Emerging Issues in Evaluation

Elizabeth G. Hicks, New England Power Service Company

DSM is constantly evolving as is evaluation. This paper will describe some new issues in the evaluation area including: evaluation and shareholder incentives, implications of new trends in program design, persistence, combination of methods, use of evaluation results to environmental credits and the future role of DSM.

Introduction

The field of DSM evaluation has come a long way since 1990 when Steve Weil called upon the evaluators at this conference to do more (Weil 1990). DSM programs have expanded from approximately $1.0B in 1990 to an estimated $2.5B in 1994 (ADSMP 1994). Evaluation activities by utilities, regulators and other entities have followed this trend. Many states have regulatory policies that provide utilities with shareholder incentives and in many cases the incentives and lost revenue payments are based on evaluation results (Schlegel et al. 1993) (MDPU - 90-251). This use of evaluation results has created the verification versus evaluation debate in some quarters. (Kushler et al. 1992)

DSM at electric utilities has continuously evolved since the 1980's. Programs are offered to all customer classes with a variety of delivery mechanisms including: rebates, direct installation, audits, educational materials or performance based. Gas DSM is also growing (ADSMP 1994).

Just as DSM is changing, so is evaluation. This paper will discuss some of the current issues in the evaluation area, and present the author's vision of the future in this area. In particular it will tie some of the changes in evaluation to program design, shareholder incentives, environmental trends, analytical trends, increased competition in the utility sector, and the role of DSM in the future.

Evaluation and Program Design

The selection of evaluation activities is closely linked to the design of the DSM programs being evaluated. DSM programs have changed greatly since the late 1980's and early 1990's. In particular there is a trend toward lower rebates (at least in the Northeast) financing, more complex measures, market transformation programs, and a resurgence of interest in innovative pricing.

As utilities move away from full-cost rebates, possibly by having customers finance part or all of the cost of DSM measures, free-ridership may become higher and possibly more difficult to measure. Also, utilities are often using DSM as a customer-retention strategy or for economic development. As of yet, little evaluation work has been done on DSM as a customer-retention strategy, so this is potentially an area that may grow in the next few years. Converting the concept of retained customers into measurable savings may be difficult.

As utilities move toward more complex measures such as HVAC, VSD's and industrial process measures, much more site-specific analysis is required than for lighting. Measuring the savings associated with these technologies is both more complex and more costly, and often the information is site-specific and not easily transferable to other sites. For example, in some of New England Electric's recent evaluation work, it costs approximately $30-40K per site to monitor VSD installations, up to $60K to measure storage cooling or liquid pressure amplifiers. There is much less information on the measurement of these types of technologies than more familiar Commercial/Industrial Lighting technologies. Utilities that have done work in this area include New England Electric, Northeast Utilities, PG&E and some of the Wisconsin utilities.

Particularly in areas where utilities have a capacity surplus, a major focus of DSM programs is on market transformation especially lost opportunity markets such as new construction and equipment replacement. Provided an efficient program design is used, DSM is potentially at its least expensive in the new construction, equipment replacement and remodeling markets (Chaisson 1992). One of the specific challenges of evaluation in new construction or equipment replacement is that there is no "before." Hence, while metering can clearly help to determine operating parameters such as hours of use, it cannot necessarily accurately determine the "delta watts. " As an illustrative example, only one of the 32 refereed evaluation papers at the 1992 ACEEE conference dealt with new construction. The source of this type of information is a good baseline survey or study of existing equipment/new construction practices. Part of a good baseline study is developing an understanding of how the underlying market for a given technology works, as well as the interaction between the customer and the distributor, the distributor and the wholesaler, and the wholesaler and the manufacturer. A good example of this can be found in (Easton 1993), where many of the New England utilities worked together to develop baselines for HVAC. This study developed a very good representation of the HVAC market in New England. It included in person and telephone interviews with market participants, review of equipment specifications and analysis.

A good understanding of the market is also critical in any efforts at estimating spillover from DSM programs. Spillover can be loosely defined as any additional efficiency increase or conservation action that can be tied to a utility DSM program, above and beyond that from the measures installed in the program. Spillover can also encompass other measures of market transformation such as reduced product costs. This type of impact can affect several groups, including participants, nonparticipants in the utility's service territory, nonparticipants outside the utility's service territory, and lastly, the supporting infrastructure of distributors, retailers, and manufacturers.

Many studies have been initiated to explore ways to measure spillover effects. They include efforts by the Wisconsin Center for Demand Side Research, the California CADMAC Study, and a consortium of New England utilities and their regulators. The New England effort, while not complete at the time of this paper will include: a definition of spillover, literature review, bibliography and discussion of potential analytical techniques to quantify spillover (XENERGY and Easton 1994). A key evaluation/load forecasting issue that evaluators will have to attempt to quantify is the interaction between spillover associated with DSM programs and the impact of new energy-efficiency standards proposed under EPACT.

As their world becomes increasingly more competitive, utilities are exploring more innovative and flexible pricing. Recent examples include real time pricing (Georgia Power, Niagara Mohawk) and green pricing (Southern California Edison). The evaluation of these types of programs is increasing. A number of evaluations or other studies of these types of programs can be found at this conference in the Innovative Pricing Session.

Role of Evaluation and Shareholder Incentives

The concept of shareholder incentives was first raised in the late 1980's. Even before this, in 1986, the Wisconsin Public Service Commission allowed WEPCO to earn an additional 1% of return on some of its conservation investment (Chamberlain et al. 1992). As of September, 1993, 28 states and the District of Columbia have approved shareholder incentives (ADSMP 1994), and 21 states and the District of Columbia have approved lost revenue recovery.

Many states require that utilities use evaluation results in some manner in their calculation of shareholder incentives and lost revenues. Examples include California and Massachusetts. In Massachusetts, Boston Edison, Western Massachusetts Electric, and Massachusetts Electric are required to file evaluation results as part of their shareholder incentive filings. No statewide protocols exist; each of the utilities has filed its individual evaluation results with the MDPU. A recent MDPU order did set a standard as opposed to a protocol that the MDPU will accept evaluations if they are deemed to be reviewable, appropriate and reliable (MDPU 1994). In California, a specific set of protocols is in place, which requires that utilities revisit their savings at specific intervals.

As with most initiatives, there are both pros and cons associated with tying shareholder incentives to evaluation results. On the con side, it usually puts evaluation on a cycle determined by regulatory requirements rather than research needs. For example, many of the Massachusetts utilities are required to file their results for 1993 on June 1, 1994. This has the potential to focus evaluation efforts on measuring kW and kWh savings, not process or market issues. It also may potentially divert resources away from other DSM activities into DSM accounting. Finally, utilities may be less likely to try new evaluation methodologies because of the perceived "risk" of tying shareholder incentives to untried methodologies.

There are many pros as well. The first is that it has caused utilities to expend more resources on evaluation and to take evaluation results seriously. It has helped the regulatory community to become more familiar with evaluation issues. It has helped to ensure that utilities close the loop, i.e. that they use their evaluation results to improve DSM program design as well as in resource planning. It has increased utility accountability for both their DSM programs and their evaluation activities. It may have also shortened the evaluation cycle, which could help to implement program design changes more quickly. Overall, the benefits of linking evaluation results to shareholder incentives exceed the negatives,

A review of the evaluation practices of the California utilities found that utilities with shareholder incentives linked to evaluation activities developed better evaluation plans, accelerated their evaluation efforts, and produced results closer to schedule. The study concluded that linking incentives to performance did improve the overall quality and scope of the evaluation efforts of these utilities (Raab and Violette 1994) (Schlegel et al. 1993).

DSM as an Environmental Strategy

Utilities are using DSM as part of their environmental strategies. For example, New England Electric includes DSM as a very important component of its environmental strategy along with renewable, reduction of NOx emissions via SNCR and other methods, greenhouse gas offsets and other initiatives such as electric vehicles (NEES 1993). As of January 1993, 26 state commissions had some sort of requirements in place for incorporating externalities in the planning process (Hashem et al. 1993).

In order to receive credits in EPA's Conservation and Renewable Reserve, a utility must either use EPA's protocol or have its state regulatory authority certify that it performs evaluation to the regulator authority's standard. To date, for example, Massachusetts Electric, Narragansett Electric and Granite State Electric have received 97, 27, and 6 SO_2 allowances credits respectively, for savings associated with their 1992 DSM programs.

Some states, such as Massachusetts, are currently developing procedures for NOx offsets for DSM (Donovan 1994). Utility DSM programs are an important component of the President's Climate Change Action Plan. A longer run component of this is that utilities may need to "tally" how much CO_2 they have saved since 1990. This could add a new wrinkle to evaluation, by more closely linking actual DSM impacts to utility dispatch or more likely, a power pool dispatch of resources.

A related area may be the comparison of competing technology emissions. An example is the environmental impact evaluation of Southern California Edison's (SCE) Residential Energy Usage Comparison Project in which measurements were taken of NOx, SO_2, CO_2, ROG's (reactive organic gases and particulars) (Smith et al. 1993). The EPRI CLEAN Project is also attempting to provide utilities with this type of information (EPRI 1994).

In some cases evaluators are taking a closer look at the environmental impacts that may come from DSM programs. For example, in Ontario Hydro's evaluation of the Espanola Power Saver's project (a community-based conservation initiative), the evaluation looked at:

1. indoor air quality impacts,
2. the risk of Legionnaires disease,
3. radon,
4. other indoor air polluters, and
5. waste management.

(MacLeod and Haites 1993)

Clearly, if the interest in DSM as an environmental strategy continues, evaluation may need to focus more on this area.

Analytical Trends

Overview

Numerous techniques can be used to evaluate DSM programs. The approach chosen will depend on the size, scope and type of program, the decisions that will need to be made, the goals of the evaluation, the methods the evaluation will use, and the level of accuracy required (Kushler et al. 1992). Kushler identified the most common impact evaluation methodologies as: engineering methods, statistical billing analysis, metering surveys and on-site visits (Kushler et al. 1992). Process evaluations review program information, and use interviews, surveys and focus groups to get input from utility staff, participating and nonparticipating customers, and implementation contractors and suppliers.

This section will briefly discuss some recent trends in the impact and process evaluation areas.

Process Evaluations

As DSM programs mature, in some cases full-scale process evaluations may not be needed every year if programs are operating relatively smoothly. A current trend is to perform "mini" process evaluations: process evaluations that focus on specific issues or programs elements. As an example, New England Electric is currently performing evaluations on specific issues within its large C/I programs. The Company is also conducting mini/process evaluations on its commissioning services and the internal process by which evaluation results are input to be the DSM data-base. To some extent this can be viewed as applying practices learned in the evaluation process to related parts of the business.

Impact Evaluations

This section will deal with two areas of impact analysis: persistence studies and the combination of multiple results. While these are clearly not the only issues plaguing impact evaluations, they are of great importance. Trends in the metering and monitoring area are discussed in detail in Goldberg, (1994), and hence are not raised here.

Persistence. As discussed in Raab and Violette (1994), "the major dilemma in the study of persistence is trying to develop a useful strategy at a reasonable cost." Persistence itself can be broken into measure persistence and program persistence. Measure persistence is really the gross savings impact of the actual measure, while program persistence is the net impact of the whole program over a period of time. Measures may not "persist" for many reasons, including measure failure, market factors such as remodeling cycles, operational factors such as O&M, and technical degradation (Jeppson and Rudman 1993). Along with measure persistence, Snapback, measure replacement, and free-ridership can impact program persistence over time (Raab and Violette 1994).

Raab and Violette's analysis found that only three states—Wisconsin, New York and California-have begun to address persistence systematically from a regulatory perspective. Both New York and Wisconsin are now requiring utilities to file annual plans on how they plan to address persistence in their DSM programs. California, in its March 1993 PUC Interim Order on ExPost Measurement and Evaluation, adopted an approach to measuring persistence that is based on gross measure impacts, with repeated measurements further out in time. A very recent MDPU ruling directed Massachusetts Electric to perform persistence studies on the majority of its DSM programs (MDPU 1994).

Experience in the actual measurement of persistence is limited. The two most common methods are longer-term billing analysis and on-site visits.

The best examples of longer-term billing analysis come from Seattle City Light. For example, Coates explored persistence of savings, using billing analysis for participants in the Commercial Incentives Pilot Program (BPA-sponsored program). In this study, savings for 1987, 1988, and 1989 were determined, using billing analysis (Coates 1992). The measures installed were predominately lighting measures. Seattle City Light's analysis of its multifamily retrofit program also used billing analysis to determine longer-term program impacts (Okumo 1992). In this case, the Company analyzed participants from 1986 and 1987 to determine longer-term savings for this program.

The New England Electric Companies have used on-site visits extensively to determine persistence of savings. Because most of NEES's programs did not start as full-scale programs until 1990 or 1991, the measurements in this analysis go back only 1 or 2 years (MECo 1992), (MECo 1993). The measures examined in these studies include C&I Lighting, window film, economizers, injection molding machines, and residential lighting.

Persistence is a major evaluation issue. Evaluation will help utilities and other parties determine for which measures or programs persistence may be a particular problem. It is a problem that may best be solved not by measurement, but by solid program design practices. Vine (1992) suggests the following strategy for dealing with persistence.

- measurement and verification plans,
- program design,
- operations and maintenance procedures,
- building commissioning,
- training and education,
- technology performance tools, and
- cooperative research projects.

These practices, along with good measurement, should help to ensure that DSM measures provide savings for a long time to come.

Combination of Results. The use of multiple studies to determine impacts from DSM is clearly becoming more commonplace. There are really two general models for combining results. The first is to combine, in some manner, the results of different evaluation studies of the same program. Techniques for this include triangulation (comparing the results of more than one method to estimate program impact), leveraging data (using data or results from one method as input into another), and baysian methods (using a systematic approach to adapt and update prior information based on results from new analysis) (Violette 1991). The other model for multiple studies is the use of meta-analysis, which is a term used to refer to a series of quantitative methodologies for synthesizing the results of multiple studies.

The objectives of these two categories of evaluation approaches are quite different. In the case of combining different studies of the same program, the objectives are usually to obtain a better understanding of divergent results, to use this knowledge to improve the analytical techniques used, and to produce overall program results that are based on multiple studies. An example of this type of analysis can be found in Karr et al., (1993), where the results from two components of the evaluation strategy for a commercial lighting program were combined—in this case, billing analysis and on-site monitoring.

The objectives of meta-analysis are several: 1) to synthesize findings or research or to combine research findings, 2) to analyze the findings of multiple researchers to determine if their findings are consistent, and 3) to find the factors that might be causing findings to be inconsistent and to quantify the impact of these inconsistencies (Greene et al. 1993).

In Greene, et al., a meta-analysis of the realization rates of a number of utilities' commercial and industrial programs were examined. This study found meta-analysis to have a great deal of promise, but also identifies some lessons learned. In particular, the authors found variations in how much of the information was presented across utilities, differences in engineering algorithms for similar measures, and incomplete information.

These sorts of analysis will become more commonplace as the literature on DSM program evaluation grows—provided, of course, that researchers in these areas continue to make their results public. However, as the utility industry becomes more competitive, utilities may not "go public" with their results.

Confidentiality of Results

As the utility industry becomes more competitive, which will be discussed in more detail in the following sections, it is likely that results of evaluation may be kept confidential. Concerns about competition from alternative fuel suppliers, increased competition among electric utilities, the potential for retail competition, and cases where integrated resource planning regulators require utilities to put programs out to bid are driving forces.

Utilities in Wisconsin already file many of their evaluations under seal. Boston Edison recently asked the Massachusetts Department of Public Utilities to issue a protective order to prevent disclosure of evaluation-related data. The hearing officer in this case recommended that the request be denied. To date, the Company has not appealed the decision (Raab and Violette 1994).

How this all will evolve is unclear. Even though the Boston Edison request was denied, Boston Gas is now asking for similar protection from the MDPU. If these kinds of requests become more common-place, the ability to share information in the evaluation area may be severely limited. It will also potentially make it more difficult to perform multi-utility studies, which may be the best approach for measuring phenomena such as spillover, or new construction baselines which address more than one utility's service area. Clearly, the extent to which evaluation results will be shared in the future will be highly dependent on how the utility industry evolves and how DSM evolves in the new environment.

DSM in a More Competitive Environment and Evaluation Implications

This section identifies the possible future of DSM programs and the potential implications for evaluation. Clearly, this is not an attempt to project the future, but an effort to present a possible framework to examine DSM and evaluation in the context of a more competitive utility market.

Hirst (1994) discusses possible forms that competition may take in generation, transmission and distribution, and retail in the electric industry (see Figure 1). His view is that the future of DSM programs is largely independent of the extent of competition in either the generation or T&D markets. He finds utility DSM programs are most likely to be affected by competition at the retail level (Hirst 1994).

As shown in Figure 1, Hirst uses three potential scenarios for the retail electric sector:

1. Franchise monopoly, obligation to serve remains intact;

2. Monopoly largely intact with some wheeling; and,

3. Competition for a certain size of customer.

In the first scenario, the franchise will hold and current electric utilities will continue to serve their existing customers. In this case some large customers will cogenerate or find alternative suppliers. In the second scenario larger customers will be able to obtain service from a number of alternative producers or other utilities. In the third case customers above a certain size would be able to choose their supplier. Utilities as well as retail brokers would be vying for customers.

Hirst asserts that under the first scenario, the impact on utility DSM programs is likely to be minimal, or at least not driven by competitive forces. On the other hand, he suggests that if there is considerable competition for end-use customers, DSM programs are likely to change dramatically, as utilities will not be able to charge all customers for the costs of DSM programs. Prices will be set in a competitive arena.

Hirst suggests that under increasing competitive pressures, utilities will focus DSM efforts more on customer service and less on system resource benefits. There will be more emphasis on capacity savings, and utilities will seek to have individual customers pay for these services. As utilities become more and more cost-conscious, they will identify ways to make programs more cost-effective.

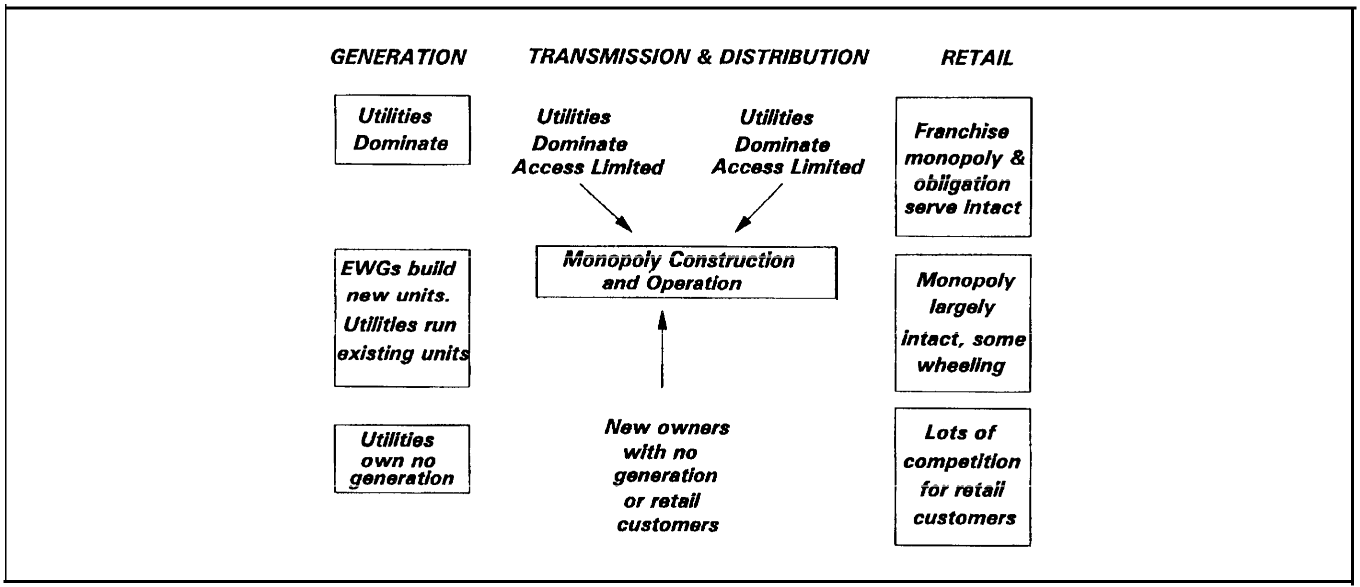

Figure 1. Possible Forms of Competition

Evaluation will be a major tool in helping utilities identify cost saving program design changes. If indeed there is more focus on reducing capacity needs to reduce lost revenues, evaluation may need to place more emphasis on measuring kW benefits. (Currently the emphasis on kW versus kWh savings varies greatly.) The move away from a resource perspective toward a service perspective may mean focusing evaluation efforts more on customer satisfaction and other more traditional market-research techniques.

Skills learned in evaluation will also be highly valuable in other areas of the utility in a more competitive market. If more focus on marketing programs occurs, there will be a continual need to evaluate these programs to help to ensure that the utilities' marketing programs are successful.

How the future will evolve is more than murky. It will be a challenge not only for evaluators but also for all involved in the electric utility industry, including the regulators.

References

Chaisson, 1992. Chaisson, J. and Coakley, S. "Setting Priorities for Utility Investment in Cost Effective DSM." *ACEEE Summer Study on Energy Efficiency in Buildings,* Asilomar, CA, 1992.

Coates, 1992, Coates, Brian. "Persistence of Energy Savings in a Commercial Conservation Program," *ACEEE Summer Study on Energy Efficiency in Buildings,* Asilomar, CA, 1992.

Donovan, 1994. Donovan, Deborah, New England Power Service Company, personal communication.

Easton, 1993. Easton Consultants, New England. "HVAC Baseline Study," 1993.

EPRI, 1994. Electric Power Research Institute, "End-Use Emission Update," March 1994, Volume 2, Issue.

Greene et al., 1993, Pamela Greene, RCG/Hagler Bailey and others. "A Meter Analysis of Commercial and Industrial Lighting Programs," 1993 *International Energy Program Evaluation Conference,* Chicago, IL, 1993.

Hashem, 1993. Hashem, J. et al. "States Report: State Requirements for Considering Environmental Externalities in Electric Utility Decision-making." Berkeley, CA Association for Demand-Side Management Professionals, March, 1993.

Hirst, 1994. Hirst, E. "Electric Utility Programs in a Competitive Market Draft Report," Oak Ridge National Laboratory, 1994.

Jeppson and Rudman, 1993. John Jeppson, RCG/Hagler, Bailey and Monica Rudman, California Energy Commission. "Measure Life Research: Options and Lessons", *1993 International Energy Program Evaluation Conference,* Chicago, IL, 1993.

Karr et al., 1993, Anthony Karr and others. "Synthesis of Findings from Two Independent Impact Studies of a C/I Retrofit Program," 1993 *International Energy Program Evaluation Conference,* Chicago, IL, 1993.

Kushler et al., 1992. Kushler, Martin and others. "The Purpose, Practice and Profession of DSM Evaluation: Current Trend-Rand Future Challenges," *ACEEE 1992 Summer Study, 1992*, Asilomar, CA, 1992.

Lagerberg, 1993, Brian Lagerberg, Vince Schueler and Paul Leistner. Washington State Energy Office, "A Meta-Analystic Approach to Estimating Natural Gas Conservation Savings," Preliminary Weatherization Program Findings, 1993 *International Energy Program Evaluation Conference*, Chicago, IL, 1993.

MacLeod and Haites, 1992. MacLeod, Gillian and Haites, Eric. "Evaluation of the Environmental Impacts of the Espanola Power Savers Project," *1993 International Energy Program Evaluation Conference*, Chicago, IL, 1993.

MDPU, 1994. Massachusetts Department of Public Utilities, Order 92-217-B, 1994.

MECo, 1993, Massachusetts Electric, "1992 Performance Measurement Report," June 1993.

MECo, 1992, Massachusetts Electric, "1991 Performance Measurement Report," June 1992.

NEES, 1993, New England Electric, "NEESPLAN 4", Westborough, MA.

Okumo, Debra, 1992. "Revisiting Multifamily Retrofit Electricity Savings," *ACEEE 1992 Summer Study on Energy Efficiency in Buildings*, Asilomar, CA, 1992.

Raab and Violette, 1994. J. Raab and D. Violette. "Regulating DSM Program Evaluation: Policy and Administrative Issues for Public Utility Commission," NARUC, Washington, D.C., 1994.

Schlegel, 1993. Schlegel, J., et al. "Do Shareholder Incentives Work?" 1993 *International Energy Program Evaluation Conference*, Chicago, IL, 1993.

Smith et al., 1993. Smith, Bruce, Dykema, John, Quantum Consulting and Delaney, Paul, Southern California Edison. "The Three E's of DSM: Evaluation, Externalities and the Environment," *1993 International Energy Program Evaluation Conference*, Chicago, IL, 1993.

Vine, 1992. Vine, Edward. "Persistence of Energy Savings: What Do We Know and How Can It Be Ensured?" *ACEEE 1992 Summer Study on Energy Efficiency in Buildings*, Asilomar, CA, 1992.

Violette, 1991. Violette, D., et al. "Impact Evaluation of Demand Side Management Programs," Volume I, A Guide to Current Practice, EPRI, February 1991, EPRI-CU-7179.

XENERGY and Easton, 1994. XENERGY, Inc. and Easton Consultants. "Draft Report, Literature and Definition Review Spillover Scoping Study," Burlington, MA, 1994.

Weil, 1990. Weil, Steve. "The Urgent Need for Verifying DSM Achievements," *ACEEE 1990 Summer Study on Energy Efficiency in Buildings*, Asilomar, CA, 1990.