The California Evaluation Framework: Moving Evaluation into a Systems Approach for Planning and Conducting Energy Program Evaluations

Marian Brown, Southern California Edison Company
Lori Megdal, Megdal & Associates
Pete Jacobs, Architectural Energy Corporation
Nick Hall, Johna Roth, TecMarket Works

ABSTRACT

The California Public Utilities Commission has ordered that a new framework be established to guide the planning and conduct of California's energy efficiency program evaluations. A primary goal of the new framework is to establish an evaluation approach that provides reliable information to help ensure meeting California's energy needs, while also supporting continued program improvements and helping to meet the information needs of policymakers and program managers.

The new Framework provides a systems approach to planning, conducting and funding evaluations of energy programs, instead of the more traditional program-specific or sector-specific approach. The Framework provides a structured decision process in which portfolio-level considerations of quality, uncertainty, and reliability directly influence decisions about what type(s) of evaluation should be conducted, when the evaluations should be conducted, and the approaches appropriate for conducting the research. It considers evaluation designs for the following types of evaluations: impact (both gross and net), metering and monitoring efforts, process, market effects studies, non-energy effects research, and information and education program evaluations. Each chapter summarizes alternative approaches and their strengths and weaknesses. The Framework also includes chapters on appropriate sample design and analysis, on identifying and dealing with sources of uncertainty, and on a market-based perspective for calculating avoided costs and conducting cost-effectiveness tests.

This paper summarizes the contents of the Framework.

CPUC Direction

The California Public Utilities Commission (CPUC) ordered in D. 01-11-066 that a new Framework be established to guide the planning and conduct of energy efficiency programs funded by California's Public Goods Charge on energy bills. A primary goal of the new Framework is to establish an evaluation approach that provides reliable information to help ensure California's energy needs are met, while also supporting continued program improvements and helping to meet the information needs of policymakers and program managers.

The project was built on an advisory group and CPUC vision established early in the project: that program administrators need an evaluation road map that serves as a decision-guidance system for determining what research should be conducted for specific types of energy programs. This is especially important given the current system of program delivery that relies on scores of utility and non-utility program implementers, with varying degrees of evaluation science understanding. The Framework is designed as a reference manual for evaluators,

providing guidance on planning, budgeting, and implementing program-specific evaluations. Additionally, the Framework can be used as an evaluation policy support document, providing policy makers with an information source from which evaluation policies and budgets can be established.

Several CPUC evaluation goals are incorporated into the Framework:

1. Provide reliable evaluation results to support energy policy and supply decisions;
2. Allow programs to be equally compared according to their energy impacts;
3. Help understand and verify program energy and peak savings;
4. Help identify and quantify market and non-energy effects;
5. Provide information needed to estimate program cost-effectiveness; and
6. Provide recommendations for program changes that help improve cost-effectiveness.

In addition to accomplishing the above high-level goals, the Framework was designed to support the following more specific objectives.

1. Increase the level of reliability of program savings impact estimates for use in resource planning forums where the uncertainty of these estimates needs to be compared against the uncertainty of other key components of the resource plan.
2. Increase the quality of feedback to program administrators from evaluation projects to both improve program designs and increase the net savings from their programs.
3. Provide guidance to program administrators on what types of evaluation are recommended and are likely to be most beneficial for documenting operations and objective accomplishments.
4. Provide guidance to program administrators on the methodological approaches and study focus needed to perform specific types of evaluations.
5. Provide flexibility that allows for the use of alternative evaluation approaches, when they can be shown to provide as reliable results as the methods presented in the Framework.

Systems Approach

Strategic timing of evaluation studies places studies in the context of the overall cycle of program planning, implementation and evaluation. In the Framework approach, evaluation results are used to make informed decisions on program improvements and future program designs and offerings. This cycle provides for a continuing process of program improvement, so that the programs match available market opportunities and continually improve their cost effectiveness over time.

The timing of these activities is influenced by several, often-competing considerations. The following considerations are described:

- Timely input to policy considerations and decisions;
- Early feedback to program implementers;
- Program lifecycle stage;
- Evaluation data time lags;
- Program and energy procurement portfolio planning requirements;

- Evaluation planning requirements and regulatory oversight;
- Program design/solicitation, selection, review, and implementation preparation;
- The value of dispersed timing for evaluations;
- Contract requirements for "pay for performance" programs;
- Market inertia;
- Timing needs for retention, measure life, and technical degradation analyses; and
- Regulatory oversight and review.

The Framework provides an example of an overall strategic planning process by dividing the program cycle into five major program activities, which are further described below:

I. Goal Setting - Updating and Potential Analysis;
II. Portfolio Analysis - Sector and Program Priorities;
III. Portfolio/Program Design, Selection, Review and Approval;
IV. Program Launch Preparation, Overall Evaluation Planning and Regulatory Review; and
V. Program Implementation, Evaluation, M&V, and Market Assessment & Ongoing Regulatory Oversight.

Activity I: Goal Setting - Updating and Potential Analysis

Activity I consists of establishing the high-level goals for the efficiency portfolio and estimating the achievable potential for the efficiency resource. Overall energy efficiency policy goals (such as targeted percentage reduction in per capita energy use) are restated in terms of specific energy and demand reduction targets. Overall policy goals are reconciled with load forecasts and other information on utility energy and capacity requirements.

Once the specific energy and demand goals are defined, the potential analysis helps refine the goals based on the overall availability of the conservation resource within technologies, market sectors, and geographic area. Technology and market performance information from program evaluations, market saturation studies and other overarching studies are incorporated into the efficiency resource potential estimates. Information is also assessed regarding the energy and demand goals, potential analysis, and other policy objectives, such as geographic and sector equity issues.

Activity II: Portfolio Analysis – Sector and Program Priorities

A preferred portfolio of energy efficiency programs for the upcoming program cycle is defined based on the goals and potentials established in Activity I and further review of evaluation data on program impacts, process and market effects, and past overarching studies. Lessons learned from current and past program efforts along with efficiency potentials by technology and market sector from Activity I are used to define a portfolio of programs that meet the overall efficiency resource goals. The results of the portfolio analysis are presented for public comment and regulatory review. Preliminary evaluation planning and budgeting may also take place during this activity.

Activity III: Portfolio/Program Design, Selection, Review and Approval

Designs for specific programs meeting the portfolio goals defined above are developed and/or solicited and selected. (This requires specific targeted RFPs to fill gaps rather than broadcast general calls for open ended proposals.) The proposed portfolio of programs is finalized and presented for public comment and/or approval by the responsible entity.

Activity IV: Program Launch Preparation, Overall Evaluation Planning and Regulatory Review

Once the program portfolio is designed, selected and approved, the programs are prepared for implementation. Contracts (if needed) are negotiated, trade allies and key stakeholders are notified, and materials and internal processes are developed to prepare for program introduction. Concurrently, an overall evaluation planning process takes place. This process establishes overall goals and objectives for the program evaluations, sets evaluation priorities based on perceived risks to achieving the portfolio savings objectives and other objectives, and addresses program design issues related to evaluation. Issues such as evaluation resource allocation, evaluation study consolidation, the need to evaluate a particular program during a given program cycle, and evaluation scheduling are addressed during the overall planning process. The overall evaluation plans will need to be reviewed by an appropriate body or bodies to assure that they meet the information needs of policy makers, portfolio managers, program administrators, and program implementers.

Activity V: Program Implementation, Evaluation, M&V, Market Assessment and Ongoing Regulatory Oversight

After the overall evaluation plans have been approved, more detailed evaluation plans must be developed and evaluation professionals selected to conduct the studies. The overall evaluation plan developed in Activity IV will likely schedule new or redesigned programs to initiate early process evaluation coordination and support activities as well as measurement and verification (M&V) planning during the initial program startup period. These early actions can help assure that program designs support the evaluation function, in addition to helping establish the platform from which early evaluation feedback can be initiated. These early actions can also lead to improved database designs that help support the evaluation efforts. Following these early efforts, the evaluations will then likely swing into standard evaluation activities after the program is running, but early enough in the program cycle to provide feedback and corrective recommendations to program implementers in time for the program to benefit from those recommendations. Early impact evaluation activities to support program progress tracking consist primarily of measure installation tracking and verification combined with ex-ante savings estimates by measure. Adjustments to ex-ante savings estimates may be made based on early issues identified during M&V activities.

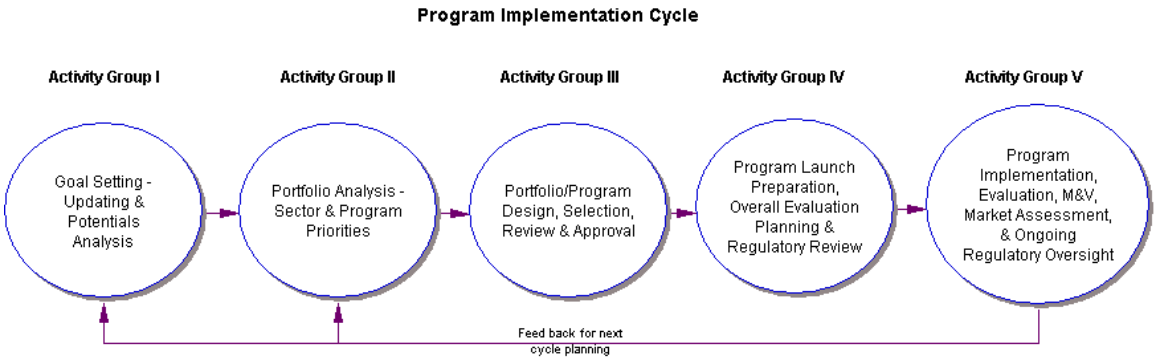
The full net impact evaluation analysis proceeds according to the schedule laid out in the program evaluation plan. Ex-post savings by measure and/or program are estimated, and the final program impacts are estimated based on the final program accomplishments. Assumptions underlying the efficiency potential analysis can then be updated based on the full net impact

analysis. These data then feed back into the goal setting and potentials analysis activities, and the cycle repeats to allow for an integrated planning process.

The Framework makes no recommendations on the organizations responsible for these efforts, but presents this example of how a system might function to allow for a coordinated process in which the steps that need to be integrated are established along a single timeline.

The steps and feedback paths for this type of an integrated planning process are shown in Figure 1.

Figure 1. Program Implementation Cycle



On top of this cycle of repeating activities, longer timescale analyses of market effects and measure persistence can be conducted to provide important information to the integrated planning process. The Framework is designed to function within an evaluation planning and implementation process that is tailored to the program contracting cycle.

The Roadmap Approach

The Framework provides a structured approach to planning and conducting evaluations of California's energy programs in multiple ways. One is developing the systems approach described above. Another is its chapters on each type of evaluation, which provide alternative acceptable methods for conducting the evaluation and their pros and cons. A third is its use of "roadmaps," for which Figure 2 below, in the Impact Evaluation section, provides an example.

Within each of the chapters dealing with a specific type of evaluation is an evaluation decision "roadmap." These roadmaps are designed to assist the program administrator with their evaluation planning process and the related program-specific evaluation decisions. The roadmaps consist of a set of decision trees or decision flow diagrams that walk through the process of determining if an evaluation is needed and what type of evaluations, methods, or steps are expected.

Framework Presentation

Nine key components of the Framework are incorporated into the Framework's design for performing program evaluations. These nine components, which consist of the types of evaluations covered by the Framework and key concepts and considerations that go into planning and conducting evaluations, are summarized below. Each is given a chapter in the Framework.

Impact Evaluation

Impact evaluations focus on estimating the net effects from the implementation of one or more energy efficiency programs. The Impact Evaluation chapter presents the fundamentals of statistical billing analysis, engineering analysis, and net-to-gross (free ridership and spillover) analysis. Within these general analysis types, it discusses several alternative evaluation approaches to document the amount of energy saved by a program, including some discussion of strengths and weaknesses and bases for choice among the methods. Roadmap figures are provided that summarize the decision and activity steps in conducting the evaluations. The roadmap for making decisions about billing analysis is presented in Figure 2 below to provide an example of the roadmaps.

The billing analysis section covers simple aggregate pre-post comparisons, experimental and quasi-experimental designs, normalized annual consumption (NAC) models, conditional demand analysis, analysis of covariance, and statistically adjusted engineering (SAE) models. Critical billing analysis issues such as model misspecification, non-random error terms, non-random measurement error and error-in-variable bias, heteroscedasticity, and autocorrelation are discussed along with quality control measures to avoid these problems.

The engineering analysis section discusses engineering and building energy simulation models. The basic forms of common engineering equations for typical energy efficiency measures are presented, along with a discussion of the terms in the equations. Links to Measurement and Verification as a data source for defining these terms are discussed. Building energy simulation models are discussed, including descriptions and applications for DOE-2, EnergyPro, Micropas, and EnergyPlus. The issue of simulation model validation is presented.

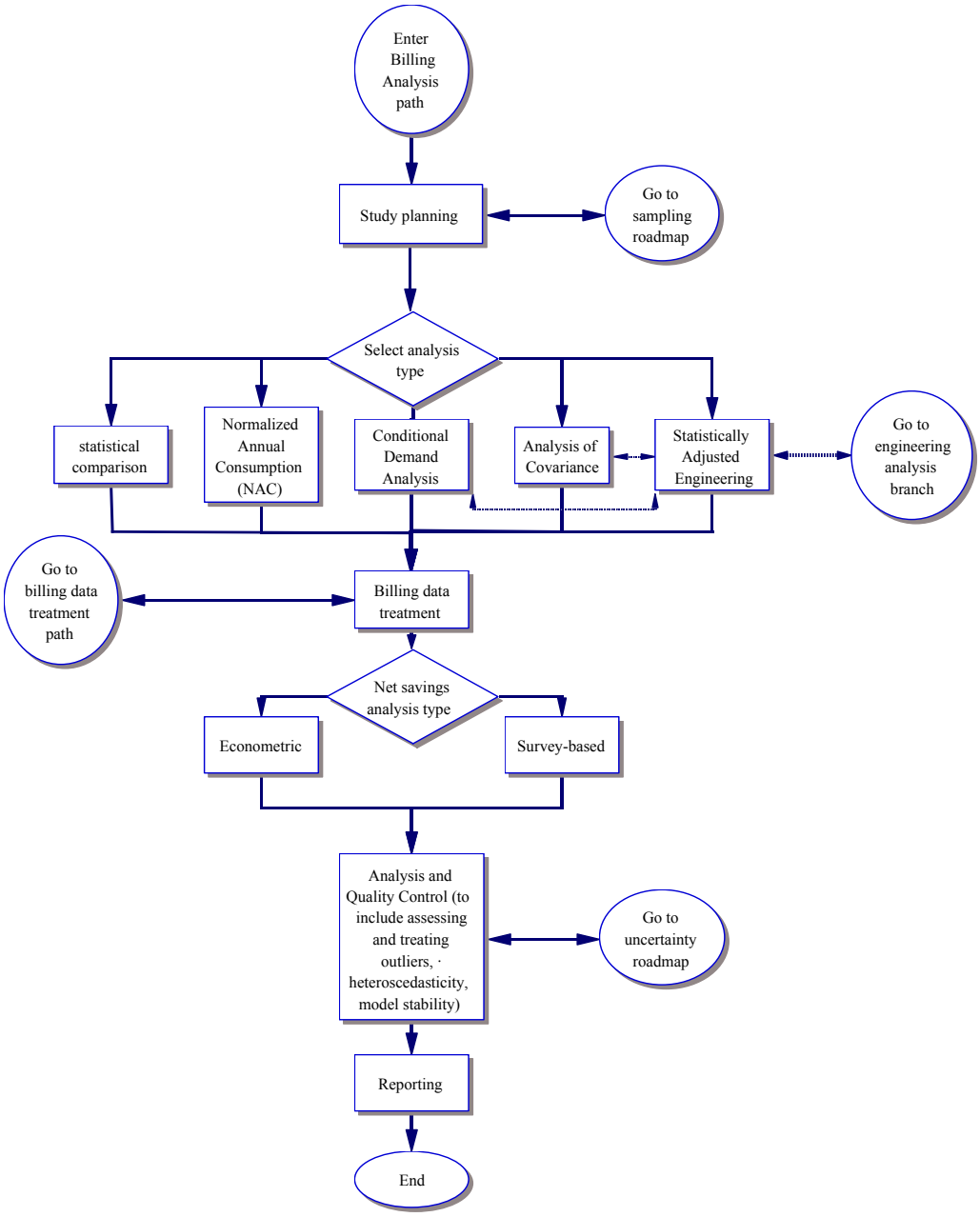
The third section describes the principles underlying net-to-gross analysis and the analytical methods that can be used to estimate net-to-gross ratios. Two main analysis paths are presented: survey-based methods and econometric methods. A variety of econometric approaches are described and issues with each are explored. For survey-based methods, the problems and sources of uncertainty are discussed. When survey-based methods are to be used, the Framework recommends that a carefully structured set of questions be used. It also recommends that the California Public Utilities Commission consider either using the standard set of questions developed through a study ordered by Massachusetts regulators or ordering a study that would examine the Massachusetts study and modify its recommendations as warranted by additional experience. The fundamental conclusion is that all methods have real problems, but that free ridership and spillover are much too important to ignore; consequently, net-to-gross analysis must be attempted.

Measurement and Verification

The impact evaluation will often employ metering, monitoring and verification studies to help accurately estimate the ex-post program savings. These efforts are typically referred to as *Measurement and Verification* (M&V). The M&V approaches typically used in impact evaluation are some form of field measurements taken to help identify how much energy is used before the program actions are taken, how much energy is being used after the actions are taken, the use conditions associated with an installed technology, or a change in behaviors that is to produce the energy savings. In many cases, M&V activities are used to define the terms in engineering equations, and reduce the uncertainty in engineering estimates. The M&V section

references the IPMVP M&V protocols, ASHRAE Guideline 14 and several IPMVP companion documents prepared for the Federal Energy Management program (FEMP). Elements of an M&V plan are presented, along with typical instrumentation and data acquisition approaches for common energy efficiency measures. The issue of M&V timing and providing early feedback to program implementers is presented. The concept of a data warehouse to review and combine the results of M&V studies across multiple programs to serve as a resource for future M&V projects is also presented.

Figure 2: Billing Analysis Roadmap



Process Evaluations

The process evaluation is a systematic assessment of an energy efficiency program for the purposes of documenting program operations and identifying improvements that can be made to increase the program's efficiency or effectiveness for acquiring energy resources. The Process Evaluation chapter discusses the intent and focus of the process evaluation and the skill levels needed to conduct these evaluations. It also identifies the variety of investigative issues associated with the process evaluation and the tools typically used in these studies. Additionally, the chapter discusses the need for process evaluations to be conducted in time for programs to benefit from the evaluation findings, and the need to establish early information feedback systems with the program administrators to allow for evaluation results to feed the program redesign process to obtain the maximum level of energy resources within the program delivery period.

Information and Education Program Evaluation

Information and education program evaluations focus on assessing the degree to which program goals are accomplished and estimating the effects of the program activities on their target markets. They can also serve as an information source for assessing the cost-effectiveness of the program. Evaluations of information and education programs in California typically have a different research goal than the evaluations of programs that have energy impact goals. Most of California's information and education programs do not have energy impact goals and are not expected to be cost-effective from an energy acquisition perspective. Instead these programs are designed to influence the ability of other programs to achieve their energy impact goals or they are focused on trying to influence the short-term or long-term decision processes associated with acquisition and use of energy-consuming technologies or behaviors. The evaluation of information and education programs within the Framework focuses on documenting the effects of the program at reaching its information transfer or educational goals.

Market Transformation Program Evaluation

The Market Transformation (MT) Program Evaluation chapter of the Framework focuses on the evaluation of program-induced market effects when the program being evaluated has a goal of making longer-term or lasting changes in the way a market operates. These evaluations examine changes within a market that are caused, at least in part, by the energy efficiency programs attempting to change that market. These evaluations are challenging, as markets are constantly in a state of change as new and competing technologies are offered or as other non-program market transformation efforts compete with the program's efforts. Two other forms of market analyses, market baseline studies and market operations studies, are also discussed in this chapter as their development and use may be required as part of completing a quality MT program evaluation effort.

Non-Energy Effects Evaluation

Non-energy effects evaluations look at the intended or unintended effects that occur in addition to the energy impacts associated with a program. This chapter identifies the conditions

for conducting a non-energy effects evaluation and provides a discussion on the variety of types of non-energy effects that have been studied and the methodologies that have been used.

Uncertainty

As a result of CPUC policies associated with this Framework, it is the explicit intent of this Framework that evaluations conducted on California energy efficiency programs be conducted in a way that provides reliable technology-specific or program-specific ex-post net energy impact findings. The Uncertainty Chapter describes how evaluations should assess and report the level of uncertainty and potential sources of bias associated with the evaluation findings and explain what actions are taken to limit uncertainty and mitigate bias. Evaluation users should to be able to determine the reliability of the study results and to determine if the results can be used for supply decisions, for public program policy-making, or for updating deemed energy savings estimates. The information is also designed for use in a summative fashion, allowing the propagation of uncertainty for a group of programs, for the PGC-funded portfolio, or the overall portfolio of energy efficiency programs.[1]

Sampling

The Sampling Chapter focuses on the use of sampling in impact estimation, rather than its uses in other types of studies. It discusses the relationships among population characteristics, sample sizes, and sample selection methods and the ability to assess ex-post effects that reliably represent the impacts that a program has achieved. Two main methods for calculating sample sizes and for allocating samples for studies are outlined: simple random sampling and stratified ratio estimation. The chapter also provides information for the calculation of relative precision, factors needed to create efficient sample designs, and how to use these within an impact evaluation's analysis phase.

Evaluation and Cost-Effectiveness

This chapter is targeted to three audiences in the area where issues overlap among program evaluations, cost-effectiveness analysis, and their uses and interpretations. First, it helps evaluators see how the results from their evaluations will be used in cost-effectiveness analysis. Second, it demonstrates to program staff and administrators who calculate or use cost-effectiveness analysis how evaluation and cost-effectiveness work together. Third, it helps policy makers understand some of the key issues involved in using evaluation results to estimate cost-effectiveness, since these tests are often used to inform a policy decision about whether to continue to invest in a program. The chapter does not establish methods for calculating avoided costs or for conducting cost-effectiveness tests. The California *Standard Practice Manual* and the *Avoided Cost and Cost Effectiveness Study* address these topics.

[1] See the Wright et al. paper in the ACEEE 2004 Summer Study Proceedings for a description of the Framework's Uncertainty and Sampling Chapters.

Key Features of the Framework

The Framework contains several key features and sections not generally found in previous documents of this type.

- **Portfolio focus.** The emphasis of evaluation on reviewing and improving a portfolio of energy efficiency programs that compete along with supply options is presented.
- **Coverage of a full range of types of evaluation.** The previous California M&E Protocols emphasized impact evaluation above other evaluation disciplines. This document presents an integrated suite of studies, designed to create both formative (program guidance) and summative (program performance) information.
- **Written for a variety of audiences.** The introduction of non-utility implementers has changed the audience for information on evaluation study methodology and results. The Framework considers a broad range of interests and backgrounds.
- **Evaluation priorities and resource allocation.** The Framework contains guidance to portfolio managers, evaluation planners, program implementers, and regulatory oversight bodies on establishing evaluation priorities and budgets.
- **Program theory/program logic models.** The Framework presents the concept of program theory and logic models as a key component to designing evaluation studies and interpreting evaluation results.
- **Ethics.** A section on evaluation ethics, referencing the guiding principles of the American Evaluation Association, is presented**.**

Issues Not Covered by the Framework

The Framework is not an evaluation protocol, but it can serve as a reference for future evaluation protocol development. Several options for addressing key evaluation issues are presented, but a precise, prescriptive path is not presented. The Framework does not address the issue of efficiency program administration or the regulatory structure within which evaluations will be designed, implemented and used. Thus, it differs from the previous California M&E protocols in two ways. First, it offers guidelines for good evaluation studies of multiple types (impact, process, market effects. etc.), rather than prescribed methods for undertaking impact evaluations and savings persistence studies only. Second, the prior California M&E protocols defined the regulatory process within which studies would be completed, reviewed, and used to determine performance incentives for the utility program administrators.

The Framework does not attempt to be a single reference source for evaluation methodology issues. Instead, it provides an overview of the issues for each type of evaluation and provides numerous references for additional information. References were selected to be ones that would be available to Framework users at little or no cost. One consequence of this approach is that several excellent evaluation guidebooks produced during the 1990s by EPRI, the Electric Power Research Institute, Palo Alto, CA, are <u>not</u> referenced in the Framework because they are not available at low cost to a large part of the intended audience for the Framework. Those with access to EPRI reports should look for these guidebooks as valuable additional sources of information.

Conclusions

The California Evaluation Framework provides evaluation guidance for a wide range of stakeholders on why and how evaluations can be conducted in California. The Framework demonstrates how the evaluation process can be placed into a strategic system under which portfolios are structured and programs are selected, implemented, and evaluated, leading to more cost-effective programs and improved portfolios. The Framework sets the stage for conducting evaluations that focus on improving the reliability of evaluation results, enabling policy makers to more effectively use evaluation studies to support the energy supply planning process. The Framework also describes the need for evaluations to identify their strengths and weaknesses within the evaluation report, so that readers can better understand the reliability of the evaluation findings. The Framework also identifies the importance of a wide range of evaluation efforts and describes the conditions under which different evaluations should be conducted.

The Framework is posted on the website of the California Measurement Advisory Council, CALMAC, at www.calmac.org. Currently it is posted in the CALMAC/Filings section as a document and appendices dated 3/1/2004. At some later time, it will probably be posted in the Searchable Database of studies.

References

CADMAC (1999). *Protocols and Procedures for the Verification of Costs, Benefits, and Shareholder Earnings from Demand-Side Management Programs.* California Public Utilities Commission. 93-05-063.

California State Governor's Office (2001). *Standard Practice Manual (SPM): Economic Analysis of Demand-Side Management Programs.* October.

Sebold, Frederick, Alan Fields, Lisa Skumatz, Shel Feldman, Miriam Goldberg, Kenneth Keating, and Jane Peters. (2001). *A Framework for Planning and Assessing Publicly Funded Energy Efficiency*. Pacific Gas & Electric Company. March 1. Study PG&E-SW040.

TecMarket Works Framework Team (2004). *The California Evaluation Framework (Final Draft)* Southern California Edison Company. March. http://www.calmac.org/calmac-filings.asp, http://www.cpuc.ca.gov/static/industry/electric/energy+efficiency/rulemaking

Wright, Roger, Tim Hennessy, Marian Brown, Nick Hall, Lori Megdal, and Ken Keating. (2004). "A Proposed New Framework for Evaluation in California: The Sampling Roadmap," *American Council for an Energy-Efficient Economy Summer Study*. Asilomar, CA. pp. forthcoming.