Establishing a Baseline in Commercial New Construction DSM Impact Evaluation – Comparison of Three Approaches

Douglas E. Mahone, Heschong Mahone Group
Taghi Alereza, ADM Associates, Inc.
Athena Besa, San Diego Gas and Electric Co.
Anne G. Lee, Pacific Gas and Electric Co.
Sharon K. Noell, Portland General Electric

Defining an accurate energy efficiency baseline for commercial new construction DSM programs is crucial to the measurement of program impacts. If the baseline is wrong, impacts will be over- or understated. Under the California measurement and evaluation protocols, impacts must be calculated on a whole building basis, so the baselines established for individual measures are, in effect, aggregated up to a whole building baseline. A consequence of this is that the efficiency of both incented and non-incented measures affects the bottom line program savings.

For reasons of convenience, the Title 24 energy code has been widely used as a baseline. It does not, however, necessarily represent current practice. Many people assume widespread cheating, which would make Title 24 better than actual practice. There is evidence, however, that current practice is actually better overall than Title 24.

In defining the true baseline, one must rely on some measure of non-participant energy efficiency. This means one must clearly distinguish program participants from non-participants. In the simplest definition, non-participants include anybody who did not receive incentives from the program. In the best definition, non-participants only include those who have had no contact with the program. If some program influences contaminate the non-participant population, this will have the effect of penalizing the program for some of its success. Effective methods still need to be developed for crediting programs with these “free-driver” effects.

The use of non-participant buildings to calculate baseline levels of efficiency presumes that these buildings represent what the participants would have done absent the program. Difficulties and cost in recruiting and analyzing non-participant buildings can make this non-participant baseline unavailable or less than ideal.

In this paper, three commercial new construction impact evaluations are compared. All three started with a Title 24 baseline and then refined the baseline through further analysis. Two used participants and non-participants, one used only participants. Net-to-gross analysis using information indicates that 20% - 25% of savings were naturally occurring, meaning that the assumed program baseline was lower than actual practice.

Introduction

Energy efficiency incentive programs for new construction are faced with a unique problem—how to define the baseline against which the savings are measured. Unlike existing building retrofits, there is no before/after comparison that can be made with a new building. The new construction baseline must either be assumed, or else a fair and practical method of calculating it must be developed. If the efficiency baseline is set too low, then the savings attributable to the program will be overstated; if it is too high, savings will be understated.

Another characteristic of new construction programs is the difficulty in isolating individual energy efficiency measures from the whole building energy efficiency.

When whole building energy analysis or metered billing data are used, they address the whole building. Disaggregating this information to the level of individual measures is an expensive and imprecise procedure. Unless disaggregated, however, measures that were not explicitly part of the incentive program may create unexpected credits or penalties to measures which were incentivized. For example, unexpectedly high lighting levels may cause higher energy consumption by the air conditioning system, in spite of an incentivized efficient package system. When measurement is focused on whole building energy assessment, the impact of these credits/penalties needs to be clearly understood.

This paper reports on the approaches used by three California investor-owned electric utilities in implementing and evaluating their commercial new construction programs. In general, program design and implementation procedures made use of the state energy code (Title 24) as the baseline. For measurement and evaluation, the utilities used protocols that have been established for California (Joint M&E Protocols 1992), but these protocols were applied in somewhat different ways. The baseline issues raised by this experience are examined, and recommendations are made for improving both program implementation and measurement.

Who/What Is in the Baseline?

The first problem for a new construction program is to decide how the baseline is to be calculated, and who and what is included in the baseline.

Defining Participants and Non-Participants

The approach required by the California protocols is to compare participant buildings to a control group of non-participant buildings. The energy efficiency of the non-participants sets the baseline against which the efficiency of the participant is measured. From the utility's perspective, this usually means that the participants' efficiency should be measured as high as possible and the non-participants' efficiency as low as possible. This leads to a sensitive accounting task, where the evaluator seeks to prevent any participant effects from contaminating the non-participant pool and vice versa. If, for example, non-participants improved the efficiency of their buildings as a result of utility influence (but didn't get credit for it as participants), then the utility program is actually penalized. The higher measured efficiency of the non-participants reduces the apparent increase in participant efficiency and so reduces the measured savings.

participants. In its most straightforward definition, a participant is a customer who received a rebate under the new construction program during the time period of the evaluation. Non-participants, then, are all other customers. For most purposes, this is satisfactory. The primary objections to this definition come from the fact that program participation/non-participation in not always a yes/no or binary state; there are many possible gray areas.

One gray area consists of participants who received a rebate after the study cut-off date but before the evaluation date. If utility records do not clearly identify these cases, then they will fall into the non-participant group. If the records are good, then the rebated measures can be controlled for and the savings will be credited in the next round of impact evaluation.

Other gray areas to be considered:

- Customers who applied for a rebate but did not complete their projects. (These customers can and should be tracked by the program so that they may be excluded from the non-participant sample.)

- Customers who were educated about efficiency by the utility but did not apply for a rebate (they could have participated in design assistance, attended training at an energy center, etc.)

- Customers who were contacted by utility personnel and given literature.

- Customers who were contacted by utility trade allies, but did not pursue incentives.

These customers could be called free-drivers to the extent they make their buildings more energy efficient than people who had no contact with the utility. They all fall into the non-participant category, and so their efforts at energy efficiency have the effect of penalizing program claims. Unfortunately, it is difficult to quantify the magnitude of their savings, and to draw a clear link between the savings and the program actions.

Non-Participants. The definition of non-participant, as mentioned, is simply anybody that built a building who is not a participant. Often, though, it is desirable to put a finer point on the selection criteria. This is because the non-participant energy efficiency is compared to that of the participants. Unless sample sizes are large enough to wash out individual building effects, it is desirable to have

similarity between the energy using characteristics of the participants and the non-participants. Here are some of the possible types of non-participants:

- Somebody whose building is identical to a corresponding participant, except that it did not participate in the program. It would be identical in terms of occupancy, size, construction, energy systems, hours of operation, climate, etc. In practice, a direct match is impossible, but for common building types in large service territories a good match can often be achieved.

- Somebody with a similar building located outside the utility's service territory, who is not eligible for the program. This has been proposed by some, but it is difficult to know how to control for the obvious differences, such as different utility rates, lack of an energy code, different local climate, different vendors and construction practices, etc.

- A pool of buildings with enough diversity in energy characteristics to match the diversity among the participant population.

Essentially, these questions address either the difference between a pooled population of non-participant buildings or a one-to-one match of participants/non-participants. The resolution is usually determined by the available population of non-participants and the analysis budget for studying their energy characteristics.

An additional factor complicating the non-participant selection process is that, as time moves on, fewer and fewer customers will be completely unaffected by the utility's programs. Those that remain oblivious to the utility programs will become increasingly self-selected as resistant to energy efficiency measures, and non-representative of the building population as a whole. The evaluation will have to rely more and more on using customers who have been exposed to the programs at some level as non-participants. This will further complicate the determination of free-driver/free-rider effects.

Incented Measures and Non-Incented Measures

People who are used to thinking in terms of equipment retrofits often find it difficult to think of the whole building that is necessary for new construction evaluation. The necessity of comparing populations of participants and non-participants using billing data and of considering a wide range of possible efficiency measures makes it hard to track program effects measure-by-measure. Most fundamentally, we are comparing the overall energy efficiency of a participant population to that of a baseline, non-participant population. In California, Title 24 is a whole building energy efficiency standard, and the Measurement and Evaluation protocols require whole building evaluation and billing analysis for new construction programs.

The question then becomes, how much effort should we expend to further break down the energy savings? Through detailed survey data and energy simulations, it is possible to distinguish between incented measure savings and non-incented measure savings. This information is obtained by separately analyzing the energy effects of the two types of measures. The non-incented savings can be positive (i.e., additional energy was saved beyond that saved by the incentives), or they can be negative (i.e., the non-incented energy systems are less efficient than the norm, and so reduce the overall building efficiency). They may or may not be attributable to program effects, and so warrant careful consideration. Comparing the two types of savings to each other and to non-participant savings can provide additional insight into free-rider and free-driver effects, and can help utility program implementers understand the impact of the other measures in the building that are not obviously affected by their programs. One might conclude, for example, that additional enforcement effort is required to bring non-program measures up to code.

Through additional analysis effort, it is possible to isolate measure-by-measure effects, but this is quite labor intensive. It requires computer simulation parametrics to be individually prepared for each measure, and the results must be tracked across several end-uses (for example, lighting measures can affect lighting, heating, cooling, and fan energy consumption).

The ability to track incented and non-incented measures separately is also important under the California protocols, which require 4th and 10th year follow-up impact evaluations to demonstrate that installed measures are continuing to perform. This requires normalizing the whole building energy analyses to account for significant changes in building occupancy patterns, or for additional energy equipment changes to the building between the time of program implementation and final evaluation.

All of this means that the impact evaluation efforts must be based upon real-world analysis of the actual buildings and the actual operation of their installed (or uninstalled) measures. It is not sufficient to analyze a paper building once there is a real building in place.

Baseline Alternatives

Predetermined Standard

As discussed above, the baseline for energy impact evaluation of new construction is a complex measure of whole building energy performance, comprising a number of individual measure baselines. Unlike a simple retrofit program, where the baseline can be established only for the affected measure(s), the new construction baseline must also be able to account for non-incented measures. (Note: Many retrofits affect multiple end-uses, and so are more like new construction as discussed here.)

Some utilities define savings from DSM participation as those occurring above an pre-defined "reasonable" standard. For example, packaged air conditioning equipment with a COP above a certain minimum level is offered a rebate. The minimum level is thus assumed to be the baseline for non-participants. In setting this minimum, the utility is making its own assessment of current practice. If carefully done, such an approach can be defensible. In some cases, the utility makes this assessment on a project-by-project basis, based upon what the building owner "would have done" for that project. If this approach is not done carefully, and the base case assumes lower-than-typical energy efficiency, the full measure of savings might not be awarded in a subsequent impact evaluation study.

Energy Code

The most convenient, well-understood baseline is an energy code (Title 24 in California). An energy code typically establishes explicit and uniform efficiency criteria both for the individual components of the building and for the building as a whole. It is convenient because code compliance is a necessary part of the building permit process for most buildings, and most people involved in the non-residential building process are familiar with it. In California, all of the new construction DSM programs use a base of Title 24 concepts and procedures.

Using the energy code as a baseline makes the tacit assumption that people build to just meet the code, no better and no worse. Conventional wisdom holds that there is widespread cheating and that most people do not meet the code. There is contrary evidence from impact studies done for two California utilities that conventional practice is actually better than the code, at least for lighting and basic mechanical systems. This is probably both because the California code is currently lagging standard practice somewhat, and also because it is actually a safer and faster compliance strategy to beat the code by a modest margin than to exactly meet it.

The energy code can provide a plausible baseline if there is no non-participant sample. In California, it has a distinct limitation in that it does not cover all building types (for example, jails and hospitals are exempt). In addition, it does not cover all energy efficiency measures, such as process equipment.

Current Practice

The ideal baseline is the efficiency level represented by current standard building practice in the participant population, absent program influences. Since this cannot be determined with certainty, we look to the current practice in either a non-participant group or in the non-residential population as a whole. As discussed above, this may be better or worse (or both) than the energy code. Furthermore, it evolves over time, so current practice must be determined concurrently with the program evaluation period.

The trick is to define what current practices are, insofar as they affect the energy efficiency measures under study. These practices are likely to vary by building type, ownership, and other factors. For example, institutional buildings may have a higher standard of energy efficiency than speculative buildings. This means that efforts are needed to control the demographic and construction characteristics of the building population used to establish current practice. Ideally, for DSM programs incenting individual measures, the definition of current practice would be built up, starting with individual measures and then aggregating them to a level of whole building energy performance, factoring in the demographic and other influences. Doing this by combining engineering calculations with econometric analysis can assure that both the physical parameters of the measures are accounted for accurately, and the behavioral parameters are controlled for adequately.

Finding Naturally Occurring Efficiency

The true baseline accounts for the level of naturally occurring energy efficiency and, thence, the degree of free-ridership among program participants. In whole building analysis, this is done by determining the overall level of energy efficiency of non-participant buildings, under the presumption that this represents what the participants would have done without the program influences. With additional effort, efficiency levels can be broken down by major end-use. It is seldom worth the effort to further break the savings down by individual measures, due to the diversity of measures included under new construction DSM programs and the difficulties of separating them. To the extent that program influences can be controlled for in the baseline determination, one can

estimate the free-driver effects. Under the present protocols, however, this is quite difficult to achieve.

Comparison of Approaches

PG&E Approach

Pacific Gas and Electric Company's impact evaluation of its Nonresidential New Construction (NRNC) program made use of both participants and non-participants. The sample included 65 participant buildings (out of a population of 114) and 49 non-participant buildings (of a population of 776 buildings).

Study Baselines. The initial baseline was the Title 24 requirements, when they were applicable. For some incented measures that did not fall under Title 24, such as refrigeration, PG&E-assumed baselines were used. Simple engineering estimates of savings, similar to those used by program applicants, were used to develop initial estimates of measure savings. These estimates included on-site verification data for measures installed. These engineering estimates were further refined with hourly simulations and end-use monitoring data. This resulted in a refined energy efficiency baseline for participants and non-participants.

Effort was expended to match non-participant building types to participants, but this was not entirely successful as there are occupancy types in each group that are not matched to the other.

Econometric analysis of billing data for participants was used to develop realization rates for the savings, which had the effect of further adjusting the baseline and actual energy efficiency to better reflect the actual operation of the buildings. Finally, the net-to-gross ratio was developed to estimate the naturally occurring energy efficiency based upon the non-participant buildings. This was the final adjustment to the baseline.

Implications for Results. The net-to-gross ratio for the overall program energy savings was 0.762, which says that approximately 24% of the energy savings (calculated against the refined program baseline) were naturally occurring savings; i.e., they would have occurred even without the program. This means that the actual baseline was higher than the program assumed. Program planners will need to adjust future program design and expectations to reflect this.

The evaluation project experienced significant difficulties in recruiting non-participants, and incurred significant costs in obtaining data about the energy efficiency of the non-participants. Because of early program recordkeeping difficulties, the non-participant group selected may have included some partial participants that could have artificially raised the baseline.

The program had efficiency thresholds for participation (i.e., participant projects had to be a certain percentage better than Title 24) but once the threshold was met the incentive used Title 24 as the baseline. To avoid the naturally occurring savings, the baseline could be raised above the Title 24 level.

SCE Approach

Southern California Edison Company's impact evaluation of its Design for Excellence commercial new construction program used only participants. The study included 114 participant buildings.

Study Baselines. The program baseline, and the initial baseline for the impact evaluation, was Title 24. It was determined for each project, considering the compliance options that were used. On-site survey data indicated actual installed measures, both incented measures and non-incented measures, Hourly energy simulations of each building were used to translate the efficiency of the baseline and installed measures into energy usage and demand estimates. The simulation models were calibrated to actual building energy use through survey data (occupancy and usage patterns) and billing data analysis. This refinement in the baseline accounted for actual conditions (rather than default Title 24 conditions).

Econometric techniques were used to calculate the net-to-gross ratio for the program. The ratio was derived from analysis of surveys that questioned decisionmakers about their standard practice and how the program influenced them to change the energy efficiency of their designs. The net-to-gross ratio calculation also looked at the efficiency of non-incented measures, using this as a surrogate for non-participant measure analysis.

Implications for Results. The net-to-gross ratio for the overall program energy savings was 0.793, which says that approximately 21% of the energy savings (calculated against the refined program baseline) were naturally occurring savings; i.e., they would have occurred even without the program, This means that the actual baseline was higher than the program assumed. Program planners will need to adjust future program design and expectations to reflect this.

The lack of non-participant data greatly expedited the conduct of the impact evaluation, but it took away an important means of calibrating the net-to-gross ratio. As a result, the final baseline for calculating savings has a somewhat lower confidence level. The use of non-incented

measures in participant buildings as a surrogate for non-participant efficiency actions was an attempt to make-up for the lack of non-participant data, but it is not clear that this data is as free of bias as true non-participant data would be.

SDG&E Approach

San Diego Gas & Electric Company's impact evaluation of its Title 24 Plus commercial new construction program is using both participants and non-participants. The study includes 32 participant buildings and 114 non-participant buildings. (It is expected to be near completion by the time of the 1994 Summer Study meeting, and available results will be presented then.)

Study Baselines. All the buildings in the program used Title 24 whole building computer analysis methods to qualify for incentives, so Title 24 was the whole building baseline. Non-participants were selected to match participants as closely as possible in terms of size, occupancy and climate zone. In general, there are at least two non-participants matched to each participant. All non-participant buildings are simulated using the same energy analysis techniques as the participant buildings.

Realization rates and net-to-gross ratios will be developed from the billing data and the non-participant data.

Implications for Results. This study has the most carefully constructed non-participant control group of the three studies. It will be interesting to compare its final results with those of the two completed studies.

CADMAC Approach

A statewide group of utility program evaluators and regulators (CADMAC) is preparing to conduct a statewide baseline study to address some of the issues raised by the individual impact study baseline efforts. This study, scheduled for 1994, will evaluate current construction practices and differences among different building types. It will include buildings subject to Title 24, as well as a sample of the building types that are exempt from Title 24 (hospitals, jails, etc.). The baseline will be established on a whole building basis rather than a measure-by-measure basis.

The intent of developing a statewide baseline is to update knowledge about energy efficiency baselines. This will then allow the utilities to discontinue their reliance on Title 24 as the assumed baseline.

Conclusions and Recommendations

The use of Title 24 as a baseline is convenient for program planning and implementation because it is widely applicable and well understood. These studies have shown, however, that the energy code does not reflect current levels of naturally occurring energy efficiency. Energy savings calculated using the code as a baseline have overstated the actual savings by roughly 20% - 25% overall. Net-to-gross ratios were applied to compensate for this problem in the impact evaluations. Program planning and future projections of expected savings will have to account for this information.

Determining new construction savings on the basis of individual measure savings is only a first step to overall program impact estimation. The whole building energy performance must be identified and then used as the ultimate measure of the program; this includes non-incented measures as well as incented measures. As a consequence, the baseline must include all measures in the building.

Determination of the actual efficiency baseline—what would have happened absent DSM program influences-is a critical task in accurate program impact evaluation. It requires clear definitions of the participant and non-participant populations, good recordkeeping to keep the populations separate, and careful analysis of installed measures in both.

The use of carefully selected non-participant populations is necessary to developing an accurate baseline, despite the difficulties and expense. The proposed statewide baseline study should help to alleviate some of these problems, and by using a larger non-participant sample, should give broadly representative results.

If the baseline is contaminated by program participation effects, it will lead to overestimates of naturally occurring efficiency and will penalize the DSM program, If the baseline is too low (as the Title 24 baseline was shown to be), then program savings will be overestimated.

Acknowledgments

The authors wish to acknowledge the three utilities, PG&E, SCE and SDG&E, whose commercial new construction impact evaluation studies provided the background for this paper. These studies were managed by ADM Associates (Mr. Mahone (formerly) and

Mr. Alereza). The authors have shared many days of interesting debate in grappling with the issues discussed here, and we look forward to further progress in meeting the challenge of making buildings more energy efficient.

References

Joint M&E Protocols: *Protocols and Procedures for the Verification of Costs, Benefits, and Shareholder Earnings from Demand-Side Management Programs.* October 29, 1992. Pacific Gas & Electric, San Diego Gas & Electric, Southern California Edison, Southern California Gas, California Energy Commission, Division of Ratepayer Advocates (CPUC), Natural Resources Defense Council.

Pacific Gas and Electric Company (PG&E). October, 1993. *Non-Residential New Construction Impact Evaluation Study.* Prepared by ADM Associates, Inc. and Regional Economic Research.

Southern California Edison Company (SCE). November, 1993. *Design For Excellence Commercial New Construction Incentive Program Impact Evaluation.* Prepared by ADM Associates, Inc.