted stream of program benefits is compared to the discounted stream of program costs over the life of the program. The program is deemed to be cost effective if:

- The net present value (NPV) of the discounted stream of benefits minus the discounted stream of costs over the life of the program is greater than or equal to zero; or
- The benefit-cost ratio (i.e., the discounted stream of benefits divided by the discounted stream of costs) is greater than or equal to 1.0.

The levelized cost (per kW installed or per kWh) is also sometimes calculated to compare the energy efficiency resource to supply-side resources.[16]

[16] The levelized annual cost per kWh of an energy efficiency program, for example, is the levelized annual discounted payment amount for each year in the life of the program divided by the annual kWh saved.

Although these tests were originally developed in the 1980s, they remain essentially unchanged and are in wide use today.

Issues with the Standard Cost-Effectiveness Tests

The test most frequently used today is the total resource cost (TRC) test. A few utilities continue to use the RIM test while some utilities use a combination of tests. In general, there is no "right" test because each test reflects a different measurement, as described below.

- The participant test measures the benefits and costs of an Energy efficiency program to the participant and does not consider the impact on the utility. This test is an indicator of the attractiveness of the program to the customer.
- The RIM test measures the benefits and costs to the utility and does not consider the impact to the participant. This test evaluates the Energy efficiency program from the point of view of rates. It is also referred to as the "non participant" or "no losers" test.
- The utility cost test measures the net costs to the utility as a result of the energy efficiency program. This test measures the change in revenue requirements. Since total costs to the utility (i.e., revenue requirements) are equal to total revenues (i.e., total energy bills paid), this test also measures the change in total energy bills paid. The only difference between the utility cost test and the RIM test is that the utility cost test does not include revenue gains or losses.
- The TRC test includes all of the costs and benefits to the utility and its ratepayers as a whole. This test measures the change in the average cost of energy services to utility customers. In other words, is it cheaper to meet energy needs by saving a kWh or by supplying the kWh from a power plant. The TRC test is sometimes referred to as the "all ratepayers" test. Dollar amounts that flow between the utility and the ratepayers are ignored since they are simply transfer payments. The TRC test combines all of the benefits from the Utility test and costs from both the participant and Utility tests.
- The Societal test includes all of the costs and benefits to the utility and its ratepayers as a whole including externalities. This test measures the change in the total cost of energy services to a utility customer. The only difference between the societal test and the TRC test is that the TRC test does not include externalities.

Which is the Right Test?

There is no easy answer to this question. Each utility must make its own determination, recognizing the trade-offs involved in any such comparison. The following rules of thumb are widely used in the profession:

1. Energy efficiency programs that pass the RIM test should be implemented because they are cheaper than the supply side alternative, rates will go down, and such programs will benefit all customers.
2. Energy efficiency programs that fail the TRC test should be eliminated because they are more expensive than the supply side alternative.

3. Energy efficiency programs that pass the TRC test but fail the RIM test (i.e., the TRC net benefits are positive but the RIM net benefits are negative) are not so straightforward because non-participants will experience bill increases. However, if all customers participate in an energy efficiency program that passes the TRC test, the average cost of energy services will go down and all customers' cost of energy services will go down (even though rates may increase).[17]
4. All programs should pass the participant test; otherwise the program will not be attractive to customers.

Impact of Performance Incentive Mechanisms on Cost-Effectiveness Tests

Energy-efficiency program cost recovery and lost revenue recovery are the two mechanisms that remove the "disincentives" for utilities to invest in energy efficiency. The cost effectiveness tests assume that the utility will recover its costs and that its rates will be adjusted for changes in revenues so these mechanism do not affect the tests. However, the performance incentive mechanism which provides an "incentive" for utilities to invest in energy efficiency will affect all of the tests except the participant test. This type of incentive will increase revenue requirements and is therefore a cost in the TRC, RIM, utility cost, and societal tests.

As shown in Figure 8-1 developed by ACEEE[18], several states currently have performance incentives for energy efficiency programs. For these states, the cost of the performance incentives should not be ignored.

Figure 8-1
Status of Utility Regulatory Mechanisms for Energy Efficiency Programs in 25 States

---

[17] A. Faruqui and J. Chamberlin, "Principles and Practices of Demand-Side Management." EPRI Final Report, TR-1025556. August 1993.

[18] American Council for an Energy-Efficient Economy (ACEEE). "Aligning Utility Interests with Energy Efficiency Objectives: A Review of Recent Efforts at Decoupling Performance Incentives". October 2006. ACEEE Report Number U061. (Note that the map is limited to the 25 states covered in the report).

A Note on Consumer Surplus

Under current regulatory practice, cost-effectiveness analyses of demand response programs are based on the standard cost-effectiveness tests. Applied to demand response, these tests are inherently biased because they do not account for either the loss in consumer welfare that occurs when consumers reduce peak period usage or the gain in consumer welfare resulting from increased off-peak usage. While the topic of EM&V of demand response programs is addressed in a companion EPRI document[19], it is useful here to briefly discuss the concept of consumer surplus (CS) and its bearing on the inadequacy of the standard practice tests.

CS is defined as the difference between the value consumers derive from consumption and the amount they actually pay for that consumption. The value consumers derive from consumption, or the consumer's willingness to pay, equals the sum of the marginal utilities of the various units consumed. For a normal good, the marginal utility declines with additional units consumed, and is reflected in the familiar downward sloping shape of the demand curve. When the price of a commodity goes up, CS shrinks; likewise, when the price goes down, CS expands.

The standard cost-effectiveness tests use changes in the consumer's bill to measure changes in consumer welfare. These tests were developed this way for two reasons. First, when the tests were developed, there was concern that regulators and other policy makers would not put much faith in estimated price elasticities (which are needed to estimate CS). Second, when the tests were developed, most of the energy efficiency programs were energy conservation or load management programs, not rate programs. Hence, the price of electricity was held constant and changes in the quantity of electricity consumed were determined exogenously, often based on engineering rules of thumb. Typically, electricity usage was reduced in proportion to the efficiency change between the old technology and the new technology.

However, for rate programs, such as a time-of-use (TOU) rate or a critical peak price (CPP), price elasticities are necessary to predict the new electricity consumption that results from the new prices. Specifically, since we know the price change and the original electricity consumption, price elasticities are used to predict the electricity consumption under the new rate. Thus, bill changes can be estimated and we have the information needed to calculate CS. For example, TOU rates raise prices during the peak period and lower them during the off-peak period. Higher prices during the peak period result in lower consumption, and "consumer sacrifice;" lower prices during the off-peak period raise consumption, and result in "consumer gain." It is an empirical question whether the gain is greater than the sacrifice.

[19] *Demand Response Evaluation, Measurement, and Verification: A Synthesis of Evolving Protocols and Practices.* EPRI, Palo Alto, CA: 2008, 1018194.

9
PROCESS EVALUATION

Process evaluations are systematic reviews of program implementation policies and practices directed at improving of the efficiency of program delivery. Issues commonly examined are program administration, promotional practices, program delivery methods, incentive levels, market barriers, and data tracking. In addition, process evaluations also provide utilities greater insight into customer attitudes, perceptions, and customer requirements. Process evaluations should help utilities anticipate these needs and improve customer satisfaction as well as serving to spot both market threats and opportunities.

Fundamentally, process evaluations have three major purposes:

- Provide feedback to managers and implementers
- Document program history
- Identify implementation issues and recommend changes

Process evaluations allow utilities to learn from their own actions by offering a systematic way of reviewing the effectiveness of what they have already done with respect to a specific program over time, or across a portfolio of programs.

For regulators, the value of process evaluations lies in their ability to provide substantive evidence that the energy efficiency programs being implemented are indeed effective or, as the case may be, require changes to increase their impact. More broadly, process evaluations also help inform the entire industry about what works and what doesn’t work, the best ways to implement programs, and ways to mitigate problems that can arise.

Process evaluations typically revolve around the execution of a series of interviews, surveys, and document reviews in order to assess the performance of the energy efficiency program in question. Typically process evaluations of a program’s implementation will involve interviewing the utility staff and contractors, trade allies, and customers. The results of these interviews would inform a set of recommendations regarding changes and improvement to make the program more effective.

The focus of process evaluations can change depending upon when they are conducted. Typically, process evaluations are launched at the outset of a program and then are repeated at regular internals over the life of the program.

The initial process evaluation will likely analyze and assess the following factors:

- Preliminary feedback from customers
- Feedback from relevant trade allies
- Perceptions of the utility staff regarding the performance of the program during the launch phase
- Whether the program theory and the logic model accurately describe the situation and correctly identify the market responses
- Scalability of the program as it moves from pilot to full implementation
- Performance of the program management structure and supporting data reporting processes

During full-scale implementation, regular process evaluations typically focus on the following issues:

- Utility customers' (both participants and non-participants) perceptions of the program
- How well market barriers are being addressed
- How well messages and messaging is reaching the intended population
- Changes in perceptions, especially as regards the quality of service, or incremental progress
- How the program aligns with other programs targeting the same customers
- How increasing market saturation issues may impede the effectiveness of promotional strategies and incentive levels
- How changes in the marketplace may be impacting free ridership and spillover effects

Program Design

Process evaluation seeks to answer fundamental program design questions, including:

- What was the basic logic behind the program and the program theory that was developed to justify the program?
- What were the goals and objectives of the designers of the program, and how did they come to focus on these priorities?
- Are target markets defined and aligned with program goals?

Program Administration

Every program has to have an administrative aspect to its delivery. This "bean-counting" side of the program can also have an enormous effect on its effectiveness. Thus it is important to also examine the program forms, contracts, payment procedures, allocation of resources and coordination with utility procedures. This is also the place where one should consider both the cost and quality of the measure installation, as well as the overall cost control over its tracking mechanism.

Program Implementation and Delivery

Process evaluations should also carefully consider a program's required individual implementation steps, and assess the procedures that govern how these steps are taken. Part of this review should consider how the program is delivered to the actual customer. It should clarify and judge the roles and responsibilities of the utility staff and the implementers. This latter aspect of the evaluation helps to inform the utility staff about the performance of the third-party contractors that they have engaged.

Market Response

Since these programs do not occur in a vacuum, it is also important to track what market shifts have occurred during the time the program has been in play. One needs to consider whether the trade allies' efforts to introduce higher energy efficiency equipment have fared in the interim. Is their current performance indicative of future market reactions? Has the market changed in any significant way? If so, how has the market changed?

Approaches

Process evaluations are generally conducted through document reviews, surveys, site visits, and interviews.

Feedback from large populations, such as customers, trade allies or other market actors typically rely telephone of internet-based surveys. For smaller populations, or populations where in-depth feedback is required, two of the more commonly used methods for investigating process evaluation issues are personal interviews and focus groups. Both methods are effective in obtaining a broad range of information from customer or key players in the program delivery network. In contrast to surveys, interviews allow respondents the opportunity to respond to open-ended questions or topic areas and to explain the reasons for their responses. Typically, interviews are conducted for only a small number of people because they are relatively expensive on a per-respondent basis. Care should be taken in interviewing strategies to select the key people in the program delivery chain and to cover the range of topics of concern specified in the evaluation plan. Focus groups are useful when the evaluator wants to obtain reactions to views by other people and to identify consensus versus outlier opinions.

Document Review

The initial source of information about the program being evaluated is typically the material describing the program itself and the related documents used to set it up. These sources provide valuable insights about prior experience with the energy efficiency program. Potential sources include:

- Program files
- Regulatory documents
- Field staff files
- Database records

- Contractor files

Telephone Survey

Surveys can be an effective means to gather information from a large, representative sample of the targeted market, as long as the questions are well defined and the survey instrument is well designed. The evaluator should also be concerned with a variety of sampling and survey design strategies in applying this data collection method.

Telephone surveys are the most common way to solicit information from a large representative sample of the targeted market. A survey can be conducted using internal resources, or it can done using external contractors. Due to the fact that the interviewer has some degree of control over the process it is also a better way of getting more complete responses, and at a higher rate of response. Thus, many smaller organizations prefer to use this method for gathering information.

The survey conducted over the telephone is considered a "survey" as opposed to an interview because it combines a questionnaire with a majority of closed-ended or fixed-response questions. One of the advantages of a telephone interview is the ability to vary questions in response to the interviewer's direction, and thus it can be most effective for gathering subjective information about programs. However, it should be noted that it is important to remain within the bounds of the close-end question format so that the answers can be statistically calibrated afterwards. Certainly, it is permissible to include one or two open-ended questions designed to elicit whether the respondent has additional concerns to express.

One final note of caution: as household income increases the incidence of unlisted phone numbers increases so that this segment of the population may become underrepresented. Similarly, some households are relying only on their cell phones and no longer have "land-line" phone that are listed.

Internet Survey

Internet surveys have been used more frequently over recent years, because they allow survey respondents to visit a website and complete a survey on-line. To help increase the percentage of completed surveys it is now possible for on-line survey takers to save the half-completed survey and return later to complete it. Typically, the targeted market segment is notified of the survey by means of an email notification, postcard, or letter. Sometimes the announcement process will also include a telephone reminder in addition to the mailed notification. This is a relatively inexpensive way of gathering information. It is not labor intensive, set-up procedures are also easy to accomplish, and once the instrument is set up it takes limited staff to conduct the interviews.

Commercial Internet survey software packages are available, and are becoming more sophisticated and user friendly, but most utilities and evaluation contractors find it more practical to engage a firm that specializes in survey work to conduct the survey, as they can more readily assist with the design of the instrument and the presentation of the data. Most market research firms have the capability of conducting Internet surveys, but the benefit of reduced staff time may be offset by several factors.

Typically, Internet surveys have low response rates. It will require a much bigger sample size to achieve sufficient results to reach the desired confidence and precision.

By the nature of the medium, it may be harder to achieve a representative cross sample of the target population, since low income and elderly respondents may be underrepresented in this segment of the population. Furthermore, the low response rate may make it difficult to overcome the inherent population bias by increasing the sample size.

Ideally, Internet surveys are complemented by a telephone and/or mail survey. A follow-up telephone or mail survey removes the potential for response bias that is characteristic in Internet surveys.

Focus Groups

Focus groups are a particularly useful market research tool. Depending upon the design of the focus group, the effort can be reasonably low cost. Focus groups typically involve gathering people that have been "screened" to reflect the characteristics of the target market and having them engage in a group discussion about a specific topic.

The group might consider, for example, how elderly customers might react to a direct response program that might impact their ability to control heating and cooling in their homes. In this case, possible candidates for the focus group would be screened by age and by whether or not they were aware of the direct response program. Then, the moderator would present to the group a series of questions pertinent to the evaluation. The limitation of this approach is that one can typically only deal with no more than 5 main concepts.

Because of the more intense concentration on a limited menu of ideas this method is often employed as a precursor for other quantitative research methods. The issues revealed by the focus group can help to clarify evaluation objectives, and aid in the design of survey questions for subsequent market research. Focus groups are also well suited to explore barriers to participation, to get "live" feedback to proposed marketing concepts, and to test ideas.

One main drawback to the focus group is that the views expressed therein cannot reliably be quantitatively analyzed because of the open ended nature of the results and the small sample size.

There are three main types of focus groups to consider:

- Formal focus groups
- Informal group interviews
- Telephone focus groups

Formal groups are most effective for market research that will be very visible, is politically sensitive, or involves a difficult topic or particularly sensitive group of participants.

Informal focus groups involve interviewing a group of people at the same time rather than one-on-one. These group discussions are most appropriate for talking with people who know each other, such as building tenants or a group of residents at a senior citizen center. This method is best used if you desire more give-and-take in the discussion between the evaluators and the subjects.

Telephone focus groups are group discussions held via teleconferencing. Participants are recruited in the same way, and the discussion occurs during a pre-scheduled conference call. Phone focus groups are particularly appropriate:

- When participants wish to remain anonymous
- When participants are geographically dispersed
- When formal facilities are not readily available in the area

Representatives of large customers or other small groups of people who may know each other, or individuals in a group who may be less inclined to speak up when involved in a face-to-face situation, should be considered for a telephone focus group.

In-Depth Face-to-Face Process Interviews

Face-to-face interviews are the most effective market research tool for extracting the most high quality information from the subjects. This is due to the combined impact of both verbal and visual stimulation.

It should come as no surprise that in-depth interviews are used most often in performing process evaluations. This type of interview is especially effective when the interview subjects are those individuals most closely associated with the most important aspects of the program, such as the conceptual design, management, and operation of the program. A top candidate for a face-to-face interview would, for example, be the program manager who has a unique perspective on the program. The information gathered from these in-depth personal interviews represent specific points of view and cannot be agglomerated to represent a group of people or segment of the target group. Indeed, these interviews may present disparate and diverse perceptions of the program that cannot be reconciled.

Top candidates for these kinds of one-on-one interviews fall into two classes – program design and delivery staff, and strategic key decision makers. These decision makers may include important program participants or anyone who might affect the program in a universal way, such as an essential equipment vendor.

A minor note of caution is appropriate here. Often these kinds of interviews can elicit nearly identical views on a variety of questions or issues from the interviewees. Nonetheless, it is important to complete all interviews, even if some appear redundant. When a few central issues dominate an evaluation, this repetition can occur. However, it is important to document the full extent of uniformity of findings, as well as to validate the value of each individual's views, however duplicative they may appear to the observer.

Finally, these interviews also serve to record the history of the program. This provides institutional history that can help to inform new staff about the reasons for the program's design and record prior missteps and successes. This may serve to provide valuable direction in the design of future programs or even related initiatives for similar segments of the market.

In-Depth Telephone Process Interviews

Telephone interviews are used in a similar capacity as face-to-face interviews to address small interest groups when there is insufficient budget for in-person interviews or when the subjects are widely scattered and difficult to physically reach. Phone interviews tend to be more exploratory and more discussion-oriented than phone surveys. Unlike the reliance on close-ended questions, the phone interview will mostly rely on open-ended response questions. Typically, a discussion guide is employed to structure the conversation and the call is recorded or the evaluator relies on detailed notes taken during the calls.

Phone interviews are used, as suggested above, to canvas a widely dispersed but relatively small group (sample size less than 50) of strategic individuals. These may be in-depth interviews that last 30-45 minutes and are used to gather specific background information that is required to understand some aspect of the program. In many cases this information need may emerge from prior interviews, requiring a small effort to conduct some subsidiary research on a specific subject.

An important difference between interviews and surveys pertains to the fact that the surveys can be subjected to statistical analysis. Phone interviews are usually too subjective to allow for statistical analysis to be used. Phone surveys are also the tool of choice when statistical sampling is required for selecting respondents from a large population.

Mail Survey

Mail surveys are less frequently used, because they have very low response rates, leading to questions about the validity of results. Additionally, the design issues for mail surveys are complex, requiring a minimum of skip patterns and open-ended responses. However, there are some applications where mail surveys are still useful, including:

- Brief quality of service surveys
- Follow-up surveys of respondents previously surveyed
- Program participant surveys where telephone contacts have not been effective

On-Site Inspections

Where survey responses appear unreliable or where respondents are not able to provide the information, field observations of program data collection devices installations may be appropriate. However, the high cost of such visits may cause the evaluator to limit the number of such visits. To ensure that the actual observations of the installations are representative, and that their costs remain within the evaluation budget limits, it may be helpful to employ stratified sampling techniques.

Data Collection Research Procedures

Survey Design

Good questionnaire and survey design involves the following common sense, but necessary, steps:

- Review of objectives and research questions
- Development of specific topics
- Development of specific questions within those topics
- Logic checks to ensure that the questions map to research objectives
- Development of a draft survey or interview guide

Pre-testing of the survey instrument or interview guide is necessary to ensure the final instruments clarity, and to minimize response bias from question structure and wording.

Survey Administration and Data Entry

When survey administration is contracted to a survey research firm, it is important that the purposes and rationale for the surveys be communicated and the training of survey staff be observed. Most reputable survey firms provide these services, as well as independent reviews of survey design and content. Telephone-based surveys and interviews are usually administered using Computer Assisted Telephone Interview (CATI) software, and responses are directly entered into a database.

Some surveys and interviews - especially those involving managers and key market actors – are more likely to be administered by senior evaluators who are knowledgeable about the specific issues to be discussed. These interviews can be administered using commercially available software packages, but are as likely to use paper and work processing note-taking. Responses must then be transferred to a database for analysis. This extra step increases the chance for data-entry error and must be closely monitored.

Analyze Data

Data analysis for surveys is typically done in commercially available analysis packages such as SPSS or SAS. For small survey samples, simple counts, frequency distributions and measures of association are typically generated. For large survey samples, analysis methods are more available, and may include counts, frequencies, measures of association, measures of central tendency, regression, factor analysis, cluster analysis, segmentation, and conjoint/choice analysis.

For qualitative data, including verbatim responses to open-ended questions, systematic review of responses are typically undertaking manually by senior evaluators. Increasingly, qualitative analysis tools to assist in coding and analyzing responses may be used, depending on the amount and quality of these responses. This approach allows the evaluators to carefully read and code responses while tracking larger themes that emerge across surveys and populations.

10
REFERENCES

American Council for an Energy-Efficient Economy [ACEEE] (2007). *Potential for Energy Efficiency, Demand Response, and Onsite Renewable Energy to Meet Texas's Growing Electricity Needs*. Report No. E073. **Key topics: savings potential; demand response**

American Society of Heating, Refrigerating, and Air-Conditioning Engineers [ASHRAE] (2002). *Guideline 14 on Measurement of Demand and Energy Savings*. **Key topics: verification, energy savings, demand savings**

Baumert, K., T. Herzog, and J. Pershing (2005). *Navigating the Numbers: Greenhouse Gases and International Climate Change Policy*. Washington, D.C. World Resources Institute. **Key topics: climate change, non-energy benefits**

California Public Utilities Commission [CPUC] (2001). *California Standard Practice Manual: Economic Analysis of Demand-Side Programs and Projects*. **Key topics: program cost-benefit analysis**

California Public Utilities Commission [CPUC] (2004). *The California Evaluation Framework.* http://www.calmac.org/publications/California_Evaluation_ Framework_June_2004.pdf **Key topics: impact evaluation, process evaluation, verification**

California Public Utilities Commission [CPUC] (2006). *California Energy Efficiency Evaluation Protocols: Technical, Methodological, and Reporting Requirements for Evaluation Professionals.* http://www.calmac.org/publications/EvaluatorsProtocols_Final_ AdoptedviaRuling_06-19-2006.pdf **Key topics: impact evaluation, process evaluation, verification**

California State Governor's Office (2001). *California Standard Practice Manual: Economic Analysis of Demand-Side Management Programs*. http://www.energy.ca.gov/greenbuilding/documents/ background/07-J_CPUC_STANDARD_PRACTICE_ MANUAL.pdf **Key topics: cost-benefit analysis**

Efficiency Valuation Organization [EVO] (2007). *International Performance Measurement and Verification Protocol*. http://www.evo-world.org **Key topics: metering, verification**

Energy and Environmental Analysis, Inc. (2004). *Final Report: Distributed Generation Operational Reliability and Availability Database.* http://www.eeainc.com/dgchp_reports/FinalReportORNLDGREL.pdf **Key topic: demand response**

Goldstein, D.B. (2007). *Saving Energy, Growing Jobs.* Berkeley: Bay Tree Publishing.

International Council for Local Environmental Initiatives [ICLEI] (2003). *Clean Air and Climate Protection Software*. http://www.cacpsoftware.org/ **Key topic: Non-energy benefits**

International Performance Measurement and Verification Protocol [IPMVP] (2003). *Concepts and Practices for Determining Energy Savings in Renewable Energy Technologies Applications, Volume III*. http://www.evo-world.org Key topics: **verification**

ISO New England [ISO-NE] (2007*). Measurement and Verification of Demand Reduction Value from Demand Resources*. http://www.iso-ne.com/rules_ proceds/isone_mnls/index.html **Key topics: verification, impact assessment**

Nadel, S., A. Shipley, and R.N. Elliott (2004*). The Technical, Economic and Achievable Potential for Energy Efficiency in the U.S.—A Meta-Analysis of Recent Studies*. Washington, DC: American Council for an Energy-Efficient Economy [ACEEE]. **Key topic: energy efficiency potential**

National Action Plan for Energy Efficiency (2006). *The National Action Plan for Energy Efficiency*. http://www.epa.gov/cleanenergy/pdf/napee/napee_report.pdf **Key topics: overview of energy efficiency**

National Action Plan for Energy Efficiency (2007a). *Guide to Resource Planning with Energy Efficiency*. Prepared by Snuller Price et al., Energy and Environmental Economics, Inc. http://www.epa.gov/eeactionplan **Key topics: resource planning; data tracking, benefit-cost analysis**

National Action Plan for Energy Efficiency (2007b). *Guide for Conducting Energy Efficiency Potential Studies.* Prepared by Philip Mosenthal and Jeffrey Loiter, Optional Energy, Inc. http://www.epa.gov/actionplan **Key topics: data tracking; deemed savings estimation, benefit-cost analysis**

New York State Energy Research and Development Authority [NYSERDA] (2003). *Energy Efficiency and Renewable Energy Resource Development Potential in New York State—Final Report, Volume One: Summary Report.* **Key topic: benefit-cost analysis**

New York State Energy Research and Development Authority [NYSERDA] (2006). *Non-Energy Impacts (NEI) Evaluation*. Prepared by Summit Blue Consulting, LLC, and Quantec, LLC. **Key topics: non-energy impacts**

Northeast Energy Efficiency Partnership [NEEP] (2006). *The Need for and Approaches to Developing Common Protocols to Measure, Track, and Report Energy Efficiency Savings in the Northeast*. Prepared by J. Michals and E. Titus. ACEEE Summer Study 2006. **Key topics: net-to-gross factors, impact evaluation, tracking energy savings; deemed savings**

Northeast Energy Efficiency Partnership [NEEP] (2005). *Economically Achievable Energy Efficiency Potential in New England*. Optimal Energy. **Key topics: benefit-cost assessment**

Ozone Transport Commission [OTC] (2002). *OTC Emission Reduction Workbook (Version 2.1)*. Prepared by Geoff Keith, David White, and Bruce Biewald. **Key topic: non-energy benefits**

Schiller Consulting (2007). *Survey of Energy Efficiency Evaluation Measurement and Verification (EM&V) Guidelines and Protocols*. **Key topics: impact evaluation, net-to-gross factors, deemed savings, non-energy benefits**

Southwest Energy Efficiency Project [SWEEP] (2002). *The New Mother Lode: The Potential for More Efficient Electricity Use in the Southwest*. Report for the Hewlett Foundation Energy Series.

State of Wisconsin, Public Service Commission of Wisconsin (2007). *Focus on Energy Evaluation: Semiannual Report. PA Government Services, Inc*. http://www.focusonenergy.com **Key topics: impact evaluation, net-to-gross factors, deemed savings, non-energy benefits**

TecMarket Works (2002). *Non-Energy Benefits Cross-Cutting Report. Year 1 Efforts.* Prepared by J. Riggert, N. Hall, and T. Talerico. **Key topics: non-energy benefits**

TecMarket Works (2003). *Non-Energy Benefits to Implementing Partners from the Wisconsin Focus on Energy Program*. Prepared by N. Hall and J. Roth. **Key topics: non-energy benefits**

TecMarket Works (2005). *A Qualitative Program Effects Status Study*. Prepared by N. Hall. **Key topics: process evaluation**

U.S. Department of Energy [DOE] (2006). *109th Congress, Report on Section 1252 of the Energy Policy Act of 2005.*

U.S. Department of Energy [DOE] (2006). *Final Report on the Clean Energy/Air Quality Integration Initiative for the Mid-Atlantic Region.* http://www.eere.energy.gov/wip/clean_energy_initiative.html

U.S. Department of Energy, Office of Energy Efficiency and Renewable Energy [DOE] (2003). *Program Management Guide: A Reference Manual for Program Management.* http://www1.eere.energy.gov/ba/ prog_mgmt_guide.html

U.S. Department of Energy, Office of Energy Efficiency and Renewable Energy [DOE] (2007). *Impact Evaluation Framework for Technology Deployment Programs*. Prepared by J. Reed, G. Jordan, and E. Vine. http://www.eere.energy.gov/ba/pba/km_portal/ docs/pdf/2007/impact_framework_tech_deploy _2007_main.pdf **Key topics: program logic modeling; impact evaluation**

U.S. Environmental Protection Agency [EPA] (1995). *Conservation Verification Protocol.*

World Resources Institute [WRI] and World Business Council for Sustainable Development [WBCSD] (2005a). *Measurement and Estimation Uncertainty for GHG Emissions.* **Key topic: non-energy benefits**

World Resources Institute [WRI] and World Business Council for Sustainable Development [WBCSD] (2005b). *GHG Protocol for Project Accounting.* **Key topic: non-energy benefits**

Other Resources

Air Conditioning and Refrigeration Center, Mechanical Engineering, University of Illinois. TEL: 217-333-3115, http://acrc.me.uiuc.edu.

American Society of Mechanical Engineers (ASME), New Jersey. TEL: 800-843-2763. http://www.asme.org.

Association of Energy Engineers (AEE), Lilburn, GA. TEL: 404-925-9558, http://www.aeecenter.org.

Boiler Efficiency Institute, Department of Mechanical Engineering, Auburn University, Alabama. TEL: 334/821-3095, http://www.boilerinstitute.com.

Center for Energy and Environmental Studies (CEES), Princeton University, New Jersey. TEL: 609-452-5445, http://www.princeton.edu/~cees.

Edison Electric Institute (EEI). Washington, DC. TEL: 202-508-5000, http://www.eei.org/resources/pubcat.

Energy Information Administration (EIA), Department of Energy, Washington, D.C., TEL: 202-586-8800, http://www.eia.doe.gov

Energy Systems Laboratory, College Station, Texas. TEL: 979-845-9213, http://wwwesl.tamu.edu.

Florida Solar Energy Center, Cape Canaveral, Florida. TEL: (407) 638- 1000, http://www.fsec.ucf.edu.

IESNA Publications, New York, New York. TEL: 212-248-5000, http://www.iesna.org.

Lawrence Berkeley National Laboratory (LBNL), Berkeley CA. TEL: 510- 486-6156, Email: EETDinfo@lbl.gov, http://eetd.lbl.gov.

National Association of Energy Service Companies (NAESCO), Washington, D.C. TEL: 202-822-0950, http://www.naesco.org.

National Renewable Energy Laboratory (NREL), Boulder, Colorado, TEL: (303) 275-3000, http://www.nrel.gov.

National Technical Information Service (NTIS), U.S. Department of Commerce (This is repository for all publications by the Federal labs and contractors), Springfield Virginia. TEL: 703-605-6000, http://www.ntis.gov.

Oak Ridge National Laboratory (ORNL), Oak Ridge, Tennessee, Tel: (865) 574-5206, http://www.ornl.gov/ORNL/BTC.

Pacific Northwest National Laboratory (PNNL), Richland, Washington, Tel: (509) 372-4217, http://www.pnl.gov/buildings/.

A
GLOSSARY

A

Additionality - approach that recognizes that avoided emissions cannot be credited to programs if they would have happened independent of the project activities. For example a wood products company could not claim a carbon reduction credit for the planting of trees. While the planting of trees does absorb carbon emissions, the timber company would likely plant these trees anyway. This activity cannot be counted towards reducing carbon gas emissions since it does not represent an incremental increase in carbon reduction.

Adjustments - modifications to an assumed baseline energy or demand value to account for variables in the reporting period (i.e.- higher air conditioning load would increase demand during periods with higher than usual temperatures).

Allowances - under a cap and trade program, polluters are allowed to emit a certain amount of pollution (their “allowance”). Higher emissions would necessitate purchase of emissions credits, while emissions under the allowance would allow the producer to sell emissions credits.

Advanced Metering Infrastructure (AMI) - also known as “smart metering”, represents the capacity for automated, two-way communications between advanced utility meters and utility IT business systems. AMI facilitates the collection and distribution of metering data to customers, consultants and to the utility itself.

Analysis of covariance (ANCOVA) model - a type of general linear regression model with one outcome variable and one or more factors.

Assessment boundary - the boundary within which all primary and secondary effects associated with a project or program are evaluated.

B

Baseline - energy consumption, demand, emissions, and any other associated conditions that would have occurred if an energy efficiency project or program had not been implemented.

Baseline period - the period of time selected to represent demand, energy usage, or facility operations before an energy efficiency measure is implemented.

C

California Measurement Advisory Council (CALMAC) - CALMAC was created by the parties involved in measurement and evaluation in California as an unofficial forum for discussing and directing post-Protocol MA&E efforts. CALMAC’s mission is to provide a forum for development, implementation, presentation, discussion, and review of market assessment and evaluation (MA&E) studies for energy programs within California.

Cap and trade - an administrative approach to controlling pollution by providing economic incentives for achieving reductions in the emissions of pollutants. Under this approach, a limit (cap) is put on the amount of emissions a nation, state, or region is allowed to produce, and credits are divided among polluters through an auction or distributed based on historical emissions. Credits can then be bought and sold (or traded) in order to meet each individual polluter's emissions generation.

Carbon footprinting - a measure of the impact that human activities have on the environment in terms of the amount of greenhouse gases produced.

Co-benefits - positive impacts of energy efficiency other than the associated energy or demand savings (i.e.- emissions reductions, etc.).

Coincident demand - metered demand of an appliance, circuit, home or other building during the time of peak demand in a utility's system-wide load.

Comparison group - a group of consumers who did not participate in the energy efficiency measure, but who closely resemble the participant group in all other ways.

Conditional Savings Analysis (CSA) - regression analysis that models the change in consumption under different conditions, i.e.- the presence or absence of a particular energy efficiency measure.

Confidence - a measure of how well an evaluation has captured the true impacts of an energy efficiency program or project.

Cost-effectiveness - an indication of the estimated benefits produced by an energy efficiency program compared to the estimated total cost to achieve those benefits.

CRIS (Climate Registry Information System) - is a non-profit organization established in early 2007 to calculate and publicly report greenhouse gas emissions in a common, accurate and transparent manner consistent across industry sectors and borders.

D

Database for Energy -Efficient Resources (DEER) - a California Energy Commission and California Public Utilities Commission (CPUC) sponsored database designed to provide well-documented estimates of energy and peak demand savings values, measure costs, and effective useful life estimates.

Deemed savings - an estimate of gross savings for a single unit of an energy efficiency program developed by using data sources and widely accepted assumptions of measure savings.

Demand - the time rate of energy flow, usually referred to as MW or kW.

Demand Side Management (DSM) program - a method of planning, implementing, and monitoring activities and programs offered by electric utilities that are designed to encourage consumers to modify their level and pattern of electricity usage.

Direct emissions - changes in emissions at the site where a project takes place.

E

Effective useful life - the assumed number of years that energy efficiency measures will last.

End-use metering - a method of measurement that usually requires installing meters pre and post equipment change outs in order to determine the savings associated with an energy efficiency project or program. End-use metering is among the most accurate, and expensive, ways to estimate both energy and demand impacts.

Energy efficiency - the use of less energy to perform the same or improved function.

Energy efficiency measure - installation of equipment, subsystem, with the purpose of reducing energy use or demand to perform the same or improved function.

Energy savings - Energy savings are generally defined as a result of a specific gross kW savings associated with a particular end use over its hours of operation (or full load hours).

Engineering model - engineering calculations used to determine energy savings and/or demand reductions.

Error - a measure of the estimated difference between the observed or calculated value of a quantity and its true value.

Evaluation - determining the effects of an energy efficiency program or project

Ex ante savings estimate - predicted savings used for program planning.

Ex post evaluation estimated savings - energy savings estimates reported after an energy efficiency impact evaluation has been completed.

F

Free driver - a non-participant who implemented a particular energy efficiency measure as a result of an energy efficiency program.

Free rider - a program participant who would have implemented the program measure in the absence of the program.

G

GHG - Greenhouse Gas - a group of gaseous constituents of the earth’s atmosphere which have been identified as absorbing and emitting radiation at specific wavelengths within the spectrum of thermal infrared radiation emitted by the Earth’s surface, the atmosphere itself, and by clouds.

GHG Inventory - the process of summing all of the greenhouse gases produced by the operations and activities of a particular group.

GRI (Global Reporting Initiative) - a large multi-stakeholder network of thousands of experts, in dozens of countries worldwide, who participate in GRI's working groups and governance bodies towards a society where reporting on economic, environmental, and social performance by all organizations is as routine and comparable as financial reporting.

Gross savings - change in energy consumption or demand directly attributable to an energy efficiency program.

I

Impact evaluation - examination of the direct changes that occur after implementing an energy efficiency or demand reduction program.

Independent variables - factors influencing energy use and demand that cannot be controlled.

Indirect emissions - changes in emissions that occur at the emissions source.

Interactive factors - changes in energy use or demand occurring beyond the measurement boundary of measurement and verification analysis.

IPCC - Intergovernmental Panel on Climate Change. Established to provide decision-makers and others interested in climate change with an objective source of information about climate change. The IPCC does not conduct any research nor does it monitor climate related data or parameters. Its role is to assess on a comprehensive, objective, open and transparent basis the latest scientific, technical and socio-economic literature produced worldwide relevant to the understanding of the risk of human-induced climate change, its observed and projected impacts and options for adaptation and mitigation.

ISO 14000 - a system of environmental management standards and best practices developed by the International Organization for Standardization.

K

Kyoto Protocol - The Kyoto Protocol is an international agreement linked to the United Nations Framework Convention on Climate Change. The major feature of the Kyoto Protocol is that it sets binding targets for 37 industrialized countries and the European community for reducing greenhouse gas (GHG) emissions. The Kyoto Protocol was adopted in Kyoto, Japan, on 11 December 1997 and entered into force on 16 February 2005. 183 Parties of the Convention have ratified its Protocol to date.

L

Load profile - a graph of the variation of electric load over time (usually presented with MW on the y-axis and time on the x-axis).

Load shape - (see Load profile)

M

Market assessment and evaluation (MA&E) - See Market effect evaluation.

Market effect evaluation - evaluation of the change is structure or function of a market or the behavior of a group of consumers, based on an energy efficiency or demand reduction program.

Market transformation - a lasting change in the structure or function of a market, or the behavior of a group of consumers, that persists after an energy efficiency program is no longer offered.

Measurement - procedure for assigning a number to an observed object, event, program, or project.

Measurement and verification (M&V) - data collection, monitoring and analysis associated with the benefits and goals of an energy efficiency project.

Measurement boundary - the physical area that comprises an energy efficiency impact evaluation.

Metering - direct measurement of the energy used by an appliance, circuit, household, or building.

Monitoring - gathering of relevant energy consumption data.

N

Net savings - refers to the total change in load that remains attributable to an energy efficiency program, after accounting for the effects of free drivers, free riders, energy efficiency standards, changes in the level of energy service, and other causes of change in energy consumption or demand.

Net-to-gross ratio (NTGR) - a factor representing the difference between net program savings divided by gross program savings.

Non-energy benefits - benefits occurring due to an energy efficiency program or project which might include environmental impacts from reduced emissions, economic impacts from job creation and consumer savings, and payment impacts from reduced energy bills and increased ability of low-income consumers to pay those bills.

Non-participant- any consumer or customer who was eligible but did not participate in the energy efficiency project or program being evaluated.

Normalized annual consumption (NAC) analysis - annual energy consumption that has been adjusted to reflect usage levels under typical weather conditions, as opposed to consumption under the actual conditions experienced during that specific annual period.

P

Participant - any consumer or customer who was eligible and participated in the energy efficiency project or program being evaluated.

Payment impacts - co-benefit most likely to occur among low-income populations. As their homes are weatherized and their energy bills reduced, their ability to make payments increases. This leads to reduction in bad debt, termination, forced mobility, and collection costs. These benefits can be quite significant to the utility, participants and society at large.

Peak demand - the maximum amount of metered demand during a specified period.

Persistence - refers to the changes in the savings level of a measure over time. The performance of an energy saving measure may degrade over life due to many factors including consumer behavior, removal, burn out, and poor performance due to aging of equipment.

Portfolio - a collection of similar programs addressing the same market, technology or mechanisms conducted by an organization.

Potential studies - study conducted to assess market baseline and the potential for future savings achievable from energy efficiency programs.

Precision - the accuracy and repeatability of measurement.

Primary effects - intended effects of an energy efficiency or demand reduction program.

Program - a group of projects with similar scope and characteristics, installed in similar applications.

Project - an activity or course of action seeking to achieve improvements in energy efficiency. A subset of a Program, which encompasses many projects using the same measures.

R

Rebound effect - a change in energy usage behavior that results in increased energy consumption, even though an energy efficient measure has been implemented (i.e.- a customer who installs a CFL light bulb to save energy, but subsequently leaves the light turned on much longer than previously used).

Regression analysis - a statistical measure of the relationship between a dependent variable and specified independent variables.

Reliability - the consistency of a measure’s performance.

Reporting period - the time after implementation of an energy efficiency program in which savings are determined.

Resource acquisition program - energy efficiency and demand reduction programs that seek to achieve direct savings.

Rigor - level of expected precision and confidence.

Regional Greenhouse Gas Initiative (RGGI) - the first mandatory, market-based effort in the United States to reduce greenhouse gas emissions. Ten Northeastern and Mid-Atlantic states (Connecticut, Delaware, Maine, Maryland, Massachusetts, New Hampshire, New Jersey, New York, Rhode Island, and Vermont) will cap and then reduce CO2 emissions from the power sector 10% by 2018. States will sell emission allowances through auctions and invest proceeds in consumer benefits: energy efficiency, renewable energy, and other clean energy technologies. RGGI will spur innovation in the clean energy economy and create green jobs in each state.

S

Secondary effects - unintended impacts of an energy efficiency program or project.

Self-reporting - customer or consumer information obtained from direct elicitation of the consumer or customer. Self-reporting results may produce skewed information, as survey respondents may give the answers that they think are expected or "appropriate" instead of what actually happened. That is why the design of the survey is critical to producing reliable estimates.

Snapback effect - (See Rebound effect)

Spillover effect- reductions in energy consumption or demand caused by an energy efficiency program beyond the gross savings of the program participants. Spillover effects can come from participants and non-participants (i.e., a non-participant who chooses to install an energy efficiency measure because of advertising associated with an energy efficiency program; or a participant who chooses to install energy efficiency measures above the scope of an energy efficiency program).

Statistically adjusted engineering (SAE)- statistical analysis that incorporates engineering estimation of savings as a dependent variable.

Stipulated values - (see Deemed savings)

T

Takeback effect - (See Rebound effect)

Transmission losses - energy lost in the transmission of energy from the source to the consumer.

Triple Bottom Line- a method of accounting which seeks to incorporate the social and environmental impact of an organization's activities, in a measurable way, to its economic performance in order to show improvement or to make evaluation more in-depth.

U-V-W-X-Y-Z

Uncertainty- the evaluation process that arises from the difficulty in measuring the absence of energy use.

Verification audit – a review of a firm's GHG inventory to confirm the completeness of the GHG emissions, the reliability of the management systems, and conformity with program rules.

Western Climate Initiative (WCI)- is a collaboration of seven U.S. governors and four Canadian Premiers (Arizona, British Columbia, California, Manitoba, Montana, New Mexico, Ontario, Oregon, Quebec, Utah, Washington) created to identify, evaluate, and implement collective and cooperative ways to reduce greenhouse gases in the region, focusing on a market-based cap-and-trade system.

Export Control Restrictions

Access to and use of EPRI Intellectual Property is granted with the specific understanding and requirement that responsibility for ensuring full compliance with all applicable U.S. and foreign export laws and regulations is being undertaken by you and your company. This includes an obligation to ensure that any individual receiving access hereunder who is not a U.S. citizen or permanent U.S. resident is permitted access under applicable U.S. and foreign export laws and regulations. In the event you are uncertain whether you or your company may lawfully obtain access to this EPRI Intellectual Property, you acknowledge that it is your obligation to consult with your company's legal counsel to determine whether this access is lawful. Although EPRI may make available on a case-by-case basis an informal assessment of the applicable U.S. export classification for specific EPRI Intellectual Property, you and your company acknowledge that this assessment is solely for informational purposes and not for reliance purposes. You and your company acknowledge that it is still the obligation of you and your company to make your own assessment of the applicable U.S. export classification and ensure compliance accordingly. You and your company understand and acknowledge your obligations to make a prompt report to EPRI and the appropriate authorities regarding any access to or use of EPRI Intellectual Property hereunder that may be in violation of applicable U.S. or foreign export laws or regulations.

The Electric Power Research Institute (EPRI)

The Electric Power Research Institute (EPRI), with major locations in Palo Alto, California; Charlotte, North Carolina; and Knoxville, Tennessee, was established in 1973 as an independent, nonprofit center for public interest energy and environmental research. EPRI brings together members, participants, the Institute's scientists and engineers, and other leading experts to work collaboratively on solutions to the challenges of electric power. These solutions span nearly every area of electricity generation, delivery, and use, including health, safety, and environment. EPRI's members represent over 90% of the electricity generated in the United States. International participation represents nearly 15% of EPRI's total research, development, and demonstration program.

Together…Shaping the Future of Electricity

Printed on recycled paper in the United States of America

1016083

Electric Power Research Institute
3420 Hillview Avenue, Palo Alto, California 94304-1338 • PO Box 10412, Palo Alto, California 94303-0813 • USA
800.313.3774 • 650.855.2121 • askepri@epri.com • www.epri.com