Net Savings in Nonresidential New Construction:
Is a Market Based Approach the Answer?

Catherine Chappell, Heschong Mahone Group Inc., Fair Oaks, CA
Douglas Mahone, Heschong Mahone Group Inc., Fair Oaks, CA
Marian Brown, PhD, Southern California Edison Co., Rosemead, CA
Kenneth Keating, PhD, Portland, OR
Lori Megdal, PhD, Megdal & Associates, Acton, MA

Abstract

The goal of this paper is to review the past methodologies for estimating the net impact of the California nonresidential new construction (NRNC) program Savings by Design (SBD) and to provide recommendations for improving the evaluation of the nonresidential new construction market in California.

This paper begins with a brief overview of the Savings By Design program and highlights the key elements that present challenges to the net savings evaluation. The net-to-gross methodology review discussion summarizes the approaches that have been used since the mid-1990s, presents the net-to-gross ratios from those studies and outlines some of the on-going issues for evaluation of this sector. Finally, we discuss the issues from the broader market perspective, and present the argument for measurement of the efficiency of the new building stock, from a market-centered approach, rather than a program participant-centered approach.

Introduction

Evaluators have spent the last 25 years testing methods and paradigms to try to separate out program impacts from other impacts in a participant-centric way. What motivated participants? Who would have done the efficiency anyway? Where did the architects get their ideas from? How important were the cash incentives or the design assistance to participants in overcoming barriers? What other spillover actions do participants add to their project practices, or in their practices outside the program? What effect did the program have on the practices of non-participants? How does spillover work when trained architects and engineers change firms, move, or become advocates among their colleagues? When has a program paid for measures after they have become common practice? As will be discussed in this paper, sophisticated methods have been developed and refined in trying to answer these questions. Nevertheless, there is no perfect solution to understanding and quantifying the energy impacts of peoples' mysterious motivations.

This paper summarizes the strengths and weaknesses of the methods that have been used in past California nonresidential new construction evaluations and suggests what we believe are the most useful, defensible, transparent, and cost-efficient ways to answer these types of questions for the Savings By Design program.

Table 1 presents the net-to-gross (NTG) approach and resulting NTG ratios from prior NRNC studies performed since the mid-1990s. NTG is the ratio of the program-induced program savings divided by the observed gross changes in consumption. These results indicate that there is no correlation between the method and the resulting net-to-gross ratio. The wide divergence in results for different years and different programs is not necessarily an indication of a problem. There may be variations in the composition of the building population and study sample and there will be historical events (such as

the 2001 energy crisis) that influence construction practice and the market actors decision-making. The advantages and disadvantages of these previous net savings methods are discussed in this paper.

Table 1. Historic NRNC Net to Gross Ratios

NRNC Study	Approach	Net to Gross
2002 BEA	Self Report	60%
1999-2001 BEA	Self Report Difference of Differences	59%
1998 SCE	Difference of Differences	62%
Pre-1998 PG&E	Difference of Differences Econometric Choice	47%
1996 PG&E	Difference of Differences Econometric Choice	47%
1996 SCE	Difference of Differences Econometric Choice	62%
1996 SDG&E	Difference of Differences Econometric Choice	69%
1994 PG&E	Econometric	80%
1994 SCE	Econometric	50%
1994 SDG&E	Billing Regression*	
Average		**60%**

* Net to Gross Ratio not calculated

CA NRNC Program Background

Savings By Design is a statewide program for commercial, industrial and agricultural customers that encourages energy-efficient building design and construction. The program is administered by California's four investor-owned utilities under the auspices of the California Public Utilities Commission (CPUC). SBD offers building owners and their design teams a variety of services, including design assistance, owner incentives and design team incentives.

SBD seeks to improve the energy efficiency of nonresidential new construction in the state by:

- Directly encouraging owners to build buildings that are more energy efficient than the Title 24 energy code in California, by providing direct incentives for measures and for integrated building design.
- Increasing knowledge and awareness among design professionals, such as architects and engineers, about energy efficient design and measures.
- Helping to develop and promote design standards including corporate design standards and standard promulgated by programs such as the California High Performance Schools (CHPS) program, the LEED government building program, and San Francisco's Green Building Ordinance. All of these standards reach beyond Title 24 and often overlap with the standards of SBD itself.
- Helping move Title 24 to new levels by demonstrating that more advanced measures and requirements are effective, affordable, and acceptable in the market place.
- As indicated by the program description, SBD uses a multi-faceted approach to promoting energy efficiency in nonresidential new construction. The SBD program works closely with other compatible programs (including LEED and CHPS) in a synergistic relationship in order to increase the visibility and effectiveness of all the programs. While the cohesive approach

utilized by Savings By Design is beneficial to all programs and to the customers they serve, it presents challenges to evaluators attempting to sort out both the direct and indirect impacts of the program from other influences.

The complexity of the NRNC market and the networks that operate in this market are important elements in how decisions are made. Key actors in the market include owners, designers (architects and engineers), builders (contractors and subcontractors) and equipment manufacturers and distributors. These groups are inter-related in the market in a variety of possible relationships.

All of the nuances of the dynamic interactions of these market actors have not been completely incorporated in the survey-based methods to-date in California evaluations. This is a consequence of the policy environment, which expects single program impact estimates and unique points of influence. In fact from a market perspective, it may not be feasible or desirable to incorporate all of these interactions at the project level. Instead it may be worthwhile to conduct a market-level analysis that looks at these issues.

Review of Previous Net Savings Methods

Table 2 presents the list of California impact evaluations on non-residential new construction programs. The table also categorizes the type of free ridership and spillover methodologies upon which each of these studies relied.

There are a few observations that can be gathered from this table alone. All nine use at least one method to estimate free ridership. Five of the nine contain approaches to estimate spillover. Three of the nine included two methods for estimating free ridership.

A second and related observation is that the majority of these evaluations were conducted on efficiency programs run during the 1994-1997, which is known as the Protocol Period. The M&E Protocols were adopted by the California Public Utility Commission to provide rules for impact evaluations which were used to determine the energy savings achievements of programs for which shareholder earnings were awarded. The shareholder earnings potential was quite significant. This along with other concerns of various stakeholders created a strong environment to encourage conservative energy savings estimates. This perspective helped to create a much greater concern with estimation of free ridership (which drives down the savings estimate) than that of spillover (which increases the estimate of savings induced by the program). This is one of the reasons why there are several studies with no estimation of spillover while all studies have one or more methods devoted to estimating free ridership. Both are components of a NTG ratio.

Table 2. Approaches Used to Estimate Free Ridership and Spillover

Evaluation Study*	Free Ridership Estimation Approach				Spillover	
	Self-Report	Difference of Differences	Econo-metric Choice	Billing Regression	Self-Report	Econo-metric Choice
2002 BEA	X				X	
1999-2001BEA	X	X			X	
1998 SCE NRNC Eval.		X			X	
Pre-1998 PG&E NRNC Eval Carry-over		X	X			X
1996 PG&E NRNC Eval		X	X			X
1996 SCE NRNC Eval		X	X			X
1996 SDG&E NRNC Eval		X	X			
1994 PG&E and SCE NRNC Eval			X			X
1994 SDG&E NRNC Eval				X		

An overview of the final adopted net-to-gross rates from the California NRNC studies is provided in Table 1. There is significant variation across these estimates, ranging from 47% to 80%. Comparing the NTG ratios in Table 1 with the methods shown in Table 2 show no correlation between method and NTG ratio nor to whether spillover is included or not. We recognize that a wide divergence in results for different years and different programs is not necessarily an indication of a problem. There may be variations in the composition of the building population and study sample and there will be historical events (such as the 2001 energy crisis) that influence construction practice and the market actors decision-making.

Self Report Approach

Survey-based stated intentions, or "self-reports," is a method of estimating free ridership by asking participants directly a series of questions on what they would have done in the absence of the program.   Often these questions allow a scale of responses concerning the probability of the respondents' taking the same action, the degree of efficiency their alternative action would have had, or when action would have occurred.

Estimating "what participants would have done" can use a variety of questions to evaluate the decision-making process to identify where the participant was in the process when they learned about the program and to ascertain how the program could have influenced this process. These inquiries can address the timing of decision processes versus program intervention, information dissemination, other decision priorities, and general criteria for commercial investments. These methods are particularly useful for customers that have difficult and time-consuming processes, such as government and large commercial customers. The actual wording, sequencing, and probability or point assignments used in the survey and analysis, however, are critical in obtaining an estimate that proves defensible and reliable when tested against alternative probing or methods.

Similar surveys can also be used to estimate participant and non-participant spillover. Yet, the same issues of social desirability and construct validity apply with using surveys for these elements as they did for free ridership. Program influence on non-participants is also likely to be more indirect and, therefore, difficult to inquire after in the survey or even to identify how this link occurred. As such,

often only the most direct program influence with significant recall gets picked up in self report survey research for non-participant spillover.

The market transformation caused by prior program actions can be misconstrued as free ridership in the current program. As the program matures the lines between the measurements of program impacts become blurred with the measurement of market transformation. This touches on one of the biggest complaints about the self-report free ridership approach, because it does not adequately address program efforts aimed at changing standard specifications. For this broader, market perspective a different evaluation approach is required.

The distinction between free-ridership and spillover definitions used in the self report approach and market transformation begins to show the complicated issues between NTG analyses in a static timeframe versus seeing program participation affecting individuals over time. This leads to the much bigger issue for both the free ridership analysis and spillover analysis which is how the program changes the overall NRNC market. NRNC markets are complex with decisions often involving many parties and slowly evolving over time. A program that is active for some time, particularly where there may be limited sets of market actors, can have significant effects on the market and the behavior of the market actors and their interactions. Not considering the consequences this has on the market, the market transformation that may be occurring, can significantly underestimate the program impacts. It blurs the line between free-ridership and spillover.

Difference-of-Differences Approach

The difference-of-differences (DOD) approach generally accepts the non-participants as an indicator of the efficiency levels in new buildings that would be expected in the absence of the program. The net savings is the difference between the efficiency savings seen for participants and those seen for non-participants, where these ratios are the as-built whole building energy usage compared to the baseline whole building energy usage for each group. The net savings is the product of this difference times the participant baseline usage.

The methodology depends upon a matched sample of non-participants as a good control group for the participant savings. This methodology assumes that spillover among non-participants is off-set by the free ridership among participants. There is no way to evaluate this assumption within the DOD methodology, since neither is measured. The probability that they are exactly equal in every evaluation and even as the program matures and affects the market overall seems quite small. In addition, the DOD ignores the self-selection problem by which the most motivated market actors will be found disproportionately in the program, as opposed to in the non-participant population.

Econometric Approaches

Two different categories of econometric methods have been used in the past for evaluation of California's NRNC programs: billing regression, and econometric choice. Both methods are based upon comparing the change in participants to the change in non-participants.

The billing regression analysis used to evaluate NRNC program savings directly estimated net savings and no derivation of a NTG ratio, free ridership or spillover were produced. The method was a purposefully simple two-step process using billing data from participants and non-participants. It allowed customer-specific regression which ensures fit at the customer level and did not attempt to add sophistication to "make the model work."

The econometric choice method was used in several California NRNC impact evaluations in the mid-1990s. The methods used were generally fairly sophisticated and much of the work in California in the NTG field during this time period pushed the technical work done in the U.S. In summary, the

evaluation approach drew upon simulated baseline and actual energy usage, and a prediction of the efficiency choice based on economic and decision-maker characteristics. A double inverse Mill's ratio was used to correct for the selectivity bias.

The econometric methods allow for factors to be controlled isolating the program impact from the many other changes occurring and effects on energy usage, savings, the decision to install/invest in efficiency, and/or the decision to participate. There are few alternatives available to be able to do this. Perfect matching on all factors is practically impossible and there is no method other than econometric to test how close the match really is on all important factors. The mathematics used and statistics provided offer a sense of scientific rigor and at least a perceived sense of accuracy.

The econometric choice models that include control for many factors and sophisticated treatment for self-selection bias are scientifically appealing and are designed to meet several of the problems seen in the DOD and simple billing regression approaches. These are the primary reasons behind the expense and effort placed on these methods.

The econometric choice methods used in the NRNC studies are pretty sophisticated. As such, there is a significant amount of work involved and high level of expertise needed. Even with this, subsequent analyses have historically continued to find issues creating ever increasing levels of econometric sophistication. This type of work then requires large sample sizes with a significant degree of information for both participants and non-participants. All of this increases the costs for these methods.

However, each subsequent, more sophisticated method found problems with the prior method being used such that there is no absolute golden method available. These econometric methods have been shown to be sensitive to the model specification and influential observations. They can be less stable, and when this occurs they are less defensible.

The biggest issue, even with the econometric choice methods, is that of whether the baseline for non-participants are truly unaffected by the program given market changes that could be induced by the long standing program. This possibility is not addressed within even the most sophisticated econometric models performed to-date for NTG.

Understanding the market and the programs impacts throughout the market over time could become an important part of trying to "clean up" the econometric framework to account for market influences.

Assessment of Previous Net Savings Methods

Based on an assessment of past methodologies, we conclude that the Self-Report methodology was the most transparent and appropriate approach for evaluating this complex and diverse program and market, given what the BEA study was trying to accomplish and the needs of the evaluation audience[1]. The SBD evaluation methodology has been evolving and improving continuously over the years. The most recent self-report methodology is sound and, if there is a continued need to calculate net to gross ratios, should be continued with some improvements. Improvements include taking a more expanded look at the various actors involved in the buildings' decision-making (or at least for the largest expected savers).

However, this program-based assessment may not be the best approach for assessing the overall savings achieved in the NRNC market. We recommend that the evaluation effort consider alternative, market-centered approaches, as discussed in the following section.

---

[1] Heschong Mahone Group, Inc (HMG), Chappell, Catherine, Mahone, Douglas; , Brown, Marian; Keating, Ken; Megdal, Lori. 2005. *Methods for Estimating Net-to-Gross Ratio for Non-Residential New Construction White Paper.*

Net Impacts from the Market Perspective

All energy efficiency programs operate within markets. New construction efficiency programs are as intertwined in a complex market as we could possibly envision. Changes in new construction practices, equipment choices, design, and services promoted by utility programs are but a small fraction of the activity and interactions in a large and dynamic market. In the non-residential construction market, the influences at work on equipment selection, building design, and construction practices are very numerous. In California, the utilities team with the California Energy Commission, and interested parties to change the efficiency of all of the building stock, especially at the time of construction or renovation. The utilities fund three interacting activities with this ultimate goal of energy efficiency – new construction programs, codes and standards initiatives, and infrastructure that supports improvements in practices.

Seldom is the case as obvious and strong for operating the programs together, setting targets that encompass the whole market, and evaluating those targets based on market targets as with non-residential new construction initiatives. Targets based on markets need measurement based on markets, not on individual programs. All of the program activities take place in a market where energy prices and costs of doing business may be driving at least an awareness of the need for efficiency.

In an ideal world, the goal of energy efficiency in the whole state should result in targets based on the whole market – improvements in consumption per square foot over several years in the whole targeted market – and lead to evaluations that measure whether or not those improvements are being met as a result of all the factors at work, for and against the improvements. The programs should be evaluated together, along with the efforts of the CEC, and the utilities should be rewarded for the success of the portfolio. The cohesive portfolio could be run as either a part of public benefits or the procurement efforts, but the portfolio should be designed and planned as one program.

In the absence of an ideal world, there are ambitious programs, that in the current regulatory structure, often may end up attempting to claim credit individually that should be shared by (a) new construction programs, with their incentives, (b) the infrastructure programs with their training, awareness raising, and professional support, and, (c) codes and standards programs that provide technical support and advocacy that builds on the changes in practice established by the first two, to advocate for lasting code changes. From the utility perspective, it is even worse that evaluations may end up taking away credit from one program, or all programs, for what are actually impacts of the programs. Simply stated the spillover from earlier efforts of a program or concurrent efforts of other programs are going to be identified as naturally occurring or free-riding tendencies and reduce the perceived net value of the program.

The non-residential new construction program does not operate in a vacuum. It is logical to plan, fund, implement, evaluate and reward these programs as a "portfolio" within a market. This leads to a NTG evaluation approach based on a market perspective. It recognizes that the ultimate goal of the program is to save energy and peak power resulting from the ever more efficient new buildings. This end-state is the goal setting and policy perspective that justifies the programs.

The advantages of the market perspective on estimating net program impacts include logical consistency with the market and simplicity. From a market perspective, the measurement of the efficiency of the new building stock, compared to a prior measurement provides the most direct indication of the savings. It is congruent with the market in that it does not distinguish participants from nonparticipants, because the whole market is the target for the portfolio of programs. This is a market-centered approach to evaluating the programs, rather than a participant-centered approach. In its most fundamental formulation the net impacts of the program are the changed consumption intensity in the market beyond the forecasted baseline.

To operationalize the market approach, consider that the impact at Time 2 is the observed changes from Time 1 minus the changes in the baseline. This structure considers that the sampling, modeling and measurement is an ongoing, repetitive process so that when a measurement is taken at Time 1, it not only establishes a benchmark for where practice is in the market in terms of energy intensity, but includes a forecast of where it would be in the future absent any further intervention. When measured at Time 2, the same process is repeated to estimate where practice might continue to grow in Time 3 in the absence of any further market intervention. This provides the incremental values required for performance tracking. However, remembering that the impacts from interventions in Time1-2 will still be part of what is later assumed to be baseline in Time 3-4, the ultimate measure of the success of the portfolio of NRNC and code programs, with the supporting institutions will be the difference between Time 1 and Time 5, with some effects trailing off into the future. The changes in the shorter time periods, when measured across all new construction – the true target of the portfolio – may be small, but the long term impacts should be substantial and be judged against all of the costs of all the investments in the portfolio: the NRNC programs, the technology and demonstration centers, and the Codes and Standards Programs.

This approach measures the end-state and directly parallels the way that the statewide efficiency goals are set. In its simplicity it can stand above the endless questions surrounding the psycho-sociological intent of the market actors – if those actors can be identified, and if they recognize what influenced their decisions. A market approach measures the gross changes in consumption and subtracts out a pre-determined naturally occurring dynamic baseline. The nuances of self-selection, market transformation influencing perceptions of free-ridership, nonparticipants' behavior as a surrogate for that of participants, and efforts to turn linked and simultaneous decisions into discrete choices are no longer confounding. In this way the many issues discussed earlier about self-selection, statistical biases, unstable models, poorly matched nonparticipants, and self-report biases that lead to an endless "do-loop" of research are avoided. Rather than estimating multiple influences on the net-to-gross ratio, there is only a single unknown that needs to be estimated, the baseline – that which would be expected to happen in the absence of the portfolio or its programs.

However, the unknown – and possibly unknowable – baseline creates a big uncertainty.   In estimating a net program impact in this market, the baseline must subsume all of the interacting influences on the market. These include the inertia of building practices, the fashions in building design, the normal adoption of new technologies, the adoption curves whereby certain building types lead and others lag, and the innovation stimulated by prior program experiences. We don't automatically include codes in this list. Whether codes are considered part of the net program impact or part of the baseline will be determined by how an evaluator/regulator views the role of codes on the NRNC market as: (1) an exogenous baseline influence, (2) a part of the net results of the program portfolio, or (3) the future goal of the current program. The market perspective would clearly view previous, concurrent, and future code impacts as part of the programmatic result of the utilities' NRNC portfolio.

The problems that were encountered in other evaluation approaches do not disappear with a market perspective. Some of the old problems overlap with the newer construct. These include trying to establish a hypothetical alternative reality, which would be part of the forecast market baseline, and discerning how market actors would have behaved in the absence of the program effects.

The other disadvantage is that it is an approach that doesn't integrate well with the current regulatory and program structures. The market based perspective on net-to-gross makes strong evaluation and market sense. The reality of the market is that programs take years to have their full impacts in this sector. Programs that promote code changes build on the work of NRNC programs that push building practice, train market actors, target influential networks and sort out good measures from weak ones. Code changes are logically seen as a result of years of incentive-based program efforts. The measurement of the joint impacts is a multi-year effort on a continuum, rather than a specific annual or

biannual net impact. Nevertheless, there are regulatory and evaluation traditions that seek to know the time-limited, incremental affect of each year's investments. In addition, desires for program efficiency drive the need to know whether program dollars are being spent where they aren't necessary in a given year – when a participant with prior positive experiences does not pass up a standing incentive offer who is then identified as a free-rider. There are regulatory expectations that put a high priority on separating out spillover, interactive market effects, free-riders etc., rather than subsuming them into the program effects and baseline categories. The market based evaluation methods can provide gross and net impacts, but standard research would need to be added to find out if efficiencies in the program could be made through identifying whether there are free-riders among those who took incentives.

Measuring Energy Impacts

The measurement of observed changes in energy consumption within the whole new building is generally referred to as "gross savings." In the NRNC sector, the definition of the baseline has become a core issue for the net savings as well as gross savings. Measuring change requires knowing "change from what?" The studies referenced on the history of NRNC evaluations reflect a clear evolution of thinking that was unique to new construction and code programs. The whole building had to be considered the unit of measurement. This was an important step up from looking at building construction as a series of individual measures.

The whole-building approach led to defining the baseline in terms of matched "nonparticipants"-- whole buildings of the same function, built at the same time in the same climate zone, of the same general size. In the difference of differences (DOD) approach the non-participants buildings were considered surrogates for how the program buildings would have been built. They, in essence became the baseline. The baseline was sought by modeling whole buildings with and without the program measures. This became the gross savings based on a gross baseline. Then the self report or econometric analyses attempted to determine what the real baseline was for each participant. This resulted in an estimate of net savings (without consideration of the participant and nonparticipant spillover). In all cases there were simplifying assumptions necessary to estimate the baseline, and that estimated baseline was essential to the definition of net savings. For the DOD model, the assumption was that the efficiency of the nonparticipants' buildings represented exactly how the participants would have built their buildings, even though they were different people building different buildings. For the modeling approaches, the assumptions were related to the validity and reliability of self-reports, the equating of probabilities of participation and action as the same as free-ridership, the modeling assumptions and specifications, and other issues discussed in the earlier chapters.

From a market perspective, the next step up in defining the baseline is to consider the whole *market* of new buildings instead of simply making distinctions between participant and nonparticipants buildings, or the buildings actually built with the program and the intended or likely building in the absence of the program. Sampling methods to get a representative cross-section of new construction in the market are the same as those used for other research questions. The big difference in sampling for DOD and a market approach is that the DOD tries to match a fraction of the market *to the participants* rather than to use a representative sample of the whole *market*. The non-DOD approaches either try to control for the differences between the participants and nonparticipants statistically or simply concentrate on the participants and their motivations.

The issue of the baseline remains. It is a forecast, because it really isn't just a point in time estimate. If it were, then the findings at Time 1 would be the baseline against which Time 2 is measured. But that would deny the dynamic aspect of a market. Efficiency will change in the absence of the portfolio efforts, one way or another. In a vacuum, it could be hypothesized that the building

practices would fall back to minimum code or worse if there were no further utility efforts. It is more likely that the effects of the programs will continue on with some persistence among those directly affected by the program – a program effect, but one for which further investment is unwarranted. Various ways of modeling this type of short-term market effects of incentives, training, and awareness can be found in Sebold et al., (2001)[2].

Nevertheless, a forecast of market changes is still a forecast. How a defensible forecast can be made for use as an *a priori* moving baseline is a difficult, but not insuperable challenge. Forecasts are used for a lot of serious public policy, and are usually off the mark unless it is a deterministic situation. Markets are surely not deterministic. Therefore judgment is necessary. This can be accomplished through a Delphi method, but as noted in the *Framework*, the technique's defensibility depends to a great deal to the careful way that it is carried out with an understanding of "why" different forecast elements are suggested[3] At Time 1 the evaluators have a point estimate of energy use intensities and penetration of specific measures. For future defensibility, it is better to get expert opinion up front of where the market would go in the future without further intervention rather than to ask later for a retrospective interpretation. As noted in other literature on survey methodology, what actually has happened will modify the respondent's version on what would have been forecast to happen. In hindsight, a lot of things look like they would and should have been forecast.

The type of people who an evaluator would want in a Delphi group would include code officials, CEC forecasters, leading architects and engineers, experienced evaluators, and developers. Utility program people may be very knowledgeable and may have followed the market better than many, but their objectivity would always be shadowed by their direct involvement.

Market Analysis Technical Approach

The challenge in doing a market-based evaluation approach to the nonresidential new construction market sector is to develop a broad-based but reasonable metric of building efficiency. The current evaluation (BEA) approach to determining the efficiency of individual buildings entails a detailed on-site survey of building characteristics, followed by the development of a building-specific DOE-2 model which is (sometimes) calibrated to billing data, and ending with a series of parametric runs that compare the as-built efficiency of the building and its end-use systems with a model of the Title 24 baseline efficiency for the building. Clearly this approach is too resource-intensive to do for a market-wide study.

The logical approach to market-wide assessment of building efficiencies would be to start with the billing data for a representative sample of buildings. Such a sample would have to account for differences in climate zone, building type, building size, and perhaps other factors. It would likely be a large sample, probably in excess of 600 buildings. By contrast, recent BEA studies have had sample sizes of 150 buildings, including both participants and nonparticipants. There would also have to be access to 12 or more months of billing data for these buildings.

Billing data alone will not serve the purpose, however. The energy use intensity (EUI) of a building, as derived from billing data, needs to be informed by other, building-specific information to be useful. Of course the square footage of the building must be known, since EUI is energy use per square foot. Beyond that, however, it is important to know about the energy uses that are attached to the billing meter, and whether there are multiple meters serving the building. It is also important to characterize the energy uses.

---

[2] *A Framework for Planning and Assessing Publicly Funded Energy Efficiency*, 2001.
[3] *California Evaluation Framework*, p. 268.

If the EUI characterization of individual buildings could be informed by this type of end-use information, then it would be far more useful than general EUIs by building type and climate zone. This information could be obtained from on-site surveys, but they would much simpler than the highly detailed surveys currently performed for the BEA study.

In addition to its usefulness in characterizing the energy efficiency of a larger population of buildings, this market-based EUI approach could also have broader application beyond describing utility program effectiveness. This approach would provide direct feedback on California's progress toward the goal of reducing per capita energy consumption. It would provide a superior method for characterizing the relative efficiency of commercial buildings than the current EPA Energy Star Buildings method. The approach could also be adapted to the real estate appraisal of relative energy efficiency in commercial buildings, which has been limited by the availability of comparable efficiency data among buildings.

Of course, there are many details that would have to be worked out in the study design and implementation of the EUI approach outlined here, but we believe that a practical methodology could be developed. It would be a substantial advance in current evaluation practice for the nonresidential new construction market, and it could ultimately be adapted for other market sectors as well.

Conclusion

Specifically, from a market perspective, the measurement of the efficiency of the new building stock, compared to a prior measurement, provides the most direct indication of the savings achieved market-wide. It is congruent with the market, in that it does not distinguish participants from nonparticipants, because the whole market is the target for the portfolio of programs. This is a market-centered approach to evaluating the programs, rather than a participant-centered approach. In its most fundamental formulation the net impacts of the program are the changed consumption intensity in the market beyond the forecasted baseline. A market-based approach requires a logic model that describes how all of the programs, infrastructure, and initiatives within the market are designed to operate as a market portfolio to achieve the end-state objectives.

The biggest disadvantage is that the biggest uncertainty is unknowable – the baseline. As discussed previously, the baseline is not a static point estimate, but a trajectory of expected change over time. This is especially true in a market as diverse and dynamic as NRNC. In estimating a net program impact in this market, the baseline must subsume all of the interacting influences on the market. The other disadvantage is that it is an approach that may not integrate well with the current regulatory and program structures. It would, however, integrate well with the portfolio or unified approach to setting targets, implementing and evaluating programs, and rewarding program accomplishments. Adoption of such a market based approach is an agenda item for future policy direction.

The advantages of the market perspective on estimating net program impacts include logical consistency with the market. The market based perspective on net to gross makes strong evaluation and market sense. The reality of the market is that programs take years to have their full impacts achieved in this sector. A long-term, integrated perspective can provide a more realistic picture of the market changes.

There are several reasons for evaluating the programs within the market as a unified portfolio, judging all the non-baseline effects against the sum of all of the contributing costs. We believe that it is desirable to plan, fund, implement, evaluate and reward these programs as a portfolio within a market. Given the new proposed paradigm of crediting program activity in terms of delivered energy savings, we argue that now is an opportune time to implement this expanded market evaluation approach.

References

Brost, Matt , Roger Wright, Ramona Peet, Cathy Chappell, Douglas Mahone and Pete Jacobs (2002). "Measuring Accomplishments of Energy Efficiency in California's Non-Residential New Construction Market." *American Council for an Energy Efficient Economy Summer Study*: Asilomar, CA. pp. 10.35-10.45.

Heschong Mahone Group, Inc (HMG), Chappell, Catherine, Mahone, Douglas; Brown, Marian; Keating, Ken; Megdal, Lori. 2005. Methods for Estimating Net-to-Gross Ratio for Non-Residential New Construction White Paper. Prepared for Marian Brown, Southern California Edison Co. in support of the Statewide Nonresidential New Construction Market Assessment and Evaluation activity.

Pacific Gas & Electric Company. (1997) *Impact Evaluation of Pacific Gas & Electric Company and Southern California Edison 1994 Nonresidential New Construction Programs.* Prepared for Pacific Gas & Electric Company and Southern California Edison by RLW Analytics, Inc., March. PG&E Study ID Number 323, SCE Study ID Number 523.

Pacific Gas & Electric Company. (2000) *Pre-1998 Non-Residential New Construction Impact Evaluation Carry-Over*. Prepared for Pacific Gas & Electric Company by RLW Analytics, Inc. with Architectural Energy Corporation, Eskinder Berhanu Associates, ASW Engineering, and GeoPraxis, March. PG&E Study ID Number 400.

RLW Analytics, Inc. (1998) *Southern California Edison 1996 Non-Residential New Construction Evaluation*. Prepared for *Southern California Edison* by RLW Analytics, Inc., Feb. Study ID Number 543.

RLW Analytics, Inc. (1998) *Pacific Gas & Electric Company 1996 Non-Residential New Construction Evaluation*. Prepared for Pacific Gas & Electric Company by RLW Analytics, Inc. with Architectural Energy Corporation, July. Study ID Number 389.

RLW Analytics, Inc. (1999) *Southern California Edison 1998 Non-Residential New Construction Evaluation*. Prepared for *Southern California Edison* by RLW Analytics, Inc. with Architectural Energy Corporation and ASW Engineering, December. SCE-06.

RLW Analytics, Inc. (2003) *1999-2001 Building Efficiency Assessment (BEA) Study: An Evaluation of the Savings by Design Program*. Prepared for Pacific Gas & Electric Company, San Diego Gas and Electric, and Southern California Edison by RLW Analytics, Inc. with Architectural Energy Corporation and Eskinder Berhanu and Associates, April.

RLW Analytics, Inc. (2004) *2002 Building Efficiency Assessment Study: An Evaluation of the Savings by Design Program*. Prepared for Pacific Gas & Electric Company, San Diego Gas and Electric, and Southern California Edison by RLW Analytics, Inc. with Architectural Energy Corporation and Eskinder Berhanu and Associates, July.

San Diego Gas & Electric (1996) *1994 Nonresidential New Construction Program: First Year Load Impact Evaluation*. Principal Investigators: Patrick Kirkland, Dean Schiffman, Rob Rubin, and Leslie Willoughby from Market Programs & Planning, February. MPAP-94-P52-935-604, Study ID Number 935.

San Diego Gas & Electric (1998) *1996 Nonresidential New Construction Program: First-Year Load Impact Evaluation*. Prepared for San Diego Gas and Electric by RER. Study ID Number 1004.

Sebold, Frederick. D., Alan Fields, Lisa Skumatz, Miriam Goldberg, Kenneth Keating and Jane Peters (2001) A *Framework for Planning and Assessing Publicly Funded Energy Efficiency*. Pacific Gas and Electric, PG&E -SW040.

TecMarket Works Framework Team (2004). *The California Evaluation Framework*. Southern California Edison Company. Study ID K2033910. Prepared for the California Public Utilities Commission and the Project Advisory Group, June, prepared by a team of 11 consultant firms. http://www.calmac.org/calmac-filings.asp