Learning from the Past and Predicting the Future: Linking Program Evaluations to Energy Efficiency Planning Studies

Mike Ting and Mike Rufo, Itron Inc., Oakland, CA USA

ABSTRACT

This paper characterizes and assesses the major sources of uncertainty in energy efficiency (EE) potential and planning studies and describes how an important subset of those uncertainties can be reduced by leveraging timely, evaluation-based data to yield up-to-date, observation-based estimates of measure saturations, measure costs and savings, and customer adoption behavior. We then outline a few of the key challenges associated with building more active and direct linkages between program evaluations and EE potential and planning studies, such as timeliness in a dynamic world and balancing evaluation priorities with the needs of planning studies. Finally, this paper offers a roadmap of initiatives that can be pursued in the near-term to better leverage current evaluation activities to improve planning studies in meaningful and important ways.

Introduction

Energy efficiency potential studies are experiencing a new wave of attention in the US and around the world as utilities and policy makers race to establish programs and savings targets sufficient to meet the challenges of a climate-constrained world. Indeed, potential studies are now taking center stage in policy and resource planning activities that go well beyond the scope and objectives of the first wave of utility-sponsored planning studies conducted in the 1990s. Given the increasing importance, scope, and frequency of potential studies in today's world, it is critical to assess the quality of the tools and key data used in potential studies, identify key uncertainties, and implement strategies to reduce these uncertainties on an ongoing basis going forward.

In many ways, program evaluations are mirror images of potential studies. In order to develop ex-post savings estimates from a particular measure, one must develop baseline technology data, estimate eligible customer populations, and estimate program participation rates. Potential studies endeavor to estimate similar quantities but in an ex-ante fashion. Importantly, however, potential studies face the continual challenge of trying to accurately characterize measure-level technology and participation data across the entire spectrum of EE technologies and programs in order to project future participation and savings. As a result, potential studies face a host of uncertainties, particularly with respect to technology markets and customer adoption behaviors that are very dynamic in nature. Program evaluations are perfectly positioned to fill some of these key data gaps and hold the promise of being an important vehicle to reduce key uncertainties in potential studies on a continual and ongoing basis, if explicitly designed to do so.

Research Objectives

Drawing from the extensive experience of the authors in conducting program evaluations and potential studies over the past two decades, this paper will characterize and assess the sources of uncertainty in energy efficiency forecasting studies and identify the subset of those uncertainties that can be best addressed by leveraging timely, evaluation-based data. This paper will then assess and describe the key challenges associated with actively linking program evaluations to forecasting studies. Finally, this paper will provide a roadmap and recommendations for addressing these key challenges and promoting more active and ongoing linkages between program evaluations and EE potential and planning studies.

Brief History and Basics of EE Potential Studies

EE potential studies were popular throughout the utility industry from the late 1980s through the mid-1990s. This period coincided with the advent of what was called least-cost or integrated resource planning. EE potential studies became one of the primary means of characterizing the resource availability and value of energy efficiency within the overall resource planning process. In this section, we provide a brief overview of the definitions and conceptual frameworks typically used in potential studies and describe the main input data and analytic elements common to all potential studies.

Definitions and Conceptual Framework

Like any resource, there are a number of ways in which the EE resource can be estimated and characterized. Definitions of EE potential are in similar to definitions of potential developed for finite fossil fuel resources like coal, oil, and natural gas, where resources are typically characterized along two primary dimensions: the degree of geologic certainty with which resources may be found and the likelihood that extraction of the resource will be economic. Somewhat analogously, EE planning studies have defined several different types of energy efficiency potential. Among the most common types of potential defined are *technical, economic, achievable, program, and naturally-occurring* potential. These potentials are shown conceptually in Figure 1 and described below.

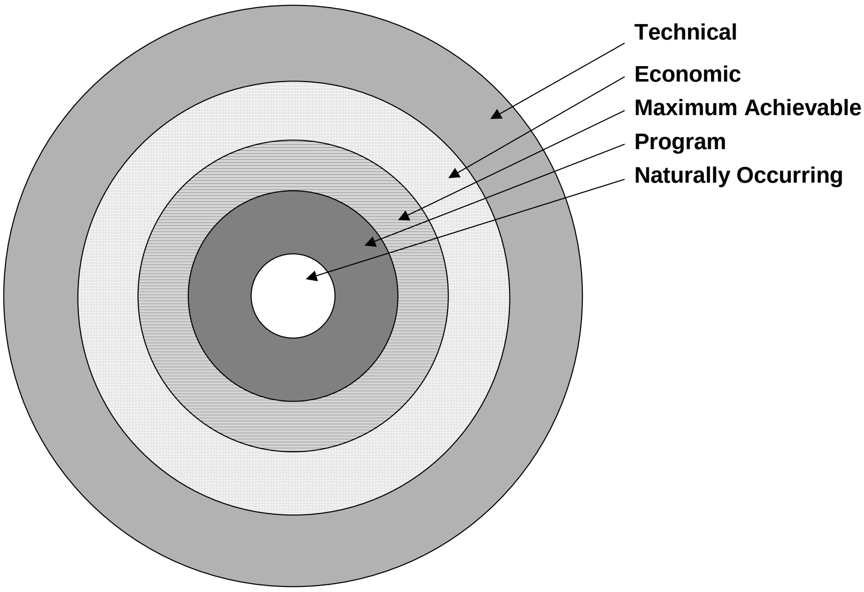

Figure 1. Conceptual relationship among EE potential definitions

Technical potential is often defined as the complete and instantaneous penetration of all EE measures analyzed in applications where they were deemed technically feasible from an engineering perspective. Total technical potential is developed from bottom-up estimates of the technical potential of individual measures as they are applied to discrete market segments (e.g. specific industries or types of residential or commercial buildings).

Economic potential is typically refers to the technical potential of those EE measures that are cost effective when compared to either supply-side alternatives or the price of energy. Economic potential takes into account the fact that many EE measures cost more to purchase initially than do their standard-efficiency counterparts. The incremental costs of each efficiency measure are compared to the savings delivered by the measure to produce estimates of energy savings per unit of additional cost. These estimates of EE resource costs are then compared to estimates of other resources such as building and operating new power plants.

Achievable potential refers to the amount of savings that would occur in response to specific levels of program funding and measure incentive levels over time. Because achievable potential will vary significantly as a function of the specific type and level of measure incentives and program marketing applied, it is usually developed for multiple, specific program funding scenarios (e.g. "business as usual" funding, "increased funding", and "maximum funding") and is thus sometimes referred to as *achievable potential, market potential,* or *program potential*.

Naturally-occurring potential refers to the amount of savings estimated to occur as a result of normal market forces over time, that is, in the absence of any utility or governmental intervention going forward. In this respect, net savings associated with achievable potential are savings that are projected beyond those that would occur naturally in the absence of any market intervention.

Main Input Data and Analytic Elements of EE Potential Studies

Functionally, while potential studies conducted by different investigators will differ in particular aspects of data development and model specification, most comprehensive potential studies share a common set of input data and analytic elements. In this section, we provide a brief review and description of each of these common elements.

Baseline data on end-use energy consumption. Most studies of EE potential start with an analysis of current energy use at a level relevant to proposed program interventions in a given service territory. This analysis involves constructing a bottom-up characterization of energy use at the end-use and technology level in the particular market segments of interest, e.g. existing single-family homes, office buildings, grocery stores, or metal fabrication facilities. The key data necessary to establish the bottom-up modeling baselines required for energy efficiency potential studies are: 1) end-use technology saturations, 2) end-use technology densities, 3) end-use energy intensities, and 4) end-use load shapes. Residential baseline analyses also requires data on the number of households by building type (e.g. single-family detached homes vs. multi-family buildings) in order to scale and calibrate residential end-use estimates to total residential sales and peak demand. Similarly, commercial baseline analyses requires data on commercial floor space by building type (e.g. offices, retail stores, hospitals, or schools) in order to scale and calibrate commercial end-use estimates to total commercial sales and peak demand. Table 1 provides a summary of the key types of baseline data required for potential studies and the common sources of each type of baseline data.

Table 1. Summary of key baseline data required for potential studies

Data Type	Units	Common Sources
Units of consumption	• Number of households or kWh sale (residential) • Square feet of floor space or kWh sales (commercial) • kWh sales (industrial)	• Utility billing data • CIS data • Regulatory commissions
End-use technology saturation	• Share of households with technology installed (residential) • Share of floor space with technology installed (commercial) • Share of load with technology installed (industrial)	• Self-report surveys • On-site surveys • Market tracking studies
End-use technology density	• Cost units per consumption unit (e.g., lamps/home, tons cooling/square foot, motor horsepower/kWh)	• Self-report surveys • On-site surveys
End-use energy intensity	• Annual kWh/household (residential) • Annual kWh/square foot (commercial) • None or kWh/unit of production or kWh/value of shipments (industrial)	• Building simulations • Engineering estimates • End-use metering studies • Utility load research • Econometric studies (e.g. conditional demand analysis)
End-use load shapes	• Distribution of end-use energy consumption across times of the day, days of the week, and season	• Building simulations • End-use metering studies • Utility load research

Since the results of the baseline analysis determine the amount of energy use and peak demand that can ultimately be affected by the set of EE measures being considered, the quality of both the primary data and the data development process associated with estimates of baseline end-use consumption greatly influences the credibility and accuracy of efficiency potential forecasts.

Measure data. Along with baseline data on current energy use, the other key input data required for potential studies are data that describe the EE measures being considered in the analysis. The key measure data required are measure costs, measure savings, measure feasibility, and measure saturation. Measure costs are expressed as either full costs or incremental costs, depending on whether the measure is a retrofit (full cost, including any labor costs associated with installation) or replace-on-burnout measure (incremental first cost, relative to standard efficiency replacement). In many studies, measure costs are also normalized to “cost units” in order to allow reasonable scaling of measure costs across segments that have different technology densities and equipment capacities (e.g. $/ton of cooling capacity). Compared to savings, measure costs have not been empirically well studied throughout the history of the energy efficiency field. Measure savings can be expressed as percentage savings relative to the base technology or in terms of kWh, kW, or therms. Measure saturation is defined as the share of total consumption units (e.g. households or commercial floor space) where a given measure is already installed. Measure feasibility is typically defined as the share of households, commercial floor space, or industrial load where a given measure is technically and practically feasible. Examples of barriers that limit measure feasibility include color requirements that limit the use CFLs as replacements for incandescent lamps and the use of constant-volume heating, ventilation, and air conditioning (HVAC) systems that limit the use of variable frequency drives with fan motors. Together, these two variables serve to avoid gross overestimates of efficiency potential by explicitly taking into account practical and technical barriers to particular measures and limiting the analysis to the share of the market where given efficiency measures have not yet been installed. Table 2 provides a summary of the key measure data required for potential studies and lists the common sources of each type of measure data.

Table 2. Summary of key measure data required for potential studies

Data Type	Units	Common Sources
Measure costs	• $/cost unit (e.g. per lamp, per ton of cooling capacity, per square foot of insulation)	• Measure cost studies • Market tracking studies
Measure savings	• Savings relative to base case technology at equivalent level of service	• Measure impact evaluations (e.g. billing analysis, M&V) • Engineering analysis
Measure saturation	• % of households with measure installed (residential) • % of floor space with measure installed (commercial) • % of load with measure installed (industrial)	• Self-report surveys • On-site surveys • Market tracking studies (including supply-side analyses)
Measure feasibility	• % of eligible households where measure is technically and practically feasible (residential) • % of eligible floor space where measure is technically and practically feasible (commercial) • % of eligible load where measure is technically and practically feasible (industrial)	• Engineering judgment

Interacting baseline estimates of end-use consumption with data on measure savings, measure feasibility, and current measure saturation produces estimates of technical potential. Figure 3 shows an example of how baseline end-use and measure data interact to produce estimates of technical electric energy savings potential in the commercial sector. Figure 2 shows that, as is the case with baseline consumption data, the quality of both the primary data and the data development process associated with the characterization of energy efficiency measures greatly influences the credibility and accuracy of efficiency potential forecasts.

$$\begin{array}{c}\text{Technical}\\\text{Potential}\\\text{(GWh)}\end{array} = \overbrace{\left(\begin{array}{c}\text{Units of}\\\text{Consumption}\\(10e6\ ft^2)\end{array}\right)\left(\begin{array}{c}\text{End-use Tech}\\\text{Saturation}\\(\%)\end{array}\right)\left(\begin{array}{c}\text{Base Tech}\\\text{EUI}\\(kWh/ft^2)\end{array}\right)}^{\textit{Baseline Data}}\overbrace{\left(1-\begin{array}{c}\text{Measure}\\\text{Saturation}\\(\%)\end{array}\right)\left(\begin{array}{c}\text{Measure}\\\text{Feasibility}\\(\%)\end{array}\right)\left(\begin{array}{c}\text{Measure}\\\text{Savings}\\(\%)\end{array}\right)}^{\textit{Measure Data}}$$

Figure 2. Example calculation of the technical potential for electric energy savings in commercial buildings

Economic data. The key economic inputs utilized in potential forecasts are summarized in Table 3 and include retail electricity and natural gas rates, avoided electricity costs, discount rates, and inflation rates. Together with measure costs and savings data, these data represent the key inputs to the cost-benefit calculations used in potential studies to estimate economic potential and model measure adoption.

Table 3. Summary of key economic data required for potential studies

Data Type	Units	Common Sources
Retail energy rates	• $/kWh (electric) • $/therm (gas)	• Utility sales and revenue data • Utility and regulatory forecasts
Avoided generation energy costs	• $/kWh	
Avoided generation capacity costs	• $/kW-yr	
Avoided T&D capacity costs		
Environmental adder	• $/kWh	
Utility discount rate	• %/yr	
Inflation rate		
Consumer discount rate		

Cost-effectiveness calculations. To estimate economic potential, it is necessary to develop a method by which it can be determined that a measure or program is cost effective. Most studies use the total resource cost (TRC) test to assess cost effectiveness, which is a form of societal benefit-cost test that measures the net costs of a demand-side management program as a resource option based on the total costs of the program, including both the participants' and the utility's costs. Other tests that are sometimes used in analyses of program cost-effectiveness include the utility cost test (UTC), the ratepayer impact measure (RIM) test, and participant tests.[1]

Adoption modeling. Estimating technical and economic potentials are necessary steps in the potential forecasting process from which important information can be obtained. However, the end goal of the process is to better understand how much of the efficiency potential can be captured in programs, whether it would be cost-effective to increase program spending, and how program costs can be expected to change in response to measure adoption over time.

Whether as a result of natural market forces or program intervention, the rate at which measures are adopted is typically modeled as a function of several distinct factors, including but not limited to: 1) the availability of the adoption opportunity as a function of capital equipment turnover rates and changes in building stock over time, as well as market availability over time; 2) customer awareness of the efficiency measure; 3) the cost-effectiveness of the efficiency measure from the customer perspective, and 4) market barriers associated with the efficiency measure.

Simulating availability rates requires a stock accounting algorithm that handles capital stock turnover and stock decay over the study period (due to both measure adoption and decay in the building stock), depending on whether the measure is a retrofit, replace-on-burnout, or new construction measure. Current measure awareness levels can be partially derived from market assessment studies and self-report surveys, but in cases where such data do not exist, analyst judgment is often required. Changes in measure awareness going forward can be modeled in several different ways. For example, incremental increases in awareness can be modeled as a constant growth rate or a direct function of administrative and/or marketing budgets over time. It is also possible to differentiate the "effectiveness" of different marketing approaches on customer awareness (e.g. custom C&I programs versus mass market programs) and explicitly account for different rates of information retention and awareness decay over time.

---

[1] For a more detailed overview and discussion of these and other cost-effectiveness tests used for planning and evaluation of efficiency programs, see CPUC 2001.

The final and most critical step in adoption modeling is to estimate the fraction of the eligible, available, and aware market that adopts each efficiency measure in each year of the study period. Compared with all the other analytic elements involved in potential studies, this dimension of potential forecasting remains the area of greatest challenge and least consensus among analysts and modelers. As such, analysts have developed a variety of specifications for the adoption function itself. Despite this diversity, however, most adoption models currently used in potential studies estimate adoption as a function of at least two key factors – the cost-effectiveness of the measure (from the customer perspective) and the market barriers associated with the measure. Where adoption models differ is exactly how these factors are represented and the degree and manner in which they ultimately impact customer adoption (e.g. the sensitivity of customer adoption to changes in measure incentives, marketing, etc).

Sources of Uncertainty in EE Potential Forecasting

There are two principal classes of uncertainty underlying the results of EE potential studies. The first area is uncertainty associated with estimates of the current characteristics of end-use electricity consumption and energy efficiency measure data (hereafter, "current market" uncertainty). The second area concerns estimates of the future potential for energy efficiency, which is affected by the uncertainty in the first area, as well as uncertainty in future energy prices and electric load forecasts, changes in market and energy efficiency measure characteristics over time, and forecasts of customer adoption of measures as a function of program interventions, among other factors (hereafter, "forecast" uncertainty). While there is considerable overlap in the underlying data associated with both types of uncertainty, it is useful to separate these classes of uncertainty since the types of research necessary to reduce these two types of uncertainties are significantly different.

Current Market Uncertainty

Estimating EE potential involves a process of modeling the substitution of existing energy-using equipment and systems with energy-efficient equipment and systems. As such, this process starts with estimates of current equipment characteristics and energy use by end use and market segment, i.e. baseline end-use data. These baseline data typically are provided as inputs to energy efficiency potential studies and are, in the best of cases, developed from up-to-date and statistically accurate studies that involve detailed collection of technology market shares and comprehensive modeling of end-use energy consumption and peak demand. When these baseline data are absent, outdated, or inaccurate, the uncertainty in estimates of current equipment shares and associated consumption and peak demand directly impacts the accuracy of energy efficiency potential estimates since energy efficiency potential can vary significantly by equipment type and market segment.

Current market uncertainty is also associated with uncertainties in measure data. For measure costs and savings, uncertainties exist to varying degrees across individual technologies. In general, new measures (e.g., those that have been on the market for two years or less) have somewhat greater uncertainty in costs and savings than measures that have been on the market for longer periods (e.g., 3 years or more). Dynamic markets for existing energy efficiency measures can also lead to substantial declines in incremental measure costs, making it difficult to maintain up-to-date information. High-efficiency lighting technologies are prime examples of dynamic technology markets.

Measure feasibilities can also represent significant sources of current market uncertainty, since they are generally derived from engineering judgment and experience, rather than comprehensive sets of observed data. However, the uncertainty of feasibility estimates again varies significantly across measures and tends to be lowest among the measures that make the largest contributions to technical, economic, and achievable potential (e.g. CFLs, premium T8 lamps, high-efficiency residential AC and commercial packaged AC). Measures whose current feasibility estimates are most uncertain include evaporative coolers, whole-house fans, tankless water heaters, perimeter dimming, variable frequency drive controls, energy management systems, and commercial cool roofs.

The most significant uncertainties in the measure-level data, however, are associated with the measure saturation data, particularly those derived from self-report customer surveys. While self-report surveys can produce fairly accurate saturation estimates for certain types of measures (e.g. CFLs, ENERGY STAR appliances), self-reported saturations of many measures suffer from self-report bias and high levels of misreporting. This is particularly true with residential building shell measures such as floor and wall insulation where most renters and indeed many homeowners lack the information necessary to give accurate responses.

Measure useful lives are another important area of uncertainty in measure data. Useful life affects the persistence of savings, that is, the number of years over which the savings will occur. The estimated number of years of savings is an important input to the benefit-cost analysis in EE potential studies. Over- or underestimating useful lives will concomitantly over- or underestimate the measure’s benefit-cost ratio. Useful life uncertainty includes both uncertainties in the percent of measures that will be retained by customers within a short period after installation (often referred to as short-term retention) and the length of time the measures will last across all customers on average (usually referred to as average effective useful life). Short-term retention rates are sometimes less than one for measures that either: 1) have a high early failure rate, 2) are immediately disliked by customers (perhaps because they do not believe the service level is equivalent to the less-efficient alternative), or 3) cease to operate because of a major change in the customer’s business (e.g., the business shuts down and no new tenant has arrived to utilize the equipment). Short-term retention rates are typically quite high for most measures. However, some measures have more risk of retention rate reduction compared to other measures. For example, a CFL applied in a poor application or a retro-commissioning controls measure that can be easily overwritten by a building operator that prefers the convenience of the previous control strategy. Average effective useful life is affected by the physical life of the measure (which can be measured and estimated in terms of a survival function) and longer-term changes in customer facilities that may lead to removal of measures before the end of their physical life (again, primarily changes in physical equipment or control strategies associated with tenant changes and renovation cycles).

Forecast Uncertainty

Forecasts of EE potential are directly affected by current market uncertainty. In any forecasting process, one wants to begin with as accurate an assessment of current conditions as possible; errors in estimates of current conditions are otherwise carried forward and exacerbated. However, even with perfect data on current market conditions, forecasts are subject to their own uncertainties by their very nature. The key areas of forecast uncertainty include, but are not limited to, the following: 1) uncertainty in future levels of end-use energy service demand, 2) uncertainty in future cost-effectiveness of measures, 3) uncertainty in future customer adoption preferences and behavior, and 4) uncertainty in interactions between current measure portfolios with new measures, future codes and standards, and other future demand-side management (DSM) programs and initiatives. Each of these four areas of forecast uncertainty is discussed in more detail below.

Embedded in all forecasts of EE potential are assumptions about the future levels (and relative shares) of end-use energy service demand, e.g. the volume and temperature of heated space in homes. These assumptions are usually designed to be internally consistent, at least in aggregate, with those predicted in utility (or government agency) load forecasts. However, since EE potential varies considerably across end uses, building types, and sectors, deviations from the predicted levels of future end-use energy service demand can significantly affect the size, character, and cost-effectiveness of the EE resource.

The cost-effectiveness of EE measures, whether from a societal, utility, and customer perspective, is determined primarily by the cost and savings characteristics of the measures themselves. Changes in these measure characteristics over time, therefore, directly impact the cost-effectiveness outcome for any given measure. In potential forecasts, such changes in cost-effectiveness can impact the calculation of both economic potential (by changing the TRC ratio) and achievable potential (by changing the benefit-cost ratio to the customer and the associated adoption rates). While the probability of significant changes in measure cost and savings over 10 years (the forecast horizons of most potential studies) is small, the same is not true for a host of newer technologies. CFLs are a perfect case in point, having experienced dramatic cost decreases over the past 10 years (Itron, 2006). Going forward, similar cost declines are expected for LED lighting technologies and other high-efficiency “emerging” technologies. The unpredictability of such changes in measure characteristics is therefore a significant source of forecast uncertainty in any EE potential forecasting study.

It should be noted, however, that changes in measure cost and savings characteristics are not the only factors that impact measure cost-effectiveness. Future energy prices carry their own set of uncertainties associated with future fuel costs, generation technology costs, capacity costs (particularly in constrained areas), and the value of environmental externalities. Since energy prices directly influence the cost-effectiveness calculation for efficiency measures, the assumed future trends in energy prices therefore directly impact the results of potential studies, and uncertainties in future energy prices contribute directly to the overall uncertainty of the potential forecast results.

In the absence of changes to the empirical cost-effectiveness of efficiency measures going forward, customer adoption of efficiency measures can also change as a result of changes to program design and delivery, changes in customer awareness, and/or changes in customer preferences. While changes in program design can, in theory, be accounted for directly in potential forecasts, given that the utility can accurately anticipate such changes in its future offerings. However, such changes in program design typically occur as the result of lessons learned from ex-post evaluations and not from ex-ante planning activities. Even if such program design changes could be accurately predicted by program planners, customer adoption preferences are also affected by a host of cultural and behavioral factors that are difficult to predict by nature, such as customer acceptance of new technologies, customer awareness of energy efficiency measures and programs, and public perceptions about climate change, to name a few.

Finally, there are a host of policy and program interactions that also introduce significant forecast uncertainty, particularly from the perspective of utility program planners. For example, any time building energy codes or equipment efficiency standards are strengthened or expanded, the result is often a relative decrease in the achievable EE resource available to be captured by voluntary, incentive-based utility programs going forward over the short term, especially in the absence of an influx of new, cost-effective efficiency measures and technologies. As state and federal governments move towards more and more aggressive trajectories for codes and standards, it is therefore becoming increasingly important to account for these interactions in utility program and resource planning studies. However, it is difficult to accurately predict both the scope and timing of future codes and standards, especially over the longer term. Conversely, EE potential can also interact in positive ways with other programs and initiatives, such as policies to promote advanced metering infrastructure (AMI). Pilot studies of homes provided with real-time information on energy consumption and costs suggest that access to such information can significantly increase energy conservation and the adoption of EE measures (Faruqi, Sergici & Sharif 2009). However, the magnitude of these observed effects varies tremendously across the limited number of studies conducted to date (3-13%), and thus the true impact of AMI initiatives on longer-term EE potential remains highly uncertain.

The Role of Evaluation-based Data in EE Planning Studies

Given our assessment of the key uncertainties in current estimates of technical, economic, and achievable potential, below we present and describe the subset of those uncertainties that can be reduced by leveraging timely, evaluation-based data and the challenges associated with building more active and meaningful linkages between program evaluations to EE potential and planning studies.

Measure saturation estimates

As discussed previously, the most significant uncertainties associated with current estimates of technical and economic potential are due to uncertainties in measure saturation estimates. Self-report surveys often play a central role in program evaluation activities due to their relatively low cost and large sample sizes and can be easily designed to estimate measure saturation levels. Increasingly, self-report surveys conducted for program evaluations are being augmented by on-site verification activities in order to reduce self-report bias and increase the accuracy of evaluation results. For technologies and measures most subject to self-report bias (e.g. wall and ceiling insulation), the results of such on-site verification studies can be leveraged directly to help reduce current market uncertainty in EE potential and planning studies.

Audit programs also represent a largely untapped source of observation-based measure saturation estimates. Many utilities and government agencies conduct home and business energy audits as part of stand-alone audit programs or prerequisites for custom rebate programs, but the data from these audits are often not integrated into larger baseline or measure data development efforts. Moreover, evaluations of audit programs often focus on determining the energy savings impacts of those programs, rather than constructing a systematic, detailed characterization of the customer base. In this sense, if audit programs or evaluations of audit programs can be designed to establish systematic reporting of the presence of efficiency measures and standardized, central databases of audit information, these activities could yield a wealth of on-going, observation-based measure saturation estimates that would greatly reduce the current market uncertainty of EE potential and planning studies. It should be noted, however, that measure saturation estimates derived from audit data are subject to self-selection bias and would likely need to be validated by periodic survey studies.

Measure cost and savings estimates

For measures whose costs and savings are either difficult to accurately quantify by nature (industrial measures, commercial HVAC) or are particularly dynamic (e.g. lighting), uncertainties in their measure costs and savings estimates is also an important source of overall uncertainty in EE planning studies. As part of the application process for program rebates, customers and/or installers are often required to submit unit price information for eligible energy efficiency measures. Such applications are a natural, low cost, and on-going source of up-to-date measure cost data and can be used to update and refine incremental cost estimates, benefit-cost analyses, and other metrics relevant to tracking economic potential, estimating future program participation, and other program planning issues. In particular, applications for custom retrofit programs could serve as an important source of measure cost data for commercial lighting and HVAC measures that are generally procured for customers by contractors. As in the case of audit data, increasing the scope of evaluation activities to include the estimation and reporting of contractor-reported, average measure costs based on program applications could yield a wealth of up-to-date measure cost data, particularly in the commercial and industrial sectors, that would greatly reduce the uncertainty in EE potential and planning studies.

While the vast majority of current efficiency programs and portfolios are based upon individual measures, going forward it is likely that more and more cost-effective opportunities for energy savings will come from integrating building designs with particular packages of technology choices rather than from individual technologies (e.g. passive lighting designs matched with advanced lighting controls). Such integrated approaches are often afforded only limited analysis in EE potential studies and/or analyses based exclusively on building simulation results. As such, potential estimates for integrated measures often suffer from aggregation bias in measure cost data and significant uncertainty in measure savings data due to a severe lack of real-world, evaluation-based impact estimates with which to benchmark energy and peak demand savings derived from building simulation models. In this sense, designing evaluations to explicitly assess the costs and savings of such integrated measures implemented through custom retrofit programs or new construction programs (despite their relatively minor role in most program portfolios) would greatly reduce the uncertainties in current cost and savings estimates and improve the quality and defensibility of downstream potential estimates.

Customer adoption behavior

The adoption dynamics in most EE potential forecasting models are based primarily on the results of a very limited number of customer adoption studies, many of which were conducted in the mid-1990s. The lack of more recent data on customer behavior relative to the adoption of EE measures is one of the most significant sources of forecast uncertainty in current estimates of achievable potential. As such, developing revised data on customer adoption behavior relative to various measure incentive levels at a level that is reflective of the scope and structure programs currently in utility program portfolios represents an important opportunity to reduce a significant source of forecast uncertainty going forward.

Although conjoint and double-bounded choice studies have been the traditional means to quantify customer adoption preferences in the past, an alternative approach to developing revised and relevant data on customer behavior would be to leverage the diversity of current programs being offered across different service territories and regions as large scale “natural experiments”, the results of which could feed into revealed preference analyses. For example, there are a multitude of local government programs in California that offer similar measures but at a variety of incentive levels and with a variety of delivery mechanisms. Importantly, these programs have also had varying levels of success in terms of customer adoption rates. Given enough information on program structures and activities and the customers who have adopted measures through these programs, one could leverage the subtle differences in measure incentives and adoptions that exist to conduct revealed preference analyses in a uniform framework. Such an approach would require program administrators and implementers to actively record and track customer adoption data for their respective programs in a manner which allows the data to be compiled and analyzed in a uniform revealed preferences framework. The primary advantage of this approach compared to traditional conjoint studies is that it would enable customer adoption to be analyzed for many individual measures simultaneously and on an on-going basis. One potential disadvantage in using behavioral data from “natural experiments” is the increased potential for spurious correlation that comes from lack of strict study controls.

Another alternative approach for collecting and developing customer behavior data on an on-going basis is to integrate follow-up procedures for all customers that have participated in an energy audit, either through audit-only programs or as a prerequisite for custom retrofit programs. Integrating such follow-up procedures either into the program design itself or the evaluation of those programs and systematically recording and tracking related measure adoption decisions would provide a wealth of valuable quantitative information on customer adoption preferences that would not only help optimize portfolio and program design over the near term but also vastly improve the accuracy and defensibility of longer-term efficiency adoption forecasts on an on-going basis. As in the case of using audit data to develop measure saturation estimates, however, it should be noted that customer adoption behavior data developed from audit follow-ups is likely subject self-selection bias and would likely need to be periodically validated by more controlled discrete choice studies.

Challenges with Integrating Evaluations and Planning Studies

Building the types of explicitly-designed linkages between program evaluations and EE planning studies described above, however logical they appear, faces a number of distinct challenges. Key challenges include, but are not limited to, the following: 1) timeliness in a dynamic world, 2) whole-market versus utility program perspectives, 3) balancing evaluation priorities and resources with planning study needs. Below we describe each of these key challenges in more detail.

Timeliness in a dynamic world. Program evaluations are typically time-intensive processes. Programs are operated on 2-3 year cycles, and ex-post evaluations can take 6-12 months, sometimes longer, to complete following the end of a program cycle. Final evaluation results, therefore, often reflect the market conditions from 3-4 years previous. As described earlier, an important source of current market uncertainty in EE planning studies is the lack of up-to-date baseline and/or measure data, particularly in the case of technology markets that have proven to be particularly dynamic over short time periods (e.g. lighting). In this sense, one of the key challenges in effectively leveraging evaluation-based data for EE planning studies is ensuring timely availability of evaluation results.

Whole-market vs. utility program perspectives. Evaluations traditionally focus, for obvious reasons, on assessing impacts from measures supported directly through utility programs. In contrast, EE potential and planning studies attempt to characterize the entire suite of measures commercially available in a given market. In this sense, the measure scope (and related data needs) of EE potential studies are by definition larger than for evaluation studies. However, we have argued that program implementers and evaluators are perfectly positioned to collect and process certain type of key data for non-program measures that could prove to be enormously valuable to EE planning studies (e.g. measure saturation data from audit programs and measure adoption data from audit follow-ups/evaluations). Another key challenge, therefore, is strategically re-defining program implementation and evaluation activities and requirements beyond only measures directly supported by utility programs in a manner that does not hinder the effectiveness of program delivery or introduce significant additional implementation or evaluation costs.

Balancing evaluation priorities with planning study needs. The whole-market vs. program-only challenge described above is actually one specific example of a more general set of challenges associated with expanding scope of evaluations to address planning study needs. This more general challenge is related to properly balancing evaluation objectives and priorities (e.g. measuring program impacts and efficacy) with the production of other data and information designed to feed EE planning studies. Exactly how much scope can be added to evaluation activities without jeopardizing the primary evaluation objectives?

Roadmap for Going Forward

Overcoming the challenges associated with actively linking program evaluations with EE planning studies in the ways described above is clearly a tall order and realistically cannot occur overnight. In this sense, the evaluation-planning integrations described above are more reasonably interpreted as longer term goals. Nonetheless, there are a host of initiatives that can be pursued in the near-term to better leverage current evaluation activities to improve planning studies in meaningful and important ways. Below, we present and describe four specific activities meant to provide a short-term road map for such initiatives.

Coordinated phasing of evaluation and planning studies

As described previously, program evaluation results often require 6-12 months, sometimes longer, to complete. Similarly, EE potential and planning studies also require 6-12 months to complete. Given these long project timeframes, it is essential that EE planning studies be planned to start soon after, but not before, the latest round of program evaluation results become available. Since many of the outputs of evaluation studies serve as direct inputs to EE potential and planning studies, when these activities conducted concurrently, planning studies are forced to use older data. These types of “lost opportunities” can be systematically avoided by simply sequencing evaluation activities to feed planning studies in an organized, pre-determined, and timely manner.

Testing experimental evaluation approaches

Clearly, some of the most promising opportunities for leveraging evaluation-based data for use in EE planning studies will require current evaluation activities to expanded and/or redesigned. A reasonable and low-risk first step in pursuing such changes to current evaluation activities is to conduct smaller-scale tests of experimental evaluation approaches designed to incorporate the production of specific types of data needed for EE planning studies in addition to the data needed to estimate program impacts and efficacy. The results of these experimental evaluation approaches can then be used to derive first estimates of the additional time, costs, and operational issues associated with larger-scale expansion/redesign of evaluation activities and assess the quality of the additional planning-related data derived from these new activities and processes.

Evaluating compliance with building codes and appliance standards

In EE potential and planning studies, it is often taken as a given that compliance with codes and standards (or even labeling programs like the U.S. EPA's ENERGY STAR) is at or near 100%. However, the limited number of compliance studies that have been conducted to date tend to suggest otherwise.[2] Given the aggressive outlook for codes and standards in many parts of the world and the importance of interactions between codes and standards and utility programs, evaluating compliance with codes and standards will become increasingly important going forward, not only for utility program planning but also for load forecasting and resource planning. Such compliance evaluation activities need not be integrated with or affect utility program evaluation activities per se. In this sense, evaluation of codes and standards compliance levels represents a completely separate activity from utility program evaluation that need not cannibalize or otherwise create tensions for scarce utility program evaluation resources.

Increasing the use of scenario analysis in EE potential and planning studies

While all of the evaluation-planning linkages described outlined above would serve to significantly reduce the uncertainty associated with current forecasts of technical, economic, and achievable potential, such forecasts will always be, by nature, uncertain. In these types of longer-term planning studies where future relationships between many key variables are largely unknown and unpredictable, scenario analysis is a particularly useful, and under-utilized, tool. Well-designed scenario analyses provide bounded contextual frames for exploring the future from different perspectives that can facilitate organizational learning and generate critical insights into strategic decision making (Ghanadan & Koomey 2005). Most recent potential studies use scenario analysis to explore the potential impacts of increased incentive levels on customer adoption of EE measures by including a "base case" or "business-as-usual" funding scenario, where incentive levels and marketing budgets are set to align with current or baseline utility programs, and an "increased" or "advanced" funding scenario, where incentive levels and marketing budgets are increased significantly beyond business-as-usual levels. However, those scenario analyses are extremely narrow in scope and are typically not framed to examine potential outcomes associated with variations in assumed retail energy rates,

---

[2] See, for example, Khawaja et al. 2007 and GAO 2010.

avoided costs, technology learning, or changes in the level and structure of energy service demand, among other important variables that contribute to forecast uncertainty. Given energy efficiency's central role in both resource procurement and climate change mitigation, it is critical to bound forecasts of EE potential under a wider variety of possible futures, rather than focusing solely on outcomes related to increasing incentives.

References

California Public Utilities Commission (CPUC). 2001. *California Standard Practice Manual: Economic Analysis of Demand-Side Programs and Projects*. San Francisco, CA: California Public Utilities Commission.

Faruqi, A., S. Sergici and A. Sharif. 2009. *The Impact of Information Feedback on Energy Consumption – A Survey of the Experimental Evidence*. San Francisco, CA: The Brattle Group, Inc.

Ghanadan, R., and J. Koomey. 2005. "Using energy scenarios to explore alternative energy pathways in California." *Energy Policy* 33 (2005) 1117–1142.

Khawaja, M.S., L. Lee, M. Levy, and L. Benningfield. 2007. *Statewide Codes and Standards Market Adoption and Noncompliance Rates*. Final Report CPUC Program No. 1134-04. Prepared for Southern California Edison. Portland, OR: Quantec, LLC.

United State Government Accountability Office (GAO). 2010. *Report to the Ranking Member, Committee on Homeland Security and Governmental Affairs, U.S. Senate: ENERGY STAR PROGRAM – Covert Testing Shows the Energy Star Program Certification Process Is Vulnerable to Fraud and Abuse*. GAO-10-470. Washington, DC: United State Government Accountability Office.