DSM IN CONTEXT: UNDERSTANDING THE VALUE OF DSM AND THE VALUE OF DSM PROGRAM EVALUATION

RICHARD SONNENBLICK, HADI DOWLATABADI & LESTER LAVE
DEPARTMENT OF ENGINEERING AND PUBLIC POLICY,
CARNEGIE MELLON UNIVERSITY
PITTSBURGH, PENNSYLVANIA 15213

15. Annual North American conference, International Association of Energy Economics, Seattle, WA

Abstract

Over the past two decades, demand-side management programs have emerged as a major element of electric utility integrated resource plans; EPRI and several PUC commissioners have concluded that DSM has saved utility and ratepayers many dollars. While DSM holds great potential as a means of abating greenhouse gas emissions and reducing acid rain, many programs have been ineffective and inefficient. If DSM is to serve the public welfare and attain its potential, extensive evaluation is required of the effectiveness of the programs and of their effects on the net energy bills of participants and non participants, as well as of the resulting short-and long-term changes in customer energy use. Since evaluation is expensive, a critical decision concerns the extensiveness of the evaluation. We discuss several benefits of evaluation, e.g., reducing the variance of demand forecasts and thus the need for new capacity, in order to determine the optimal level of the program.

We model the effect of a commercial DSM program on the price of energy and use short term elasticities to estimate energy consumption for program participants and non participants. We compare these consumption estimates to estimates of program savings to assess the magnitude of this effect and the importance of choosing the appropriate evaluation method.

Introduction

The oil price shocks and 1973 and later made evident the need for increasing the efficiency of energy use in the USA. Thousands of energy conservation programs have been created and billions of dollars have been invested in demand-side management (DSM) programs.[1] Initially, the perceived need to reduce energy use negated the desire for evaluation. However, by the early 1980s, DSM programs were sufficiently extensive and expensive that evaluation could no longer be ignored. Early evaluations actually produced no data on whether the DSM program saved kilowatt hours or were managed well.[2]

The last decade has seen an infusion of techniques from economics, social psychology, organizational behavior, and engineering into DSM planning, implementation, and evaluation. This deluge of methods and perspectives has

DISCLAIMER

expanded the array of tools available, but there has been little effort to be systematic in using the tools where they would contribute the most.

Evaluation is expensive. A utility and their PUC must determine how much of the DSM resources should be allocated to evaluation. Of those utilities that evaluate their programs, many utilize a rule-of-thumb and spend 5% to 10% of DSM resources on evaluation.[3] But this rule-of-thumb is not based on a systematic assessment of each program's evaluation needs, or on the value of the information obtained through evaluation.

Deciding on evaluation expenditures requires an analysis of what information is to be gained and how much reduction in uncertainty is likely to come from more extensive evaluation. Determination of the appropriate level of evaluation for a DSM program requires an understanding of the uses of evaluation information. In this paper, we analyze two uses for program evaluation information in order to examine the impact of using different evaluation methods and varying levels of evaluation rigor.

There are two inter-related dimensions of uncertainty with which utility and DSM professionals must be concerned: Uncertainty surrounding program performance and uncertainty surrounding the evaluation of program performance. In this paper, we focus on the information content of evaluation programs in order to determine what level of evaluation is needed. Using information from the Database of Energy Efficiency Programs (DEEP) at Lawrence Berkeley Laboratories, we will eventually incorporate statistical data on program performance into the model so that we can compare the magnitudes of and relationships between program performance uncertainties and program evaluation uncertainties.

Calculating Rate Impacts

DSM attempts to reduce the need for new capacity, or to shut existing capacity, through programs that provide consumers with information, conduct audits to determine where savings are profitable, or subsidize the purchase of energy efficient equipment are activities that require resources which, other factors held constant, will add to utility costs and increase electricity price.

Only a portion of rate payers desire or receive DSM services. Those who do may benefit handsomely. Customers who do not receive the services or who do not find them helpful must bear the costs with no direct benefits. Nonetheless, these customers may receive indirect benefits. For example, if all electricity is supplied by a hydroelectric installation at a cost of $0.01 and a price of $0.04, the low price will stimulate extensive usage with little thought for high efficiency. However, if new capacity is required at a cost of $0.07 and a price of $0.10, all customers will face a large price increase when the new plant is added. If, instead, the utility can generate "negawatts" at a price of $0.05, even those who don't benefit from DSM directly will benefit indirectly.

Some care needs to be taken in interpreting this point. Nonusers would still find their electricity price has increased, but it would increase less than if the DSM program had not been initiated -- and the new generating facility were built. Thus, in the sense of the next best alternative, higher electricity prices due to DSM may leave even nonparticipants better off. Participants presumably would be better off, despite a higher price, for two reasons. The first is that they would be using fewer kilowatt hours and so their bill might decline. The second is that price would have increased still more if the DSM program had not been implemented.

Some utilities and PUCs have taken this logic to imply that any DSM program is beneficial, even to nonparticipants. This conclusion is fallacious. If the DSM program is ineffective (does not produce negawatts), costs rise with no offsetting benefits. Alternatively, the program can produce negawatts, but at costs so high that it would be cheaper to build the new capacity.

An undesired effect (at least according to the utility) of a DSM program would be to reduce demand below current levels. If this occurs, the utility will lose the net revenue that represents the difference between price and marginal generation cost. For example, if the DSM program were so successful that it reduced net revenue to the utility by $10 million, price would have to rise to cover this difference.

This logic suggests a DSM evaluation criterion shown in table 1. If the benefits exceed the sum of the costs, society gains by the DSM program. If not, the program should not be implemented. Unfortunately, the DSM program can be worth while for society, but still represent a net-cost to nonparticipants. For example, the cash incentives to participants represent a transfer from the nonparticipants to the participants. This transfer can changes the distribution of benefits and can change them sufficiently to make nonparticipants worse off.

Table 1. Costs and benefits used in rate impact calculation

Costs	Benefits
DSM Management & Implementation	Avoided incremental capital costs
Subsidies to participants	Environmental quality gains, etc.
Possible reduced coverage of fixed costs	

Massachusetts Electric Companies' Enterprise Zone Small Commercial and Industrial Program operated from Fall, 1985 to Spring, 1987. The program audited and installed energy efficient lighting measures free of charge to eligible commercial and industrial customers. Approximately 1400 customers received energy audits, and 775 had new lighting equipment installed. Program evaluation was extensive. Billing data, building data, and end-use characteristics were compiled from program participants.[4] Nadel assessed program savings using several different methods:[5]

1. An engineering estimate based on laboratory measures of old and new lighting equipment efficiencies.
2. A cross-section comparison of pre- and post-program electricity usage by participants and matched non participants using a two-sample t-test.
3. A time-series comparison of pre- and post-program electricity usage by participants, using survey information to exclude from the comparison buildings that had undergone changes in utilization.
4. A conditional demand analysis that regressed the presence of individual energy end-uses (obtained for each building during the audit) against pre- and post-program energy use in participating buildings.

The time series comparison was complicated by data coordination and survey problems, and thus the results obtained are of dubious significance. The other three methods provided estimates within 6% of each other. Table two lists the per customer annual savings estimates together with rate impacts, short run changes in annual energy use for program participants and non participants due to price impacts, and long run changes in annual energy use for the average customer.

Table 2. Savings estimates from the Enterprise Zone Program

Evaluation Method	Average Annual Savings for DSM H/holds (kWh)	Total Program Savings for DSM H/holds (kWh)	Rate Impact (mils/kWh)	Change in Customer Demand due to price effect (%)	Total Price-Induced Demand Reduction (kWh)
Engineering	6600	$5.1x10^6$	1.4	-0.47%	$4.7x10^6$
Cross-section comparison	6900	$5.3x10^6$	1.4	-0.48	$4.8x10^6$
Time-series comparison	2300	$1.8x10^6$	1.3	-0.44	$4.4x10^6$
Conditional Demand	6500	$5.0x10^6$	1.4	-0.47	$4.7x10^6$

Note a) The price fluctuations that occur when conservation program costs and lost revenue are ratebased affects customer consumption behavior. For this analysis, we are utilizing the short run price elasticities of -0.28 for the commercial sector.[6] We utilize the Allen own-price model to calculate final energy consumption.

Rate impacts do not differ significantly across evaluation methods because the rate calculations are dominated by the cost of the program (approximately $1.2M). Using the short run price elasticities of demand, we estimate that the change in energy consumption due to price changes is much less than program savings estimates themselves, which were 9% of participant energy consumption.

Because all ratepayers, not just program participants, are affected by the change in price, the overall decrease in demand for the customer base is on the order of the total annual program savings. Although this model is simplistic in that it does not account for regulatory lag in ratebasing conservation costs, or other factors that effect consumption such as weather, this result demonstrates the potential of DSM programs to create significant second order effects. In this case, however, these second order effects could be estimated with any of the evaluation methods explored.

Characterizing the Value of Program Evaluation Information

Costs of Evaluation

The data gathering and analysis requirements of evaluation can be expensive. The most expensive component of any evaluation is data acquisition. DSM programs, unlike supply-side investments, are highly decentralized. The collection of qualitative and quantitative data is a painstaking, labor-intensive process. Pacific Gas and Electric's estimates of their per-participant costs for implementing surveys are shown in table four.[7]

Table 4: Estimates of data acquisition costs[8]

Survey Type	Per Participant Cost ($1991)
Mail or Phone Survey	$65-100
On-site Survey	$135-675

These estimates of evaluation cost are used to estimate the costs of evaluations described in the next section. Evaluation costs will vary from program to program since management, program, and market characteristics affect data acquisition and analysis costs. Foresight and planning can often reduce data collection costs. Auditors can collect pertinent data while auditing customer's structures. Evaluation planners can coordinate with utility database departments to ensure that the data already being collected is usable for program tracking and evaluation analyses. Sample sizes in evaluations range from 10 to 2000 depending on the objective of the evaluation and the type of program being evaluated.

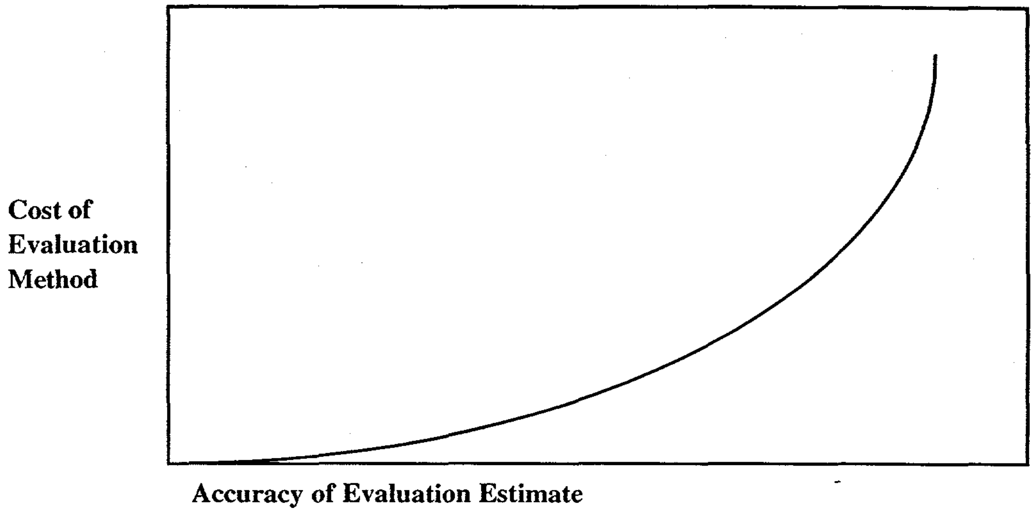
RCG/Hagler Bailly estimates the cost of evaluating a medium sized program for the commercial sector (i.e., 2,500 participants and a design to save 10% of total energy use) weatherization program to be as follows:[9]

Table 5. Estimates of evaluation cost

Evaluation Description	Approximate Cost
Back-of-the-envelope engineering calculation	$15,000
Adding billing information to compare pre- and post-program consumption	$75,000
Multivariate regression with socio-economic indicator variables and controls for self-selection (requires more data, surveys, site visits)	$150,000
Adding extensive load research and end-use metering	$300,000

For most programs, the techniques in table five increase in accuracy as they increase in cost. One can conceptualize a cost of reducing uncertainty curve. Points on the curve represent different evaluation methods. Uncertainty in each method could be measured using the variance of statistical models, sample error, or by comparison with the most robust method's result. Such a curve is sketched in figure one.

Figure 1. Hypothetical cost of evaluation accuracy curve



Evaluation methods that combine information from several types of evaluations (sometimes called triangulation methods) may provide savings estimates superior to any single technique. Due to the myriad evaluation methods available, both singly and in combination, a cost of accuracy curve that listed every possible evaluation method or combination of methods would be impossible to produce.

Any constructed cost of accuracy curve would be a step function, reflecting the discrete and nonlinear nature of different levels of evaluation. If this curve were calculated for different programs (by performing numerous evaluation techniques and plotting their uncertainty and cost in a stepwise manner) it would reveal that different programs achieve different improvements in accuracy from the same evaluation method. This is because different programs are subject to different combinations of weather and price fluctuations, free-ridership, etc. The marginal accuracy of successive evaluation techniques is determined by the existence and magnitude of factors for which each technique can control.

By combining the cost of reducing uncertainty curve with a representation of the benefits gained from a reduction in energy demand uncertainty, the optimum level of evaluation expenditures can be calculated. In this section, the optimum level of evaluation expenditures is calculated for a residential weatherization retrofit project.

Forecasting Demand

Load and capacity demand forecasts are adjusted to reflect estimates of DSM program savings. Typically, residential and commercial sector conservation program contributions to demand reductions are estimated using end-use engineering models of electricity use. Adjustments are made to these models based on information gained through program evaluations.

An alternate method is to base estimates of future demand reductions from DSM on evaluations of current full-scale and pilot DSM programs. Integrating DSM results into the planning process allows one to examine the impact of different evaluation methods on demand forecasts and plans for new capacity.

Electricity load and capacity planners are risk averse; underestimating demand and underbuilding capacity can lead to brownouts and unreliable service. In contrast, overestimating demand can lead to overbuilding system capacity and a disallowance of cost recovery mechanisms by government agencies.[10] For this case study, we make the simplifying assumption that in order to avoid underbuilding, planners will plan to the ninth decile of demand estimates.[11] Stated differently, they plan for capacity and load that will meet demand at the 90% upper confidence level of their demand estimate. Thus, the variance associated with an evaluation's estimate of demand reduction would directly affect planning decisions. If the standard deviation associated with a demand reduction estimate (obtained using a statistical multivariate regression model) is unnecessarily large, planners will plan for excess capacity and risk overbuilding.

Assessing Programs with Incomplete Evaluation Data

Although evaluation is best planned and performed concurrently with the DSM program and its pilot level implementations, this has not always been done. We can simulate a retrospective analysis to get estimates of the value of various evaluation and other variables. The objective of this exercise is to see which data are of greatest value in the evaluation so that the information content of a variable can be compared to its cost of collection in deciding whether to collect this information in future evaluations. In this approach, we would look at each variable and combination of variables in a regression to measure the bias and precision of the estimates. These two measures, bias and reduction of error variance, characterize the value of each variable.

Without complete program data, one could use data from the regression performed during a billing or conditional demand analysis to approximate this regression. This technique assumes one can estimate the change in variance when a regression variable is dropped by using the relation between t-values and partial correlation coefficients to calculate the R2 of the model with the omitted variable. The technique requires the coefficient means and their variances, the means of the dependent and independent variables, the residual sum of squares, the standard error of the estimate, and the coefficient of determination. Its key assumption is that explanatory variables are uncorrelated. Each variable contributes explanatory power orthogonal to the other explanatory variables. Thus, when a variable is dropped, the coefficients of the other variables do not change. By dropping variables from the regression, we also assume that the residual is unbiased.

In this section, a statistical technique is used to estimate the increase in the standard error of a regression model predicting energy consumption when explanatory variables are dropped from the regression model. A larger standard error will increase the value of the upper 90% confidence interval and guide utilities to plan for more capacity than is necessary. Table six describes two reduced forms of a time-series regression model used to calculate energy savings from a residential weatherization program. The costs of each evaluation level are based on estimates of data acquisition costs given earlier.

Table 6. Full and reduced forms of time-series regression models

Variables retained	Controls for	Cost of evaluation
POST, HDD, MPELECT	Weather, price of electricity	$30,000
POST, HDD, MPELECT, NUMRES, AGE25_34, EDUCHH, AGEHH	above plus demographic characteristics	50,000
POST, HDD, MPELECT, NUMRES, AGE25_34, EDUCHH, AGEHH, SQFTHT, WOOD, STOCK	above plus dwelling characteristics (full model specification)	75,000

Variable legend: POST—pre or post program time period, HDD—heating degree days, MPELECT—marginal price of electricity, SQFTHT—number of sq. ft. heated, WOOD—one if house uses wood stove, 0 if other, STOCK—number of major electric appliances, NUMRES—log(number of residents+1), AGE25_34—number of residents ages 25-34, EDUCHH, AGEHH—education level (four-way variable) and age of head of household.

These numbers are only rough estimates of evaluation costs. As discussed earlier, the cost of an evaluation depends on the degree of coordination between program planners, evaluators, and utilities. Actual costs would be less if, for example, program employees were instructed to compile the data during each home energy audit.

By dropping variables from the regression model, we approximate a regression model that omitted those variables. Such a model might be used in a less comprehensive evaluation due to a lack of data acquisition resources. Thus, we can calculate the effects of different levels of evaluation on planning decisions. A detailed description of the statistical method used to simulate a regression using the analysis of variance is available from the authors.

Table seven reports the results of using this technique on the time-series regression model used to estimate energy savings due to a residential audit program in Portland, Oregon, sponsored by the Portland General Electric Utility.[12] Each group of variables retained represents a distinct level of evaluation. The table presents the estimated energy demand (represented by an upper 90 or 95% confidence level) for the program population of 5,000 residents. If we assume that the model incorporating all available explanatory variables is the most accurate, then the difference in the estimated energy demand calculated by that model and any other model is unnecessary capacity, representing a loss to the utility. The magnitude of this loss is dependent on the utilities' current capacity and its cost of acquiring additional capacity.

Table 7. Change in confidence level when variables are omitted[13]

Variables retained	Root variance term of the upper confidence level estimate of energy consumption (kW)	
	90%	95%
Weather, billing, price	621	743
and demographic vars	605	723
and dwelling vars	583	697

Notes: Adding the root variance term to the mean capacity demand, 9525 kW, yields the actual confidence level estimate

The next section calculates estimates of additional generation costs due to increases in demand estimate variance using the information in table eight.

Estimating the Costs of Excess Capacity: The Value of Information

We assume that an increase in estimated demand results in a need to increase the size of available capacity. Thus, benefits of a reduced variance in demand estimates are a function of the incremental cost of peaking generation additions. If the lead-time of constructing a peaking generation plant is five years, the discounted value cost of the facility (which includes interest payments during the construction period) as of the date it begins generating electricity should be computed.[14] Table eight describes the three plants considered for incremental cost of peaking generation addition.

Table 8. Possible sources of additional generation[15]

Plant type	Capacity (MW)	Capital Cost ($/kW)	Design to Startup time (Years)	Adjusted Capital* ($/kW)
Gas turbine	40	443	< 1	532
Pulverized coal-spray dryer FGD	300	1368	five	1286
Integrated gasification/ Combined cycle-bituminous	200	1845	five	1735

Notes: *Adjusted cost is adjusted for reserve margin, and present valued if startup time is > 1 year (5% discount rate).

Gas turbine supply options might be used if capacity additions are required in the near-term, and the coal-based supply options might be used if capacity is required in several years. Table nine lists the calculated capacity requirements and resulting capital costs based on the 90% upper confidence estimates of demand.

Table 9. Estimates of additional capacity requirements and costs for three plant types

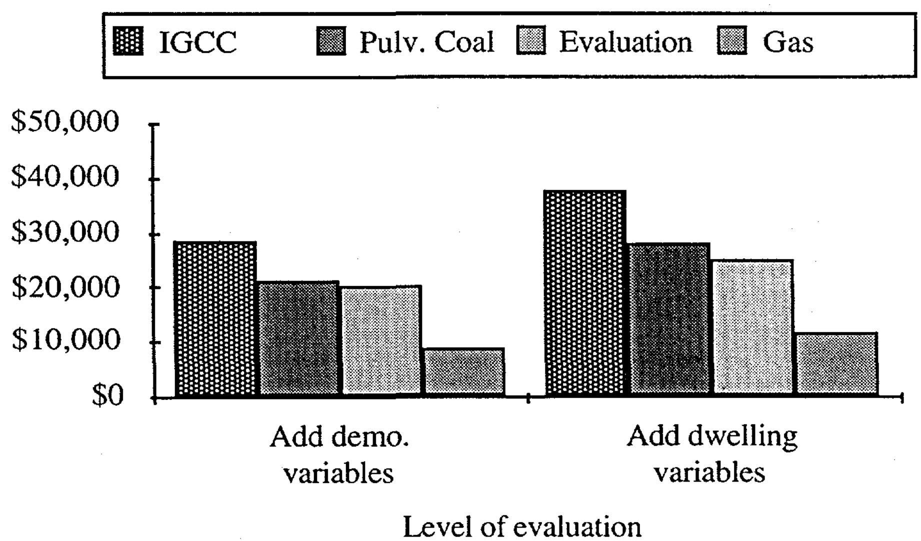
Variables retained	Root variance term of 90% upper confidence estimate (kW)	Cost of necessary additional capacity to meet 90% UCI ($,000's)		
		Gas Turb	Pulv. Coal FGD	IGCC
Weather, billing, price	621	$330	$799	$1,077
and demographic vars	605	$322	$777	$1,049
and dwelling vars	583	$310	$750	$1,011

The information in table nine and the cost of evaluation estimates in table six allow us to plot the marginal evaluation cost and marginal evaluation benefit (that is, the avoided cost of construction of additional capacity).[16]

Comparing Evaluation and Additional Generation Costs

Figure two shows the marginal cost of evaluation and marginal benefit of capacity reductions for all three types of power plants when planning to the 90% confidence level. Lines representing the supply facilities describe the reduction in facility construction costs due to an improvement in the accuracy of the forecasting model.

Figure 2. Marginal costs and benefits of evaluation: Planning at the 90% confidence level of demand



In figure two, the marginal cost of both levels of evaluation falls between the marginal benefit curves for the pulverized coal and gas turbine supply options.

Thus, the question of whether evaluation benefits exceed costs is dependent on the supply option used. Based on this use of evaluation information, evaluation provides net benefits when either coal-based supply option is planned.

Planning at the 95% level increases the benefit of evaluation for all three supply options. However, the marginal benefits gained if a gas turbine is being planned continue to be less than marginal evaluation costs.

The net benefit accrued from a given level of evaluation is dependent on the size of the DSM program being evaluated. Evaluation costs will be similar for a five million dollar DSM program and a fifty million dollar DSM program, but the resulting increase in accuracy of demand forecasts will be greater for the larger program. The program investigated here had 5000 participants. If 10,000 residents had participated, evaluation benefits would have exceeded evaluation costs for all three supply options.

The method outlined here is useful to determine what level of evaluation is best for a DSM program. Since there are other uses of evaluation information (improving subsequent program iterations, cost- effectiveness screening), even if the marginal costs of an evaluation exceed the marginal benefits due to reductions in planned capacity additions, a comprehensive evaluation may still be justifiable.

Conclusion

Myriad ineffective DSM programs, billions of dollars devoted to DSM, and the potential of more effective DSM to lower the need more new capacity are good reasons for more comprehensive and effective DSM evaluation. All evaluation is not the same: There are vast differences in the cost and information content of available evaluation programs. Utilities and PUC need to think carefully about the costs and benefits of each evaluation alternative.

As utilities progress from using bottom-up (engineering estimates) of program savings, they face challenging issues. The more advanced methods allow control for extraneous factors and can combine results from different evaluation techniques (engineering estimates, billing analyses, and end-use metering). But these advanced methods can be much more expensive than engineering judgment. The decision about which evaluation method to use requires a careful analysis of the information content of each method and the value to the utility of this information. Utilities should not leap from Chevettes to Cadillacs without a careful look at what they are getting for their money and whether the additional information will generate net savings.

Acknowledgments

We thank the Office of Policy at the U.S. Department of Energy for partial support of the research reported in this paper. The authors are responsible for any errors and opinions.

[1]Nadel, S.M., "Critical review of utility DSM programs", *Annual Review of Energy*, 1992, p. 508.

[2]Hemphill, R.F., Meyers, E.A., "Electric utility conservation programs: progress and problems", in *Energy Conservation*, Sawhill, J.C., Cotton, R., eds., Brookings, Washington, DC, 1986, p.152.

[3]Wirtshafter, R.M., Baxter, L.W., "Establishing priorities for future evaluation efforts", Proceedings of the 1991 International Energy Program Evaluation Conference, Chicago, Ill., pp. 137-142.

[4]Information on the program's process evaluation is in Hicks, Nadel, Mystakides, and White, *Evaluation Report on Massachusetts Electric Company's Enterprise Plan, Executive Summary*, Westborough, MA, New England Power Service Co., 1988.

[5]Nadel, S., "Electricity savings from a small C&I lighting retrofit program: Approaches and Results", Proceedings of the 1989 Energy Program Evaluation Conference, Chicago, Ill, pp. 107-112.

[6]Bohi, D., Zimmerman, M.B., "An Update on Econometric Studies of Energy Demand Behavior", *Annual Review of Energy*, 9, 1984, pp. 105-154.

[7]Callahan, S.J., "Energy audits—beyond the customer report", in *Meeting Energy Challenges*, Proceedings of the Great PG&E Energy Expo, Pergamon Press, 1985, p.507.

[8]Inflated from 1985 dollars (and rounded to the nearest $5) using the Services Cost index from the *Economic Report of the Office of the President.*

[9]RCG/Hagler, Bailly, Inc., *Impact Evaluation of Demand Management Programs*, EPRI CU-7179s vol 1., 1991.

[10]The cost of overbuilding is incurred by society regardless of the outcome of prudence review hearings. If cost recovery is allowed, customers bear the costs. If cost recovery is disallowed, the utility shareholders bear the costs.

[11]The selection of an appropriate confidence interval is dependent on a utilities' current supply and demand characteristics. The examples later in the paper calculate both 90% and 95% confidence intervals.

[12]It is not standard practice to report variable means, covariance matrices, and analysis of variance when describing the results of regressions evaluating energy conservation. Often, only point estimates are reported, without any supporting information whatsoever.

[13]The residual R^2 remaining when all regression variables were dropped was zero (10^{-6}). This suggests that the assumption of orthogonally-related regression variables may be permissible for this model.

[14]This computational scheme follows conventions established by the EPRI Utility Planning Methods Center, *EPRI Technical Assessment Guide*, EPRI P-6587-L, 1989, p. 8-3.

[15]EPRI Technical Assessment Guide, p. 7-56, 7-19, 7-41, 1988. All prices inflated to 1991 levels using the Nonresidential production of durable equipment index from the *Economic Report of the Office of the President.*

[16]To simplify the calculation, we calculate the benefit of not adding capacity rather than the benefit of deferring capacity. An example of the appropriate deferral calculation is in Busch, J.B., "Determining the value of conservation to Thailand's electric utility", Proceedings of the IEEE/PES International Joint Power Generation Conference and Exposition, San Diego, October 1991.