A REPORT FROM THE DATABASE ON ENERGY EFFICIENCY PROGRAMS (DEEP) PROJECT

**LBL-38201**
**UC-1322**

# Where Did the Money Go? The Cost and Performance of the Largest Commercial Sector DSM Programs

*Joseph Eto, Suzie Kito, Leslie Shown, and Richard Sonnenblick*

Energy & Environment Division
Lawrence Berkeley National Laboratory
University of California
Berkeley, California 94720

December 1995

The work described in this study was funded by the Assistant Secretary of Energy Efficiency and Renewable Energy, Office of Utility Technologies, Office of Energy Management Division of the U.S. Department of Energy under Contract No. DE-AC03-76SF00098. The DEEP Project has also received funding from the New York State Energy Research and Development Authority, the Bonneville Power Administration, the Rockefeller Family and Associates, the Electric Power Research Institute, and in-kind support from the Energy Foundation.

# Contents

Tables . . . . . iii

Figures . . . . . v

Abstract . . . . . vii

Acknowledgments . . . . . ix

Acronyms and Abbreviations . . . . . xi

Executive Summary . . . . . xiii

Chapter 1: Introduction . . . . . 1

Chapter 2: The Largest Commercial Sector DSM Programs . . . . . 3
- DSM Program Spending . . . . . 3
- DSM Program Characteristics . . . . . 5

Chapter 3: The Cost and Performance of Commercial Sector DSM Programs . . . . . 9
- Measuring the Total Resource Cost and Utility Cost of Energy Savings . . . . . 9
- The Total Resource Cost and Utility Cost of Energy Savings . . . . . 13
- Program Cost Effectiveness . . . . . 17
- Explaining Differences in the TRC of Energy Savings . . . . . 20

Chapter 4: Developing Consistent Program Cost Information . . . . . 25
- Measure Costs . . . . . 25
- Utility Nonmeasure Costs . . . . . 30
- The Treatment of Shareholder Incentives . . . . . 32

Chapter 5: Measuring the Energy Savings from DSM Programs . . . . . 33
- Classifying Methods for Measuring Annual Energy Savings . . . . . 33
- Estimating the Economic Lifetime of Savings . . . . . 38
- Measuring Free Riders . . . . . 42

Chapter 6: Summary . . . . . 45

References . . . . . 47

Appendix A: Program Selection and Data Collection . . . . . . . . . . . . . . . . . . . . . . . . . . 49
Appendix B: On Treating the Benefits and Costs from Early Replacement and Normal Replacement Retrofit Programs Consistently . . . . . . . . . . . . . . 53
Appendix C: The Transferability of Realization Rates . . . . . . . . . . . . . . . . . . . . . . . . . 55

# Tables

Table 2-1. 1992 Utility DSM Budgets . . . . . 4
Table 2-2. Summary of DSM Spending . . . . . 5
Table 2-3. DSM Program Characteristics . . . . . 7
Table 3-1. Cost and Savings Components of the TRC and UC of Energy Savings . . . . 10
Table 3-2. The Total Resource Cost and Utility Cost of Energy Savings . . . . . 14
Table 3-3. Comparison of the Most and Least Expensive Programs . . . . . 21
Table 3-4. Regression Equations for Total Resource Cost . . . . . 24
Table 4-1. Full vs. Incremental Measure Costs . . . . . 29
Table 4-2. Utility Administrative Cost Categories . . . . . 30
Table 5-1. Summary of Annual Energy Savings Methods . . . . . 34
Table 5-2. DSM Program Measure Costs as a Function of Savings Evaluation Method . . . . . 38
Table 5-3. Measure Lifetimes and Free-Ridership Rates . . . . . 40
Table 5-4. The Effect of Assuming Standard Measure Lifetimes on the TRC of Energy Savings . . . . . 42
Table A-1. Overview of DSM Program Selection Process . . . . . 50
Table B-1. Retrofit Program Evaluation: Normal vs. Early Replacement . . . . . 54

# Figures

Figure ES-1. The Total Resource Cost of Commercial Lighting Programs . . . . . . . . xiv
Figure 2-1. 1992 Utility Energy Efficiency DSM Spending . . . . . . . . . . . . . . . . . . . . . . 3
Figure 2-2. DSM Program Types vs. Savings by End Use . . . . . . . . . . . . . . . . . . . . . . 6
Figure 2-3. Utility and Participant Measure Costs Contributions . . . . . . . . . . . . . . . . . . 8
Figure 3-1. The Total Resource Cost of Commercial Lighting Programs . . . . . . . . . . . . 15
Figure 3-2. The TRC vs. UC of Energy Savings . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 16
Figure 3-3. The Cost Effectiveness of 1992 Commercial Sector DSM Programs . . . . . 18
Figure 3-4. TRC and RIM Benefit-Cost Test Ratios . . . . . . . . . . . . . . . . . . . . . . . . . . 20
Figure 4-1. Measure Costs vs. Incremental Measure Costs . . . . . . . . . . . . . . . . . . . . . 27
Figure 5-1. Economic Lifetime of Savings as a Function of the Mix
of End-Use Savings . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 41

# Abstract

We calculate the total resource cost (TRC) of energy savings for 40 of the largest 1992 commercial sector DSM programs. The calculation includes the participating customer's cost contribution to energy saving measures and all utility costs, including incentives received by customers, program administrative and overhead costs, measurement and evaluation costs, and shareholder incentives paid to the utility. All savings are based on post-program savings evaluations. We find that, on a savings-weighted basis, the programs have saved energy at a cost of 3.2 ¢/kWh. Taken as a whole, the programs have been highly cost effective when compared to the avoided costs faced by the utilities when the programs were developed. We investigate reasons for differences in program costs and examine uncertainties in current utility practices for reporting costs and evaluating savings.

# Acknowledgments

This report could not have been prepared without the information provided by and cooperation of the following individuals and their companies or agencies:

Sheldon Switzer, Robert Howard, and Marcie Forster - *Baltimore Gas and Electric*
Jerry Greer and Bruce Gretter - *Boston Edison*
Don Schultz - *California Public Utilities Commission*
Frank Congedo - *Central Hudson Electric and Gas*
Phil Hastings and Margie Force - *Central Maine Power*
Mark Quinlan - *Connecticut Public Utilities Commission*
Jane Reilly - *Consolidated Edison*
J. Latif Baig and Alecia Parker - *Consumers Power*
Dave Grimason and Todd Presson - *Green Mountain Power*
Phillip Moroney - *Long Island Lighting*
Richard Parker - *Maine Public Utilities Commission*
Marty Kushler - *Michigan Public Service Commission*
Donna O'Neil - *Montana Power*
Dorothy Conant - *New England Electric System*
Jim Gallagher, Bill Mills, and Harvey Tress - *New York State Department of Public Service*
Dorinda Noll - *New York State Electric and Gas*
Tim McClive - *Niagara Mohawk Power*
Peter Skywark - *Orange and Rockland*
Susan Buller and Ed Mah - *Pacific Gas and Electric*
Mark Kumm - *Potomac Electric Power*
Bill Hopkins - *Puget Sound Power and Light*
Linda Wilson - *Rochester Gas and Electric*
Marty Katz and Joan Rothmann - *Sacramento Municipal Utility District*
Kitty Corbin-Teora and Maureen Pflum - *San Diego Gas and Electric*
Cindy Preciado and Marian Brown - *Southern California Edison*
Leszek Kaspariowisz - *Texas Public Utilities Commission*
Bruce Blackburn and Charles Craven - *Texas Utilities Electric*
John Amalfi and Paul Kuehn - *Western Massachussetts Electric*

In addition, we would like to thank the following individuals for their comments on the draft of this report: Eric Hirst, Ted Flanigan, Chuck Goldman, Jon Koomey, Mark Kumm, Theo MacGregor, Peter Miller, Steve Nadel, Diane Pirkey, Ralph Prahl, Don Schultz, Harvey Tress, and Ed Vine. We would also like to thank Mithra Moezzi for helping us set up the regression analysis presented in Chapter 3.

The work described in this study was funded by the Assistant Secretary of Energy Efficiency and Renewable Energy, Office of Utility Technologies, Office of Energy Management Division of the U.S. Department of Energy under Contract No. DE-AC03-76SF00098. The DEEP Project has also received funding from the New York State Energy Research and Development Authority, the Bonneville Power Administration, the Rockefeller Family and Associates, the Electric Power Research Institute, and in-kind support from the Energy Foundation.

# Acronyms and Abbreviations

| | |
|---|---|
| DEEP | Database on Energy Efficiency Programs |
| DSM | Demand-side management |
| EIA | Energy Information Administration |
| EPRI | Electric Power Research Institute |
| GWh | Gigawatt-hour |
| HVAC | Heating, ventilation, and air conditioning |
| PUC | Public Utility Commission |
| SAE | Statistically-adjusted engineering |
| TRC | Total resource cost |
| UC | Utility cost |

# Executive Summary

Utility demand-side management (DSM) activities are at a crossroads. After five years of unprecedented growth, during which aggregate DSM spending increased nearly fourfold to almost $3 billion in 1994, utilities and public utility commissions are reexamining their roles and responsibilities in improving customer energy efficiency. Many issues need to be considered, including the magnitude and value of uncaptured energy efficiency opportunities, the extent of utilities' obligations to serve, and the maturity of the energy services infrastructure. There are also concerns that historic utility investments in energy efficiency have not been cost effective. They are based in part on the drop in avoided costs since the early 1990s. They are also based on concerns that DSM programs have cost more than originally anticipated.

This report presents findings from a major U.S. Department of Energy project to address these latter concerns.[1] We examine three central questions regarding DSM program performance: What have they cost? Have they been cost effective? What explains differences in cost? This report answers these questions by looking closely at the performance of 40 of the largest 1992, commercial-sector, DSM programs. Taken together, utility spending on the 40 programs accounted for nearly one third of total industry spending on energy efficiency in 1992. Despite rapid evolution in the designs of DSM program, we find many important lessons with continuing relevance for today's DSM programs.

Our primary measure of DSM program performance is the total resource cost (TRC) of energy savings. The TRC includes both utility and customer-paid contributions to the acquisition of an energy efficiency resource. It also includes the cost incurred by utilities to measure savings and any incentives received by the utility for the successful operation of its programs. All savings are based on some form of post-program savings evaluation.[2] We express the TRC in units of ¢/kWh so that it can be directly compared to a utility's avoided cost to determine the cost effectiveness of a program.

We find that, on a savings-weighted basis, the TRC for the programs is 3.2 ¢/kWh (see Figure ES-1). Our results confirm the importance of including customer cost contributions in order to determine the full cost of energy savings. Customer cost contributions account for 31% of the TRC.

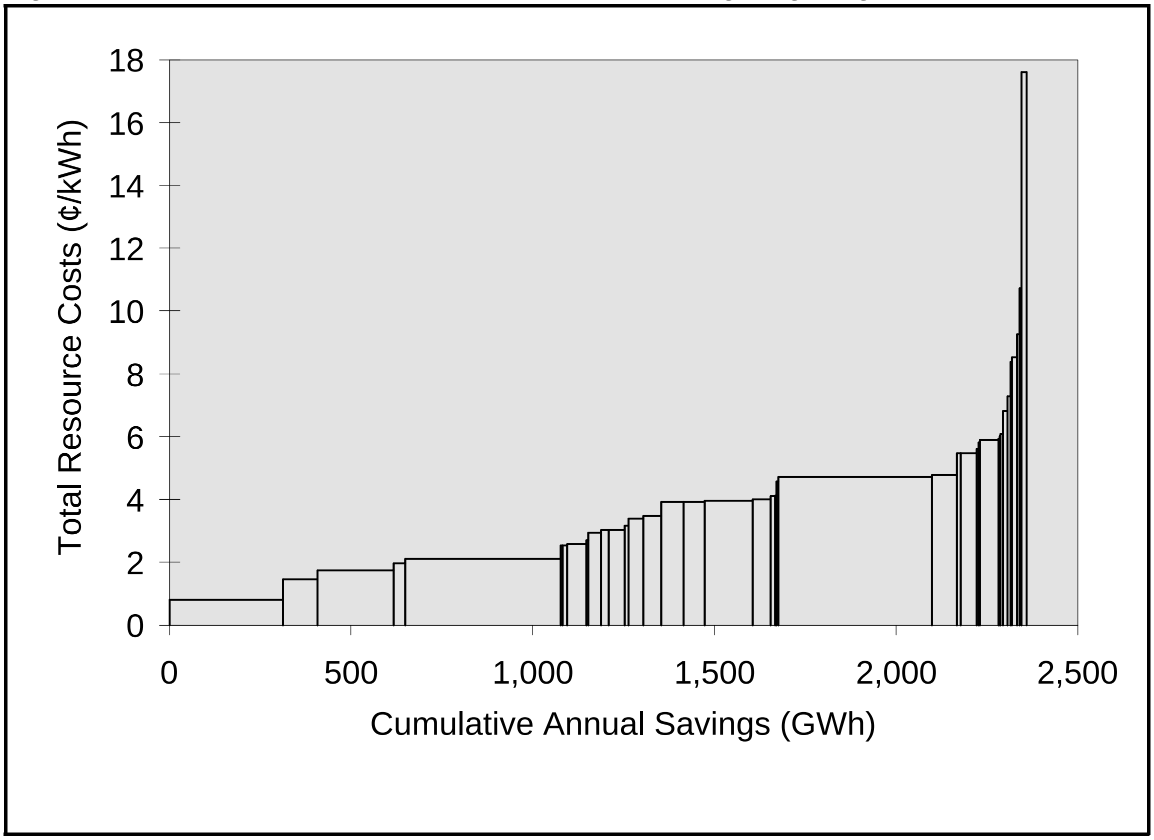
---

1 Other reports from the DEEP project include "The Cost and Performance of Utility Commercial Lighting Programs," Eto et al. (1994) and "Utility Residential New Construction Programs: Going Beyond the Code," Vine (1995).

2 We selected 1992 because it was the most recent year for which post-program savings evaluation results were consistently available for use in our study.

When compared to the direct costs avoided by the utilities (i.e., not including environmental externality adders), the savings-weighted TRC benefit-cost test ratio exceeds three (3.2), indicating that, taken as a whole, the programs have been highly cost effective. Of course, avoided costs have changed dramatically since the time when these programs were implemented. Yet, even considering avoided costs 50% lower than those in place when the programs were developed, we still find that the savings from the programs, taken as a whole, have been cost effective (although a substantial number of individual programs would not be considered cost effective). We conclude that, from an overall societal perspective, the ratepayer and participant monies used to acquire energy savings through these programs have been well spent.

**Figure ES-1. The Total Resource Cost of Commercial Lighting Programs**




We also looked closely at programs that might be representative of future DSM programs. These programs offer lower financial incentives to DSM program participants in an effort to reduce the rate impacts associated with the programs. We find that the decision to increase required customer contributions to the cost of energy saving measures has had little or no effect on the total cost of energy saved by the programs. We conclude that there is no reason to expect that future DSM programs, which rely on increased customer cost contributions, will either cost more or be less cost effective than programs offering larger financial incentives.

Nevertheless, we find wide variations in the cost of energy saved by individual programs; in particular, many smaller programs have not been cost effective in comparison to the sponsoring utilities' avoided costs. However, total spending on these programs was only 12% of the total represented by our 40 programs.

We systematically examined program cost variations in two steps. First, we conducted detailed reviews of each program cost element and of the various methods used to evaluate savings for each program. Our goal was to determine to what extent our (or the utilities') methods introduced bias in our findings. We found that our methods for treating reported and imputing missing cost and savings information were conservative. We also examined the use of standardized assumptions for several known-to-be-uncertain quantities reported by utilities (the economic lifetime of savings and free riders), but found that they had little discernable effect on our findings.

Second, we conducted exploratory statistical analyses to examine the correlation between various program features and the TRC. Other things being equal, direct installation programs are more expensive, while larger programs were less expensive. We also found that savings evaluation method and program start date were not statistically significant factors in explaining differences in the TRC.

No one knows the future of utility DSM programs. However, we feel strongly that discussions about this future should be based on unbiased and critical assessments of the performance of past programs. The goal of the DEEP project is to contribute information to this end.

*CHAPTER 1*

# Introduction

Utility demand-side management (DSM) activities are at a crossroads. After five years of unprecedented growth, during which aggregate DSM spending increased nearly fourfold to almost $3 billion in 1994, utilities and public utility commissions are reexamining their roles and responsibilities in improving customer energy efficiency. Many issues need to be considered, including the magnitude and value of uncaptured energy efficiency opportunities, the extent of utilities' obligations to serve, and the maturity of the energy services infrastructure. It is our belief that evidence on the actual performance of utility DSM programs should be an integral part of the discussion. Ideally, this evidence will help us answer the questions: What have utility-sponsored energy efficiency DSM programs cost? Have they been cost effective? What explains differences in program costs? This report describes the results from a major research project to address these questions.

The goal of our project is to develop consistent and comprehensive information on the cost of energy efficiency delivered through the nation's largest DSM programs. We have focused on the commercial sector because the energy efficiency opportunities there are thought to be large and highly cost effective. As a result, commercial sector programs often represent the largest single element in a utility's portfolio of DSM programs. We focus on 1992 programs because post-program evaluations for 1992 were the most recent ones consistently available when we began our study.

Developing consistent and comprehensive information on the total cost and measured performance of utility DSM programs is difficult. As Joskow and Marron (1992) document, utilities' reporting and savings evaluation practices differ tremendously. Customer costs are frequently omitted, utilities' overhead allocation practices vary, and measurement and evaluation costs are generally incurred in years subsequent to the program year being studied. In addition, savings evaluation practices range from simple extractions from program tracking databases (which may be augmented with substantial, site-specific information, such as metered hours of operation) to sophisticated econometric analyses of billing information (which may also include detailed, site-specific information).

However, we do not agree with Joskow and Marron that variations in reporting and savings practices alone create such large uncertainties regarding the total cost of energy efficiency that reliance on DSM as an energy resource is unwarranted. We believe that systematic treatment of differences in reporting and evaluation methods along with careful examination of utility evaluations and annual filings corroborated by extensive discussions with utility staff to verify interpretations can result in meaningful comparisons of DSM program performance. The challenge is to represent differences precisely, document all data treatments clearly, and assess critically the biases that the analysis may introduce.

This report builds upon previous work that analyzed 20 commercial sector lighting DSM programs (Eto et al. 1994). This report differs in several ways. The earlier work was based on a convenience sample of programs for which information on the total cost of energy savings was readily available. In the present work, we sought information on only the largest commercial sector DSM programs so that our results would capture a substantial fraction of utility DSM spending in 1992. In some cases (14 of the original 40 programs), our selection criterion (a budget of $1 million or more) meant only that we had to update information from our previous report to include the nonlighting program elements or to replace older data with information for the 1992 program year. However, we eliminated six programs that did not meet our selection criterion and added 26 new programs not previously reported on.

This report consists of five chapters following the introduction. In Chapter 2, we provide an overview of the 40 programs. We document the portion of total industry spending represented by the programs and summarize general characteristics of the programs. In Chapter 3, we define and present findings on the total cost of energy savings delivered by the programs and assess the programs' cost effectiveness. We also attempt to explain differences in program costs, by referring to features of both the sponsoring utilities (e.g., avoided cost) and the programs (e.g., program size, program type, etc.). In Chapter 4, we document important data collection and data treatment issues that we addressed in developing consistent information on the cost of the programs. In Chapter 5, we describe the evaluation methods used to measure savings from the programs. In both Chapters 4 and 5, we quantify the effects of key uncertainties on our estimates of the costs of energy savings. In Chapter 6, we summarize our main findings. Three appendices follow the list of references. In Appendix A, we describe the DSM program selection and data collection process. In Appendix B, we propose a method for accounting for cost and savings consistently from programs that target normal replacements rather than early replacements. In Appendix C, we describe technical considerations affecting the transferability of program evaluation results.

# The Largest Commercial Sector DSM Programs

In this chapter, we provide an overview of the 23 utilities and 40 utility DSM programs examined in this report. We document the large proportion of industry DSM spending represented by the programs and compare the sponsoring utilities' overall DSM spending to industry averages. We then describe distinguishing features of the programs; Chapter 4 draws on these features to help explain differences in program costs.

## 2.1 DSM Program Spending

Utilities spent about $380 million on the 40 programs in our sample, which represents nearly a third of total 1992 industry spending on energy efficiency DSM programs ($1.2 billion); see Figure 2-1. The programs account for more than half of their sponsoring utilities' energy efficiency DSM program budgets ($720 million).

**Figure 2-1. 1992 Utility Energy Efficiency DSM Spending**

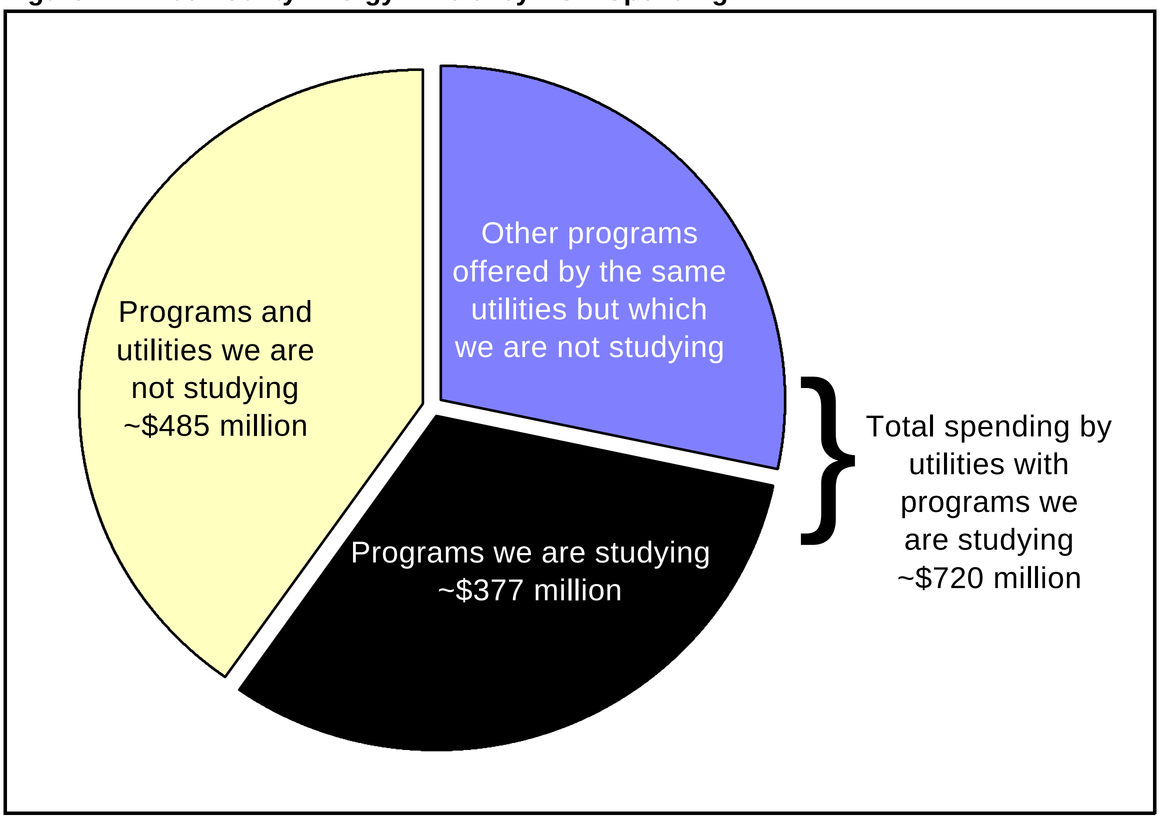


**Table 2-1. 1992 Utility DSM Budgets**

| Utility ID[3] | Programs Evaluated/ Total Energy Efficiency Expenditures | Total Energy Efficiency Expenditures/ Total DSM Expenditures | Total DSM Expenditures/ Electric Revenues |
|---|---|---|---|
| A | 107%[4] | 77% | 2.3% |
| B | 81% | 82% | 2.6% |
| C | 36% | 60% | 1.6% |
| D | 33% | 81% | 1.2% |
| E | 41% | 56% | 4.6% |
| F | 74% | 46% | 2.5% |
| G | 15% | 70% | 3.9% |
| H | 19% | 90% | 5.6% |
| I | 40% | 85% | 1.8% |
| J | 47% | 59% | 1.5% |
| K | 71% | 76% | 4.1% |
| L | 17% | 86% | 2.1% |
| M | 49% | 100% | 3.2% |
| N | 75% | 64% | 4.0% |
| O | 75% | 69% | 2.9% |
| P | 48% | 54% | 2.6% |
| Q | 27% | 80% | 4.1% |
| R | 217%[5] | 76% | 1.8% |
| S | 33% | 55% | 2.9% |
| T | 32% | 90% | 6.1% |
| U | 34% | 65% | 3.9% |
| V | 22% | 70% | 1.2% |
| W | 38% | 57% | 0.5% |
| Spending-Weighted Mean | 52% | 67% | 2.4% |
| Average | 54% | 72% | 2.9% |
| Standard Deviation | 43% | 14% | 1.5% |

---

3 By agreement with the utilities providing information for our report, we do not report information that would allow identification of specific utilities or programs. A list of the utilities whose programs are included in the report can be found in the Acknowledgments. See Appendix A for additional discussion of this decision.

4 In this case, the utility has not reported all of its spending on this program under the EIA category of energy efficiency.

5 In this case, we are working with utility program spending spread over two years.

Overall, total spending on DSM energy efficiency programs represents about 70% of the sponsoring utilities' total DSM program budgets[6] (see Table 2-1). Thus, the programs we are examining (about one-half of the energy efficiency spending by these utilities) account for slightly more than one-third of the sponsoring utilities' total DSM budgets.

DSM spending by our 23 utilities ranges from less than 0.5% of electric revenues to more than 6%. Weighted by spending, total DSM spending by our utilities averages 2.4% of revenues, which is significantly higher than the industry average of 1.4%.

The sponsoring utilities are the industry's leading DSM providers. Although they collectively accounted for less than 30% of total electric industry revenues in 1992, their total DSM spending (i.e., energy efficiency plus other DSM programs) represents nearly half of total industry DSM spending in 1992 (see Table 2-2).

**Table 2-2. Summary of DSM Spending ($ millions)**

| | Utility Expenditures on Programs Evaluated | Total Utility Energy Efficiency Expenditures | Total Utility DSM Expenditures | Utility Electric Revenue |
|---|---|---|---|---|
| Utilities we are studying | 377.1 | 720.0 | 1,081.5 | 46,028.1 |
| All utilities reporting to EIA | | 1,204.7 | 2,243.3 | 158,753.6 |
| Utilities we are studying as % of all utilities reporting to EIA | | 60% | 48% | 29% |

## 2.2 DSM Program Characteristics

The goal of our project is to examine the largest 1992, commercial sector, energy efficiency DSM programs. Each program individually accounted for more than $1 million of utility spending.[7] Several were among the largest single DSM programs operating in that year.

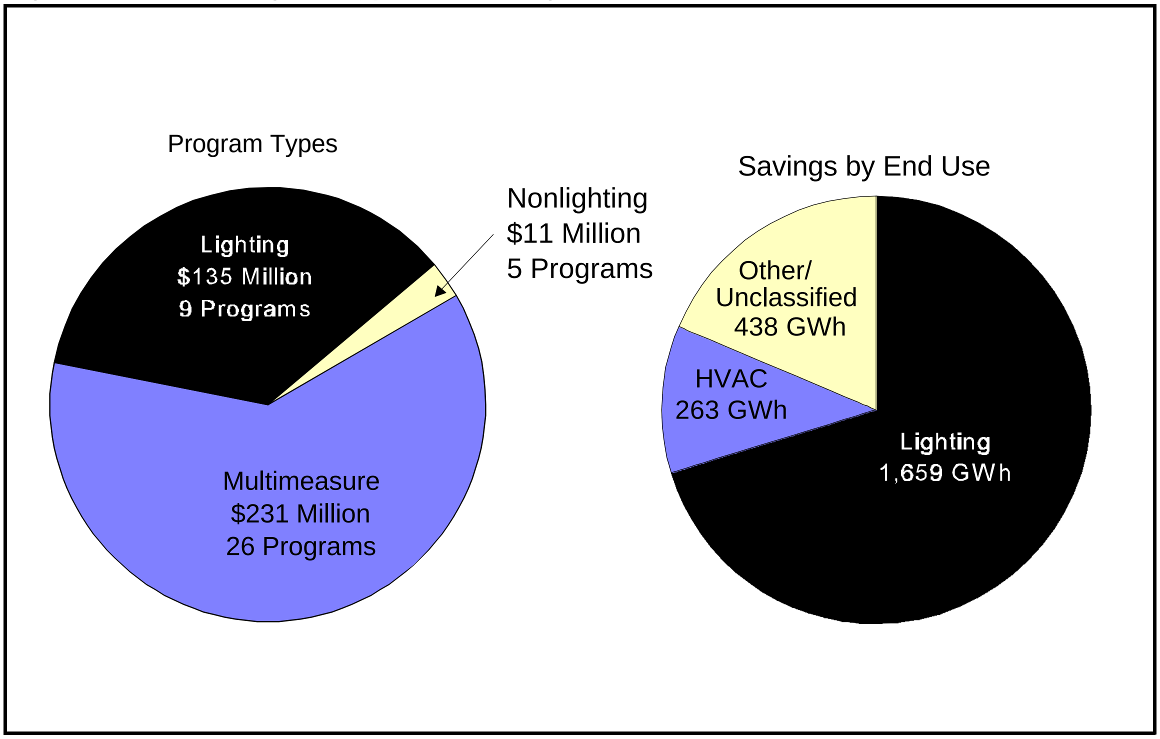
We categorized the majority of programs as multimeasure programs; the next largest number are lighting only programs (see Figure 2-2). All major commercial sector end uses were targeted by some programs, including lighting, heating, ventilating and air conditioning

6 DSM refers to a variety of utility-sponsored programs designed to influence customers' use of energy. In addition to energy efficiency, DSM also includes load management and load building programs.

7 We included one program that spent slightly less than $1 million.

(HVAC), motors, shell, refrigeration, water heating, process, and other (see Table 2-3). Nevertheless, lighting measures account for the majority of the savings from the programs.

**Figure 2-2. DSM Program Types vs. Savings by End Use**




The lighting technologies promoted by the programs were quite similar. For 30 of the 35 lighting and multimeasure programs for which we had information, 26 promoted compact fluorescents, electronic ballasts, and either T-8 or T-12 fluorescent lamps; 24 promoted reflector systems; and 22 promoted lighting controls and high intensity discharge lamps.

The programs are all full-scale (as opposed to pilots) but vary in maturity. Five only began full-scale operation in 1992 while three began full-scale operation prior to 1986 (see Table 2-3).

The majority of programs (29) offered rebates, but there are also a number of direct installation programs (11). Many of the rebate programs were linked to utility-sponsored audit activities.[8] Several rebate programs also featured loan or financing options although rebates constituted the bulk of the programs' activities.

[8] However, we could not consistently identify whether audits were formally considered a part of or were separate from rebate programs.

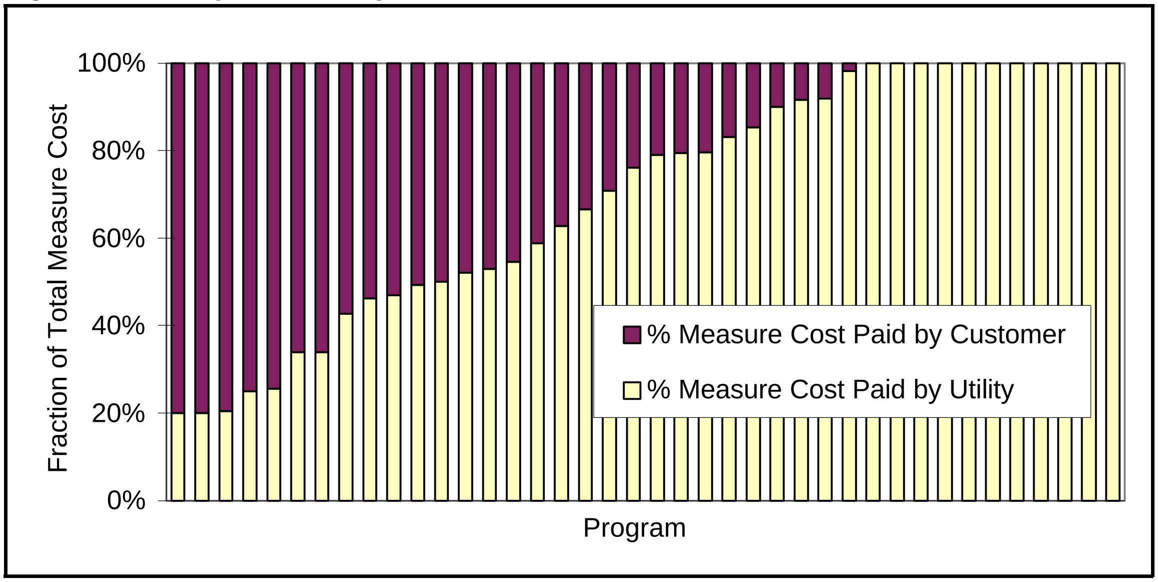
**Table 2-3. DSM Program Characteristics**

| Program ID | Program Start Date | Incentive Type | Measures Offered |
|---|---|---|---|
| 1 | 1990 | Rebate | L |
| 2 | 1991 | Rebate | H |
| 3 | 1990 | Rebate | L, H, M, O |
| 4 | 1991 | Direct Installation | L |
| 5 | 1986 | Rebate | L, H, S, M, R, O |
| 6 | 1990 | Rebate | L, H, S, M, R, O |
| 7 | 1991 | Rebate | L |
| 8 | 1990 | Rebate | H |
| 9 | 1991 | Direct Installation | L, H, R, P |
| 10 | 1990 | Rebate | L |
| 11 | 1990 | Rebate | H, M |
| 12 | 1989 | Rebate | L |
| 13 | 1990 | Rebate | H |
| 14 | 1987 | Rebate | L, H, S, M, R, W, P, O |
| 15 | pre-1986 | Rebate | L, H, S, R, P |
| 16 | 1991 | Rebate | L |
| 17 | 1989 | Direct Installation | L, H, S, M, W |
| 18 | 1990 | Direct Installation | L, H, M, W |
| 19 | 1988 | Direct Installation | L, H, S, M, R, W, P, O |
| 20 | 1989 | Rebate | L |
| 21 | 1989 | Rebate | H, S, R, P |
| 22 | 1992 | Direct Installation | L, H, M, R, W, P, O |
| 23 | 1991 | Rebate | L, H, S, M, R, W, P, O |
| 24 | 1989 | Rebate | L, H, S, M, R, W, P, O |
| 25 | 1990 | Direct Installation | L, H, W |
| 26 | 1989 | Rebate | L, H, S, M, R, W, P, O |
| 27 | 1990 | Direct Installation | L, H, W |
| 28 | 1991 | Direct Installation | L, H, S, W |
| 29 | 1992 | Rebate | L, H, S, M, R, W |
| 30 | pre-1986 | Direct Installation | L, H, S, R, W |
| 31 | 1988 | Rebate | L, H, O |
| 32 | pre-1986 | Rebate | L, H, S, M, R, W, P, O |
| 33 | 1990 | Rebate | L, H, S, M, R, W, P, O |
| 34 | 1990 | Rebate | L |
| 35 | 1989 | Rebate | L, H, M, R, W, P, O |
| 36 | 1991 | Rebate | L, H, S, M, R, W, O |
| 37 | 1990 | Rebate | L, H, M, R, P, O |
| 38 | 1992 | Direct Installation | L, H, W |
| 39 | 1992 | Rebate | L, H, M, R, O |
| 40 | 1992 | Rebate | L, H, S, M, R, W, P, O |

Key: lighting (L), HVAC (H), motors (M), shell (S), refrig. (R), water heating (W), process (P), other (O)

Most of the rebate programs required some form of customer cost contribution toward the purchase and installation of energy saving measures. The percent of full measure costs contributed by the customer ranged up to 80% (see Figure 2-3). Increasing customer cost contribution is an important strategy for utilities seeking to reduce the rate impacts of their DSM programs. For this reason, we pay special attention to the cost of energy savings from programs with high customer cost contributions in the next chapter.

**Figure 2-3. Utility and Participant Measure Costs Contributions**



*CHAPTER 3*

# The Cost and Performance of Commercial Sector DSM Programs

In this chapter, we present findings from our examination of 40 of the largest 1992 commercial sector DSM programs. We first define two measures of performance, the total resource cost and the utility cost of energy savings, and briefly review the inputs to the calculation.[9] We then devote the bulk of the chapter to answering three questions: How much have the programs cost? Have they been cost effective? What explains differences in program cost?

## 3.1 Measuring the Total Resource Cost and Utility Cost of Energy Savings

There are many ways to measure the performance of utility DSM programs; each depends on the objective of the program. Nadel (1991) has written extensively on a variety of DSM program performance indicators, such as the percent of revenue spent on DSM programs, DSM program participation rates, and the cost of conserved energy. Wellinghoff and Flanigan (1992) define 20 measures of program performance for commercial lighting DSM programs. We focus on two particular measures of performance: the total resource cost of energy savings or TRC, and the utility cost of energy savings or UC.[10] We believe that these two measures capture the most important dimensions of the resource-acquisition objectives that underlie DSM programs. The TRC is the more comprehensive in scope and measures the total cost of energy savings from a societal perspective. The UC is more limited in scope and measures the direct costs borne by utility ratepayers. Recent interest in the UC has increased as utilities have become more concerned about the rate impacts of their DSM programs. Table 3-1 summarizes the cost and savings components for both the total resource cost and utility cost of energy savings.

---

9 We reserve detailed discussion of these inputs for later chapters.

10 The reader is cautioned that our choice of terms, TRC and UC, refer to costs, as measured in ¢/kWh. When we use these terms in the context of DSM benefit-cost tests (from which they were derived, but from which they differ slightly as described in Table 3-1), we will explicitly label them as TRC or UC *benefit-cost test ratios*, which are dimensionless.

**Table 3-1. Cost and Savings Components of the TRC and UC of Energy Savings**

| | Included in Total Resource Cost of Energy Savings | Included in Utility Cost of Energy Savings |
|---|---|---|
| **Cost Elements** | | |
| Participant-paid measure costs | Yes | No |
| Utility-paid measure costs | Yes | Yes |
| Participant- or utility-paid measure costs associated with free riders | Yes[11] | Yes |
| Utility administrative costs (including overhead and measurement and evaluation) | Yes | Yes |
| Utility shareholder incentives | Yes | Yes[12] |
| Changes in customer operating costs (+/-) | No | No |
| **Savings Elements** | | |
| Savings from free riders | Yes[13] | No |
| Savings from non-free riders | Yes | Yes |
| Savings recaptured through takeback | Sometimes[14] | Sometimes |
| Savings due to participant and nonparticipant spillover[15] | No | No |

11 Our treatment of free riders in the TRC differs from standard practice. See discussion in Section 3.1.1.

12 Our inclusion of shareholder incentives in the UC differs from standard practice, which does not include them. Our definition is, thus, a more comprehensive measure of the direct ratepayer costs of energy savings.

13 As noted previously, our treatment of free riders in the TRC differs from standard practice. See discussion in Section 3.1.1.

14 Takeback refers to savings that are "recaptured" by program participants, typically through increased energy services. A common example in the commercial sector is the installation of efficient security lighting in areas that were formerly unlit. See discussion of takeback in Chapter 5.

15 Spillover refers to savings from measures installed as a result of the program, but not through the program. It can include additional measures installed by program participants outside of the program or measures installed by nonparticipants as a result of the program. See discussion of spillover in Chapter 5.

In developing the TRC and UC for our 40 DSM programs, we have adopted two conventions for expressing our results. First, both the TRC and UC are calculated by dividing the levelized cost of a program by annual energy savings.[16] Second, all levelizations are performed using a common real (i.e., net of inflation) discount rate of 5%. Thus, the units of the TRC and UC are expressed as a cost per kilowatt-hour of savings (¢/kWh).

### 3.1.1 Total Resource Cost of Energy Savings

The total resource cost of energy savings consists of two types of costs, measure and nonmeasure costs. Measure costs may be borne by the participating customer, the utility, or, more typically, by both. The measure costs borne by the customers participating in a DSM program are a major cost element that is often missing from discussions of the cost of energy savings. However, including these costs is essential for assessing the total cost of energy savings to society.

Nonmeasure costs refer to all costs incurred by the utility in operating its DSM program except for the utility's contribution to the cost of the measures delivered by the program. Nonmeasure costs include direct costs in the form of program staff, advertising expenses, and administrative support; and indirect costs in the form of departmental overhead and measurement and evaluation expenses. DSM shareholder incentives paid to a utility for its performance in implementing a DSM program are also included in our calculation of the TRC and UC.

Chapter 4 describes the data underlying our calculation of program costs. In that chapter, we focus on two classes of issues that could threaten the robustness of our findings: (1) the possibility that measure costs may be overestimated because utilities sometimes report costs as representing full measure costs when, in some cases, incremental measure costs may be more appropriate; and (2) the importance of including overhead, measurement and evaluation, and shareholder incentives in the cost of energy savings. With regard to shareholder incentives, there are unresolved conceptual and measurement issues associated with determining the extent to which these incentives are truly net societal costs or simply transfers between ratepayers and shareholders. Our calculation of the TRC is conservative because we include shareholder incentives.

All measurements rely on some form of post-program savings verification, ranging from on-site inspections and customer surveys to sophisticated regression analyses with customer

16 Levelization is an engineering/economic technique that spreads costs in equal nominal amounts over the lifetime of a program so that the present value of nominal amounts is left unchanged. See, for example, EPRI (1991). Levelization is more appropriate than simply dividing total costs by lifetime savings because levelization accounts for the time-value of money. The importance of accounting for the time-value of money increases as savings extend farther into the future or when discount rates are high.

billing information and end-use metering. The evaluation methods used to verify energy savings from the programs are discussed in Chapter 5. Chapter 5 also examines the impact of the following uncertainties on our findings: (1) differences in evaluation methods used to estimate annual energy savings; (2) differences in the economic lifetime of savings; and (3) differences in the effects of free-rider estimates on the UC of energy savings.

Our treatment of the savings and costs associated with free riders differs from standard practice. The standard definition of the total resource cost of energy savings does not include the measure costs and savings from free riders. We include them. Our reason for including them is based on our review of current practices in estimating free riders (see Chapter 5). While methods for estimating the savings from free riders are maturing, we are concerned that they cannot be applied reliably to the measure costs associated with free riders. Yet current practice appears to simply use the same adjustment factor for both. Rather than continue this practice, we have chosen to address the issue of free ridership by minimizing the number of adjustments to our data involving free riders. Thus, we do not reduce savings or costs to remove the effect of free riders in the TRC, while we do remove the savings associated with free riders for the UC. Eto et al. (1994) show that the inclusion of free-rider savings has little material effect on the TRC as long as the levels of free ridership and the size of nonmeasure costs are low.

The TRC, in principle, also does not take credit for takeback and includes all savings resulting from participant and nonparticipant program spillover. As described in Chapter 5, we found that, in practice, takeback was generally minimal or assumed to be controlled for in billing analyses. Also as described in Chapter 5, only two programs claimed explicit credit for spillover. In view of the current controversy over the measurement of spillover, we have not included these estimates in our calculations.

### 3.1.2 Utility Cost of Energy Savings

Our calculation of the UC differs from the TRC in three ways. First, it does not include measure costs borne by participating customers. Because these costs are not recovered by the utility, they have no rate impact. Second, the UC includes all shareholder incentives paid to the utility. Whether or not these costs are costs to society or just transfers, they are recovered from ratepayers. Third, the savings attributable to the programs are decreased by the utility's free-ridership estimates.[17] Some of the savings associated with a DSM program would have occurred without the DSM program. These savings cannot be attributed to the DSM program and therefore must be removed.

[17] Chapter 5 describes the measurement of free ridership by the utilities and the sensitivity of our findings to these measurement assumptions.

## 3.2 The Total Resource Cost and Utility Cost of Energy Savings

The savings-weighted mean TRC and UC for our 40 programs is 3.2 ¢/kWh and 2.7 ¢/kWh, respectively (see Table 3-2). Utility nonmeasure costs, which include utility administration, overhead, measurement and evaluation, and shareholder incentives) account for 0.8 ¢/kWh or 25% of the TRC. Measure costs, split between utility and participants, account for 44% (1.4 ¢/kWh) and 31% (1.0 ¢/kWh), respectively, of the remaining savings-weighted TRC of energy savings.

The large fraction of the savings-weighted TRC accounted for by participant-borne measure costs (31%) highlights the importance of including these costs in a full accounting of the total societal cost of energy savings. Ignoring these costs would make the apparent cost of energy savings one third less expensive than they actually are.

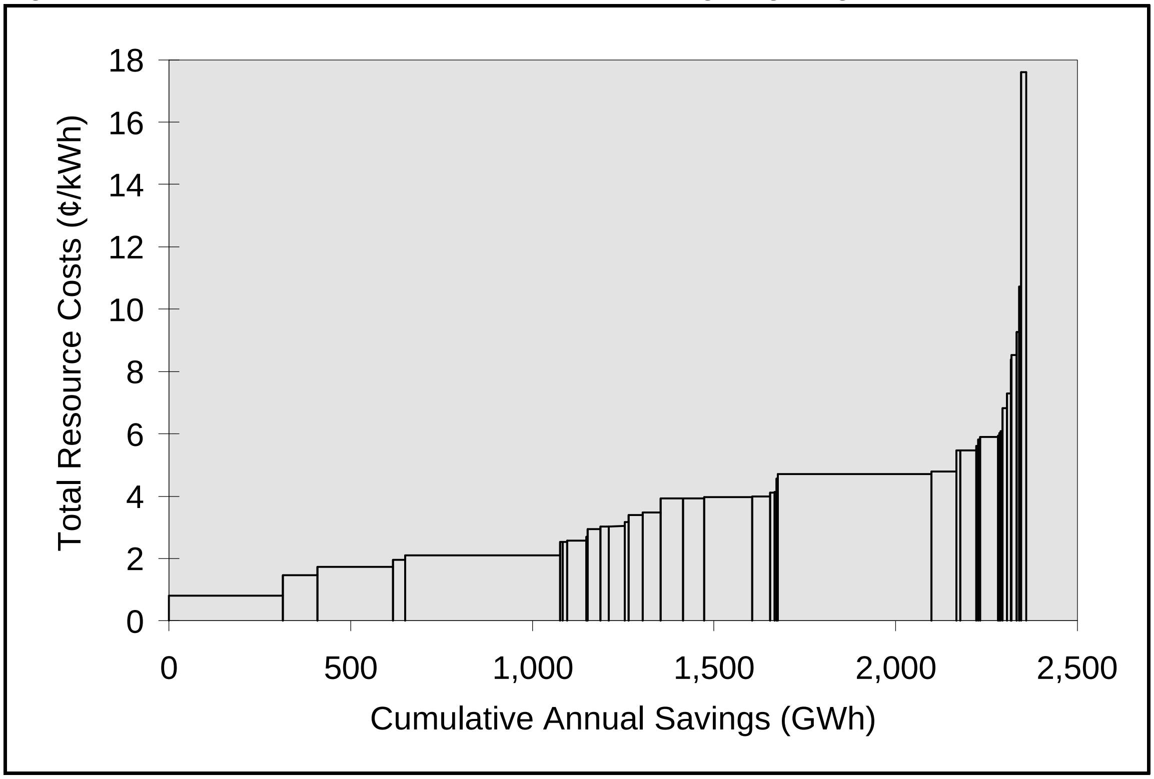
Figure 3-1 arranges the DSM programs from the least expensive to the most expensive and plots them sequentially against energy savings; the "width" of each program along the x-axis represents the savings accounted for by each program. This form of presentation shows that the savings-weighted average is dominated by several very large and inexpensive programs, and that the most expensive programs are comparatively small in size. For example, 28% of the savings have cost less than 2 ¢/kWh and 50% have cost less than 3 ¢/kWh. At the same time, only 1% have cost more than 9 ¢/kWh.

The savings-weighted TRC of energy savings (3.2 ¢/kWh) is almost 20% lower than previously reported DEEP project findings for 20 commercial lighting programs, which presented a savings-weighted TRC of 3.9 ¢/kWh (Eto et al. 1994). Moreover, the previous findings did not include shareholder incentives. We believe the reason for the difference in findings can be traced to two sources. First, as indicated in Figure 3-1, the results for our sample are strongly affected by the presence of large, inexpensive programs. The inclusion of large programs was a conscious element of the program selection criteria for the current report, which was not pursued in the earlier report. Second, for the programs that were included in both the earlier report and this report, we are generally relying on information from a more recent program year (that is, 1992 program information versus 1991 or earlier program information). Several of these programs have reduced the cost of acquiring energy savings. In the final subsection of this chapter, we will attempt to use both factors (program size and program maturity) to help explain differences in the costs of the programs.

**Table 3-2. The Total Resource Cost and Utility Cost of Energy Savings (¢/kWh)**

| Program ID | Nonmeasure Costs (Admin. and M&E) | Shareholder Incentives | Utility-Paid Measure Costs | Participant-Paid Measure Costs | Utility Costs | Total Resource Costs | Avoided Costs | TRC Ratio | RIM Ratio |
|---|---|---|---|---|---|---|---|---|---|
| 1 | 0.4 | 0.7 | 1.9 | 1.7 | 3.5 | 4.7 | 6.6 | 1.4 | 0.4 |
| 2 | 3.2 | 0.6 | 1.8 | 0.0 | 7.1 | 5.6 | 11.2 | 2.0 | 0.6 |
| 3 | 0.3 | 0.2 | 1.9 | 0.6 | 2.7 | 3.0 | 8.2 | 2.7 | 0.7 |
| 4 | 0.3 | 0.8 | 3.5 | 0.0 | 4.7 | 4.6 | 4.0 | 0.9 | 0.2 |
| 5 | 0.4 | 0.5 | 1.0 | 4.1 | 2.0 | 5.9 | 5.3 | 0.9 | 0.3 |
| 6 | 0.4 | 0.6 | 1.7 | 6.7 | 2.6 | 9.3 | 5.5 | 0.6 | 0.3 |
| 7 | 0.1 | 1.6 | 1.6 | 1.5 | 3.8 | 4.8 | 6.6 | 1.4 | 0.5 |
| 8 | 0.9 | 0.3 | 7.5 | 2.0 | 10.7 | 10.7 | 6.6 | 0.6 | 0.3 |
| 9 | 21.8 | (10.4) | 33.8 | 2.9 | 45.2 | 48.1 | 6.6 | 0.1 | 0.1 |
| 10 | 0.5 | 0.1 | 1.1 | 1.3 | 2.3 | 3.0 | 3.1 | 1.0 | 0.3 |
| 11 | 1.6 | 0.1 | 1.0 | 0.0 | 2.7 | 2.7 | 3.7 | 1.4 | 0.3 |
| 12 | 0.1 | 0.4 | 1.1 | 0.2 | 2.0 | 1.7 | 5.1 | 2.9 | 0.4 |
| 13 | 0.6 | 0.1 | 1.0 | 3.8 | 2.2 | 5.5 | 5.2 | 1.0 | 0.4 |
| 14 | 0.2 | 0.0 | 1.3 | 1.7 | 1.6 | 3.2 | 3.0 | 1.0 | 0.5 |
| 15 | 0.2 | 0.0 | 1.9 | 1.3 | 2.1 | 3.4 | 4.0 | 1.2 | 0.5 |
| 16 | 0.4 | 0.0 | 1.7 | 3.4 | 2.2 | 5.5 | 8.9 | 1.6 | 0.8 |
| 17 | 1.2 | 0.1 | 5.5 | 0.0 | 7.4 | 6.8 | 10.7 | 1.6 | 0.7 |
| 18 | 3.9 | 0.1 | 12.5 | 1.2 | 23.7 | 17.6 | 9.8 | 0.6 | 0.3 |
| 19 | 0.3 | 0.1 | 2.6 | 0.0 | 3.5 | 3.0 | 7.9 | 2.7 | 0.6 |
| 20 | 0.4 | 0.0 | 1.8 | 0.4 | 3.3 | 2.5 | 4.5 | 1.8 | 0.4 |
| 21 | 0.3 | 0.0 | 1.5 | 0.7 | 1.8 | 2.5 | 4.5 | 1.8 | 0.4 |
| 22 | 1.5 | 0.0 | 4.3 | 0.0 | 5.8 | 5.8 | 12.1 | 2.1 | 0.9 |
| 23 | 2.4 | 0.0 | 2.0 | 1.6 | 5.0 | 5.9 | 12.1 | 2.0 | 1.0 |
| 24 | 0.5 | 0.4 | 2.4 | 0.6 | 3.7 | 4.0 | 6.7 | 1.7 | 0.5 |
| 25 | 0.9 | 0.6 | 7.1 | 0.0 | 8.7 | 8.5 | 10.1 | 1.2 | 0.6 |
| 26 | 0.6 | 0.2 | 2.6 | 0.7 | 3.6 | 4.1 | 7.1 | 1.7 | 0.5 |
| 27 | 1.1 | 0.3 | 7.0 | 0.0 | 8.5 | 8.4 | 10.0 | 1.2 | 0.6 |
| 28 | 0.5 | 0.0 | 5.6 | 0.0 | 6.6 | 6.1 | 5.9 | 1.0 | 0.4 |
| 29 | 0.5 | 0.0 | 1.2 | 2.3 | 1.8 | 3.9 | 5.2 | 1.3 | 0.6 |
| 30 | 0.9 | 0.6 | 4.5 | 0.0 | 6.0 | 6.0 | 4.8 | 0.8 | 0.3 |
| 31 | 0.4 | 0.6 | 2.2 | 0.9 | 3.2 | 4.1 | 5.4 | 1.3 | 0.4 |
| 32 | 0.1 | 0.0 | 0.3 | 0.4 | 0.9 | 0.8 | 7.7 | 9.6 | 0.6 |
| 33 | 0.3 | 0.7 | 0.7 | 0.4 | 2.3 | 2.1 | 7.0 | 3.3 | 0.5 |
| 34 | 0.3 | 0.8 | 1.4 | 0.0 | 3.3 | 2.6 | 10.4 | 4.1 | 0.9 |
| 35 | 0.3 | 0.6 | 0.3 | 0.8 | 1.2 | 2.0 | 17.6 | 9.0 | 1.7 |
| 36 | 0.1 | 0.0 | 0.3 | 1.0 | 0.5 | 1.5 | 3.1 | 2.1 | 0.5 |
| 37 | 1.9 | 0.0 | 0.8 | 0.8 | 3.4 | 3.5 | 4.4 | 1.3 | 0.4 |
| 38 | 0.8 | 0.8 | 5.7 | 0.0 | 7.7 | 7.3 | 5.6 | 0.8 | 0.4 |
| 39 | 0.6 | 0.3 | 1.5 | 1.5 | 2.6 | 4.0 | 5.6 | 1.4 | 0.6 |
| 40 | 1.1 | 0.4 | 2.1 | 0.2 | 3.7 | 3.9 | 5.6 | 1.4 | 0.6 |
| Weighted Average | 0.4 | 0.4 | 1.4 | 1.0 | 2.7 | 3.2 | 6.6 | 3.2 | 0.5 |
| Mean | 1.3 | 0.1 | 3.5 | 1.1 | 5.4 | 6.0 | 6.9 | 1.9 | 0.5 |
| Standard Deviation | 3.4 | 1.7 | 5.5 | 1.4 | 7.6 | 7.5 | 3.1 | 1.9 | 0.3 |

**Figure 3-1. The Total Resource Cost of Commercial Lighting Programs**



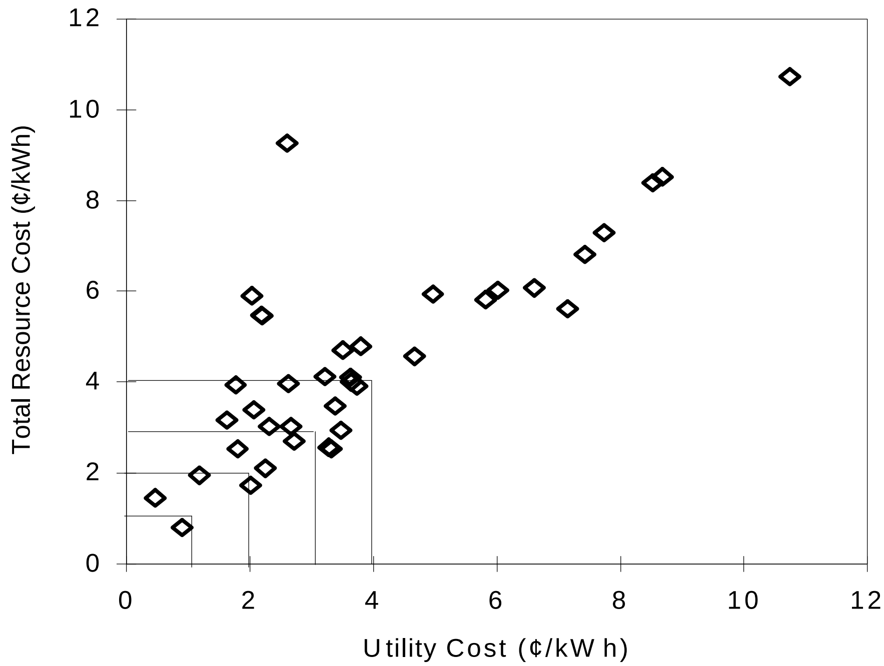

Note: Does not include Program #9.

The unweighted means for the 40 programs are 6.0 ¢/kWh (standard deviation = 7.5 ¢/kWh) and 5.4 ¢/kWh (standard deviation = 7.6 ¢/kWh) for the TRC and UC, respectively. However, they are strongly affected by one very high-cost program, which has a TRC and UC of 48.1 ¢/kWh and 45.2 ¢/kWh, respectively. The cost of this program is more than four standard deviations higher than either mean. Without this program, the unweighted means drop 20% to 4.9 ¢/kWh (standard deviation 3.1 ¢/kWh) and 4.4 ¢/kWh (standard deviation 4.0 ¢/kWh) for the TRC and UC, respectively.[18]

Figure 3-2 compares the TRC to the UC for each program. The figure indicates that several programs minimized both the TRC and UC of energy savings. For example, ten programs have TRCs of energy savings that are less than 3 ¢/kWh. Of these, seven programs also have UCs of energy savings that are less than 3 ¢/kWh, and four programs have UCs of energy savings that are less than 2 ¢/kWh.

[18] In view of this finding, we will selectively exclude this program from subsequent discussions in which its influence dominates the comparisons being made.

**Figure 3-2. The TRC vs. UC of Energy Savings**




In thinking about current trends in DSM program design, concerns about DSM program rate impacts are leading utilities to reduce the size of rebates offered to customers. The average fraction of measure costs paid by the utility was 70% (see Figure 2-3). However, five programs offered rebates that paid about 25% or less of the cost of the measures. By examining results for these programs, we can gain some insight into the effect of this program design decision on the costs of energy savings.

Two of these five programs were among the least expensive with a TRC of less than 2 ¢/kWh; one program was one of the most expensive with a TRC of over 9 ¢/kWh. The unweighted mean TRC for the five programs is 4.8 ¢/kWh with a standard deviation of 3.2 ¢/kWh. These five programs cost less on average than the average for the entire group of programs (nevertheless, the result is not statistically significant given the high standard deviations associated with both means). Based on this limited sample, there appears to be no inherent reason why DSM program designs with high customer cost contributions should cost any more than DSM programs with lower customer cost contributions.

## 3.3 Program Cost Effectiveness

In this section, we describe program cost-effectiveness using two standard DSM benefit-cost tests, the TRC benefit-cost test ratio and the ratepayer impact measure (RIM) benefit-cost test ratio.

### 3.3.1 The Total Resource Cost Benefit-Cost Test Ratio

The societal value, and hence cost effectiveness, of DSM programs is measured by the resource costs they allow the utility to avoid. Avoided electric supply costs are the primary measure of the resource value of DSM programs offered by electric utilities. They depend on the economic circumstances of a particular utility and on the load shape impacts and economic lifetime of savings from a particular DSM program. Definitions of what cost components are avoided and what methods are used to estimate them differ among utilities.[19]

We worked from each utility's benefit-cost ratios for its DSM programs in order to develop a "top-down" estimate of program-specific avoided costs.[20] For example, from a TRC benefit-cost test ratio of two and a levelized estimate of the denominator (i.e., total utility and incremental participant costs) of 4 ¢/kWh, we can conclude that the effective avoided cost or numerator is 8 ¢/kWh (i.e., 2 x 4 ¢/kWh). By estimating avoided costs in this fashion, we bypassed the need to know the specific avoided costs faced by a particular utility, as reported in time-of-day-, seasonal-, and annual-differentiated avoided costs for energy and capacity. As a consequence, the avoided costs we report (see Table 3-2) differ from those that might be published by the utility, for example, in a tariff sheet of payments to qualifying facilities. In addition, the avoided cost for two programs from the same utility may differ because the assumed load shape impacts and lifetimes differ. Most important of all, even though we present avoided costs in units of ¢/kWh, they include avoided capacity costs implicitly. Finally, without commenting on their appropriateness, we chose to eliminate environmental externality adders in an effort to ensure greater comparability among utilities.[21]

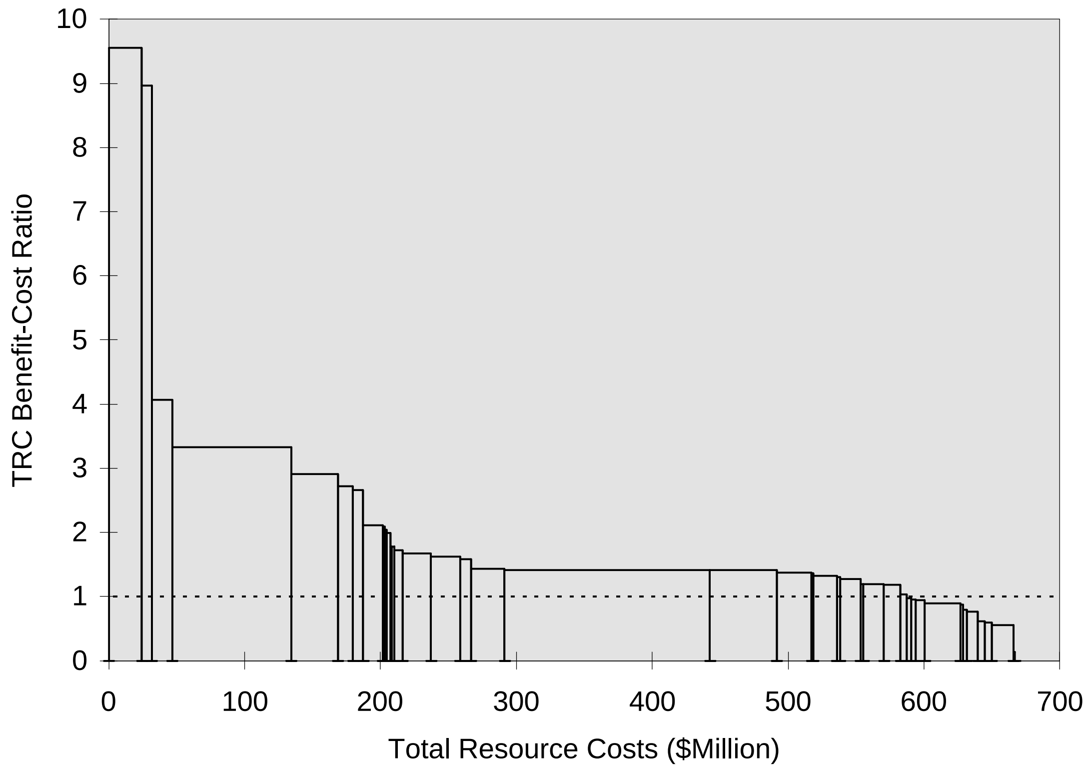
Figure 3-3 presents, in descending order, the TRC benefit-cost ratio for each program. The total resource cost associated with each program is plotted horizontally across the x-

[19] See Busch and Eto (1995) for a summary of current practices in estimating avoided costs for use in measuring the resource value of DSM programs.

[20] In order to increase comparability among program-specific avoided costs, we also normalized them using both the weighted average cost of capital reported by the utilities and the same real discount rate used to estimate the TRC and UC: 5%.

[21] We were also able to determine that a few utilities included estimates of avoided transmission and distribution costs in their overall estimates of avoided costs.

**Figure 3-3. The Cost Effectiveness of 1992 Commercial Sector DSM Programs**




axis.[22] The savings-weighted TRC benefit-cost test ratio of avoided costs to program costs is 3.2, indicating that, taken as whole, the programs are highly cost effective.

The unweighted mean of the TRC benefit-cost test ratios is 1.9 with a standard deviation of 1.9. Since the savings-weighted TRC benefit-cost test ratio is higher, we can conclude that some of the largest programs are also the most cost effective. The high standard deviation also indicates that some programs are not cost effective; 11 of the programs have TRC benefit-cost test ratio of less than 1.0. This should not be too surprising because there are several extremely high-cost programs. As indicated on Figure 3-3, however, the 11 programs that are not cost effective account for only 12% of the total resource costs of all of the programs.

The most critical issue for our estimates of program-specific avoided costs is that they are based on a forecast of the future and hence are inherently uncertain. For many utilities, avoided costs have dropped significantly since the time when they were first developed. In

[22] This form of presentation is similar to Figure 3-1. In Figure 3-3, we use the total resource cost for the x-axis, rather than annual energy savings.

particular, the program planning estimates for our 1992 programs were for the most part based on estimates of avoided cost developed in 1991.

In view of this situation, it is useful to consider how lower avoided costs would affect our findings. If we assume that avoided costs are 50% lower than those originally reported, TRC benefit-cost test ratios drop below unity for an additional 19 programs. However, the savings-weighted TRC benefit-cost test ratio would be 1.6. We conclude that dramatically lower avoided costs can have a dramatic effect on the cost effectiveness of individual programs. Nevertheless, taken as a whole, the majority of savings from the programs remain cost effective.

### 3.3.2 Ratepayer Impact Measure Benefit-Cost Test Ratio

The ratepayer impact measure (RIM) benefit-cost test ratio measures the relationship between the supply costs avoided by a DSM program, and the DSM program costs and revenue losses incurred by the utility (CPUC/CEC 1987). Assuming perfect ratemaking practices, a RIM test ratio of less than one indicates that the program is likely to increase retail rates. However, the RIM test is an imperfect measure of the actual rate impacts of DSM measures. Among other things, the exact schedule of future avoided costs, the ratemaking treatment of DSM program costs, on-going changes in the utility's cost structure, and, most important of all, prevailing rate design practices, must be accounted for to determine the exact impact on retail rates.[23]

The savings-weighted mean RIM benefit-cost test ratio for our programs is 0.5. The unweighted mean is virtually identical (0.5 with a standard deviation of 0.3). The simple conclusion we draw is that avoided costs are lower than retail rates.[24]

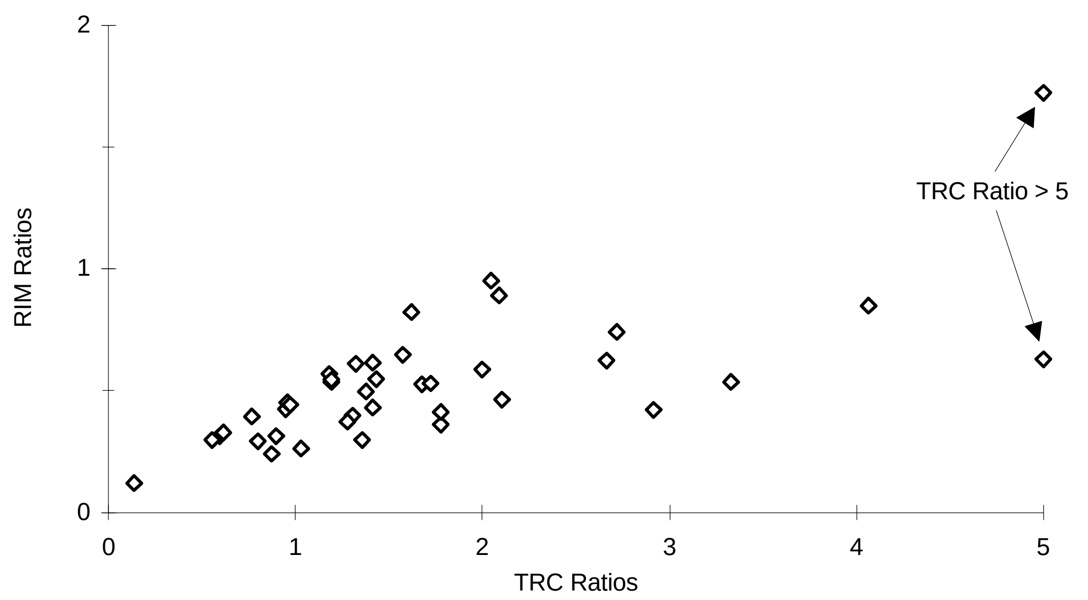
Figure 3-4 presents the TRC and RIM benefit-cost test ratios for each program. The figure suggests that high TRC benefit-cost test ratio programs tend to have higher RIM benefit-cost test ratios. We believe this relationship is driven primarily by the relationship between avoided costs and retail rates. High avoided costs (leading to higher TRC benefit-cost test ratios) eventually exceed retail rates and drive the RIM ratio upward.

---

[23] See Hirst and Hadley (1994) for a discussion of the rate impacts of DSM programs; see Pye and Nadel (1994) for a summary of studies that have attempted to measure the actual rate impacts of DSM programs.

[24] We estimated revenue losses using the average retail rate for the commercial class of each utility, as reported in EIA (1993 and 1994a).

**Figure 3-4. TRC and RIM Benefit-Cost Test Ratios**




## 3.4 Explaining Differences in the TRC of Energy Savings

What makes some programs more costly than others? In addressing this question, it is useful to begin by simply reviewing the differences between the best and the worst programs. Table 3-3 summarizes key features for the five least and five most expensive programs as well as for the entire sample of programs.[25]

Starting with program costs and measure and nonmeasure costs, it is clear that the least expensive programs were run with substantially lower administrative or nonmeasure costs. These programs were either run more efficiently (even including, or perhaps because of, shareholder incentives) or they were able to spread fixed, nonmeasure costs over a larger base of energy savings (see below).[26] It is also clear that high measure costs account for most of the costs of the more expensive programs. Either the measures installed were very costly or the installations were such that comparatively less energy was saved per installation (e.g., facilities receiving the measures had few hours of operation).

---

[25] Program #9, the very expensive outlier described earlier, has not been included in this comparison.

[26] The inclusion of audits in the administrative cost of a program would also tend to drive cost upward. Unfortunately, we were not able to identify whether audit costs were included consistently for all programs.

**Table 3-3. Comparison of the Most and Least Expensive Programs**

| | Five Least Expensive Programs | Five Most Expensive Programs | All Programs[27] |
|---|---|---|---|
| TRC (¢/kWh) | 1.6 (0.5) | 10.9 (3.9) | 4.9 (3.1) |
| Nonmeasure Cost (¢/kWh) | 0.5 (0.4) | 1.8 (1.2) | 1.1 (0.9) |
| Shareholders Incentives | 4 of 5 | 5 of 5 | 28 of 39 |
| Measure Cost (¢/kWh) | 1.1 (0.3) | 9.1 (2.7) | 3.8 (2.6) |
| Avoided Costs (¢/kWh) | 8.1 (5.6) | 8.4 (2.2) | 6.9 (3.1) |
| Measure Costs Paid by Utility (%) | 49 (26) | 78 (34) | 69 (28) |
| Program Size (GWh/yr) | 215.3 (159.5) | 8.7 (5.0) | 60.4 (104.8) |
| Participants (per year) | 4,721 (4,626) | 796 (1,095) | 1,691 (2,563) |
| Savings/Participant (MWh) | 106.3 (120.1) | 48.2 (80.9) | 71.8 (124.8) |
| Lighting Fraction of Total Savings (%) | 53 (33) | 73 (42) | 64 (39)[28] |
| Program Type (Rebate = 0, Direct-Install = 1) | 0 of 5 Direct Install | 3 of 5 Direct Install | 10 of 39 Direct Install |
| Program Start Date | 1987 (5) | 1990 (0) | 1989 (3) |
| Economic Lifetime of Savings (Years) | 14.4 (1.2) | 11.3 (2.2) | 13.1 (3.1) |
| Measure Cost Reporting[29] (Full = 1, Incremental = 0) | 1 of 5 Full | 4 of 5 Full | 25 of 35 Full |
| Savings Evaluation Method[30] (Billing-Metering = 1, Tracking = 0) | 3 of 5 Bill-Meter | 3 of 5 Bill-Meter | 24 of 39 Bill-Meter |

27 Does not include program #9.

28 N = 37 for this explanatory variable.

29 See Chapter 4 for a discussion of measure cost reporting.

30 See Chapter 5 for a discussion of savings evaluation methods.

For the least expensive programs, several features stand out. First, the programs were very large, measured either by annual savings or by number of participants. Second, the programs include some of the older, possibly more mature programs in the sample. Third, because the programs tended to report incremental measure costs, they appear to have targeted normal replacements rather than early replacements.[31]

The most expensive programs also have some common features. First, the programs are quite small as measured by total savings. Second, they appear to be somewhat newer programs compared to the entire sample. Both these factors suggest that these programs are not fully mature, such that fixed administrative costs are being spread over a smaller base of savings. Finally, they include more direct-install programs, for which full measure costs would be reported.

Perhaps more interesting than the differences between the least and most expensive programs are the similarities between them. Avoided costs, the percentage of measure costs paid by the utility, lighting fraction of total savings, economic lifetime of savings, and savings evaluation methods are all quite similar to one another. These similarities have important implications for previous DEEP findings and for the potential impact of methodological differences on current findings.

Eto et al. (1994) found evidence suggesting that avoided costs were positively correlated with TRCs and concluded that avoided costs helped to explain the differences in program costs. They speculated that avoided costs could be thought of as the value standard against which utilities designed programs. In this situation, higher avoided costs led to higher cost programs. In the current situation, the explanation appears more complicated, probably because of confounding influences, such as program size, type, and maturity.

The similarity in the portion of savings attributable to lighting and the similarity in the lighting measures promoted, suggests that differences in the portfolio of technologies installed may have been less important than the savings that resulted from whatever technologies were installed. Earlier we speculated that the more expensive programs may have ended up targeting installations with lower savings (because, for example, of a small number of hours of operation). This finding of similarity in lighting savings fractions lends some credence to this hypothesis. A definitive conclusion can only be drawn by examining detailed demographic information on actual installations.[32]

---

31 As described in Chapter 4, whether the higher cost utilities targeted early replacements (rather than normal replacements) or simply reported full instead of incremental measure costs is more difficult to determine.

32 Unfortunately, as described in Appendix A, we were not able to obtain these data, except in a limited fashion for a few programs.

Similarity in the economic lifetime of savings and savings evaluation methods gives a preview of the methodological findings presented later in Chapter 5. In that chapter, we consider the extent to which bias in assumed measure lifetimes or resulting from choice of savings evaluation method influences our findings. For these programs, we find preliminary evidence suggesting that these potential sources of bias do not appear to be a significant factor in explaining the difference between high- and low-cost programs.

Although examining high- and low-cost programs can illustrate trends, an overview must include information from the entire sample and must rigorously account for the relative influences of various possible explanatory variables on the outcome. For a final look at the data, we used multiple regression techniques to conduct a series of exploratory analyses on the TRC.[33]

In Table 3-4, we present the results from two regressions of the various explanatory variables (drawn from Table 3-3) on TRC.[34] The first model, labeled "best fit," was selected by including only those variables that had the greatest explanatory power.[35] The second model, labeled "all variables," was estimated using all available explanatory variables. Comparing coefficients in the second model for variables included in the first model provides some evidence for the stability of the correlations found in the "best fit" model.

While suggestive, our regression results are by no means definitive. Taken together, the three explanatory variables account for slightly more than 30% of the observed variance in the results. Nevertheless, two variables (program type and program size) appear to be statistically significant (T-statistic greater than 2). Moreover, the coefficients appear stable between the two regressions.

With respect to our earlier findings examining low and high TRC programs, we find strong confirmation for the high cost associated with direct-installation programs and for the comparatively lower cost associated with larger programs (as measured by annual savings). Specifically, we find that direct installation programs cost about 2 ¢/kWh more than rebate programs and that programs costs go down about 1 ¢/kWh for every 100 GWh in annual energy savings.

---

33 We continue our earlier practice of eliminating the very high-cost program #9 from this analysis.

34 We excluded three types of variables from Table 3-3 in running these regressions. First, we excluded "number of participants" because it is a linear combination of "annual savings" and "savings per participant." Second, we excluded "utility contribution to measure costs" because it is highly correlated with "program type (direct versus rebate)." Third, we excluded "measure cost accounting method (full versus incremental)" and "fraction of savings accounted for by lighting measures" because this information was not available for all 39 programs.

35 Specifically, we used the automatic variable selection procedure in SAS called *Forward*.

**Table 3-4. Regression Equations for Total Resource Cost (¢/kWh)**

| | Best Fit | All Variables |
|---|---|---|
| Intercept | 6.35 (3.10) | 38.2 (0.12) |
| Program Type<br>(Direct Installation = 1 versus Rebate = 0) | 2.34 (2.28) | 2.22 (2.05) |
| Program Size (Annual kWh Saved) | -8.63 E-9 (-2.02) | -9.26E-9 (-1.95) |
| Shareholder Incentive (Yes = 1 versus No = 0) | 1.64 (1.67) | 1.67 (1.64) |
| Economic Lifetime of Savings (Years) | -2.58 E-1 (-1.85) | -2.71E-1 (-1.73) |
| Savings/Participant (kWh/participant) | -5.06 E-6 (-1.50) | -5.30E-6 (-1.48) |
| Avoided Cost (¢/kWh) | 1.45 E-1 (1.04) | 1.45E-1 (0.99) |
| Program Start Date | | -1.61E-2 (-0.10) |
| Savings Evaluation Method<br>(Billing-Metering = 1 versus Tracking = 0) | | 4.94E-1 (0.51) |
| Adjusted R-square | 0.312 | 0.273 |

Note: T-statistic in parentheses.

We find evidence of a weak, but not statistically significant, relationship between the TRC and the presence of shareholder incentives, the economic lifetime of savings, savings per participant, and avoided costs. The presence of shareholder incentives and higher avoided costs are correlated with higher program costs. Longer economic lifetimes and higher savings per participant are correlated with lower program costs. We also find that there does not appear to be a statistically significant correlation between the TRC and program start date or program savings evaluation method.[36]

[36] We will further explore the relationship between shareholder incentives, savings evaluation methods, and the economic lifetime of savings in Chapters 4 and 5.

# Developing Consistent Program Cost Information

In this chapter, we describe the development of information on the major cost components of DSM programs. The components of cost include measure costs borne by the utility and participating customers and utility nonmeasure costs, such as program administration, measurement and evaluation, and shareholder incentives. We describe our data sources, discuss issues associated with interpreting them consistently, and document our procedures for developing the final data set. The discussions focus on the key uncertainties associated with each cost component and begin to assess the possible effect of these uncertainties on our findings.[37] We also describe our concerns regarding the inclusion of shareholder incentives in the calculation of the total resource cost of energy savings.

## 4.1 Measure Costs

Measure costs are the costs associated with purchasing and installing an energy-saving measure. Developing these costs for the UC is straightforward; it simply requires identification of rebates and incentives paid to customers in a utility's records. Developing these costs for the TRC is complicated by two overlapping issues. First, measure costs may be borne by the utility, the participating customer or, more typically, partially by both. Second, depending upon the baseline situation/condition assumed by the utility (e.g., normal replacement versus early replacement), only a fraction of the total installed cost of a measure may be assignable to the energy savings from a measure. For clarity, we will refer to these assignable costs as *incremental* measure costs.[38]

### 4.1.1 Utility-Paid Incremental Measure Costs

Utility-paid measure costs (rebates and direct installation costs) were generally well-documented. However, two factors complicate the calculation of utility-paid incremental measure costs for the TRC. The first is free riders—customers who would have acquired an energy-saving measure even if the utility were not sponsoring a program promoting the

[37] Many of these assessments are based on comparisons of the unweighted means of different subgroupings of programs. For these comparisons, we do not include the cost of program #9 because of its extreme influence on the unweighted mean (see Table 3-2).

[38] As described in the text, the term incremental must be defined in comparison to a baseline; in some cases, this means that the incremental cost of a measure will be, in fact, the full cost of the measure.

measure. Conventionally, the costs and savings of free riders are eliminated from the TRC.[39] The second and more subtle factor depends on the definition of the baseline against which measure costs are estimated. The baseline depends on the decision to participate in the utility's DSM program with respect to the decision to replace the existing equipment targeted by the program.

On the one hand, if equipment is at the end of its economic life,[40] the decision to replace it is assumed to be imminent. For energy-saving measures installed in these circumstances, the incremental measure cost is the difference between the cost of the equipment that would normally replace the equipment being retired and the actual cost of the equipment promoted by the utility. Incremental measure costs may legitimately appear to be quite small; for example, there may be no additional installation costs, and additional equipment cost may be only a fraction of the total equipment cost. We call adoption of energy-saving measures at this point in the equipment lifecycle *normal replacements* (see Figure 4-1). Normal replacement is common for HVAC measures, in which equipment at the end of its useful life is replaced by new, energy-efficient equipment.

On the other hand, for customers whose equipment is not at the end of its economic life, the incremental measure cost of energy savings is the full cost of the measure, including full equipment and total installation costs. If the customer had not decided to participate in the utility program, no costs would have been incurred (as well as no savings). We call the adoption of energy savings measures in these situations *early replacements*.[41] Early replacement is more common for lighting measures than for HVAC measures. In these instances, working lighting equipment is removed and replaced with energy-efficient equipment.[42]

Although the distinction between normal replacement and early replacement is easy to make in theory, it is often quite difficult to apply in practice; nonetheless, the difference has great implications for the TRC. For both normal replacements and early replacements, savings may

---

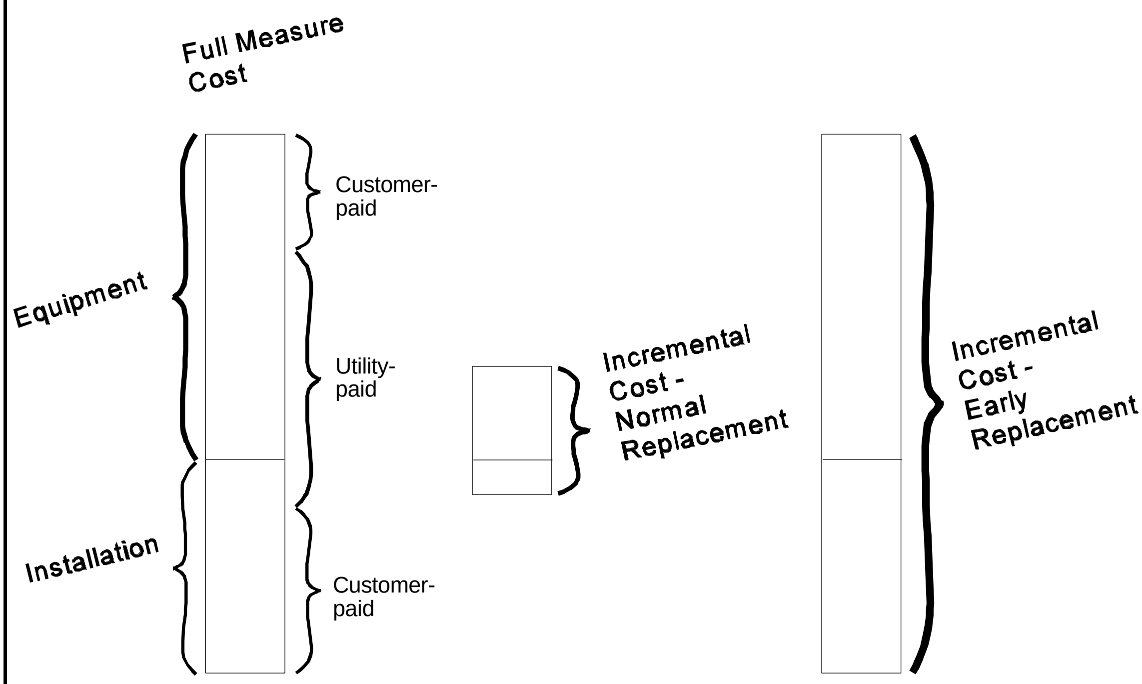
39 Chapter 3 describes how and why our treatment of free ridership for the TRC of energy savings differs from conventional practice.

40 The concept of the economic lifetime of energy savings versus the physical lifetime of energy-saving measures is discussed in Chapter 5.

41 Our choice of this term acknowledges that, at some point, all equipment is likely to be replaced or retired at some point in the future. Hence, participation in a utility's DSM program results in an acceleration of the normal replacement or retirement date.

42 As discussed further in the context of the economic lifetime of savings (see Chapter 5), removing working lighting equipment may still be classified as normal replacement, if replacement is conducted at the time of building remodeling or renovation.

**Figure 4-1. Measure Costs vs. Incremental Measure Costs**




or may not be affected.[43] However, because the costs attributable to the savings differ, the TRC will differ, even though there is no difference in the actual out-of-pocket cost of measures installed.

For example, in examining the measure cost component of the TRC for programs of two utilities with similar customer populations, we found that rebate levels offered by the programs were nearly identical. However, one utility assumed that participants were replacing existing equipment at the end of its economic life (normal replacement) and reported that their rebates covered 100% of the incremental measure costs (including both installation and equipment costs). The other utility assumed that participants were replacing existing equipment prior to the end of its economic life (early replacement) and estimated that their rebates covered only 50% of the incremental (now defined as the full) cost of the measures. Assuming identical savings, the first utility's estimate of the measure cost component of energy savings will be half of the second's.

[43] Savings are a function of the baseline. See Appendix B for a discussion of the treatment of savings from early replacement and normal replacement retrofit programs.

Careful attention to this distinction can also lead to counterintuitive results. For example, a direct-installation program by definition pays 100% of full measure costs. If the measure installed is in fact a normal replacement (and the participant is not a free rider), then the utility has paid more than the incremental measure cost. In this case, the UC could exceed the TRC.[44]

### 4.1.2 Participant-Paid Incremental Measure Costs

Participant costs are not a standard element of a utility's internal accounting system. Utility and regulatory priorities are understandably directed toward utility-paid and, hence, ratepayer-paid DSM program costs. In the best situation, participant cost information is based on customer invoices from completed installations.[45] However, information on participant costs is not reported uniformly.

We did not obtain an estimate of participant costs for 10 programs. For these programs, we worked from information on rebate design, program planning filings, and rebate levels to develop an estimate of these costs. For one program, the rebate design indicated that the rebates were intended to pay for an assumed fraction of usually full but sometimes incremental measure costs. We used the reported fraction to impute participant-paid incremental measure costs. To determine participant-paid incremental measure costs for five programs, we relied on program planning documents describing the projected TRC ratios for programs. To estimate participant costs for two programs, we used information on the rebates paid per measure along with independent estimates of the full cost of measures derived from documents provided by other utilities. This is a conservative estimate in that it assumes all measures were early replacements. For two programs, we examined a sample of rebate applications and customer invoices to determine the fraction of full measure costs paid by the participant; this too is a conservative assumption.[46]

For the remaining 30 programs, some estimate of participant cost was provided, but it was often difficult to determine when the reported costs represented incremental measure costs, as defined in Section 4.1.1. Costs were described in several different ways. Sometimes full versus incremental costs were explicitly labeled. Within the categories of full rather than incremental costs, equipment *and* installation or equipment *but not* installation costs were

---

[44] This is an example of a more general phenomenon which occurs whenever the utility's rebate exceeds the incremental cost of the measure. It can occur whenever rebate designs are based on full measure costs, when, in fact, the program baseline is normal replacement (rather than early replacement).

[45] Customer invoices are the best in this instance because they represent a very tangible form of cost documentation. However, they typically represent full equipment and total installation costs. Whether it is appropriate to consider these to be incremental costs depends on the baseline condition/situation assumed.

[46] We address the issue of just how conservative our assumptions are in Section 4.1.3.

sometimes reported. In most cases, we had to ask our utility contacts to clarify exactly what their reported costs referred to.

In summary, when participant costs were explicitly described as representing incremental measure costs (true for 6 of the 30 programs), we assumed that the utility had made some accounting for the program baseline (i.e., whether the measure was adopted as a normal replacement or an early replacement). When measure costs were termed “full measure cost” (true for 19 of the 30 programs), we could not determine whether this meant that the utility considered all measures early replacements or simply that the utility had not paid rigorous attention to the difference between measure costs for normal replacements versus those for early replacements. We made the conservative assumption that all measures were early replacements for the purposes of estimating incremental participant costs. That is, we did not adjust any reported measure costs. If this assumption is not correct, our estimate of the total resource cost is biased upwards. Finally, for five programs, we could not determine whether measure costs represented full or incremental measure costs.[47]

### 4.1.3 Assessing Uncertainty in Measure Costs

We have made conservative assumptions and interpretations in developing information on measure costs. One way to assess the effect of this conservatism is to examine the difference in the measure cost component of cost of energy savings between programs that report incremental versus full costs. To control for differences in the costs of different types of measures (e.g., HVAC versus lighting), we considered only those programs in which lighting accounted for more than 90% of savings. The results suggest that the differences between the two reporting approaches can have a large effect on the measure cost component of the cost of energy savings (see Table 4-1). Given the small number of programs (3) that reported incremental measure costs, however, the statistical significance of this finding is at best only suggestive.

**Table 4-1. Full vs. Incremental Measure Costs**

| | Mean (¢/kWh) | Standard Deviation (¢/kWh) | Number |
|---|---|---|---|
| Incremental Measure Costs | 2.2 | 0.9 | 3 |
| Full Measure Costs | 4.9 | 1.9 | 12 |

Note: These costs do not include program #9.

We can conclude, however, that our efforts to impute missing participant costs based on full measure costs are likely to be conservative and thus overstate the TRC.

[47] Overall, we classified measure costs as “incremental” for 10 programs, as “full” for 25 programs, and as “unknown” for 5 programs.

## 4.2 Utility Nonmeasure Costs

Utility nonmeasure costs include both direct and indirect costs. Direct costs include the costs associated with running a DSM program, such as advertising and marketing expenses; program recruitment; and administering the payment of rebates (but not the rebates themselves, which are considered to be part of measure costs). Indirect costs include an allocation of general administrative overhead, and measurement and evaluation costs. Although less tangibly connected to the delivery of a program, indirect costs are legitimate elements of the cost of acquiring energy savings. Shareholder incentives are also an indirect cost. However, we discuss them separately in Section 4.3.

### 4.2.1 Direct Nonmeasure Costs

Direct costs were readily available for all programs although the components of direct cost were categorized in a variety of ways (see Table 4-2). Although we observed some general similarities among direct cost categories, we did not understand the accounting systems and procedures of each utility sufficiently to assign costs consistently to a standard set of categories.[48] Hence, we aggregated all categories of direct nonmeasure costs to a single total.

**Table 4-2. Utility Administrative Cost Categories**

| Utility #1 | Utility #2 | Utility #3 | Utility #4 | Utility #5 |
|---|---|---|---|---|
| ○ Implementation Labor | ○ Promotion | ○ Development | ○ Direct Administration | ○ Administration |
| ○ Administrative Labor | ○ Contracted Administration | | | |
| ○ Contract Employees | ○ Program Administration | ○ Implementation | ○ Overhead | |
| ○ Contract Services | ○ Internal Administration | | | |
| ○ Transportation | (for 3 programs | ○ Marketing | | |
| ○ Materials | combined) | | | |
| ○ Tele/Info Services | | | | |
| ○ Graph. Arts/Postage | | | | |
| ○ Employee Expenses | | | | |
| ○ Direct Charges | | | | |
| ○ Miscellaneous Direct | | | | |
| ○ Charges | | | | |
| ○ Overhead | | | | |

### 4.2.2 Indirect Nonmeasure Costs

Indirect costs, consisting of program overhead and measurement and evaluation costs, were also available for most programs. Concerns in analyzing these two types of costs differed.

Overhead was sometimes reported separately and sometimes included as part of direct costs. As with the other direct cost categories, we wanted to assure ourselves that some allocation

48 See Berry (1989) for a comprehensive discussion of issues associated with utility reporting of administrative costs.

of overhead was included. Overhead costs were reported separately for nine programs. For another 29 programs, program documentation or utility staff indicated that overhead was already included in reported direct costs. For another two programs, we could not locate an explicit overhead cost category or determine whether it was already included in the direct costs reported.

Developing information on measurement and evaluation costs presented other challenges. First, measurement and evaluation costs were sometimes not separately reported but were included in other program cost categories. This was especially true for programs whose primary source of estimated savings information was program tracking databases (see Chapter 5). Second, measurement and evaluation costs reported in program year 1992 generally referred to measurement and evaluation activities conducted to estimate savings from a prior program year. Third, when measurement and evaluation costs were separately reported, they were commonly reported as an aggregate total for all measurement and evaluation activities for a given program year.

Our approach to measurement and evaluation costs was as follows. We generally attempted to identify and report measurement and evaluation costs expended to evaluate savings for the 1992 program year by searching records to locate the future year in which they were incurred (we found them for 14 programs). More commonly, we simply relied on 1992 expenditures on measurement and evaluation as a reasonable proxy for the measurement and evaluation costs associated with evaluating the load impacts of the 1992 program (we did this for 23 programs). For three programs, measurement and evaluation costs were not reported separately, but were included in another cost category.

### 4.2.3 Assessing Uncertainty in Indirect Nonmeasure Costs

Some insight into the effect of including or excluding overhead, and measurement and evaluation costs can be gained by examining the subset of programs for which these costs are explicitly reported. For the nine programs that reported explicit overhead costs, overhead costs averaged 4% (standard deviation 4%) of total utility costs (measure + nonmeasure costs). For the 37 programs with identified measurement and evaluation costs, measurement and evaluation averaged 3% (standard deviation 2%) of total utility costs. We conclude that, while inclusion of these costs is important for completeness, neither represents a significant fraction of the UC (and hence, TRC).

## 4.3 The Treatment of Shareholder Incentives

As described in Chapter 2, we include shareholder incentives in our estimates of both the TRC and UC. After discussing our procedure for developing estimates of these costs, we discuss our reasons for including them in the TRC.

For the utilities that receive DSM shareholder incentive payments, we were generally able to locate these payments in regulatory filings. However, because of the design of the incentives, the filings typically contained a single amount reflecting the utility's total reward for DSM activities in a given program year. The designs of shareholder incentive mechanisms include bonuses, rate-return adjustments, shared-savings, and hybrids combining two or more of these individual incentive types.[49] When program-specific incentives were not available, we chose to allocate a portion of total incentive payments to our programs based on the energy saved by each program as a fraction of the total energy saved by all of the utility's DSM programs.

Although there is no question that shareholder incentives are an element of the UC (because they will be recovered from ratepayers), there are differences of opinion about whether they should be included when estimating the TRC. Some argue that shareholder incentives are no more than transfer payments from ratepayers to shareholders and, therefore, are not a cost to society. However, others argue that shareholder incentives are a cost to society like management fees and therefore should be included in the TRC. The difficulty in assessing these positions is that there is no standard for an appropriate management fee for utility delivery of energy savings. For example, Stoft et al. (1995) argue that one must posit the existence of "hidden utility costs" in order to justify and establish the appropriate level for DSM shareholder incentives. At the same time, they concede that there are substantial practical difficulties in estimating hidden costs with precision. Moreover, they speculate that the range in current shareholder incentive payments likely exceeds the range of hidden cost. In economic terms, some of these payments are just transfers.

### 4.3.1 Assessing the Impact of Including Shareholder Incentives in the TRC

We have included shareholder incentives in our calculation of the TRC. Because some fraction of these costs may be transfers, our decision to include them is a conservative one. As with the other uncertainties we have considered, it is useful to consider the effect of assuming instead that utilities do not bear any hidden costs and that all shareholder incentive payments are simply transfers. If we exclude shareholder incentives for the 27 programs that receive them, the mean TRC of programs falls by about 7% from 4.9 ¢/kWh (standard deviation 3.1 ¢/kWh) to 4.6 ¢/kWh (standard deviation 3.1 ¢/kWh). From the standpoint of cost effectiveness, only one program becomes non-cost effective when shareholder incentives are included.

[49] See, for example, Stoft et al. (1995) for a review of DSM shareholder incentive mechanisms.

*CHAPTER 5*

# Measuring the Energy Savings from DSM Programs

In this chapter, we discuss three quantities underlying the measurement of energy savings from DSM programs: annual energy savings, economic lifetime of energy savings, and free riders. The first two directly affect the TRC; all three directly affect the UC. We rely on utility reports and contacts as final and have not made adjustments to the savings information we received. At the same time, we recognize that all three quantities are difficult to measure with precision. In the following sections, we survey current practices and explore the effects of key uncertainties on our findings.

## 5.1 Classifying Methods for Measuring Annual Energy Savings

We classify methods for measuring annual energy savings into three broad categories: (1) tracking database methods, (2) billing analyses, and (3) end-use metering.[50] Most programs use more than one of these methods. For example, all utilities maintain a tracking database of some sort to record information on program participants. Most utilities, however, augment their tracking databases to increase the reliability of their savings estimates. For example, the statistically-adjusted engineering or SAE method reconciles a preliminary estimate of savings from a program's tracking database through a regression on customers' bills. Similarly, end-use metering is often used to refine estimates of hours of operation and, in some cases, changes in connected load. Thus, one can think of the various methods as part of a continuum that starts from a tracking database. The object of our survey is to characterize important differences among and within the three types of methods and to indicate which method or combination of methods was used to develop the savings reported in Chapter 3. Table 5-1 summarizes the annual energy savings methods used by the 40 programs.

### 5.1.1 Tracking Database Methods

Tracking database methods are often referred to as engineering estimates; however, we feel that this name is inaccurate because almost all evaluation methods involve some amount of engineering, so the name should not be applied only to tracking database methods. In addition, the word "engineering" obscures the fact that substantial post-program evaluation information is often incorporated into the estimate. This information ranges from the simple verification of program installations to detailed end-use metering of affected electrical circuits.

[50] See, for example, Sonnenblick and Eto (1995) for detailed discussion of the three approaches and of the relationships among them.

**Table 5-1. Summary of Annual Energy Savings Methods[51]**

| Tracking Database Methods | Number of Programs (% of 40) |
|---|---|
| Verification of Measure Installation | |
| On-Site - Sample | 17 (43%) |
| On-Site - All | 18 (45%) |
| Self-Report - Sample | 4 (10%) |
| Self-Report - All | 2 (5%) |
| Hours of Operation | |
| On-Site - Sample | 17 (43%) |
| On-Site - All | 3 (8%) |
| Self-Report - Sample | 9 (23%) |
| Self-Report - All | 9 (23%) |
| Based on Previous Study | 6 (15%) |
| Billing Analyses | |
| Bill Comparison | 1 (3%) |
| Bill Comparison w/Comparison Group | 3 (8%) |
| Bill Regression w/Comparison Group | 3 (8%) |
| SAE Regression w/Comparison Group | 12 (30%) |
| End-Use Metering | 11 (28%) |

In its simplest form, the basic tracking database equation for energy savings consists of three terms:

Energy Savings = Number of measures installed * Changes in connected load * Hours of use

We distinguish among tracking database methods by the way in which information is introduced into this equation. Starting with the first term, measure installations can be verified by either on-site inspections (as in 35 of our programs) or customer self-reports (six programs). On-site inspections may be conducted by utility staff, contractors to the utility, or both. Customer self-reports include information reported to the utility on a rebate

51 For 11 programs, some methods were applied to only a subset of the energy saved by a program. For five programs, more than one method was used simultaneously to estimate savings for the program. The "number of programs" could add up to more than 40 because some programs used more than one method to verify installations or assess hours of operation.

application or through responses to telephone or mail surveys administered by the utility. The methods are applied either to all participating customers (21 programs) or to a sample of them (20 programs).

Changes in connected load are typically read from engineering tables that compare the connected load of the removed or replaced equipment to that of the more efficient replacement equipment. End-use and short-term, spot metering, described below, are the only alternatives for measuring this quantity directly.

Hours of operation can also be measured with end-use metering. More commonly, they are estimated either through on-site inspections (20 programs) or customer self-reports (18 programs). In this case, on-site inspections are in fact no more than on-site interviews of customers. They may, however, be augmented by inspections of the premises to collect operating information for different zones within a premise. Again, either all (26 programs) or only a sample (12 programs) of participating customers may be surveyed.

For six programs, hours of operation were determined through tables that list "standard" hours of operation for specific end uses (such as lighting or chillers), usually with separate entries for different commercial building types (such as offices or schools). These tables, though often based on previous metering studies, are often unreliable (Sonnenblick and Eto 1995). For five of our programs, these estimates were later either augmented by end-use metering or superseded by an SAE billing analysis. Use of these methods decreases, but does not eliminate concerns regarding the error that could be introduced by relying on look-up tables.

### 5.1.2 Billing Analyses

Billing analyses, in contrast to "bottom-up" tracking database methods, are a "top-down" approach for estimating savings. They are based, at a minimum, on monthly or annual billing information from participating customers, collected both prior to and after the installation of DSM measures. Billing information can be analyzed using a simple differencing approach that directly compares pre-program to post-program consumption, with the individual bills sometimes first weather-normalized (as in 4 programs); they can also be analyzed using multivariate regressions (15 programs). The accuracy and reliability of estimates of net savings can be improved by including billing information from a comparison group of nonparticipating customers (done in 19 programs). A recent, very popular class of regression methods, called the statistically adjusted engineering or SAE method, relies on a preliminary estimate of savings (used by 12 programs). The coefficient emerging from the SAE regression measures is interpreted as a measure of the percentage of previously estimated savings that the regression model is able to confirm.[52]

### 5.1.3 End-Use Metering

End-use metering is often regarded as the most accurate savings evaluation method because it measures the quantities most directly related to energy savings.[53] It is also the most expensive evaluation method to implement because the cost of data collection is high. As a result, it is usually implemented for only a small sample of participating customers. For the 14 programs that relied on end-use metering, the fraction of the population of participating customers metered ranged from less than 1% to 12%. In absolute numbers, nine programs metered fewer than 40 customers, and two metered more than 50 customers. All of the metering studies are classified as short-duration studies, in which the metering periods generally last from two to four weeks.

### 5.1.4 Measuring Takeback

Takeback refers to increases in amenity or energy service resulting from adoption of a DSM program. Its existence has been documented in residential settings, for example, when a customer increases air-conditioning use following purchase of a more efficient air conditioner (see Nadel 1993).

Changes in energy service have not been studied systematically for the commercial sector.[54] Pre-/post-billing analyses can implicitly pick up the energy use impacts of amenity changes resulting from program participation. However, the effect is usually impossible to isolate. Ten programs attempted to identify changes in energy service levels through customer

---

[52] Sonnenblick and Eto (1995) demonstrate that the realization rate coefficient estimated by these models is subject to a well-understood but generally unacknowledged bias and imprecision resulting from errors in the preliminary estimates of energy savings. These errors, which are pervasive in the tracking databases used by SAE models, compromise the straightforward interpretation of the realization rate as representing the fraction of the preliminary estimate of savings “confirmed” by the SAE model.

[53] At the same time, as Sonnenblick and Eto (1995) demonstrate, estimating energy savings from end-use metering is subject to a number of uncertainties, including omission of HVAC/lighting interactions, problems of short-duration metering such as the ability to account for seasonal operating changes, and compromises to the representativeness of the sample resulting from the selection of metered circuits within a premise.

[54] From the standpoint of social welfare maximization, takeback can be socially desirable even though more energy may be used. This can be seen intuitively when an industrial firm modernizes its equipment making it more efficient, yet then uses more energy because it has increased productivity. The most generally recognized instance of takeback in the commercial sector is the installation of energy-efficient outdoor security lighting where there was previously no or only very poor lighting.

surveys. Five concluded that there was no evidence of takeback. Two estimated small amounts of takeback for specific end uses, usually less than 10%. Another five programs made provisions in their savings equations for takeback, but then assumed it was negligible. As noted, the utilities relying on billing analyses (19 programs) all assumed takeback was accounted for implicitly.

### 5.1.5 Measuring Participant and Nonparticipant Spillover

Participant and nonparticipant spillover represent savings that are caused indirectly by the actions of a utility's DSM program. Participant spillover refers to additional energy saving actions taken by a participant outside of the utility's DSM programs (e.g., installing the same or different measures without a rebate). Nonparticipant spillover refers to energy savings actions taken by nonparticipants (i.e., those receiving no rebate) as a result of the program.

Evaluation methods for measuring spillover are in their infancy.[55] Only two utilities made an explicit attempt to incorporate spillover in their estimates of program savings. The evaluations for 14 programs included survey questions on the subject of spillover. In several of these, the survey results were used to develop estimates or ranges of spillover savings. However, while reported, they were not included in the savings reported by the utility. Thus, to the extent that there are spillover effects from the programs, they are not accounted for in either the TRC or UC.

### 5.1.6 Assessing Uncertainties in the Measurement of Annual Energy Savings

There are no generally accepted methods for measuring annual energy savings. All methods are subject to bias and imprecision. There is anecdotal evidence that the simplest forms of tracking database estimates of savings are biased upwards.[56] However, there is little information to judge bias and precision independently.

We conducted a preliminary examination of our programs to see whether the methods used to estimate annual savings were systematically related to the resulting savings. Considering the 24 programs in which lighting accounted for more than 60% of savings, we compared the measure cost component of the TRC for the 15 programs that relied either on billing analyses or end-use metering to estimate savings to the nine programs that relied on a tracking database to estimate savings. As seen in Table 5-2, the mean measure cost of the programs with savings based on tracking databases is slightly lower than the mean for programs with savings based on either billing analyses or end-use metering. Nevertheless, the standard

---

55 See Violette and Rosenberg (1995) for a recent summary of proposals for measuring spillover.

56 See Nadel and Keating (1991) for the first published discussion of this phenomena.

deviations of the two means overwhelm the modest differences in means. Our data, therefore, do not support the existence of a statistically significant correlation between measurement method and annual energy savings. This is consistent with the multiple regression results presented earlier in Chapter 3.

**Table 5-2. DSM Program Measure Costs as a Function of Savings Evaluation Method**

| | Mean (¢/kWh) | Standard Deviation (¢/kWh) | Number of Programs |
|---|---|---|---|
| Billing Analysis or End-Use Metering | 4.4 | 3.2 | 15 |
| Tracking Database w/o End-Use Metering | 4.0 | 2.0 | 9 |

Although this simple examination is by no means definitive, it suggests that there may be more important influences that affect savings other than evaluation methods. We have shown that broad conclusions regarding the bias in various evaluation methods cannot be substantiated by our sample. As indicated, tracking databases vary greatly in the degree and quality of information they incorporate on actual installations. We conclude, in particular, that simple adjustments, such as the application of realization rates developed for one program to adjust the savings from another, cannot be justified without a more detailed understanding of the evaluation methods involved and the populations to which they were applied. See Appendix C for a detailed discussion of this issue.

## 5.2 Estimating the Economic Lifetime of Savings

The economic lifetime of savings is required to establish cumulative energy savings. However, because many DSM technologies are new to the market, studies of the full lifecycle of savings from measures have only been completed for most short-lived measures (those lasting less than five years).

More commonly, utilities have conducted short-term persistence studies to determine measure retention, removal, and failure for one to four years following installation (Wolfe et al. 1995). In our sample, eight programs had completed measure persistence studies that included the 1992 program year. (Typically, the studies include other program years as well.) The studies generally found high rates of persistence for most measures. Notably, several of these studies found low renovation rates in offices, restaurants, and retail premises in contrast to earlier, well-reported findings of high (25% or more per year) renovation rates in these types of premises.[57]

---

[57] See, for example, Skumatz and Hickman (1992), which found comparatively higher rates of business turnover, renovation, and remodeling.

Despite emerging information on the persistence of measures, we found no formal attempts to incorporate this information into reported measure lifetimes. Instead, measure lifetimes were based on a combination of manufacturer's specifications (i.e., equipment lifetime, which can differ from economic lifetime in both directions depending on installation-specific conditions), expert judgment, and, for five programs, negotiated agreements between the utility and the regulator. For several programs, both expert judgment and negotiated agreements were described as taking into account factors such as persistence and pre-mature equipment retirement through building renovation. Nevertheless, the manner in which these factors were accounted for necessarily involves subjective judgments.

Information on the lifetime of savings was generally reported separately for each measure or as a savings-weighted aggregate for all measures (33 programs). We did not receive information on measure or savings lifetimes for seven programs. We developed estimates for three of these programs by constructing a weighted average based on the largest contributors (weighted by either savings, measures, or participants) to savings. For two programs, in which savings were not reported by measure or participant, we made an estimate based on lifetimes reported for programs offering similar measures. For two programs, we used the lifetimes for the popular measures installed.

Measure lifetimes are reported in Table 5-3. They range from six to 18 years. The simple average is 13.0 years with a standard deviation of 3.1 years.

### 5.2.1 Assessing Uncertainties in the Estimation of Economic Lifetimes of Savings

The estimation of the economic lifetime of savings remains a critical source of uncertainty in the measurement of energy savings from utility DSM programs.[58] It will be several years before it is possible to conduct definitive studies to determine the long-term persistence and economic lifetime of savings from many of the most popular DSM measures.

It is straightforward to calculate the sensitivity of the TRC or UC of energy savings to different savings lifetimes. For a program with a savings lifetime of 13 years and a TRC of energy savings of 4 ¢/kWh, a decrease in savings lifetime to 10 years increases the TRC by 22%, and an increase in savings lifetime to 16 years decreases the TRC by 13%.

The economic lifetime of savings from a DSM program also depends on the mix of measures installed. Generally, lighting efficiency measures are assumed to be shorter-lived than HVAC and motor efficiency measures (see Figure 5-1 and Table 5-3).

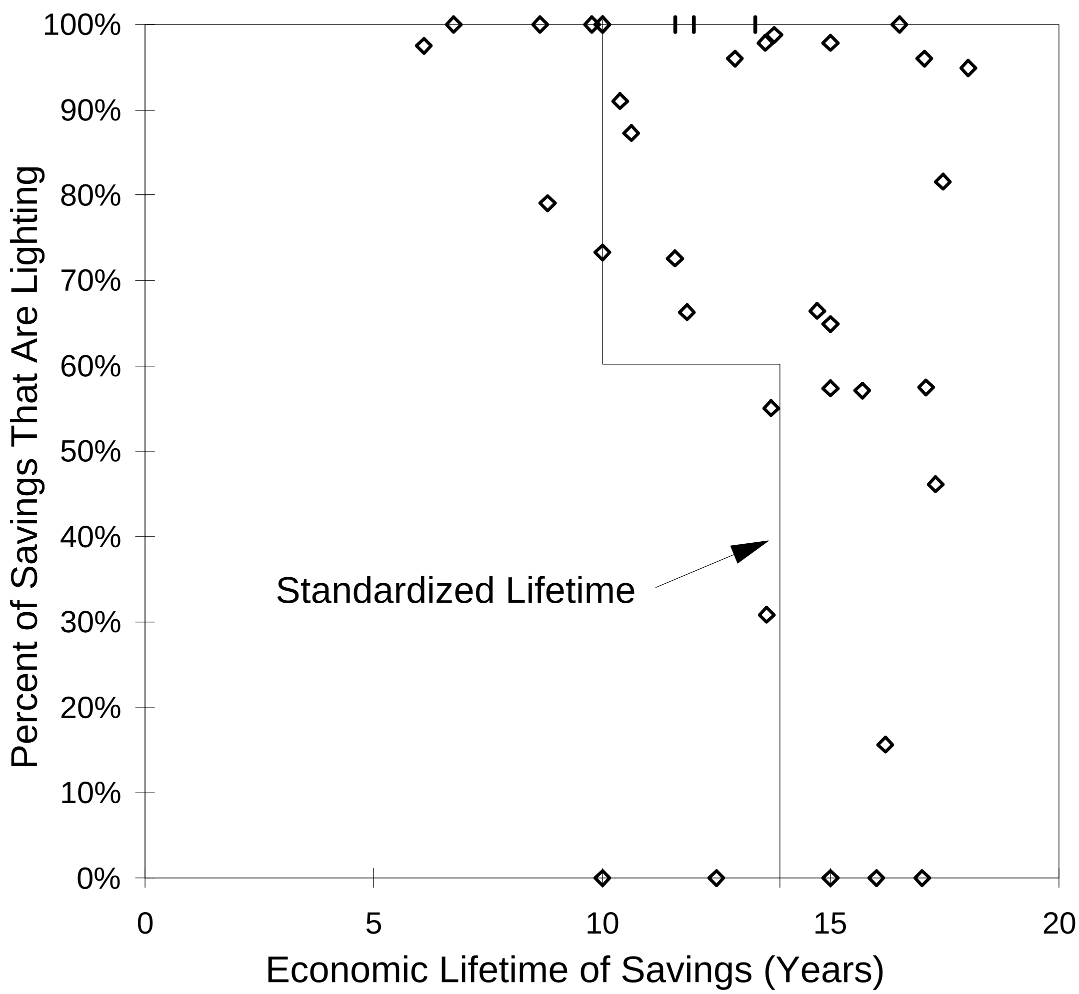
---

[58] Sonnenblick and Eto (1995) demonstrate that the imprecision in savings estimates is typically dominated by imprecision in economic lifetimes rather than imprecision in annual energy savings.

**Table 5-3. Measure Lifetimes and Free-Ridership Rates**

| Program | Term (Years) | Free Ridership (%) | Lighting (%) |
|---|---|---|---|
| 1 | 9.8 | 8% | 100% |
| 2 | 16.0 | 21% | 0% |
| 3 | 10.6 | 9% | 87% |
| 4 | 11.6 | 2% | 100% |
| 5 | 11.9 | 9% | 66% |
| 6 | 10.4 | 0% | 91% |
| 7 | 10.0 | 12% | 100% |
| 8 | 10.0 | 19% | 0% |
| 9 | 10.0 | 0% | 100% |
| 10 | 8.6 | 23% | 100% |
| 11 | 15.0 | 1% | 0% |
| 12 | 13.3 | 23% | 100% |
| 13 | 17.0 | 26% | 0% |
| 14 | 15.0 | 9% | n/a |
| 15 | 17.1 | n/a | 57% |
| 16 | 12.0 | 4% | 100% |
| 17 | 15.0 | 8% | 98% |
| 18 | 8.8 | 31% | 79% |
| 19 | 10.0 | 13% | 73% |
| 20 | 6.8 | 38% | 100% |
| 21 | 12.5 | n/a | 0% |
| 22 | 6.1 | 0% | 97% |
| 23 | 14.7 | 13% | 67% |
| 24 | 15.7 | 7% | 57% |
| 25 | 13.8 | 2% | 99% |
| 26 | 17.5 | 5% | 82% |
| 27 | 13.6 | 2% | 98% |
| 28 | 10.0 | 8% | n/a |
| 29 | 10.0 | 6% | n/a |
| 30 | 17.1 | 17% | 96% |
| 31 | 17.3 | n/a | 46% |
| 32 | 13.6 | 49% | 31% |
| 33 | 13.7 | 25% | 55% |
| 34 | 16.5 | 9% | 100% |
| 35 | 16.2 | 4% | 16% |
| 36 | 15.0 | 0% | 65% |
| 37 | 11.6 | <21% | 73% |
| 38 | 18.0 | 8% | 95% |
| 39 | 12.9 | 17% | 96% |
| 40 | 15.0 | 12% | 57% |
| Average | 13.0 | 12.2% | 69.8% |
| Standard Deviation | 3.1 | 11.4% | 35.2% |

**Figure 5-1. Economic Lifetime of Savings as a Function of the Mix of End-Use Savings**




We examined the sensitivity of our findings to the economic lifetime of savings by replacing reported lifetimes with a standard set of assumptions. For programs in which lighting savings accounted for more than 60% of savings, we assumed a lifetime of 10 years (24 programs). For the remaining programs (in which lighting accounted for less than 60% of savings), we assumed a lifetime of 14 years (12 programs).[59]

[59] We did not obtain information on the fraction of savings accounted for by lighting for three programs. We also excluded program #9 from our comparison due to its extremely high TRC. See discussion in Chapter 3.

We find that the use of standard measure lifetimes increases mean TRCs by about 10%, but that the increase is not statistically significant (see Table 5-4). In particular, use of standard lifetime estimates does not reduce variance in TRC results. We conclude that uncertainty in TRCs due to reliance on necessarily estimated lifetimes is not materially reduced through the use of standard assumptions.

**Table 5-4. The Effect of Assuming Standard Measure Lifetimes on the TRC of Energy Savings**

| | Mean | Standard Deviation (n = 36) |
|---|---|---|
| TRC - Calculated using Reported Measure Lifetimes | 4.9 | 3.2 |
| TRC - Calculated using Standard Measure Lifetimes | 5.5 | 3.3 |

## 5.3 Measuring Free Riders

Free riders directly affect the UC because they reduce the savings attributable to the effects of a utility's DSM program.[60] Because the utility cannot take credit for the savings from free riders, it must spread the costs it incurs over a smaller base of savings. As documented in our previous study (Eto et al. 1994), there are three general approaches for measuring free ridership: surveys, survey-based models of customer choice, and billing analysis with a comparison group. The first two approaches yield a direct estimate of free ridership. The third, in principle, controls implicitly for free ridership.

More than three quarters of the programs (33) conducted surveys to develop an independent estimate of free ridership. Two of these used the surveys to estimate models of consumer choice. One program relied on a billing analysis to control for free riders.[61] Five programs reported free-ridership factors that were based on agreements reached between the regulator and the utility. One program did not conduct a formal evaluation of free ridership and assumed that there was no free ridership in its program. Among the survey-based approaches, nine reported free ridership based solely on participation, and the remainder reported free ridership weighted by program savings. The free-ridership estimates are also reported in Table 5-3. They range from 0% to nearly 50%. The simple average is 12.2% with a standard deviation of 11.4%.

[60] As described in Chapter 3, they also affect the TRC, but to a lesser degree.

[61] In point of fact, many programs indicated that they had controlled for free riders in their billing analyses. However, all but one also developed an independent estimate of free ridership based on survey analysis.

### 5.3.1 Assessing Uncertainties in the Measurement of Free Riders

Evaluation experts have raised a number of questions regarding the accuracy of free-ridership estimates. These questions include the ability of survey methods to determine free ridership based on potentially biased customer responses, the adequacy of billing analysis methods to identify free ridership implicitly (Train 1994), and the ability of either method to capture what are known as deferred and "X" free riders (Nelson 1995). The general conclusion drawn by these commentators is that free ridership may be understated.

To see the potential effect of bias in free-ridership estimates, we recalculated the UC using a common assumption of 20% free ridership. (The mean from our 40 programs is 12.2%.) With this assumption, the mean UC increases slightly from 4.4 ¢/kWh to 4.7 ¢/kWh (standard deviations go from 4.0 ¢/kWh to 3.8 ¢/kWh).[62] We conclude, as was found for lighting, that the potential reduction in bias from the use of standardized assumptions for free ridership is offset by the large variance inherent in the original free-ridership rates. Thus, standardized assumptions have only a modest and not statistically significant impact on UCs.

[62] Program #9 was not included in this calculation.

# Summary

We have calculated the total resource cost and utility cost of energy savings for 40 of the largest 1992 commercial sector DSM programs. The TRC includes the participating customer's cost contribution to energy saving measures. The TRC and UC include program overhead, and measurement and evaluation costs, as well as shareholder incentives. All savings are based on post-program savings evaluations.

We find that, on a savings-weighted basis, the programs have saved energy at a cost of 3.2 ¢/kWh. Several of the least expensive programs rely on significant customer cost contributions. Thus, we find no reason to believe that future DSM programs, which rely on these contributions to minimize rate impacts, will either be more costly or less cost effective.

Taken as a whole, the savings from the programs have been highly cost effective when compared to the avoided costs used in first developing the programs. Moreover, the majority of savings remain cost effective even when compared dramatically lower avoided costs, which are more representative of the avoided costs currently faced by utilities. Nevertheless, a substantial number of individual programs would not be considered cost effective under these lower avoided costs.

The results are dominated by several large and inexpensive programs; some programs, albeit small in absolute size, do not appear to be cost effective. We conducted exploratory analyses to determine what factors help to explain variations in program cost. We found program type and program size to be statistically significant factors; our overall regression equations explained about 30% of the variance in the TRC of energy savings.

Measuring the cost of energy savings delivered by utility DSM programs is difficult because accounting practices and conventions differ among utilities. Information on participant costs, which are of critical importance to the TRC, is especially difficult to collect, both because it is not normally a part of a utility's accounting system and because it depends on the assumed program baseline. We demonstrated that including these costs are important; they account for almost a third of the TRC of energy savings. Overhead and measurement and evaluation costs (and hence concerns about their potential omission) are smaller in comparison to participant costs.

We include shareholder incentives in the UC and the TRC. Including them meaningfully in the TRC requires an assessment of so-called "hidden costs," which are difficult to measure. Even a generous interpretation of the magnitude of these costs, however, does not pose a threat to the cost effectiveness of the programs.

The science of measuring annual energy savings has progressed to the point that the differences among methods are less discernible than they used to be. In particular, savings based on tracking databases now appear to incorporate substantial after-the-fact performance information. At the same time, new questions have been raised challenging the reliability of more sophisticated methods. Our decision not to adjust savings is based on this improved understanding of the strengths and limitations of current approaches. This is not to say that evaluations methods are free from bias and imprecision; they most certainly are not. However, categorical statements regarding bias and imprecision are not supportable without detailed examination of assumptions, methods, and underlying data. Moreover, we demonstrated that differences in savings evaluation methods were not statistically correlated with changes in program costs.

We remain concerned about the accuracy of the estimated economic lifetime of measures because it is still inherently a forecasted quantity. We found, however, that the effect of standardizing measure lifetimes had little measurable effect on the TRC. The measurement of free riders, too, is another area in which differences in estimates appear to reflect the choice and application of evaluation method as well as differences in free ridership. Once again, we found that standardized assumptions had little effect on our results.

No one knows the future of utility DSM programs. However, we feel strongly that discussions about this future should be based on unbiased and critical assessments of the performance of past programs. The goal of the DEEP project is to contribute information on program costs and cost effectiveness, and on the measurement of program costs and savings, to this end.

# References


Berry, L. 1989. “The Administrative Costs of Energy Conservation Programs.” Oak Ridge, TN: Oak Ridge National Laboratory. ORNL/CON-294.

Blumstein, C. and J. Bushnell 1994. “A Guide to the Blue Book: Issues in California’s Electric Industry Restructuring and Reform.” *Electricity Journal,* 7(7):18-29. September.

Busch, J. and J. Eto 1995. “The Estimation of Avoided Costs for Electric Utility Demand-Side Planning.” *Energy Sources* (forthcoming).

California Public Utilities Commission and California Energy Commission (CPUC/CEC) 1987. Standard Practice Manual, Economic Analysis of Demand-Side Management Programs. P400-87-006. Sacramento, CA. December.

Electric Power Research Institute (EPRI) 1991. “Technical Assessment Guide.” EPRI P-2410-SR, Palo Alto, CA. May.

Energy Information Administration (EIA) 1994a. “Financial Statistics of Major Publicly-Owned Utilities 1992.” DOE/EIA-0437(92)/2. Washington, D.C. January.

Energy Information Administration (EIA) 1994b. “Electric Power Annual 1992.” DOE/EIA-0348(92). Washington, D.C. January.

Energy Information Administration (EIA) 1993. “Financial Statistics of Major Investor-Owned Utilities 1992.” DOE/EIA-0437(92)/1. Washington, D.C. December.

Eto, J., E. Vine, L. Shown, R. Sonnnenblick, and C. Payne 1994. “The Cost and Performance of Utility Commercial Lighting Programs.” LBL-34967 Lawrence Berkeley National Laboratory, Berkeley, CA.

Hirst, E. and S. Hadley 1994. “Price Impacts of Electric-Utility DSM Programs.” ORNL/CON-402. Oak Ridge National Laboratory, Oak Ridge, TN.

Joskow, P. and D. Marron 1992. “What Does a Negawatt Really Cost? Evidence from Utility Conservation Programs.” *The Energy Journal,* 13(4):41-74.

Nadel, S. 1993. “The Takeback Effect: Fact or Fiction?” *Proceedings*. 1993 International Energy Program Evaluation Conference. Chicago, IL.

Nadel, S. 1991. “Use of Simple Performance Indices to Help Guide Review of DSM Program Performance.” *Proceedings*. NARUC/DOE National Conference on Integrated Resource Planning. National Association of Regulatory Utility Commissioners, Washington, D.C.

Nadel, S. and K. Keating 1991. “Engineering Estimates versus Impact Evaluation Results: How Do They Compare and Why?” *Proceedings*. 1991 International Energy Program Evaluation Conference. Chicago, IL.

Nelson, D. 1995. "Deferred Free Riders: The Interaction of Persistence and Free Ridership." *Proceedings.* 1995 International Energy Program Evaluation Conference. Chicago, IL.

Pye, M. and S. Nadel 1994. "Rate Impacts of DSM Programs: Looking Past the Rhetoric." *Proceedings.* Fifth National Conference on Integrated Resource Planning. National Association of Regulatory Utility Commissioners, Washington, D.C.

Skumatz, L. and C. Hickman 1992. "Measure Life Study: The Effect of Commercial Building Changes on Energy Using Equipment." *Proceedings*. ACEEE Summer Study on Energy Efficiency in Buildings, 3:3.281-3.292. American Council for an Energy Efficient Economy. Washington, D.C.

Sonnenblick, R. and J. Eto 1995. "A Framework for Improving the Cost Effectiveness of DSM Program Evaluations." LBL-37158. Lawrence Berkeley National Laboratory, Berkeley, CA.

Stoft, S., J. Eto, and S. Kito 1995. "DSM Shareholder Incentives: Current Designs and Economic Theory." LBL-36580. Lawrence Berkeley National Laboratory, Berkeley, CA.

Train, K. 1994. "Estimation of Net Savings from Energy Conservation Programs." *Energy—the International Journal,* 19(4):423-441.

Vine, E. 1995. "Utility Residential New Construction Programs: Going Beyond the Code." LBL-36603. Lawrence Berkeley National Laboratory, Berkeley, CA.

Violette, D. and M. Rosenberg 1995. "Setting A Research Agenda for Assessing Market Transformation and Spillover." *Proceedings.* 1995 International Energy Program Evaluation Conference. Chicago, IL.

Wellinghoff, J. and T. Flanigan 1992. "Commercial Lighting Programs, The Keys to DSM Success." IRT Environment, Inc., Aspen, CO.

Wolfe, A., M. Brown, and D. Trumble. 1995. "Measuring Persistence: A Literature Review Focusing on Methdological Issues." Oak Ridge, TN: Oak Ridge National Laboratory. ORNL/CON-401.

# Program Selection and Data Collection

To obtain information for the current project, we had to address an important new data collection issue: the impact of the California Public Utilities Commission "Blue Book" order. Utility concerns regarding a host of issues loosely labeled "competition" were addressed in this order.[63] The prospect that the monopoly franchise may disappear has led many utilities to adopt a defensive position about sharing information on ratepayer-funded DSM programs.

We began our project by identifying the utilities with the 50 largest DSM program portfolios, as measured by total 1992 DSM energy efficiency program spending reported to the Energy Information Administration on Form EIA-861.[64] We then made preliminary phone calls to verify that each utility had a commercial sector program that spent more than $1 million in 1992. Forty had programs that appeared to meet this criterion (see Table A-1), the Department of Energy and the Electric Power Research Institute sent a formal letter of introduction describing the project to upper-level staff, generally vice presidents or director/managers at these 40 utilities. The letter described our proposed treatment of data, our two-stage verification and review process, and our guarantee not to present information so that it could be attributed to an individual utility.

The letter was successful in enlisting initial participation from 31 utilities with a combined total of 52 commercial sector, energy efficiency DSM programs.[65] We began working with these utilities by requesting that they send us readily available information on their programs. We received information in a variety of forms, including regulatory filings, annual DSM program summaries, and impact and process evaluation reports. We used this information to complete what we could on a detailed data collection form. Some utilities offered to complete the form for us. In several cases, these offers were made because formal documents that would have allowed us to complete the forms were not available. In one case, the offer was because of a corporate policy of not releasing DSM program information.[66]

---

63 The Blue Book order called for, among other things, providing customers with direct access to generation markets (see Blumstein and Bushnell 1994).

64 EIA collects information separately on energy efficiency, load management, and load building DSM programs (see EIA 1994b).

65 Our discussions resulted in only four formal refusals from the original 40 utilities contacted. Five additional utilities did not have programs that met our size criterion (spending of greater than $1 million in 1992) or had programs that were targeted solely to new construction.

66 Ultimately, we were unable to include this utility in our analysis because the same corporate policy also precluded the utility from providing cost or savings information on its program.

**Table A-1. Overview of DSM Program Selection Process**

| | Utilities | Programs |
|---|---|---|
| Largest DSM utilities as measured by 1992 energy efficiency program spending, as reported to EIA | 50 | |
| Based on preliminary information, appeared to have a commercial sector efficiency DSM program, excluding new construction, larger than $1 million in 1992; sent letter soliciting participation | 40 | |
| Sent or agreed to provide program information | 31 | 51 |
| Final data set based on evaluation of completeness, ability to obtain additional information, and/or availability of information from other sources | 23 | 40 |

At this point, we ran headfirst into a major stumbling block. We had planned on an extensive review process with each utility in order to clarify our interpretations of the information provided and to obtain important missing information. However, several utilities indicated that they were not in a position to provide any further assistance to us in data collection, either in verifying the interpretations of the material previously sent to us or in providing the additional information needed to complete the data collection form. In some cases, relevant staff had left the department or the utility; in others, the information we sought was not readily available or had never been collected. In several cases, utilities cited recent cutbacks in staff or related staffing constraints.

We then closely examined the information provided by the utilities, assessed what information we might still be able to obtain either from the utility or from other sources such as commission staff, and made a final list of utilities and programs for inclusion in our report. We revised our data needs and chose to proceed with only those programs for which we felt confident that we could develop a meaningful estimate of the total resource cost of energy savings. At a minimum, this meant that we needed to have or be able to obtain sufficient information on both customer cost contributions and the methods used to estimate savings.[67]

Common features of the data analysis process include: (1) treatment of confidentiality; (2) treatment of data veracity; and (3) treatment of missing data.

In view of concerns about confidentiality that utilities expressed to us as we were developing the project, we have opted not to mention utilities or programs by name. Information is either

[67] Originally, we hoped to develop comprehensive information on: the number of measures installed by technology type, the distribution of savings by end use and premise type, and demographic information on participating customers. We learned that many utilities either do not record these program data or do not record them in a form from which this information can be readily provided.

presented in aggregate or in a form that preclude identification of individual programs or utilities.

Utility reporting and savings evaluation practices differ markedly. For the purposes of our analyses, we take the information provided by the utilities as final. We do not question their veracity or introduce independent judgments to adjust them. However, we recognize that opinions differ regarding the accuracy of the data. In some cases, the documents we reviewed had been the subject of intense regulatory scrutiny, typically as part of a cost recovery or DSM incentive proceeding. In general, we base our information on the most recent sources available in order to reflect utilities' decisions to update previously reported costs and savings.

We have approached the subject of uncertainty in our data humbly. While it would have been straightforward to make adjustments, we concluded that we could not make them confidently without substantially more information on the assumptions underlying the data.[68] Instead, we have attempted to bound the effect of uncertainties by assessing their impacts separately for important cost items or aspects of savings in Chapters 4 and 5. In each case, the objective of our assessments is to understand how potential biases might compromise our findings. In other words, we attempt to confirm the extent to which our findings are driven by differences among programs versus differences in the ways utilities report information on their programs.

For several categories of information, notably customer cost contributions, avoided costs, and shareholder incentives, many utilities were either unable to provide information or unable to provide it in the form required for our analyses. The procedures we developed to estimate or impute these data are also described in Chapters 4 and 5.

68 Appendix C provides an example of the strict conditions that must be satisfied in order make these adjustments confidently.

# On Treating the Benefits and Costs from Early Replacement and Normal Replacement Retrofit Programs Consistently

In Chapter 4, we described how the definition of incremental measure cost depends on the timing of the equipment retrofit decision. We identified two situations, normal replacement and early replacement, which called for the use of different baselines in measuring incremental measure costs. Although we did not discuss the issue in Chapter 5, the same general issue regarding the definition of a program baseline also arises in measuring energy savings. In this Appendix, we propose a unified framework for treating costs and savings consistently for these two retrofit situations.[69]

The gist of our approach is the recognition that early replacement represents an acceleration of an equipment replacement decision that would have taken place at some point in the future. Framed in this manner, the issue becomes one of characterizing how much the replacement decision has been accelerated (in time) and how (if at all) the decision has been changed from that which would have been made in the absence of the program. Thus, for early replacement decisions, there are two periods of interest: period A - the current year through the (future) year when the equipment would "normally" be replaced; and period B - the period after the (future) normal replacement year until the time of the next normal replacement (or retirement).[70] For normal replacement decision, only period B applies.

The treatment of costs and savings for the two replacement decisions is presented in Table B-1. For early replacement, in period A, the baseline consists of the energy use and costs associated with the former (now, replaced equipment). That is, in the absence of the program, this is the equipment that would be using energy. For both early replacement and normal replacement, in period B, the baseline consists of current practice, which we define as the equipment that would typically be installed if there were no DSM program promoting a more efficient equipment option.

Depending on current practice, the program baseline in period B might be represented by the minimum efficiencies called for in a state building code or state/federal product efficiency standard. For normal replacement, the efficiencies called for in existing codes or standards

69 We are indebted to D. Schultz, California Public Utilities Commission, for pointing out the need for this clarification and for his initial thinking in guiding the development of this approach.

70 For simplicity, we will not discuss end-effects beyond the initial replacement decision nor will we discuss treatment of the salvage value of equipment whenever it is replaced. We will also not discuss the incorporation of net-to-gross effects in the calculation.

would be used. For early replacement, the efficiencies expected to be in place at the time of the (future) normal replacement would be used.

**Table B-1. Retrofit Program Evaluation: Normal vs. Early Replacement**

| Parameter Definition | Normal Replacement | Early Replacement |
| --- | --- | --- |
| *Baseline Reference* | Current practice, potentially referencing applicable minimum product efficiencies or practices called for in state/federal standards | Period A: Pre-existing operating condition<br><br>Period B: Current practice at time of (future) normal replacement, potentially referencing a state/federal standard |
| *Incremental Measure Cost* | Full cost of replacement equipment minus current cost of baseline replacement equipment | Full cost of replacement equipment minus present value of the future cost of baseline replacement equipment |
| *Energy Savings* | Post-consumption adjusted for baseline (e.g., minimum product efficiencies or practices called for in state/federal standards) | Period A: Pre-consumption minus post-consumption<br><br>Period B: Post-consumption adjusted for future baseline (e.g., minimum product efficiencies or practices called for in future state/federal standards) |

Period A: Current year through (future) year of normal replacement.
Period B: Year of (future) normal replacement through time of next normal replacement.

# The Transferability of Realization Rates

Because of the cost and complexity of program evaluation, some of the programs in our sample have used the results from one program evaluation to revise their tracking database estimate from another program or from the same program but for another program year's participants. In most cases, a ratio of the ex post evaluation savings estimate to the program savings estimate in the tracking database is used to adjust the unevaluated program's tracking database estimate of savings. Although we applaud the use of as much information as possible in estimating savings and appreciate any attempts to reduce the costs of evaluation, we see potential pitfalls in the indiscriminate use of this technique. In the following paragraphs we explain the difficulties we see with the transferring of realization rates among program participants or programs.

First, we note that all evaluation techniques are susceptible to error, and recommend that this error should be reported with the evaluation result. Generally error is characterized as imprecision around the evaluation result, and the imprecision is assumed to be normally distributed (i.e., with a bell-shaped curve) and reported as a symmetric confidence interval around the point estimate.[71] Imprecision should describe the uncertainty of the result based on the practical and theoretical limitations of the evaluation technique(s) used. For example, techniques that sample only a segment of the participant population are subject to some uncertainty based on the size and variability of the sample relative to the entire population. Calculation of imprecision can also involve subjective judgments, as in the case of persistence of savings throughout a measure's assumed lifetime: A subjective estimate of imprecision, based on expert judgment of a program's designer and evaluation regarding persistence of savings over time could be used to bound the annual savings estimate. What is important is that an effort be made to quantitatively estimate and communicate the limitations of the evaluation methods used. Recognizing that an estimate *is thought to* be accurate to +/-5% is different from accepting that the same estimate *is* accurate to +/-50%.

Second, we assert that transferring a realization rate from the population of participants for which it was calculated to another population of participants must involve an increase in error. This is because there are always some differences between the two programs' tracking databases, participant populations, and program characteristics, and these differences may

[71] Evaluation methods are also susceptible to bias. Bias includes any systematic errors which may be present in the evaluation methods used, resulting in a savings estimate that under- or overstates actual savings. Evaluation bias is difficult to uncover and is often assumed to be insignificant. In fact, an underlying (yet, to our mind, unproven) tenet of both end-use metering sampling and ordinary least-squares (OLS) regression in billing analyses is that the intermediate calculations, and thus the results, are unbiased. See Sonnenblick and Eto (1995) for further development of this issue.

alter the relationship between the tracking database estimate and the savings actually achieved by the participants. Minimally there is an increase in the imprecision of the resulting savings estimate, representing the uncertainty about the homogeneity of all characteristics between the two programs. The increase in imprecision may be compounded by significant bias if there are systematic differences related to energy consumption patterns between the two populations.

The following list describes the key areas that we believe introduce error when realization rates are transferred:

*Differences in the methods used to compile the tracking database information.* Sonnenblick and Eto (1995) find large variation in the accuracy of tracking database estimates of savings, depending on the sources of the information input into a database and the sophistication and flexibility of the database. For example, programs with tracking databases that incorporate information from rebate applications will probably not provide estimates as accurate (i.e., as unbiased and precise) as those from a tracking database incorporating information from site inspections of each participant's facility. Because the realization rate is based on the ratio of the ex post evaluation savings estimate to the tracking database estimate, any differences in tracking database organization and data collection between two programs could hinder one's ability to transfer a realization rate from one program to the other.

*Differences in participant characteristics that affect energy consumption and program savings.* A realization rate asserts that, on average, some percentage of the tracking database estimate of savings is actually saved by program participants. The extent to which this ratio is the same for another program is dependent on the homogeneity of the participants by rate class, by geographical location, by climate, by financial circumstances, etc. If these characteristics vary, the realization rate from one program may not accurately represent the percentage of verified savings from another program's tracking database.

*Differences in program measures, program delivery, and program rebates.* Energy conservation supply curves demonstrate that different energy efficiency measures possess different costs and benefits. From this fact it is a small step to understand that different energy efficiency measures may also save more or less energy than expected. A program's marketing and delivery characteristics may also effect the percentage of tracking database savings realized by customers.

How much imprecision does the transfer of a realization rate add to a savings estimate? Answering this question requires information on the explicit differences between the two programs and the extent to which these differences might affect the ratio of ex post evaluation results to tracking database savings estimates. An upper-bound for the imprecision could be estimated based on expert judgment and a basic understanding of the range of possible realization rates. However, this upper bound could encompass such a wide range of values that the results could be of negligible importance beyond the tracking database estimate itself.

In summary, we are skeptical of the propriety of transferring realization rates to other populations unless the implications for estimate precision are considered. Even though consideration of estimate precision may not be possible beyond the conceptual level (i.e., the level of expert judgment) we believe such an exercise is essential to ensure the integrity of the resulting information and its appropriate use in resource planning, program screening, and cost-recovery/incentive activities.