How Should We Treat Factors Contributing to Uncertainty in Measurement and Evaluation of DSM?

Susan M. Buller and William C. Miller, Pacific Gas and Electric Company

In the past, proponents of demand-side management (DSM) followed a 'more is better' philosophy without worrying too much about the accuracy of energy savings estimates. Indeed, energy savings estimates of existing and potential DSM programs have varied widely.

However, the advent of incentives for utilities, coupled with reliance by utility planners on DSM as a substitute for building new capacity, has reduced the level of uncertainty that different parties can accept. Consequently, efforts to accurately measure DSM impacts have increased substantially across the nation. One of the most problematic aspects of impact estimation has been the concept of 'net-to-gross'.

In California a net-to-gross ratio is defined as a way to account for 'what would have happened without the program'. However, it is mainly applied to estimates of energy savings in order to account only for 'free-riders'. In the future, as measurement results appear, it will begin to include other risk factors not previously accounted for, such as free-drivership, customer removal of efficient products after installation, changes in energy use resulting from business expansion, etc.

It is the position of this paper that it is time to discuss how the many factors that contribute to DSM impact uncertainty should be explicitly dealt with in energy savings calculations. Issues affecting uncertainty that are discussed in this paper include:

- Free riders and free drivers.
- Calculation of energy savings when there is no 'base-case', such as when a customer installs energy efficient measures at the same time as a change in business use.
- Who should bear the risk of uncertainty impacts? Utilities, ratepayers, program participants, or evaluation consultants?

Discussion of these issues is elaborated within the context of California's regulatory process, shareholder value calculations, integrated resource planning mechanism, and internal utility electric resource planning. The paper will include examples of the role risk factors play in actual evaluations.

Introduction

We are observing an evolution of the meaning of net savings. It has been primarily associated with gross savings net of free ridership. More recently, net savings has been defined as gross savings minus 'naturally occurring' savings (Violette, et al. 1990).

The concept behind measurement of net savings is simple. It is the dream of all measurement and evaluation professionals to own a time machine.[1] The only way to get a true picture of the effect of a DSM program is to measure all your customers for a year or two, then reverse the clock. Now you can institute your program and send customers through the same economic, demographic, and climatological stimuli with the single addition of the DSM program. A simple comparison of the two metered results yields the effect due solely to the program.

The challenge facing the measurement and evaluation professional is that the above is impossible. All efforts to measure net savings are, of necessity, estimates of 'what would have happened without the program'. There are several components that make up the effect of a DSM program, and several general methodologies that have varying degrees of success at estimating these components.

It is our view that eventually net savings will operationally be defined as measurement results times a net-to-gross ratio. The net-to-gross ratio will include other factors not accounted for in measurement and evaluation and could eventually be termed the 'last adjustment factor'. Particularly in the context of shared savings schemes, the actual numerical result that is used as this adjustment factor will be determined as the result of a negotiated, regulatory or political process. It will fall within a range that leaves the parties comfortable with the result.

Net-to-gross should be viewed as the 'adjustment' factor applied to estimates of a program's effectiveness that accounts for our uncertainty about that estimate. The choice of the appropriate net-to-gross estimate will depend on how many of the individual components have already been included in the savings estimate to be adjusted. Before beginning a discussion of the correct treatment of uncertainty, we need to examine the components that comprise the savings estimate of a program and the factors that contribute to our uncertainty about whether that estimate is the same as 'what happened solely due to the program'.

Following a discussion of the components that comprise the uncertainty surrounding program impact measurement, this paper discusses the affect of measurement and evaluation methodologies on either increasing or decreasing uncertainty.

Components Comprising an Estimate of Program Effects

Since we cannot wind history back in order to measure the true effects of a DSM program, we are forced to estimate the factors that contribute to, or subtract from, program savings. Those factors that reduce or offset program savings include:

Gross energy savings. This is the difference between a customer's usage before participation in the DSM program and their energy use the first full year after participation, *all other things being equal.* The emphasis is important, as will be more fully explored below.

Persistence. This is a measure of how long the energy savings difference will last. Usually program implementers assume that deterioration of energy efficient equipment parallels that of standard equipment, i.e., is the same as *measure life.*

Measure Life. The life of the equipment.

Free Drivers. Free drivers are people who installed energy efficiency measures without participating in the program, but who were influenced by the program's existence. There is at least anecdotal evidence that the impact of free drivers can be substantial.[2]

There are also components that reduce program savings. These include:

Rebound. Often a customer will capture some of their financial savings from reduced energy consumption by enhancing their quality of life. For example, if residential customers insulate their homes, they may reduce the size of the resulting savings by turning up their thermostats.

Free Riders. Free riders are those participants who would have implemented energy savings decisions without the existence of the program. This is the component most frequently confused with the term net-to-gross. In fact it is only one part of the uncertainty surrounding estimation of program effects.

There are also factors that can either increase or decrease program savings, but that at the least will confound measurement efforts. These are not necessarily components but are factors that affect energy use and that operate independently of the program. They include:

Other Changes at the Site. It is not uncommon, especially in the nonresidential sector, for energy efficient improvements to occur while the customer is doing something else, such as expanding or remodeling.

Behavioral Effects. Customers may change their normal behavior, such as failing to replace burnt-out lights because a retrofit is scheduled.

External Effects. Other factors such as weather or economic upturns or downturns can affect energy use independently of an energy efficiency program.

Next we will address the effect the measurement has on the component being measured.

How Measurement Techniques Influence Measurement Results

All measurement and evaluation methodologies (except the time machine as mentioned in the introduction) will have drawbacks. Some measurement methods are very good at measuring some program effects, but not others. Other methods can measure program impacts, but cannot separate them from factors simultaneously measured. For example:

Metering and Billing Analysis will measure the impacts of energy efficient measures and will include the effects of rebound, but will not be able to separate the two. In addition, for some measures (like lighting) metering might underestimate usage before the retrofit (because the customer has put off replacing burnt-out bulbs prior to retrofit) and overestimate usage after the retrofit (because more than the normal number of bulbs are working). Program savings are therefore underestimated.

Metering and billing analysis also fail to correctly measure savings if the retrofit takes place during remodelling. In both cases, there is no appropriate pre-retrofit usage to apply as a base case. Consequently, when a customer remodels concurrent with installation of energy efficient equipment, no true before/after comparison is possible with 'real time' devices such as meters.

Engineering Estimates do not have the problems with a non-existent base case that metering and billing analysis may have. The engineering estimate produces a hypothetical base case. On the other hand, engineering models will not, on their own, estimate rebound effects. Nor will engineering models identify discrepancies between manufacturers estimates of equipment efficiency and true operating efficiency.

Customer surveys of participants and nonparticipants can be used to identify free ridership and free drivership. However, it has been shown[3] that participants, when looking back on their decision making process, tend to overestimate their desire to install energy efficient equipment. On the other hand, surveys of vendors and manufacturers could have the opposite effect. Vendors, aware that their sales increase during a energy efficiency incentive program, could tend to underestimate the number of people who would have purchased energy efficient equipment anyway.

Free drivers are much harder to identify, especially when a new technology is being rebated. The utility's program affects market movement and product availability. Some nonparticipants would not have been able to purchase equipment without the program, yet, when surveyed, customers cannot reveal this fact. Nonparticipants, whose purchasing behavior was affected by the program, could correctly report that they never heard of the rebate program. Also, some program participants who are identified as free riders might have been unable to purchase the equipment without the program's effect on the market. In fact, they are not free riders at all.

The above examples are not intended to exhaust the discussion of strengths and weaknesses of different measurement techniques. What we want to demonstrate is the complexity of measuring net savings and the impossibility of arriving at a detailed definition of net-to-gross that could apply to all programs.

In general, the net-to-gross number by which a program's estimated results are increased or decreased should include all components of uncertainty that are not specifically addressed in the estimated results. In some cases, this will depend on the measurement method used to estimate program impacts. In other cases, where program savings are pre-determined by engineering estimates, the net-to-gross number will include those components not included in the engineering estimates.

Who Should Bear the Risk?

Before coming to recommendations about how one should respond to uncertainty in measurement and evaluation of program impact savings, one more factor needs to be considered about uncertainty. By now, we hope that the reader recognizes that independent measurement of <u>all</u> the components affecting program impacts is extremely difficult. We are left with the dilemma that even though we try to estimate the separate components of program impact, we cannot eliminate uncertainty.

A good question to ask ourselves at this point is: 'What are the consequences of this remaining uncertainty?' Candidates include utility shareholders, ratepayers, and future utility customers. Before addressing this question, however, we need to clarify that the cost of uncertainty in some cases is different from the cost of uncertainty in other cases. For example, program impacts need to be known to an exactitude sufficient to promote correct decision making.

For electric resource planners, the degree of uncertainty must promote adding necessary additional power and not adding unnecessary additional power. Supply side decisions can have a long lead time, which complicates the decision. On the other hand, long term planning on PG&E's 20,000 MW system can't even see effects smaller than 25 to 50 MW. It isn't at all difficult to estimate program impacts to within 25 MW.

When a utility is receiving some sort of incentive to invest in energy efficiency, there could be a perception that program impacts must be much more exact. After all, when ratepayer dollars are used to reward shareholders, the tendency is to micromanage calculation of the 'correct' amount. In reality, the appropriate degree of accuracy is that which promotes correct decision making. The shareholder incentive must be accurate enough to promote energy efficiency when that is the 'best' investment decision.

Probably almost everywhere today a substantial amount of cost-effective energy efficiency is the 'best' investment decision for utilities. It would be very hard to overstimulate investment in energy efficiency by utilities. On the other hand, consistent over- or underestimation of program impacts could easily lead to incorrect decisions and installation of unnecessary supply-side resources.

One clear choice is to target the shareholder as the entity to bear the risk of uncertainty. This would argue for consistently conservative interpretations of estimates of the components contributing to program impacts. As indicated above, however, there is a cost. It is important to recognize that in a desire not to overpay the shareholder, we must not incorrectly characterize programs or measures as not cost effective.

Our focus in this paper is uncertainty surrounding measurement and evaluation of DSM impacts. There are many other issues that are not included here, but we will mention that similar (and in some cases worse) issues surround uncertainty of supply-side resources. There are three possible outcomes of misjudging the appropriate investment in DSM:

1) There is an over-investment in DSM and the DSM performs as predicted. This will result in excess capacity. Generally, ratepayers bear the costs of excess capacity.

2) The DSM does not deliver as expected. This will result in a shortfall and the necessity for a crash supply program. It is not clear whether ratepayers or shareholders will bear the costs in this situation.

3) There is an under-investment in DSM. The consequent over-investment in supply will mean higher rates. If the supply-side construction has passed a prudency review, then the ratepayer will bear the costs in this situation.

Uncertainty and Net-To-Gross

The link between uncertainty and net-to-gross is in changing net-to-gross into the 'last adjustment factor' discussed above. Specifically, we mean the incorporation of <u>all</u> factors that contribute to uncertainty into a 'final adjustment factor'. This final adjustment will be the result of a political process played out in the regulatory arena where utilities will be representing the interests of the shareholders, the regulators will be representing the interests of ratepayers and the interveners will be representing their own interests. The final adjustment factor will be litigated in an arena where the result will represent the appropriate slice that divides the risk of uncertainty between the parties.

Some of the following conclusions are obvious and others will be disputed. We, of course, invite comment. This paper is not intended to represent a crystallization of the uncertainty issue. We want to spark a more complete discussion than has heretofore been popular, to clear up some persistent misunderstandings, and to bring some structure to the debate. In furtherance of these goals, we suggest the following:

1) Stop using a net-to-gross ratio as if it only corrected for free riders. There are many components that affect program impact estimates and free ridership is only one of them. <u>Any</u> component not included in the estimate must be included in the net-to-gross ratio or 'final adjustment factor'. This leads us to:

2) Whenever conducting measurement and evaluation of program impacts, or when estimating program savings for another purpose (program planning, calculation of shareholder incentives, etc.), be quite clear about which components are being included, and which much still be accounted for. This of course means:

3) All program planners, program evaluators, regulators, and other professionals engaged in estimating program impacts <u>must</u> become familiar with the components that affect estimation of program impacts. Further:

4) All program planners, program evaluators, regulators, and other professionals engaged in estimating program impacts <u>must</u> be familiar with the measurement tools available to estimate the above components. Especially, professionals must understand what components a given measurement method will measure, what components a given method will not measure, and what components a method will measure but will not be able to separate from other effects. Finally,

5) Almost all program estimation is sufficient for utility long-term planning decisions, at least for the next 5 - 10 years, and:

6) For other purposes, it is probably best to be slightly conservative when estimating program impacts. However, it is important to recognize that consistent underestimation of program impacts has a cost. This is especially true for those programs that do not clearly have a benefit-cost ratio of over 1.0.

Endnotes

1. The authors are grateful to David Goldstein of NRDC for the time machine metaphor.

2. Retailers in PG&E's service territory say that during our residential refrigerator rebate program, nearly all the floor models qualify for the rebate. Almost all customers purchasing a refrigerator during the program will buy an energy efficient model, whether or not they receive the rebate.

3. A recent study by PG&E (PG&E 1992) gathered contradictory evidence from participants that indicates an overestimate of free-ridership.

References

PG&E. 1992. *Net-to-Gross Ratios for PG&E's CIA Rebate Program, Study A - Participant Survey.* Pacific Gas and Electric Company, San Francisco, California.

Violette, D., M. Ozog, M. Keneipp, and F. Stern. 1990. *Impact Evaluation of Demand-Side Management Programs: Volume I - A Guide to Current Practise.* EPRI CU 7179, Palo Alto, California.