Recent Program Evaluations: Implications for Long-Run Planning

Lester W. Baxter, Oak Ridge National Laboratory
Donald K. Schultz. California Public Utilities Commission

Demand-side management (DSM) remains the centerpiece of California's energy policy. Over the coming decade, California plans to meet 30 percent of the state's incremental electricity demand and 50 percent of its peak demand with (DSM) programs.

The major investor-owned utilities in California recently completed the first round of program impact studies for energy efficiency programs implemented in 1990 and 1991. The central focus of this paper is to assess the resource planning and policy implications of Pacific Gas and Electric (PG&E) Company's recent program evaluations.

The paper has three goals. First, we identify and discuss major issues that surfaced from our attempt to apply evaluation results to forecasting and planning questions. Second, we review and summarize the evaluation results for PG&E's primary energy efficiency programs. Third, we change long-run program assumptions, based on our assessment in the second task, and then examine the impacts of these changes on a recent PG&E demand-side management forecast and resource plan.

Introduction

PG&E recently completed several program evaluations that may be applied to revise long-term DSM forecasts and resource plans. Our goal is to review these studies and assess the implications the study results have for long-term resource planning and energy policy. As we attempted to move from evaluation results to forecasting and planning application, we struggled with many issues and questions. In this paper, we discuss those issues and questions we believe characterize the general problem of using evaluation results in forecasting and planning. We then present summary results from PG&E's evaluations and examine their resource planning implications.

Evaluation Issues

In this section, we identify and discuss five issues that are of particular concern in the use of evaluation results to revise forecasts and resource plans: Evaluation Availability; Evaluation Quality; Historical or Future Program Focus; Reconciliation; and Application.

Evaluation Availability

We first identified the available evaluation studies that might be suitable for use in long-term planning. We focused on PG&E's recent program impact studies. These studies report a wealth of information on program participation levels, total program energy and peak impacts, and net-to-gross ratios. While we did not consult any in our review, process evaluations should also be of interest to forecasters and planners. A process evaluation may suggest changes to the current program that will affect future program savings. For example, a process evaluation may reveal that improved training for program delivery staff will substantially increase the program's net-to-gross ratio. If the utility commits to implement the training, forecasters will be ill advised to directly apply the estimated net-to-gross ratios from the program's impact evaluation.

Evaluation Quality

Assessing the quality of program evaluations is critically important for forecasting and resource planning applications. Yet judging the quality of many evaluation studies is difficult. Evaluation teams draw on many disciplines, including statistics, economics, sociology, engineering, and computer science and we find that comprehensive evaluation reviews require use of all these disciplines. Rarely can a single reviewer bring to an evaluation review sufficiently deep expertise from more than one or two of the appropriate disciplines. Thus, comprehensive reviews require review teams that resemble in disciplinary make up the characteristics of the team that produced the original evaluation. These team evaluation reviews typically focus on the following elements of an evaluation:

- representativeness of the study sample to the population of interest to planners;
- accuracy and precision of energy and demand impact results;
- appropriate use of control samples; and
- use of alternate methods or sensitivity analyses to derive recommended impact estimate.

Historical or Future Forecast

Forecasters must also decide if evaluation results will be applied to forecasts of historical program impacts or future program impacts. This distinction is critical because a program evaluation by necessity yields results for DSM programs already contained in the demand forecast. In contrast, a forecast of impacts from future program may be based on a program design quite different from the program evaluated.

Despite this important distinction, we find common issues in the application of evaluation results to historical program and future program forecasts. For example, both applications require energy savings impacts at the appropriate level of detail for the forecasting model. The forecasting model can be more or less detailed than the results from the evaluation. Technology-choice driven models will require much more detail than evaluations generally provide. One solution to this mismatch is to aggregate individual technologies in the forecasting model, such as 4-ton air conditioners with SEER of 12.0, to more general categories, such as high-efficiency air conditioners. End-use forecasting models may require less detail than available from the evaluation. The evaluation may have energy impact estimates for residential high-efficiency air conditioners, but the end-use model may include air conditioners in the broader category of cooling equipment, which also will include heat pumps and evaporative coolers.

Another important issue shared by historical and future forecasts is matching units of savings between forecasting models and evaluation results. At the DSM program level, evaluations typically express energy impacts in annual kilowatt-hours (kWh) or kWh per square foot of affected floorspace. Forecasters face difficult translation problems when their demand or DSM forecasting models do not support similar units. For example, we are familiar with a DSM model that expresses savings for high efficiency air conditioner measures as kWh savings per ton of air conditioning capacity per the change in coefficient of performance.

Finally, historical program and future program forecasts must be concerned with questions of measure life, savings persistence, and changes to net-to-gross ratios over time. California is approaching the measure life and savings persistence questions with long-term, joint-utility-sponsored studies that are coordinated by regulatory staff. How net-to-gross ratios change over time remains an important forecasting question, but the DSM community has not yet achieved consensus on the feasibility or usefulness of analytically resolving this question.

Historical and future program forecasts differ in at least two important ways. First, future program forecasts must reflect the penetration of existing and new DSM technologies over time. Second, future program forecasts must be sensitive to total DSM measure costs (which include utility and participant costs) and conditions that affect program cost effectiveness over time. Examples of such conditions include rising utility costs as market segments become more highly saturated and changes in electricity prices.

Reconciliation Issues: Determining if Evaluation Results Can Be Used in Planning

Reconciliation issues center on determining whether evaluation results can be used for forecasting and planning. We identify two types of reconciliation issues that we call plausibility and transferability.

Plausible evaluation results must be believable and derived using credible and logical methods. Assessing the plausibility of an evaluation result involves subjective judgement, but this judgement need not be arbitrary. In many cases, we can assess the evaluation result by viewing it from several perspectives. We list four perspectives below.

- Does including the result in the demand forecast make any elements of the demand forecast implausible? One way to answer this question is to see how the result changes the demand forecast at the sector and end-use levels. [1] For example, does including results from a commercial high efficiency space conditioning program change the kWh per square foot for the commercial sector in ways that are unreasonable?

- How does the result compare to separate estimates derived from engineering or statistical analysis used to support the demand forecast? The forecast may already include performance characteristics of different technologies, such as air conditioners with different energy efficiency ratings, or the energy savings associated with moving from R-11 to R-19 ceiling insulation. Forecasters can make use of this information to assess program evaluation results.

- Is the evaluation result the product of more than a single evaluation approach? Evaluators often take more than one approach to assess large DSM programs. PG&E's evaluation of its Commercial, Industrial, and Agricultural Retrofit Rebate program used engineering analysis of metered data, statistical analysis of billing data, and decision and survey analysis of customer decision data (PG&E 1993a). PG&E compared results from these different approaches and also applied statistical techniques, where appropriate, to integrate estimates from studies that used different approaches before publishing final results for this program.

- Is the result consistent with results from similar programs at other utilities? Interutility comparisons of DSM evaluation results are often difficult for many reasons. Differences in program measures, incentive levels, customer characteristics, geography, climate, and the built environment can confound interutility comparisons. Of course, the greater the similarity between utilities and their DSM programs the more relevant are comparisons between their evaluation results.

Plausibility is also related to evaluation quality. Much of an evaluation can be of high quality and still contain certain results that are not plausible. For example, PG&E's Commercial, Industrial, and Agricultural Retrofit Rebate program evaluation estimates peak demand savings for refrigeration measures that are more than double the original program estimates. Our review of this part of the evaluation showed that the refrigerator analysis is based in part on metered data for only a single installation. While many other estimates from the evaluation are plausible, in our judgement PG&E's refrigerator estimate is not.

Even if the study results are plausible, the next question is to decide if these results should be used to update cost or benefit estimates in a resource plan. If the results of a study are plausible, the long-term impacts of historic year programs should be revised when the initial program estimates and evaluation results differ. What is not clear, however, is when to use the results from a study of impacts from an historic program year to depict the cost or benefit characteristics of future programs. This consideration illustrates our second reconciliation issue, which we call transferability.

The transfer of evaluation results can take place over time and location. Transfer over time usually involves a single utility that wants to apply evaluation results from a historical program to revise its long-run forecast for a similar program operating in the future. What may initially appear to be a straightforward application can quickly become complicated. Among the more important complications are adjusting for inconsistencies in baseline energy use, technologies marketed, energy units, and cost accounting. We illustrate a few of these complications with specific examples.

Southern California Edison (SCE) recently completed an impact evaluation of a low-income residential lighting program (SCE 1993). The evaluation indicated that the program reduced lighting energy use, but the savings were much lower than expected. SCE forecasters could not directly transfer the results of this study to even their short-term DSM forecast for two reasons (M. Brown, SCE, personal communication, April 1994). First, SCE had already revised the baseline lighting use estimate assumed in the evaluation. This revised baseline estimate is the one used by SCE forecasters. Second, the lighting technologies marketed by the program under evaluation have been replaced by improved lighting technologies in more recent program years.

We discussed the importance of achieving consistency between energy units earlier, but have not yet discussed the importance of consistent cost categories. DSM forecasters are interested in getting cost information at both a detailed level, for measures, for example, and at an aggregate level, for programs. Important cost categories are often tracked at different levels of detail. A typical mismatch is between participant costs (the incremental cost of high efficiency appliances, for example) and administrative costs (used in the Total Resource Cost test and other tests). Participant costs are often recorded at the measure level while administrative costs are tracked only at the program level or higher. In addition, attempts to link energy impacts, which are frequently measured at the whole-building or end-use level, to costs are complicated by differences in the level of accounting between energy and costs.

Transferring evaluation results across locations, from one utility to another, raises similar issues as transferring results across time. The issues are often more difficult to resolve, however, because different utilities are more likely to market different programs and use different forecasting methods. Participants in integrated resource planning proceedings often recommend applying evaluation results from utilities in one part of the country to utility DSM forecasts in another part (see Independent Energy Producers 1993, for example). The study by Nadel and Keating (1991) is frequently cited in this regard. Yet to make the transfer a meaningful exercise, much care must be taken to insure that the programs are comparable and that the DSM forecasts under question are comparable to the historical programs evaluated.

Application Issues: Moving from Evaluation Results to Planning Inputs

Application issues arise when we actually translate evaluation results to planning inputs. We discuss four different application issues: consistency between baseline energy use and savings; program participant characteristics; cost effectiveness; and changes in program design. A major concern for forecasters is consistency between the characteristics of baseline energy use and the savings impacts in the evaluation and the forecast.[2] Unless the baseline efficiency levels of buildings or equipment from the evaluation are consistent with the baseline efficiency levels used in the forecast, the forecasters must attempt to translate the evaluation results.

Consistency problems can surface at many levels, including energy units, the level of analysis, hours of equipment or building operation, equipment efficiency, connected load, and weather. The first, the use of consistent energy units to measure energy impacts in the evaluation and the forecast, we discussed earlier. We emphasize that the translation problems are simplified if evaluations and forecasts use identical or similar units, like kWh per square foot, to express energy impacts.

Matching the level of analysis is also an important concern. Evaluations and forecasts can be designed at the program, end use, measure, or technology level. An evaluation result at the program level may be difficult to translate to the technology level. While forecasters tend to focus on this problem for energy impacts, we note the same problem exists if costs are tracked at different levels in evaluations and forecasts. Of course, any cost accounting inconsistency will complicate using the evaluation results to revise the DSM cost-effectiveness analysis in the resource plan.

Other consistency problems are more often application-specific questions. For example, forecasters must determine if the baseline assumptions about hours of equipment operation, connected load, and weather are consistent with those used to characterize the baseline in the forecast.

Forecasters must also consider how the characteristics of program participants may change from the evaluation period to the forecast period. Participants in a new incentives program may differ from those who participate when the program is more mature. These changes in participant characteristics can have implications for both forecasted program savings and costs. For example, later participants may operate equipment differently then early participants (affecting hours of operation, for example) or may be less likely to change equipment size or capacity as part of the installation decision. These later participants may require higher incentive payments to get them to participate.

We discussed earlier the importance about matching the level of analysis between evaluations and forecasts with regard to cost categories, such as administrative costs and tracking measure costs. The assumptions about program benefits, or utility avoided costs, must also be examined. The utility avoided costs used in the cost-effectiveness analysis as the program was designed and implemented may differ from the avoided costs consistent with the resource plan used to develop the DSM forecast. Thus, even if other characteristics of the historical program are completely consistent with the forecast of the future program, the cost effectiveness of the two programs may differ.

Lastly, forecasters must also consider how to include possible changes in program design, undertaken in response to the evaluation of an historical program, in future program forecasts. Our earlier example is relevant here. If an evaluation indicates program performance is not up to expectation, the utility will either choose to eliminate the program, or more likely, modify its design. In this case, a direct application of evaluation results to the forecast is not appropriate. Instead, program designers, evaluators, and forecasters must work together to identify the likely change in design and assess the effects of the revision on program performance. In the absence of good information about the effect of revisions on performance, predicting this effect will involve considerable judgement.

California regulations partially address these application issues. For years, the California Public Utilities Commission (CPUC) and the California Energy Commission (CEC) have attempted to establish common terminology and definitions for DSM programs and for cost and benefit elements (CPUC 1994). Common forecasting methodologies for resource planning have been the order of the day for more than a decade.[3] The DSM measurement protocols recently adopted by the CPUC devote a specific section to resource planning protocols that should lead to a more uniform and effective means of reporting evaluation study results.[4]

The CEC's Data Base for Energy Efficiency Resources (DEER) also has the potential to facilitate the transfer of evaluation results to long-run planning applications (CEC 1994). DEER, which contains detailed data on the universe of cost-effective DSM measures, is the result of a collaborative effort between California utilities, the CPUC, and the CEC. Many of the application issues we address here have received considerable attention as DEER was constructed. The CEC has only recently released a draft of the data base. As a result, DSM planners, forecasters, and resource planners have little experience working with DEER at this point.

Evaluation Results

PG&E has completed about 40 DSM evaluation studies covering the 1990-92 period. We reviewed 11 studies for this paper that PG&E completed by the end of 1993. The studies cover most of PG&E's major DSM programs for the 1990-92 period. The only major DSM programs not covered by our review, or by available PG&E evaluations, include certain miscellaneous residential and commercial programs and all dispatchable programs.

Table 1 is a summary of the evaluation results. Our review highlights participation levels, gross and net energy impacts per participant, and net first-year energy impacts for the total program. Whenever possible, we compare program estimates (pre-evaluation estimates) and evaluation results. Where the evaluations report confidence intervals we include them in Table 1. The table summarizes energy impacts only. Eight of the 11 studies report peak load impact results; we use many peak impact results in our resource planning analysis and omit them from Table 1 only due to space limitations.[5]

Our table contains many blank cells. The blanks do not mean the results are not available, but that we did not find the relevant data in our review of the evaluations. We believe the blanks suggest a need to implement reporting standards for DSM evaluations to make the key results readily accessible for evaluation users.

PG&E uses a program-level DSM forecasting approach for long-run resource planning. Thus, we focus primarily on realization rates for net program impacts (see endnote 2 for our definition of realization rate). The realization rates for residential program energy impacts are all less than 1.0, which means evaluation results are less than pre-evaluation estimates. With the exception of the showerhead element of the Residential Weatherization Incentives program, the residential realization rates range from about 0.60 to 0.80. Four of the 8 residential evaluations report peak load impacts. The realization rates for peak impacts range widely from 0.32 for Residential Direct Assistance (appliance element) to 3.80 for Residential Weatherization Incentives (ceiling insulation element). PG&E reports confidence intervals for net program savings for the Residential New Construction program and the Residential Energy Management Services program (weatherization element). If we assume the same relative confidence intervals apply to the respective realization rates, only the latter program has a confidence interval that includes a realization rate of 1.0 (at the 90 percent confidence level).

The realization rates for PG&E's Commercial, Industrial, and Agricultural (CIA) Retrofit Rebate program, Commercial New Construction program, and Nonresidential Energy Management Services program are also less than 1.0. PG&E reports confidence limits for gross program savings for the CIA program. If we assume the same relative confidence intervals are directly applicable to the realization rates, the confidence interval for the commercial portion of CIA includes a realization rate of 1.0. The realization rates for the major elements of these three programs range from about 0.70 to 0.87. The rates for peak impacts are about 10-15 percent below the energy impacts, ranging from 0.62 to 0.74.

Resource Planning Implications

In this section, we examine the resource planning implications of PG&E's evaluation results from three perspectives: program cost effectiveness, long-run DSM forecasts of energy and peak savings, and systemwide capacity balance. The baseline for our analysis is PG&E's long-run DSM forecast and resource plan contained in the *1992 Electricity Report (ER 92)*— California's biennial electricity policy and planning report (CEC 1993b). Before we proceed, however, we return briefly to the evaluation issues we discussed earlier.

Addressing Evaluation Issues for PG&E

Our attempt to apply PG&E's evaluation results to long-run resource planning is the basis for our earlier discussion of evaluation issues. Our appreciation for the complexity of many of these issues deepened as we struggled to move from evaluation results to long-run planning applications. By no means have we resolved all these evaluation issues for PG&E. Nevertheless, we summarize below our considerations of the major issues.

Table 1. Summary of Recent PG&E Program Evaluation Results

Program	Program Element	Unit of Analysis	Program Estimate	Evaluation Results	Realization Rate
Res Weather	Ceiling Insulation	Participation Level		3,623 (avg)	
	elec ht/cool only	Impacts Per HH(g)	1,252 kWh	1,201 kWh	0.96
	(PG&E 1993b)	Impacts Per HH(n)	1,252 kWh	1,021 kWh	0.82
		Program Impacts(n)	4,535 MWh	3,699 MWh	0.82
	Showerhead	Participation Level	5,763	3,954	0.69
	(PG&E 1993c)	Impacts Per HH(g)	832 kWh	238 kWh	0.29
		Impacts Per HH(n)	418 kWh	20 kWh	0.05
		Program Impacts(n)	2,409 MWh	78 MWh	0.03
Res Appl Eff	Compact	Participation Level		84,709	
	Fluorescents	Impacts Per HH(g)	200 kWh	256 kWh	1.28
	(PG&E 1993d)	Impacts Per HH(n)	200 kWh	163 kWh	0.82
		Program Impacts(n)	16,960 MWh	13,807 MWh	0.81
	Central A/C	Participation Level		6,320	
	(PG&E 1992a)	Impacts Per HH(g)	528 kWh	331 kWh	0.63
		Impacts Per HH(n)	528 kWh	331 kWh	0.63
		Program Impacts(n)	3,337 MWh	2,092 MWh	0.63
Res New Constr	Aggregate	Participation Level		9,534	
	(PG&E 1993e)	Impacts Per HH(g)		692 kWh	
		Impacts Per HH(n)	934 kWh	671 kWh	0.72
		Program Impacts(n)	8,900 MWh	6,400 MWh ± 1,300	0.72
Res Dir Assist	Weatherize	Participation Level		55,815	
	(PG&E 1993f)	Impacts Per HH(g)		182 kWh ±38	
		Impacts Per HH(n)	233 kWh	167 kWh	0.72
		Program Impacts(n)	13,011 MWh	9,325 MWh	0.72
	Appliances	Participation Level	20,959	18,190	0.87
	elec only	Impacts Per Appl(g)	288 kWh	286 kWh	0.99
	(PG&E 1993g)	Impacts Per Appl(n)	288 kWh	200 kWh	0.69
		Program Impacts(n)	6,038 MWh	3,647 MWh	0.60
Res EM Services	Weatherize &	Participation Level		68,699 (Msrs)	
	CFLs	Impacts Per Msr(g)			
	(PG&E 1993h)	Impacts Per Msr(n)	173 kWh	137 kWh	0.79
		Program Impacts(n)	11,858 MWh	9,415 MWh ± 3,765	0.79
CIA Retrofit	Commercial	Program Impacts(n)	441,775 MWh	383,137 MWh	0.87±0.14
Rebate	Industrial	Program Impacts(n)	158,246 MWh	133,614 MWh	0.84±0.13
	Agricultural	Program Impacts(n)	195,349 MWh	146,512 MWh	0.75±0.23
	(PG&E 1993a)				
Comm New Constr	Aggregate	Participation Level		114	
	(PG&E 1993i)	Impacts Per Bldg(g)		227,770 kWh	
		Impacts Per Bldg(n)	246,947 kWh	173,561 kWh	0.70
		Program Impacts(n)	28,152 MWh	19,786 MWh	0.70
NonRes EM	Aggregate	Participation Level	22,940 (avg)	15,529 (avg)	0.68
Service	(PG&E 1993j)	Impacts Per Prtp(g)	4,295 kWh	5,320 kWh	1.24
		Impacts Per Prtp(n)	2,345 kWh	2,905 kWh	1.24
		Program Impacts(n)	53,802 MWh	45,114 MWh	0.84

Evaluation Availability. We include our assessments on evaluation availability in our preceding discussion of evaluation results.

Evaluation Quality. We judged all 11 evaluations to be of sufficient quality to revise long-run planning assumptions. Our judgement is corroborated by PG&E and California regulators.[6] We decided to use results from all the studies but the showerhead element of the Residential Weatherization Incentives program. Results from this study are disappointing enough to convince PG&E to discontinue this program. The remaining 10 evaluations include DSM programs, or portions of DSM programs, that are most important for PG&E's long-run DSM forecast. The exception is the Residential Appliance Efficiency program. Efficient refrigerators and air conditioners are the most important measures in this program, but the evaluations we review cover air conditioners, ceiling insulation and compact fluorescent lights. We apply the available results for this program with reluctance.

Historical or Future Forecast. We apply the evaluation results to PG&E's long-run forecast of DSM impacts from future DSM programs. The forecast period runs from 1993 to 2011. Our analysis only affects certain portions of PG&E's total forecast. The portion of PG&E's forecast we leave unchanged include impacts from future statewide building efficiency standards, non-firm rates, group load curtailment, and residential pool pumps. The part of the DSM forecast we revise is always the dominant contributor to energy savings and, from 1995 onwards, forms the largest part of the peak savings. Still, the elements we leave unchanged represent about 30-35 percent of PG&E's forecasted peak savings from future programs and are thus important to PG&E's overall peak savings forecast. Finally, we do not revise forecasted impacts from historical DSM programs implemented during the 1990-92 period covered by PG&E's evaluations.[7]

Reconciliation and Application Issues. We found the program-level results from PG&E's evaluations to be plausible. PG&E made its *ER 92* DSM forecast at the program level and this facilitated our application of the evaluation results. The major issue we faced is transferring the results over time from historical programs to forecasts of future programs. We attempted to verify baseline savings estimates between PG&E's *ER 92* DSM forecast and evaluations. We were not able to verify conclusively that the respective program savings estimates rely on consistent baselines; the pre-evaluation program estimates reported in the evaluation are not accompanied by discussion or reference to either PG&E's long-run DSM forecast or the utility's annual DSM reports to the CPUC. To move forward with our analysis, we assume the baselines are consistent. This is a major assumption and our results must be gauged accordingly. Further, we did not alter PG&E's assumptions about savings persistence already included in its forecast. Persistence is the subject of an ongoing statewide study.[8] Finally, we assume that program design is held constant over time. We assume the designs leading to the realization rates from the historical program evaluations are also applied to the future programs.

Program Cost Effectiveness

Program cost effectiveness must be reassessed in light of these evaluation results. The benefits in the Total Resource Cost (TRC) test are avoided supply costs. Avoided supply costs for energy efficiency programs should be calculated using net program savings (CEC/CPUC 1987). Realization rates different than 1.0 mean that realized net savings differ from the forecasted savings used to represent program benefits in the TRC test.

We looked at how realization rates from PG&E's evaluations affect TRC test estimates PG&E reports for its *ER 92* DSM forecast (PG&E 1992b). Our assessment suggests all the programs that are cost effective in PG&E's *ER 92* forecast continue to be cost effective after the application of the evaluations' realization rates. This result is not surprising. PG&E's major DSM programs have benefit-to-cost ratios well in excess of one. The exception is the Residential New Construction program; this program's benefit to cost ratio begins to approach one when we account for the program's realization rate.[9] On the basis of our reassessment of program cost effectiveness, we did not remove any programs from PG&E's forecast.

DSM Forecast Impacts

Figure 1 shows the overall impact on the nondispatchable portion of PG&E's DSM forecast. The realized DSM forecast is lower than the DSM forecast used in *ER 92* resource planning. The realized energy impacts are about 18 percent lower while the realized peak impacts are 24 percent lower. The realized forecast as a fraction of the original forecast remains quite constant through 2011.

By 1998, the realized DSM forecast is about 710 GWh and 260 MW less than the original forecast. By 2003, or 10 years into the planning period, these differences grow to about 1200 GWh and 460 MW. At the end of the planning period in 2011, the differences amount to 1610 GWh and 640 MW. To put these numbers in some context, PG&E's forecasted gross peak demand (which includes losses) is 19,300 MW by 2003. The difference between the original and realized DSM forecasts means PG&E's peak demand (including losses) by 2003 will be about 2.7 percent higher than forecast. This figure represents about a single year's average demand growth for the PG&E

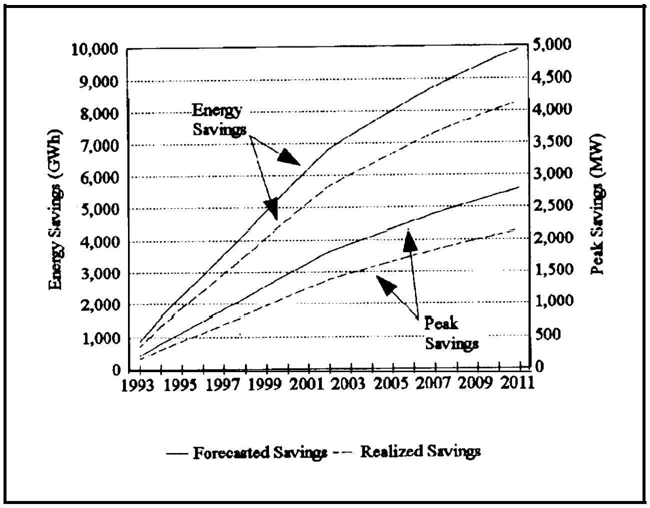

Figure 1. Original and Realized DSM Forecast for PG&E

system. We do not view these differences between the realized forecast and the original forecast as particularly dramatic. Long-run forecast differences of this magnitude, however, can affect resource decisions made at the margin. The more important issue, therefore, is to assess the effect of the realized DSM forecast on PG&E's resource plan from *ER 92.*

Capacity Balance

A sophisticated approach to examining the long-run planning consequences of a change in the DSM forecast is to include the new forecast in a capacity expansion analysis. The results of the new capacity expansion analysis are then be passed to a revenue requirements analysis. The results of both these analyses will show changes in the type, timing, and amount of resource additions and accompanying changes in revenue requirements for the new resource plan compared to the baseline.

Our exercise is not sufficiently rigorous to warrant such a sophisticated analysis, particularly due to the unresolved issue of baseline consistency. We can predict qualitatively one outcome of a more sophisticated analysis. In general, DSM is the most cost-effective resource in PG&E's resource plan. Thus, realizing less DSM will lead to either: running the existing generation system more; purchasing more energy and capacity from other power producers; or acquiring additional capacity. Any of these three options will likely lead to higher production costs for the plan with fewer cost-effective DSM resources.

We use a coarser indicator of the impact of the realized DSM forecast on long-run planning: changes in the capacity balance of PG&E's *ER 92* resource plan. Capacity balance is an important indicator of a utility system's reliability and its need for new resources to maintain reliability targets. Capacity balance is assessed against the target reserve margin a utility deems needed to reliably meet system demand and contractual export requirements. The capacity balance is the difference between total system resources and total capacity requirements. A positive capacity balance means the utility has resources in excess of its target reserve margin and, therefore, does not have a need to add resources to maintain system reliability. A negative capacity balance indicates the utility does not have sufficient resources to meet its target reserve margin. Utilities will add resources at least until the system's target reserve margin is met. We note that even if a utility has a capacity surplus, the utility may still achieve lower total system costs by adding cost-effective resources.

We examined the effect of the realized DSM forecast on PG&E's capacity balance under two conditions. In the first case, we assume PG&E follows the least-cost expansion path as indicated in its *ER 92* resource plan. In this case, PG&E begins to add new generating resources to its system by 1997 and eventually adds over 1500 MW of new capacity by 2011. PG&E's existing system will also be augmented by cost effective out-of-state power purchases, completion of pending long-run power purchases and facility upgrades, and future DSM. In the second case, we assume PG&E's system continues to be augmented by the above-mentioned resources, but does not add any new generating capacity. Thus, the first case assumes the utility will take future action to add new generation while the second case assumes the utility only takes action to follow through on its present commitments.

Figure 2 shows the capacity balance as either a surplus or a deficit. In the first case (with new additions), PG&E remains in a surplus throughout the planning period. The realized surplus is lower than the forecasted surplus, by about 48 percent in the year 2000, for example. This

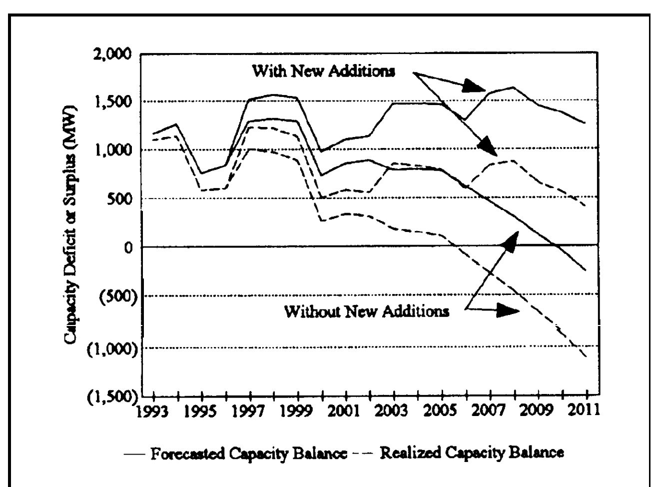

Figure 2. Effect of DSM on PG&E's Capacity Balance

suggests that PG&E's system reliability is not threatened by the reduction in future achieved DSM.

In the second case (without new additions), we note that even with the original DSM forecast PG&E has a capacity deficit by 2010. Under the realized DSM forecast, this date advances to 2006. PG&E will have to add resources 4 years earlier to maintain system reliability. By the year 2000, the surplus with the realized forecast falls below the level we observe at any time with the realized forecast in case one. This result suggests PG&E may find it cost effective to acquire new resources earlier than under the *ER 92* plan. In that plan, PG&E adds resources in 1997 (221 MW), 1998 (22.5 MW), and 2003 (435 MW). The 2003 addition, which was far enough in the future under *ER 92* assumptions that PG&E could prudently defer commitments, may be advanced enough under the realized DSM forecast to warrant more serious examination. PG&E is currently making this examination in proceedings for the *1994 Electricity Report* using revised resource characteristics and a DSM forecast that includes information from the DSM evaluations we review here.

Conclusions

Because of the importance of energy efficiency to California energy policy, utilities and regulators are committed to DSM program evaluation and, more recently, the application of evaluation results to inform planning and policy. We review 11 recent evaluations for PG&E that cover most of the utility's major DSM programs. Most of the evaluations report realized savings below the original program estimates. For residential programs, the realized energy savings are 20 to 40 percent less than originally predicted. For commercial programs, the realized savings are 13 to 30 percent below the initial program estimates. The realized peak impacts from residential program have a much wider range than do the energy impacts. The realized nonresidential peak savings are 26 to 38 percent under the initial program estimates.

Our attempt to assess the resource planning and policy consequences of these evaluation results led us to five major evaluation issues that we believe planners, forecasters, and policy makers must address: evaluation availability; evaluation quality; historical or future program application; reconciliation; and application. In our view, these issues must be considered whenever evaluation results are transferred to forecasting and planning arenas. Because we are not able to resolve all these issues for PG&E, we can only characterize our results as preliminary and suggestive of the possible long-run planning implications of the most recent evaluations. Most importantly, we are not able to resolve consistency between program estimates in the evaluations and PG&E's DSM forecast. This is an important application issue. Further, we are unable to predict the persistence of program savings over time.

Assuming that the evaluation results are directly transferable to PG&E's DSM forecast, our analysis indicates that PG&E's realized DSM forecast from *ER 92* will fall below PG&E's original forecast. The energy savings and the peak savings forecast will be 18 percent and 24 percent, respectively, below the DSM forecast in PG&E's resource plan for *ER 92*. The evaluation results do not suggest that any currently cost-effective DSM program is uneconomical. In addition, we find PG&E will experience capacity deficits sooner under the realized DSM forecast. These projected capacity deficits are still over a decade away and, therefore, pose no near-term threat to PG&E's system reliability.

Acknowledgments

We thank Haydee Hampton and Susan Buller of PG&E for their assistance in sharing information about PG&E's DSM forecast and evaluations, respectively. Any errors we have made in the application of evaluation results to PG&E's forecast are solely our responsibility. We also thank Penne Purcell and Dennis Smith of the CEC for providing material from the *ER 92* proceedings. An earlier draft of this paper benefitted from reviews by Marilyn Brown and Eric Hirst.

Endnotes

1. See pp. 2-5 to 2-9 in CEC (1993a) for an example.

2. Program savings realization rates are determined by the ratio of post-evaluation savings to pre-evaluation savings estimates. Savings estimates are always made with reference to a baseline. In most cases the baseline used to estimate savings before program implementation will be identical to the baseline observed by the evaluation. For example, existing building and equipment energy efficiency standards usually define the baseline efficiency. Program evaluations may also discover, however, that the baseline in the field is not the same as used to estimate pre-program savings. In these cases, forecasters must determine that the baseline used in the forecast is consistent with the evaluation's observations.

3. See California Public Resources Code, Sections 25000 et seq.

4. See pp. 27-31 and 59-60 of CPUC (1993a).

5. We will make a summary of peak impact results available to interested readers.

6. See Attachment 1, page 4, Table 1 of CPUC (1993b).

7. Revising impacts for programs implemented from 1990-92 will not only affect the demand forecast in these years, but in the years beyond 1992 as well. The extent of the effect beyond 1992 will depend on assumptions made in the forecast about the lifetimes of measures installed during the 1990-92 program years.

8. See Appendix B, pp. 3-4 of CPUC (1993a).

9. The implementation of new statewide building standards in California and more recent program experience suggest this program may no longer be cost effective.

References

California Energy Commission. 1993a. *Staff Testimony on Uncommitted Demand-Side Management Forecast Issues.* CEC Docket No. 93-ER-94, testimony filed November 17, 1993, California Energy Commission, Sacramento, CA.

California Energy Commission. 1993b. *Electricity Report and Appendices.* P104-92-001 and -001A, California Energy Commission, Sacramento, CA.

California Energy Commission. 1994. *Data Base for Energy Efficient Resources.* Beta Version 2.2, California Energy Commission, Sacramento, CA.

California Energy Commission and California Public Utilities Commission. 1987. *Standard Practice Manual.* P400-87-006, California Energy Commission, Sacramento, CA.

California Public Utilities Commission. 1993a. *Protocols and Procedures for the Verification of Costs, Benefits, and Shareholder Earnings from Demand-Side Management Programs.* D. 93-05-063, December 1, 1993, California Public Utilities Commission, San Francisco, CA.

California Public Utilities Commission. 1993b. *Interim Opinion on Demand-Side Management Shareholder Incentives: Threshold Issues.* D. 93-09-078, California Public Utilities Commission, San Francisco, CA.

California Public Utilities Commission. 1994. *Demand-Side Management Reporting Requirements Manual, 5th Edition.* California Public Utilities Commission, San Francisco, CA.

Independent Energy Producers. 1993. *Testimony of Douglas A. Houston on Behalf of the Independent Energy Producers Association on Uncommitted DSM Issues.* CEC Docket No. 93-ER-94, testimony filed November 17, 1993, California Energy Commission, Sacramento, CA.

Nadel, S., and K. Keating. 1991. "Engineering Estimates vs. Impact Results." *Proceedings of the 1991 Energy Program Evaluation Conference,* pp. 24-33.

Pacific Gas and Electric Company. 1992a. *Evaluation of Impact from PG&E's 1991 Appliance Efficiency Program Central Air Conditioner Rebates.* RAE-92-Z08, Pacific Gas and Electric, San Francisco, CA.

Pacific Gas and Electric Company. 1992b. *PG&E's CEC Electricity Report Demand-Side Management Forms, Revised.* CEC Docket No. 90-ER-92, February 5, 1992, Pacific Gas and Electric, San Francisco, CA.

Pacific Gas and Electric Company. 1993a. *Evaluation of the CIA Retrofit Rebate Program Final Report.* CIA-93-XO1, Pacific Gas and Electric, San Francisco, CA.

Pacific Gas and Electric Company. 1993b. *Impact Evaluation of PG&E's Ceiling Insulation Rebate Program.* RAE-93-RCG01, Pacific Gas and Electric, San Francisco, CA.

Pacific Gas and Electric Company. 1993c. *Impact Evaluation of the 1992 Energy Saver Showerhead Coupon Program.* RAE-93-H03, Pacific Gas and Electric, San Francisco, CA.

Pacific Gas and Electric Company. 1993d. *PG&E's 1992 Residential Compact Fluorescent Lighting Program Measurement Results.* RAE-93-H02, Pacific Gas and Electric, San Francisco, CA.

Pacific Gas and Electric Company. 1993e. *Residential New Construction 1992 Impact Evaluation.* RNC-93-Q01, Pacific Gas and Electric, San Francisco, CA.

Pacific Gas and Electric Company. 1993f. *Evaluation of PG&E's 1991 Energy Partners Program.* REQ-93-S01, Pacific Gas and Electric, San Francisco, CA.

Pacific Gas and Electric Company. 1993g. *Impact Evaluation of the Targeted Customer Appliance Program.* REQ-93-C02, Pacific Gas and Electric, San Francisco, CA.

Pacific Gas and Electric Company. 1993h. *Impact Evaluation of PG&E's 1991 Residential Energy Savings Plan Program.* REQ-93-Q03, Pacific Gas and Electric, San Francisco, CA.

Pacific Gas and Electric Company. 1993i. *Non-Residential New Construction Impact Evaluation Study.* CNC-93-A01, Pacific Gas and Electric, San Francisco, CA.

Pacific Gas and Electric Company. 1993j. *Impact Evaluation of 1990-1992 Non-Residential Energy Management Services Programs.* CEQ-93-A01, Pacific Gas and Electric, San Francisco, CA.

Southern California Edison Company. 1993. *Impact Evaluation of SCE's Customer Assistance Programs.* Prepared by Barakat & Chamberlin, Inc., Oakland, CA.