BEFORE THE PUBLIC UTILITIES COMMISSION
OF THE STATE OF CALIFORNIA

Order Instituting Rulemaking Regarding Policies and Protocols for Demand Response Load Impact Estimates, Cost-Effectiveness Methodologies, Megawatt Goals and Alignment with California Independent System Operator Market Design Protocols.	**Rulemaking 07-01-041**

STRAW PROPOSALS FOR LOAD IMPACT ESTIMATION AND COST EFFECTIVENESS EVALUATION OF SOUTHERN CALIFORNIA EDISON COMPANY, SAN DIEGO GAS & ELECTRIC COMPANY, AND PACIFIC GAS AND ELECTRIC COMPANY PURSUANT TO THE ASSIGNED COMMISSIONER AND ADMINISTRATIVE LAW JUDGE'S SCOPING MEMO AND RULING DATED APRIL 18, 2007

PETER OUBORG
SHIRLEY A. WOO
Pacific Gas and Electric Company

77 Beale Street, B30A
San Francisco, CA 94105
Telephone: (415) 973-2248
Facsimile: (415) 973-0516
E-Mail: SAW0@pge.com

Attorneys for
PACIFIC GAS AND ELECTRIC COMPANY

STEVEN D. PATRICK
San Diego Gas & Electric Company
101 Ash Street
San Diego, CA 92101-3017
Telephone: (213) 244-2954
Facsimile: (213) 629-9620
Email: spatrick@sempra.com

Attorney for
SAN DIEGO GAS & ELECTRIC COMPANY

MICHAEL D. MONTOYA
JANET S. COMBS
Southern California Edison Company

2244 Walnut Grove Avenue
P. O. Box 800
Rosemead, CA 91770
Telephone: (626) 302-1524
Facsimile: (626) 302-7740
Email: janet.combs@sce.com

Attorneys for
SOUTHERN CALIFORNIA EDISON COMPANY

Dated: July 16, 2007

BEFORE THE PUBLIC UTILITIES COMMISSION
OF THE STATE OF CALIFORNIA

Order Instituting Rulemaking Regarding Policies and Protocols for Demand Response Load Impact Estimates, Cost-Effectiveness Methodologies, Megawatt Goals and Alignment with California Independent System Operator Market Design Protocols.	Rulemaking 07-01-041

STRAW PROPOSALS FOR LOAD IMPACT ESTIMATION AND COST EFFECTIVENESS EVALUATION OF SOUTHERN CALIFORNIA EDISON COMPANY, SAN DIEGO GAS & ELECTRIC COMPANY, AND PACIFIC GAS AND ELECTRIC COMPANY PURSUANT TO THE ASSIGNED COMMISSIONER AND ADMINISTRATIVE LAW JUDGE'S SCOPING MEMO AND RULING DATED APRIL 18, 2007

The Assigned Commissioner and Administrative Law Judge's Scoping Memo and Ruling (ACR) established the first two topics for this rulemaking: 1) establish a comprehensive set of protocols for estimating the load impacts of demand response programs and 2) establish methodologies to determine the cost-effectiveness of demand response programs. The ACR then directed Southern California Edison Company (SCE), San Diego Gas & Electric Company (SDG&E) and Pacific Gas and Electric Company (PG&E) (together, Joint Utilities) to submit straw proposals on load impact estimation and cost effectiveness methodology for consideration in workshops set for July 19-20 and August 1-3. The Joint Utilities submit the straw proposals for load impact estimation and cost effectiveness methodology attached to this pleading in compliance with the ACR and the Administrative Law Judge's May 25, 2007 ruling (Ruling).[1/]

1/ The ACR set July 13, 2007 as the date for filing the load impact and cost effectiveness straw proposals. On July 11, 2007, ALJ Hecht granted the Joint Utilities' request for an extension to July 16, 2007 and made the extension applicable to other parties who will file straw proposals.

The Ruling states that the major focus of the straw proposals is to develop a useable overall framework and methodology, and the process should not focus disproportionately on defining any given input value. (Ruling, page 3) To define the key issues for the Joint Utilities' straw proposals and to suggest, but not limit, possible approaches for addressing those issues, the Ruling also distributed a Staff Guidance document. The Staff Guidance includes issue areas that are contentious and complex, as the Ruling acknowledges in its discussion of the cost effectiveness straw proposal. The Staff Guidance also requires the straw proposals to include new, unexplored, difficult areas, especially for load impact estimates from demand response programs.[2/]

The Joint Utilities' straw proposals cover all the issue areas identified in the Staff Guidance, but the depth and degree of sophistication in the straw proposals have been constrained by the broad span of the issues, their complexity, their novelty in some instances, and the limited time available to develop the proposals. Therefore the Joint Utilities appreciate the Ruling's acknowledgement that additional work will be needed to develop robust protocols for estimating load impact and to resolve cost effectiveness methodological issues and inputs. (Ruling, page 3.)

///

///

///

2/ Two examples of issue areas (without limitation) that involve uncharted, difficult territory are ex ante estimation for load impacts and dynamic pricing for cost effectiveness. These required topics in the Staff Guidance significantly increase the scope and complexity of the straw proposals, and may require additional time and effort beyond this part of R.07-01-041 to be developed.

The attached Joint Utilities' straw proposals are as comprehensive as possible given the tasks required and the time available. The Joint Utilities respectfully request that the parties and the Commission recognize that further developments should be expected and allowed for load impact estimation protocols and cost effectiveness methodologies in the future.

Respectfully submitted,

PETER OUBORG
SHIRLEY A. WOO

/s/

Pacific Gas and Electric Company
77 Beale Street
San Francisco, CA 94105
Telephone: (415) 973-2248
Facsimile: (415) 973-0516
E-Mail: SAW0@pge.com

Attorneys for
PACIFIC GAS AND ELECTRIC COMPANY

On behalf of SOUTHERN CALIFORNIA EDISON COMPANY AND SAN DIEGO GAS & ELECTRIC COMPANY

July 16, 2007

Joint IOU Straw Proposal on Load Impact Estimation for Demand Response

Submitted by:

Pacific Gas & Electric Co.
Southern California Edison Co.
San Diego Gas & Electric Co.

Prepared by:

Stephen S. George, Ph. D.
Michael Sullivan, Ph. D.
Josh Bode

Freeman, Sullivan & Co.

July 16, 2007

Joint IOU Straw Proposal on Load Impact Estimation for Demand Response

1 EXECUTIVE SUMMARY

California's Energy Action Plan (EAP II) emphasizes the need for demand response (DR) programs that result in cost-effective savings and the creation of standardized measurement and evaluation mechanisms to ensure verifiable savings. California Public Utilities Commission (CPUC) Decision D.05-11-009 identified a need to develop measurement and evaluation protocols and cost-effectiveness tests for demand response. On January 25, 2007, the Commission opened a rulemaking proceeding (OIR 07-01-041), with several objectives, including:[1]

- Establishing a comprehensive set of protocols for estimating the load impacts of DR programs;
- Establishing methodologies to determine the cost-effectiveness of DR programs;

In conjunction with this rulemaking, a scoping memo[2] was issued directing the three major investor owned utilities (IOUs) in California to jointly develop and submit a "straw proposal" for load impact protocols. This report is the Joint IOU Straw Proposal for Load Impact Estimation for DR Programs. It constitutes the first step in establishing a comprehensive set of protocols for estimating the load impacts associated with DR programs.

On May 24, 2007, the Energy Division of the CPUC and the Demand Analysis Office of the California Energy Commission (CEC) issued a document entitled *Staff Guidance for Straw Proposals On: Load Impact Estimation from DR and Cost-Effectiveness Methods for DR.* The Staff Guidance document indicated that straw proposals should focus on estimating DR impacts for long-term resource planning.[3]

Estimating DR impacts for long-term resource planning is inherently an exercise in ex ante estimation. However, ex ante estimation should, whenever possible, be based on ex post evaluations of existing DR programs. As such, meeting the Commission's requirement to focus on estimating DR impacts for long-term resource planning requires careful attention to ex post evaluation of existing programs. Consequently, the protocols and guidance presented here address both ex post evaluation and ex ante estimation of DR impacts.

While DR programs differ significantly across many factors, one important characteristic, both in terms of the value of DR as a resource and the methods that can be used to

[1] R07-01-041, p.1.
[2] Assigned Commissioner and Administrative Law Judge's Scoping Memo and Ruling, April 18, 2007
[3] CPUC/CEC. Staff Guidance for Straw Proposals On: Load Impact Estimation from DR and Cost-Effectiveness Methods for DR. May 24, 2007. p.10.

estimate impacts, is whether the resource is tied to a specific event, such as a system emergency or some other trigger. Event based programs include critical peak pricing, direct load control and autoDR. Non-event based programs include traditional time-of-use rates real time pricing and permanent load shifting through technology such as ice storage.

Load impact estimation protocols dictate what must be done. They could focus on the output of a study, defining what must be delivered, on how to do the analysis, or both. The protocols provided in this report focus on what impacts should be estimated, what issues should be considered when selecting an approach and what to report, not on how to do the job.

The best approach to estimating impacts is a function of many factors— program type, target market, program size, available budget, the length of time a program has been in effect, available data, and the purposes for which the estimates will be used. Dictating the specific methods that must be used for each impact evaluation or ex ante forecast would require an unrealistic level of foresight, not to mention dozens if not hundreds of specific requirements. More importantly, it would stifle the flexibility and creativity that is so important to improving the state of the art.

On the other hand, there is much that can be learned from previous work and there are significant advantages associated with certain approaches to impact estimation compared with others. Furthermore, it is imperative that an evaluator have a good understanding of key issues that must be addressed when conducting the analysis, which vary by program type, user needs, and other factors. As such, in addition to the protocols, this document also provides guidance and recommendations regarding the issues that are relevant in specific situations and effective approaches to addressing them.

The protocols discussed here describe the minimum requirements that must be met for impact estimation and consist of the following elements:

- The impact measures and associated statistics that must be produced
- The day-types and other time periods for which the measures must be reported
- Additional data that must be reported in order to understand and interpret the estimated impacts, such as weather, event characteristics and the like
- Statistical measures that must be reported in order to provide insight regarding bias and precision associated with the evaluation and sampling methodologies employed to develop the impact measures.

Separate protocols are provided for ex post evaluation of event based programs, ex post evaluation of non-event based programs and ex ante estimation for all programs, although the differences across the three categories are relatively minor. In general, the protocols require that:

- Impact estimates be provided for each of the 24 hours on various event day types for event based programs and other day types for non-event based programs

- Estimates of the change in overall energy use in a season and/or year be produced

- Uncertainty adjusted impacts be reported for confidence levels of 10, 50 and 90 percent, reflecting the uncertainty associated not only with model parameters but also uncertainty in key drivers of demand response, such as weather.

- Output be provided in a common format, as depicted in Table 1-1 for ex post evaluation. A slightly different reporting format is required for ex ante estimation.

- Impact estimates be provided for each day type indicated in Table 1-2

- Various statistical measures be provided so that reviewers can assess the accuracy, precision and other relevant characteristics of the impact estimates

- Ex ante estimates be based on relevant ex post evaluations whenever possible, even if it means relying on studies from other utilities or jurisdictions.

The protocols in this straw proposal are focused on reporting requirements for resource planning in the future and may not be appropriate or feasible for other reporting requirements of demand response load impacts.

Table 1-1
Required Reporting Format for Ex Post Impact Estimates

Hour Ending	Reference Load (kWh/hr)	Observed Load (kWh/hr)	Load Impact (kWh/hr)	Temperature (degrees F)	Uncertainty Adjusted Energy Impacts		
					10th Percentile	50th Percentile	90th Percentile
1							
2							
3							
4							
5							
6							
7							
8							
9							
10							
11							
12							
13							
14							
15							
16							
17							
18							
19							
20							
21							
22							
23							
24							
					Uncertainty Adjusted Energy Impacts		
	Reference Energy Use	Observed Energy Use	Change in Energy Use	Degree Hours (Base 75)	10th Percentile	50th Percentile	90th Percentile
Day							

Table 1-2
Day Types to Be Reported for Each DR Program Type

	Event Based Programs			Non-Event Based Programs		
Day Types	Event Driven Pricing	Direct Load Control	Callable DR	Non-event Driven Pricing	Scheduled DR	Permanent Load Reductions
Ex Post Day Types						
Each Event Day	X	X	X			
Average Event Day	X	X	X			
Average Weekday Each Month				X	X	X
Monthly System Peak Day				X	X	X
Ex Ante Day Types						
Typical Event Day	X	X	X			
Average Weekday Each Month (1-in-2 Weather Year)				X	X	X
Monthly Sysem Peak Day (1-in-2 Weather Year)				X	X	X

2 BACKGROUND AND OVERVIEW

2.1 Background

Demand response programs are an essential element of California's resource strategy, as articulated in the State's Energy Action Plan II (EAP II). EAP II has determined how energy resources should be deployed to meet California's energy needs and ranks DR programs second in the "loading order" after energy efficiency programs. The EAP II emphasizes the need for DR programs that result in cost-effective savings and the creation of standardized measurement and evaluation mechanisms to ensure verifiable savings.[4]

California Public Utilities Commission (CPUC) Decision D.05-11-009 identified a need to develop measurement and evaluation protocols and cost-effectiveness tests for demand response. That decision ordered CPUC staff to undertake further research and recommend to the Executive Director whether to open a proceeding to address these issues. Commission staff recommended opening a rulemaking, which the Commission did on January 25, 2007. The objectives of OIR 07-01-041 are to:[5]

- Establish a comprehensive set of protocols for estimating the load impacts of DR programs;
- Establish methodologies to determine the cost-effectiveness of DR programs;
- Set DR goals for 2008 and beyond, and develop rules on goal attainment; and
- Consider modifications to DR programs needed to support the California Independent System Operator's (CAISO) efforts to incorporate DR into market design protocols.

As indicated in the ruling, it is expected that the load impact protocols will not only provide input to determining program cost-effectiveness, but will also assist in resource planning and long-term forecasting.[6]

On April 18, 2007, the Assigned Commissioner and Administrative Law Judge's Scoping Memo and Ruling indicated that the three major investor owned utilities in California must jointly develop and submit a "straw proposal" for load impact protocols. This report is the Joint IOU Straw Proposal for Load Impact Estimation for DR Programs. It constitutes the first step in establishing a comprehensive set of protocols for estimating the load impacts associated with DR programs.

---

[4] R07-01-041, p.2.
[5] R07-01-041, p.1.
[6] Ibid. p.2

On May 3, 2007, the Commission held a workshop on load impact estimation protocols. At the workshop, the joint utilities indicated that there were many potential applications of impact estimates for demand response programs, including:

- Ex post impact evaluation
- Monthly reporting of DR results
- Forecasting of DR impacts for resource adequacy
- Forecasting of DR impacts for long-term resource planning
- Forecasting DR impacts for operational dispatch by the CAISO
- Estimation for customer settlement/reference level methods (e.g., payment of incentives) in conjunction with DR program deployment.

The joint utilities also indicated that the relevant issues that must be addressed and the methodologies that can be used vary across applications and attempting to address all of these issues and methods would be extremely difficult in the short time frame allowed for development of the protocols. The joint utilities asked for a more phased approach to protocol development and for guidance and clarification regarding priorities and scope.

On May 24, 2007, the Energy Division of the CPUC and Demand Analysis Office of the CEC issued a document entitled Staff Guidance for Straw Proposals On: Load Impact Estimation from DR and Cost-Effectiveness Methods for DR (hereafter referred to as the Staff Guidance document). The Staff Guidance document indicated that the schedule could not be relaxed and that the focus of the straw proposals should be on estimating DR impacts for long-term resource planning.[7]

Estimating DR impacts for long-term resource planning is inherently an exercise in ex ante estimation. As indicated in subsequent sections, ex ante estimation should, whenever possible, be based on ex post evaluations of existing DR programs, as empirical evidence, properly developed, is almost always superior to theory, speculation, market research surveys, engineering modeling or other ways of estimating what impacts might be for a specific program. As such, meeting the Commission's requirement to focus on estimating DR impacts for long-term resource planning requires careful attention to ex post evaluation of existing programs. Consequently, the protocols and guidance contained in the remainder of this report address both ex post evaluation and ex ante estimation of DR impacts.

---

[7] CPUC/CEC. Staff Guidance for Straw Proposals On: Load Impact Estimation from DR and Cost-Effectiveness Methods for DR. May 24, 2007. p.10.

2.2 Taxonomy of Demand Response Programs

There is a wide variety of DR programs that are currently in place in California (and elsewhere) and many different ways to categorize them. While DR programs differ significantly across many factors, one important characteristic, both in terms of the value of DR as a resource and the methods that can be used to estimate impacts, is whether the resource is tied to a specific event, such as a system emergency or some other trigger. Event based programs include critical peak pricing, direct load control and autoDR. Non-event based programs include traditional time-of-use rates real time pricing and permanent load shifting through technology such as ice storage.

In addition to whether or not a program is event based, there are other characteristics of interest, such as whether or not a program uses incentives or prices to drive demand response and whether impacts are primarily technology driven, purely behaviorally driven or some combination of the two. The Staff Guidance document suggested a taxonomy of DR programs consisting of six broad categories. We found these six categories to be useful, but reclassified them into two broad groups distinguished by whether or not the programs are event based.

Event based programs include:

- **Event-based Pricing—**This program category includes prices that customers can respond to based on an event, i.e., a day-ahead or same-day call. This category could include many pricing variants such as critical peak pricing or a schedule of prices presented in advance that would allow customers to indicate how much load they will reduce in each hour at the offered price (e.g., demand bidding). The common element is that these prices are tied to called events by the utility, DR administrator or other operator.

Direct Load Control—This category includes mass-market programs such as a/c cycling as well as large customer programs such as auto-DR. The common thread is that load is controlled at the customer's site for a called event period through a signal sent by an operator.

- **Callable DR—**This category is similar to direct load control but, in this case, a notification is sent to the customer who then initiates actions to reduce loads, often by an amount agreed to in a contract. The difference is that load reduction is based on actions taken by the customer rather than based on an operator-controlled signal that shuts off equipment. Interruptible and curtailable tariffs are included in this category.

Non-event based programs include:

- **Non-event based pricing—**This program category includes TOU, RTP and related pricing variants that are not based on a called event—that is, they are in place for a season or a year.

Scheduled DR—There are some loads that can be scheduled to be reduced at a regular time period. For example, a group of irrigation customers could be divided into five segments, with each segment agreeing to not irrigate/pump on a different selected weekday.

- **Permanent load reductions and load shifting**—Permanent load reductions are often associated with energy efficiency activities, but there are some technologies such as demand controllers that can result in permanent load reductions or load shifting. Examples of load shifting technologies include ice storage air conditioning, timers and energy management systems.

Tables 2-1 through 2-3 show how the existing portfolio of DR programs for each IOU map into the taxonomy.

Table 2-1
PG&E DR Programs By Category

Program Name	Target Market Segment	Event Based Programs			Non-Event Based Programs		
		Event-based Pricing	Direct Load Control	Callable DR	Non-event Based Pricing	Scheduled DR	Permanent Load Reductions
E-CPP (Voluntary Critical Peak Pricing)	C&I	X					
E-CBP (Capacity Bidding Program)	C&I	X					
E-BEC (Business Energy Coalition)	C&I	X					
E-DBP (Demand Bidding Program)	C&I	X					
TOU (Time-of-Use Pricing)	Residential / C&I				X		
E-RSAC (Residential Smart A/C Program)	Residential		X				
E-CSAC (Commercial Smart A/C Program)	C&I		X				
DR-RFP	C&I						
E-NF (Non-firm Rate Schedule)	C&I			X			
E-BIP (Base Interruptible Program)	C&I			X			
E-SLRP (Scheduled Load Reduction Program)	C&I					X	

Table 2-2
SCE DR Programs By Category

Program Name	Target Market Segment	Event Based Programs			Non-Event Based Programs		
		Event-based Pricing	Direct Load Control	Callable DR	Non-event Based Pricing	Scheduled DR	Permanent Load Reductions
CPP (Critical Peak Pricing)	C&I	X					
DBP (Demand Bidding Program)	C&I	X					
CBP (Capacity Bidding Program)	C&I	X					
TOU (Time-of-Use Pricing)	Residential / C&I				X		
RTP (Real-time Pricing)	C&I				X		
SDP (Summer Discount Plan)	Residential / C&I		X				
AP-I (Agricultural and Pumping Interruptible Program)	C&I		X				
Automated DR	C&I		X				
OBMC (Optional Binding Mandatory Curtailment)	C&I			X			
BIP (Base Interruptible Program)	C&I			X			
I-6 Large Power Interruptible Program	C&I			X			
SLRP (Scheduled Load Reduction Program)	C&I					X	
EnerNOC Contract				X			

Table 2-3
SDG&E DR Programs By Category

Program Name	Target Market Segment	Event Based Programs			Non-Event Based Programs		
		Event-based Pricing	Direct Load Control	Callable DR	Non-event Based Pricing	Scheduled DR	Permanent Load Reductions
CPP (Critical Peak Pricing)	C&I	X					
CPP-E (Critical Peak Pricing - Emergency)	C&I	X					
DBP (Demand Bidding Program)	C&I	X					
Peak Generation Program	C&I			X			
Summer Saver Program	Residential		X				
Smart Thermostat	Residential		X				
CleanGen Generator Program	C&I		X				
CBP (Capacity Bidding Program)	C&I		X	X			
OBMC (Optional Binding Mandatory Curtailment)	C&I			X			
BIP (Base Interruptible Program)	C&I			X			
Peak Day Credit Program	C&I	X					
SLRP (Scheduled Load Reduction Program)	C&I					X	

2.3 Purpose of this Document

Protocols outline what must be done. Protocols could focus on the output of a study, defining what must be delivered, on how to do the analysis, or both. The protocols provided in this report focus on what impacts should be estimated, what issues should be considered when selecting an approach and what to report, not on how to do the job. The goal is to ensure that the impact estimates provided are useful for planners and operators and that the robustness, precision, and bias (or lack thereof) of the methods employed is transparent.

The best approach to estimating impacts is a function of many factors— program type, target market, program size, available budget, the length of time a program has been in effect, available data, and the purposes for which the estimates will be used. Dictating the specific methods that must be used for each impact evaluation or ex ante forecast would require an unrealistic level of foresight, not to mention dozens if not hundreds of specific requirements. More importantly, it would stifle the flexibility and creativity that is so important to improving the state of the art.

On the other hand, there is much that can be learned from previous work and there are significant advantages associated with certain approaches to impact estimation compared with others. Furthermore, it is imperative that the evaluator have a good understanding of key issues that must be addressed when conducting the analysis, which vary by program type, user needs, and other factors. As such, in addition to prescribing the deliverables that must be provided with each evaluation, this report also provides guidance and recommendations regarding the issues that are relevant in specific situations and effective approaches to addressing these issues.

The Protocols in this straw proposal require significant effort that may not always be justified for meeting the needs of a DR report by the IOUs and may not be feasible in the near term. The IOUs should be allowed discretion and flexibility to only use those parts of the protocols that are needed to meet specific information needs (as interim approaches). For example, there may not be adequate time to do all that is specified or all of the inputs needed may not be available at a reasonable cost. These protocols provide the ideal product when time and resources allow. And they provide guidance that should be followed to the fullest extent practicable consistent with the business needs of the IOUs.

2.4 Report Organization

The remainder of this report is organized as follows. Section 3 provides an introduction to evaluation planning and the technical issues that are addressed in greater detail in subsequent sections. Sections 4, 5 and 6 contain the protocols associated with ex post evaluation for event based programs, ex post evaluation for non-event based programs, and ex ante estimation for both event and non-event based programs. These sections also contain detailed discussions of the issues and methods that are relevant to each category of impact estimation. Section 7 contains a discussion of some of the issues and challenges associated with estimation of DR impacts for portfolios of DR resources. Section 8 provides an overview of sampling issues and methods, and section 9 contains reporting protocols.

3 EVALUATION PLANNING

Deciding on the best approach to fulfilling the output and reporting requirements associated with these protocols, and other relevant internal or external needs, requires careful planning. Some things are under an evaluator's control (e.g., methodology), some may be influenced by the evaluator (e.g., budget, schedule) and some are simply given (e.g., program, population and event characteristics in ex post evaluation). The optimal plan must take all of these factors into consideration in order to do the best job of producing unbiased impact estimates given the available budget, schedule, data and other factors over which the evaluator has no control.

This section provides an overview of some of the key issues that influence impact evaluation planning.

3.1 Understanding DR Impacts and Benefits

Impact evaluations of energy efficiency (EE) programs have been conducted for decades and there is a rich literature concerning how to approach energy efficiency impact estimation. However, there are important differences between demand response and energy efficiency impact estimation, which must be understood in order to determine how best to address the challenges associated with demand response impact estimation.

One important difference between EE and DR impact estimation stems from the fact that the operation of DR programs and the impacts and value associated with DR programs can vary significantly by time of day, day of week, and season. As such, impact estimates must be time-specific. For most DR programs, impacts should be measured in each hour over some relevant period of the day or day of the week. For others, tracking impacts at sub-hourly intervals as small as 10 or 15 minutes may be necessary. Determining the impact of DR programs on overall energy use may also be important for cost-effectiveness analysis and policy making. Thus, estimating the change in energy use on a daily, monthly and/or annual basis is an important requirement of impact evaluation.

Another important difference is that demand response impacts are more behaviorally driven than are energy efficiency impacts. Many EE programs provide incentives for or otherwise influence the purchase of more efficient equipment. While impacts from implementing such measures will vary somewhat from customer to customer based on variation in usage and equipment ownership, this variation is small relative to, for example, variation in behavioral response to price signals. The fact that impacts from some DR programs are based entirely on the behavioral response of customers to price signals or other economic incentives, and that this behavioral response can vary significantly across participants in a program and across program events, has implications for program planning, sample sizes, estimation methodology (e.g., engineering based methods are much less relevant for behavioral DR estimation), persistence and other key issues.

With EE impact estimation, free riders are defined as those customers that would have implemented a measure in the absence of the EE program stimulus. A significant challenge with EE impact estimation is determining what customers would do in the absence of the program—that is sorting out the difference between gross impacts and net impacts. This type of free ridership, which is key to EE impact estimation, is not very relevant to impact estimation for most DR programs as few customers would reduce their load during DR events in the absence of the stimulus provided by the DR program. On the other hand, there is another form of free ridership that is relevant to DR impact estimation which stems from the participation of customers who do not use much electricity during DR event periods. This type of free rider is also referred to as a structural benefiter. An example of a structural benefiter is a customer who volunteers for a CPP tariff that does not have air conditioning or typically does not use air conditioning during the critical peak period or someone who operates the unit manually and only turns it on when they come home in the evening. Participation by structural benefiters can be viewed as simply reducing historical cross subsidies inherent in average cost pricing. Nevertheless, for some program types, addressing this form of free ridership in impact estimation can be both important and difficult.

One of the most important factors that must be considered when estimating DR impacts is the inherent uncertainty associated with electricity demand, and, therefore, DR impacts. As indicated above, electricity demand/energy use varies from customer-to-customer and within customer from time-to-time based on conditions that vary systematically with weather, time of day, day of week, season and numerous other factors. As such, electricity demand/energy use is a random variable that is <u>inherently</u> uncertain.

This, combined with the fact that electricity must be produced when it is needed, is one of the primary reasons why demand response is important. The inherent uncertainty associated with electricity demand poses serious challenges for system operators (who must schedule sufficient generating resources to meet projected demand in the short term) and for system planners (who are responsible for ensuring that sufficient generation, transmission and distribution resources are available in the long run to meet projected demand).

Demand response programs are designed to change the timing of electric demand/energy use (lowering it during peak periods) by transmitting changes in prices, load control signals or incentives to customers that reflect the differences in production and distribution cost that occur under different system conditions. These programs can have the effect of reducing the <u>inherent</u> uncertainty in electric demand/energy use thus producing significant benefits for system operators and reducing the cost of capital investment and system operations.

However, like electric demand, demand response is a random variable and thus it also is <u>inherently</u> uncertain. That is, the quantity of demand response that can occur varies from customer-to-customer and within customer from time-to-time based on conditions that vary systematically with weather and other factors. To take advantage of demand response as a means to reduce <u>inherent</u> uncertainty about electric loads, it is necessary to carefully describe the variation that exists for demand response impacts. Unless this is done, it is arguable that demand response programs actually could add to the uncertainty about system loads, thus making system operations and planning even more difficult.

For purposes of evaluating the usefulness of demand response in the context of system operations, it is not enough to know that customers in a given demand response program will reduce their demand during an event occurring under some set of environmental conditions by ***x*** kW on average. Instead, what the operator needs to know is: what is the load reduction that can be expected to occur with a probability of ***y*** (e.g. 90%); at a certain time ***t*** (e.g., 2 pm); within what time frame ***w*** (e.g., instantaneously, within 10 minutes, 30 minutes, two hours or next day); and over what duration ***d*** (e.g., 2 hours, four hours, etc.). Variations in load impacts resulting from ***y, w, t*** and ***d*** are all sources of uncertainty.

In light of the above, it is not sufficient to know the mean or median impact of a DR program—it is also necessary to know how much reduction in energy use can be expected for a DR event under varying conditions at different confidence levels. For ex post evaluation, uncertainty is largely tied to the accuracy and statistical precision of the impact estimates. For ex ante estimation, uncertainty also results from the inherent uncertainty in key variables such as weather and participant characteristics that influence the magnitude of impacts. A key difference between DR impact protocols and the California EE impact protocols is the need for uncertainty adjusted impact estimates.

3.2 Understanding What Is Needed

Once a decision is made to conduct an evaluation or produce an ex ante forecast, the initial requirement for evaluation planning is to understand what is needed by those who will use the results. These protocols outline the minimum requirements for ex ante and ex post estimation, defined in terms of day types, time periods and percentiles. These requirements are designed to address questions about what the impacts were over the historical evaluation period, what they are likely to be under different conditions that may or may not have been observed during the evaluation period (e.g., a 1-in-2 weather year), and what they might be as program and participation characteristics evolve over some future period. They are focused on the needs of a relatively broad audience, including regulators and managers who want to know whether impact goals were met in the past (and if not, why not) and whether program net benefits are positive or negative. The protocols are also designed to help resource planners develop cost-effective plans that optimally balance supply and demand-side options, taking into consideration the risk and uncertainty associated with each resource option.

There are other users and applications of DR impacts that may influence the evaluation plan. For example, program implementers may be concerned with how ex post impact estimates compare with estimates used for customer settlement. This need might require additional analysis that would otherwise not be necessary. It could also influence the methodological approach, although it is important that any such need not lead to methodological choices that negatively impact the primary objective of producing accurate estimates of program impacts for cost-effectiveness analysis and long-term resource planning.

In addition to assessing what is needed by a potentially diverse user community, evaluation planning must also consider the level of precision that is required. These protocols do not dictate precision levels as both the desirable and achievable levels of precision are a function of various factors, including program size (both in terms of demand reductions as well as number of participants), budget, customer mix (e.g., large versus small, homogeneity versus heterogeneity, etc.), and others.

Still another user-driven consideration is the need for estimates associated with day-types and/or hours that differ from those that are outlined below. The output requirements described below are demanding, but still try to strike a balance between the diversity of potential user needs and the work required to meet the needs of all potential users. In the ideal world, resource planners would probably prefer impact estimates for all 8,760 hours in a year under an even wider array of weather and event characteristics than those included in these protocols. For selected, very large programs, or one-off evaluations, this added specificity might be justified and achievable. The evaluator must take these possible needs into consideration when developing an evaluation plan.

Another important consideration is the potential need for geographic specificity. The magnitude of DR impacts will vary by climate zone and participant concentration, and the value of DR varies according to location-specific transmission and distribution bottlenecks and the juxtaposition of load pockets and supply resources. Program planners may want to know the relative magnitude of DR impacts by climate zone and customer characteristics so they can target future marketing efforts. Resource planners may want to know DR impacts for different geographic regions that are dictated by the layout of generation, transmission and distribution resources. Both for planning and operational purposes, the CAISO may want to know how DR impacts vary by as many as 30 regions throughout the state. The need to provide impact estimates for various climate zones or other geographic sub-regions will, at a minimum, affect the sampling strategy and sample sizes required for impact estimation. It could also influence methodology, since additional variables may need to be included in the estimation model in order to determine how impacts differ with variation in climate or population characteristics across geographic regions.

In short, evaluation planning must begin with a thorough assessment and understanding of the needs of the people who will use the impact estimates. Important considerations over and above those that are required by these protocols include any additional applications for which the impact estimates will be used, the desired level of precision, additional ex ante scenarios (e.g., day types) and geographic specificity.

3.3 Ex Post Versus Ex Ante Estimation

Another critical consideration in evaluation planning is whether or not ex ante estimates are needed. There are methodological options that are quite suitable for ex post evaluation, but have limited value in generating ex ante estimates. For example, for an event-based program, estimating a reference value based on usage on some set of prior days (referred to as a day-matching methodology) may be quite suitable for ex post evaluation. However, on its own, this method may not be useful for estimating how impacts might change on a day with weather conditions that differ from those that occurred during the historical period. Day-matching methods are also not suitable for predicting impacts resulting from changes in customer population characteristics. Ex ante estimation requires methods that correlate impacts with changes in weather and customer characteristics unless loads are not affected by these variables (in which case ex post impacts can be used for ex ante estimation purposes). Whenever possible, ex ante estimation should always begin with ex post evaluation but ex ante estimation places additional demands on the approach that aren't necessary if only ex post estimates are needed.

3.4 Program Life Cycle Considerations

Many of the current DR programs in California are in different stages of development. Some are brand new and have little or no history, while others have been in place for a number of years. Among the latter, some are relatively stable while others are rapidly expanding or evolving in other ways. Evaluation options and requirements will vary depending on how long a program has been in place, how stable it has been in terms of participation and program features, the number of participants in the program, and whether key factors are expected to change in the future. Program life cycle is a key consideration in evaluation planning.

For example, at one extreme is a brand new program. In this situation, ex post evaluation of a utility's own program participants is not an option. Consequently, the evaluator must look for similar programs, both inside and outside of the utility's service territory, that can be used as starting values. Ideally, it will be possible to adjust these borrowed estimates for differences in program or population characteristics (e.g., difference in air conditioning saturations or climate or differences in price differentials for price-based programs).

At the other extreme would be a program that has been in place for a number of years during which program and population characteristics have changed little. This type of program should provide a very robust database that would allow, at a minimum, relatively precise ex post impact estimates. However, depending on the nature of the program, even these circumstances could provide challenges for ex ante estimation if, for example, the program is only operated under emergency conditions and only a few events have been called. In this situation, one might be able to develop precise estimates of the impacts that occurred for the events that were called, but it could be difficult to develop ex ante forecasts for day types that did not occur during the historical period, as there might not be enough variation in event characteristics over the program history to allow for estimation of the relationships that are needed to provide such predictions.

In between the above extremes are programs that have seen a good deal of change over their history or for which significant changes are planned in the future. In these circumstances, it could be useful to base ex ante estimates on a combination of ex post evaluation, estimates from other programs that are more similar to where the program is headed than where it has been, and a solid dose of evaluator judgment.

In short, program life cycle—where a program has been, where it is now and where it's headed—is a key consideration in evaluation planning.

3.5 Program Size

Program size is an important consideration in terms of developing an approach that balances the level of effort and evaluation budget with the relative importance of a program in a Company's overall DR portfolio. A program that is expected to produce 200 MWs of demand response deserves more attention and budget than one that is expected to produce 2 MWs of demand response. Programs that have thousands or hundreds of thousands of participants will require more resources, and allow for more robust investigations, than one that has a few dozen participants. Program size, measured both in terms of expected impact as well as number of participants, is an important determinant of the methodology that can and should be used for impact evaluation and estimation.

3.6 Mass Market Versus Large C&I Programs

To date, most DR programs in California have been focused on large C&I customers rather than mass market residential and small C&I customers. As advanced meters are more widely deployed, the penetration of demand response among smaller consumers is likely to increase. Issues that affect program planning vary significantly across these broad customer categories.

For large C&I customers, it is often possible and almost always preferable to use data from all program participants. Most of these customers already have interval meters. For both of these reasons, technical issues arising from sampling are largely irrelevant. On the other hand, the significant heterogeneity of this customer segment introduces a number of challenges, including the possibility that load impacts from a few very large consumers can dominate program impacts, thus increasing inherent uncertainty. There may also be challenges associated with the possibility that certain tariffs are mandatory for this segment (e.g., mandatory TOU rates), which means that it may be difficult to develop a control group of customers against which load shapes can be compared in order to estimate impacts.

With the mass market customers, the need for sampling is much more likely, and there are many issues associated with sample design that must be addressed. On the other hand, the fact that you may have many more customers to work with can also be advantageous, in that it provides a robust source of data that can allow for a rich exploration of the underlying causes of demand response. It also can provide more precise estimates of DR impacts that are not subject to wide variation due to the behavioral fluctuations for a few dominant consumers.

3.7 Evaluation Today and Evaluation Tomorrow

A final consideration in evaluation planning involves understanding what is possible or cost effective today versus what may be possible or at least less expensive when the metering and communication infrastructure has evolved to the point where all consumers have interval meters that can be remotely read. As indicated in subsequent sections, there are challenges associated with the fact that load data often doesn't exist on consumers before they go on a demand response program. As such, impacts can't be measured relative to what consumer behavior was prior to joining a program. This shortcoming can be addressed by using external control groups for comparison, but there are challenges associated with picking appropriate control groups. When advanced metering is widespread, there will almost always be pre-participation data available on consumers. Furthermore, samples can be much larger when advanced metering is widely deployed as the cost of obtaining larger samples will be much less compared with the situation today where doing so often involves installing interval meters on consumers.

4 EX POST EVALUATION FOR EVENT BASED PROGRAMS

This section contains protocols and guidelines associated with ex post evaluation for event based programs. There are three broad categories of event-based programs:

- **Event-based Pricing—**This program category includes prices that customers can respond to based on an event, i.e., a day-ahead or same-day call. This category could include many pricing variants such as critical peak pricing or a schedule of prices presented in advance that would allow customers to indicate how much load they will reduce in each hour at the offered price (e.g., demand bidding). The common element is that these prices are tied to called events by the utility, DR administrator or other operator.

Direct Load Control—This category includes mass-market programs such as a/c cycling as well as large customer programs such as auto-DR. The common thread is that load is controlled at the customer's site for a called event period through a signal sent by an operator.

- **Callable DR—**This category is similar to direct load control but, in this case, a notification is sent to the customer who then initiates actions to reduce loads, often by an amount agreed to in a contract. The difference is that load reduction is based on actions taken by the customer rather than based on an operator-controlled signal that shuts off equipment. Interruptible and curtailable tariffs are included in this category.

In addition to the protocols, this section contains detailed discussions of key issues and methodologies that should be considered when producing estimates of the change in energy use resulting from a DR program over some historical period. Guidelines and recommendations for addressing these issues are provided where appropriate.

4.1 Protocols

The protocols discussed here describe the minimum requirements that must be met for ex post evaluation and consist of the following elements:

- The impact measures and associated statistics that must be produced
- The day-types and other time periods for which the measures must be reported
- Additional data that must be reported in order to understand and interpret the estimated impacts, such as weather, event characteristics and the like

- Statistical measures that must be reported in order to provide insight regarding bias and precision associated with the evaluation and sampling methodologies employed to develop the impact measures.

4.1.1 Time Periods

Event-based programs are primarily designed to produce impacts over a relatively short period of time. In addition to impacts that occurred during an event period, spillover impacts such as pre-cooling and snap back cooling might also occur in the hours immediately preceding or following the event period. Some event-based programs might even generate load shifting to a day before or day after an event.

Emergency programs, such as interruptible/curtailable tariffs and direct load control of air conditioning, are typically used only in Stage 1 or Stage 2 emergencies, and often for only a few hours in a day. Notification often occurs just a few hours before the resource is triggered or, in the case of load control, with little or no notification at all. The load impacts associated with these programs often, though not always, are constrained to the event period and perhaps a few hours surrounding the event period. For load control programs, there may be some spillover effects following the end of the event period but there is unlikely to be much impact in the hours leading up to the event.

The load impact pattern for price-driven, event-based programs may differ somewhat from emergency programs in that notification typically occurs sooner, often the day before, and a greater proportion of load reduction during the event period may result from load shifting rather than load reduction. In the residential sector, for example, the dirty laundry doesn't go away during a critical peak period. Some customers will choose to shift their laundry activity to later in the event day, the next day or perhaps even the prior day after receiving notification that the next day will be a high priced day.

Protocols 1 and 2 describe the minimum time periods for which load impact estimates must be provided for event based programs.

> ***Protocol 1: The mean change in energy use per hour (kWh/hr) for each hour of the day must be estimated for each day type and level of aggregation defined in Protocol 5. The mean change in energy use for the day must also be reported for each day type.***

> ***Protocol 2: The mean change in energy use per year must be reported for the average across all participants and for the sum of all participants on a program for each year over which the evaluation is conducted.***

4.1.2 Uncertainty

As previously discussed, uncertainty is an inherent attribute of load and demand response estimation, and an important consideration in the application of DR impacts to resource planning and dispatch operations. For ex post evaluation, uncertainty can be controlled by selecting appropriate sample sizes, careful attention to sampling strategy, model specification and other means, but it can not be eliminated completely except perhaps in very special situations that almost never occur.[8] Even if data is available for all customers, it is impossible to observe what each customer would have used "but for" the actions they took in response to the DR program. The "but for" load, referred to as the reference load, must be estimated and there will be uncertainty in the estimate regardless of what approach is used.[9]

In order for DR impact estimates to be most useful to a diverse user community, it is important to quantify and report measures of uncertainty associated with the estimates. Some users will be content to know the mean or median impact, measured with a certain level of precision, but others will want to know how much demand response is likely to occur with a probability of, say, 90 percent. To meet this need, Protocol 3 requires that uncertainty adjusted impact estimates be provided for selected points on the probability distribution of estimated load.

> ***Protocol 3: Estimates of the 10th, 50th and 90th percentiles of the change in energy use in each hour, day and year, as described in Protocols 1 and 2, for each day-type and level of aggregation described in Protocol 5, are to be provided .***

4.1.3 Output Format

Impact estimates can be developed using a variety of methodologies. A detailed discussion of the advantages and disadvantages of selected methodologies for event based programs is provided in Section 4.2. While a variety of methods can be used, two are most common: day-matching and regression analysis.[10]

With day matching, a reference value representing what a customer would have used on an event day in the absence of participation in the DR program is developed based on electricity use on a set of non-event days that are assumed to have usage patterns similar to what would have occurred on the event days. For non-weather sensitive loads that

---

[8] For example, one can imagine a program that automatically switches off pumps that otherwise are always running and pretty much drawing the same load at all times. In this situation, sub-metering the pumps would provide a highly precise estimate of what the load would have been on the event day if they had not been switched off. However, this is not the typical situation faced by DR impact evaluators.

[9] As discussed in Section 6, with ex ante estimation, uncertainty can also result from the inherent uncertainty associated with key drivers of DR impacts such as weather. If a user wants to know what impacts are likely to occur tomorrow or on a day with a specific weather profile, it is important to recognize that the temperature at 2 pm on the day of interest, for example, is not knowable. It may have a high probability of equaling 92 degrees, say, but it is more realistic to base impact estimates on some distribution of temperatures (preferably derived from historical weather data) with a mean of 92 degrees and a distribution that would indicate, for example, that the temperature has a 90 percent probability of being between 90 and 94 degrees.

[10] Other methods include a comparison of means between control and treatment groups, engineering analysis, sub-metering, etc.

vary little across days of the week or seasons, usage on almost any set of non-event days could be used to estimate the reference value. For weather sensitive loads, day matching requires picking days that have weather similar to the expected weather on typical event days. When doing ex post evaluation, suitable days might be found from days that precede or follow the event day.[11] As discussed later, adjusting initial reference values derived from non-event days for differences in usage outside the event period between prior days and event days, or using some other form of "same-day" adjustment mechanism, can improve the accuracy of the reference value. With day matching methods, impacts are measured as the difference between the reference value and actual loads on the event day.

Regression analysis is an alternative to day-matching. Like day matching, regression analysis relies on historical information about customer loads, but instead of predicting loads using the averages observed over a given number of previous days, regression analysis focuses on understanding the relationship between loads, or load impacts, during hours of interest and other predictor variables. Examples of predictor variables include temperature, population characteristics, program effects and observed loads in the hours preceding the DR event. A detailed discussion of regression analysis is contained in Section 4.2.2.

Regardless of whether day matching or regression analysis is being used, it is possible to report observed load, a reference value and impacts for each event day. For day matching methods, the impact is calculated as the difference between the reference load and the observed load. For regression methods, the impact estimates can be determined directly from the regression model. These impact estimates can be added to the observed loads in order to estimate a reference value. Protocol 4 indicates the format in which these values should be reported for event based programs.

> ***Protocol 4: Impact estimates shall be reported in the format depicted in Table 4-1 for all required day types and levels of aggregation, as delineated in Protocol 5.***

[11] When developing reference values to use for short term settlement with participating customers, only prior days are likely to be available for use in the impact calculation.

Table 4-1
Required Reporting Format for Impact Estimates

Hour Ending	Reference Load (kWh/hr)	Observed Load (kWh/hr)	Load Impact (kWh/hr)	Temperature (degrees F)	Uncertainty Adjusted Energy Impacts		
					10th Percentile	50th Percentile	90th Percentile
1							
2							
3							
4							
5							
6							
7							
8							
9							
10							
11							
12							
13							
14							
15							
16							
17							
18							
19							
20							
21							
22							
23							
24							
					Uncertainty Adjusted Energy Impacts		
	Reference Energy Use	Observed Energy Use	Change in Energy Use	Degree Hours (Base 75)	10th Percentile	50th Percentile	90th Percentile
Day							

Each variable in Table 4-1 is defined below:

- **Reference Load (Energy Use):** An estimate of the load in an hour or energy use over a period of time that would have occurred "but for" the change in their usage in response to the DR program offering

- **Observed Load (Energy Use):** Metered usage in an hour (for load) or over a period of time (for energy)

- **Load (Energy) Impact:** The impact estimate for an hour or over a period of time (e.g., day, season, year)

- **Temperature:** The average temperature in each hour, measured in degrees Fahrenheit

- **Uncertainty Adjusted Load (Energy) Impacts:** The estimated load impact value that is likely to be equaled or exceeded X% of the time. For example, if the Uncertainty Adjusted Load Impact at the 90th percentile equals 100 MW, it means that there is a 90 percent probability that the load impact will equal or exceed 100

MW or, alternatively, a 10 percent probability that the impact will be less than 100 MW.

- **Degree Hours:** The difference between temperature in each hour and a base value. For example, if the temperature is 85 degrees in an hour and the base value is 75 degrees, the number of degree hours to base 75 in that hour would equal 10.[12] If the actual temperature is below the base value, the number of degree hours in that hour is set to 0. The number of degree hours in a day is the sum of the degree hours in all hours in the day.

- **Day:** Refers to the day on which an event occurs.

It should be noted that the requirement to report temperature and degree hours in Table 4-1 is designed to allow for easier comparison of impacts across day types, programs and utilities. Inclusion of these variables in the protocols is not intended to dictate that they be used as part of the impact estimation methodology. Other variables, such as relative humidity or some other predictor of weather sensitive load may be more useful than temperature for estimating load impacts. However, a common reporting requirement will facilitate cross-event, cross-program and cross-utility comparisons.

When reporting temperatures and degree days, it is intended that the temperature be reasonably representative of the population of program participants associated with the impact estimates. If participation in a program is concentrated in a very hot climate zone, for example, reporting population-weighted average temperature across an entire utility service territory may not be very useful if a substantial number of customers are located in cooler climate zones. Some sort of customer or load-weighted average temperature across weather stations close to participant locations would be much more accurate and useful.

4.1.4 Day Types

DR impacts will vary across event days based on a variety of factors, including variation in usage patterns (often driven by variation in weather), event characteristics (e.g., timing and event duration), event participation, and other factors. In order to understand the influence of these factors on demand response, it is imperative that detailed descriptions of these influencing factors on each event day be provided along with the impact estimates. In addition, for both ex post and ex ante cost-effectiveness analysis, it is useful to have impact estimates for "typical event days".

Among the significant factors that may vary across event days and certainly over time is the number of customers enrolled in a program, the number who are notified and the

---

[12] Given the significant variation in temperature across a day in many climate zones within California, often rising from the 60s to the 90s or higher between early morning and late afternoon, degree hours may be more informative for comparison purposes across locations than are maximum daily temperature or average temperature. Degree hours are typically better predictors of daily air conditioning load than is average or maximum temperature for a day.

number who participate. There is often confusion around these terms so it is useful to define how they are used in the protocols below.

Enrollment is intended to mean the number of customers who have joined a DR program. For any program where a customer needs to take a proactive step in order to enroll, program enrollment equals all customers that have taken that step and are in the program at a given point in time. This can differ significantly from the number of customers who might actually respond during an event or even from the number who are asked to respond for a given event. For any given program, enrollment should be the largest of the three variables.[13]

At a conceptual level, the number of customers notified of an event should equal all those that have actually received the notification. This could differ from the number of notifications sent for various technical reasons (e.g., failure of notification equipment) or because the notification method is not very effective (e.g., it might use a communication channel that doesn't do a good job of reaching its target audience). In most instances, however, there is a pretty high success rate with most notification methods used for DR programs and it is typically much easier to measure the number of notifications sent than it is to measure the number actually received. As such, we define notification as the number of notifications sent out. The number of customers notified may differ from the number of customers enrolled if a program is geographically targeted and different regions are called on different days or if some other type of dispatch operation is implemented that intentionally does not include all enrolled customers.

In order to facilitate comparison of impacts across programs and/or utilities, these protocols require that ex post evaluations provide impact estimates for the following day types and levels of aggregation, which are defined below in more detail:

> ***Protocol 5: The information shown in Table 4-1 must be provided for each of the following day types and levels of aggregation:***
>
> - ***For the average across all participants notified on each day on which an event was called***
> - ***For the total of all participants notified on each day on which an event was called for***
> - ***For the average across all participants notified on the average event day over the evaluation period***

[13] There is at least one type of program where enrollment is more difficult to define, namely a peak-time rebate program such as the one outlined by SDG&E in its AMI application. The program concept in that application was that all customers would be eligible to respond to the peak time rebate offering and some subset of the entire customer base will be aware of this offer through promotional schemes. Only customers who are aware are in a position to respond. Thus, it is difficult to determine whether the number of enrolled customers for such a program is all customers or just those who are aware and, if the latter, how to measure awareness.

Day type definitions and additional reporting requirements for each day type are summarized below:

In addition to the information contained in Table 4-1, the following information must be provided for each event day:

- ***Event start and stop time***
- ***Notification lead time***
- ***The number of customers who were enrolled in the program on the event day***
- ***The number of customers who were notified on the event day***
- ***Any other factors that vary across event days that are considered by the evaluator to be important for understanding and interpreting the impacts and why they vary across events.***

Impact estimates must also be provided for an average event day for each year over which the evaluation was conducted. An average event day is calculated as a day-weighted average of all event days.[14] The number of event days that apply to each hour may vary for programs that have variable length event periods.[15] In addition to the information contained in Table 4-1, the following information must be provided:

- ***The number of actual event days included in the calculation for each hour of the average day***
- ***Average number of customers enrolled in the program over the year[16]***
- ***Average number of customers notified across all event days in the year.***

---

[14] Put another way, it is the sum of the impacts in each hour for each event day divided by the number of event days. The reason to think of this as a day-weighted average is because the weights to use when calculating the standard errors are squared.

[15] For example, if there were 10 event days, and the event was triggered from 3 pm to 5 pm on all days and between 5 pm and 6 pm on 5 event days, the average for each hour between 3 pm and 5 pm would be based on all 10 days but the average from 5 pm to 6 pm would be based on the 5 event days on which the event was triggered for that hour.

[16] Since enrollment will change over time, a day-weighted average should be calculated (e.g., if there were 2 event days in the year and there were 100 customers enrolled on the first event day and 200 on the second, the day-weighted average would be 150).

4.1.5 Statistical Measures

The final protocol that applies to ex post evaluation for event-based programs concerns the calculation and reporting of statistical measures designed to reveal the level of precision and presence or absence of bias in the method used to estimate impacts. The requirements differ between day matching and regression based methods.

With day matching methods, an unbiased reference value is essential to accurate determination of impact estimates. There are various measures of accuracy that can be used and several must be reported so that resource planners and other decision makers can assess the validity of the evaluation estimates.

In order to assess bias, estimated load using the chosen approach to reference value determination must be compared with actual load on days that are as similar as possible to the event days for which DR impact estimates are desired. If a within-subjects evaluation design is used, participants act as their own control group and usage on actual event days can not be used for validation purposes as it has been influenced by the DR event. Validity should be assessed for participants using days that are similar to event days but upon which events have not been called. Alternatively, if usage information is available from a suitable, external control group that represents customers with characteristics similar to those of participants but who do not participate in the program, using control group data for validation of the reference methodology for predicting load on actual event days is a logical choice, assuming there are a sufficient number of actual event days in the historical period for the test to be meaningful.[17] If not, the evaluator should select a sufficiently large number of event-like days on which to test the validity of the reference methodology.

> ***Protocol 6: Based on a suitable and sufficiently large number of proxy days, the following statistics should be calculated and reported for day-matching reference value methods:***
>
> - ***The number of proxy days used in the calculations below and an explanation of how the proxy days were selected***
> - ***Average error across customers and proxy days for each hour for the entire day***
> - ***Median error across customers and proxy days for each hour for the entire day***
> - ***The relative average error for each hour. This is calculated as the ratio of the average error to the average actual load that occurred in the hour.***

[17] For example, if only one or two event days occurred during the historical period, it may be preferable to select additional, event-like days to include in the accuracy assessment than to use only the one or two days on which events actually occurred.

- ***The relative median error for each hour.***

- ***The Coefficient of Alienation, which describes the percentage of the variation in actual load for each hour that is not explained by variation in the predicted load. This is calculated as follows:***

$$\sum_{i=1}^{n_1}\sum_{j=1}^{n_2}\sum_{k=1}^{n_3}\frac{\left(y_{ijk}-\hat{y}_{ijk}\right)^2}{\left(y_{ijk}-\bar{y}_{jk}\right)^2} \qquad (4\text{-}1)$$

Where:

i = the cross-sectional unit

j = the event-like day

k = the hour of the day

y_{ijk} ***= the actual observed load for i in event-like day j and hour k***

$\hat{y}_{ijk}$ ***= the predicted load for i in event-like day j and hour k***

$\bar{y}_{jk}$ ***= the average load in event-like day j and hour k***

- ***Theil's U, calculated as follows***

$$\frac{\left[\frac{1}{n}\sum_{i=1}^{n}\left(y_i-\hat{y}_i\right)^2\right]^{1/2}}{\left[\frac{1}{n}\sum_{i=1}^{n}\left(y_i\right)^2\right]^{1/2}+\left[\frac{1}{n}\sum_{i=1}^{n}\left(\hat{y}_i\right)^2\right]^{1/2}} \qquad (4\text{-}2)$$

Where:

i= the number of periods

y_i ***= the actual observed load for period i***

$\hat{y}_i$ ***= the predicted load for period i***

$\bar{y}_{jk}$ ***= the average load for period i***

Theil's U describes the accuracy of the forecast. It is calculated for a particular forecast period and individual, resulting in a distribution of Theil's U estimates. The characteristics of this distribution, including mean and median, should be described.[18]

---

[18] For examples of now Theil's U can be applied, see KEMA-XENERGY (Miriam L. Goldberg and G. Kennedy Agnew). *Protocol Development for Demand Response Calculation—Findings and Recommendations*. Prepared for the California Energy Commission, February 2003.

A different protocol is relevant for regression models. The regression protocol is designed with two goals in mind:

1. Provide qualified reviewers with sufficient transparency and information so as to enable a through assessment of the validity, accuracy, and precision of the results.

2. Provide the information necessary to enable the readers to create models that provide the load impacts and the confidence intervals under specific scenarios.

Protocol 7: For regression based methods, the following should be reported:

1. ***Adjusted R-squared or, if R-squared is not provided for the estimation procedure, the log-likelihood of the model***

2. ***Total observations, number of cross-sectional units and number of time periods***

3. ***Coefficients for each of the parameters of the model***

4. ***Standard errors for each of the parameter estimates***

5. ***The variance-covariance matrix for the parameters***[19]

6. ***The tests conducted and the specific corrections conducted, if any, to ensure robust standard errors***

7. ***How the evaluation assessed the accuracy and stability of the coefficient(s) that represent the load impact***

4.2 Guidance and Recommendations for Ex Post Impact Evaluation of Event Based Programs

Section 4.1 delineated the key requirements associated with estimating ex post impacts for event based DR programs. The protocols describe what must be provided, not how to do the job. This section discusses a variety of issues that should be considered when deciding "how to do the job" and, where appropriate, provides some guidance and recommendations concerning how these issues might be addressed.

[19] The variance-covariance matrix is needed in order to calculate the correlations between the model parameters for use in determining forecast precision and uncertainty bands.

4.2.1 Day Matching Methodologies

As described briefly in Section 4.1.3, one approach to estimating DR impacts for event based programs relies on day matching. With day matching, DR impacts are estimated as the difference between a reference value, intended to represent what load would have been had a customer not changed their behavior in response to the DR program, and actual load on an event day.

Developing reference load shapes involves either two or three steps, depending on the nature of the load. The first step involves selecting relevant days and the second involves taking an average of the load in each hour for the days that were chosen. If loads vary with weather or other observable factors, a third step that can improve the reference load shape involves making "same day" adjustments to the initial load estimates. These adjustments can be based on differences between load in hours outside the event period on prior days and load during the same hours on the event day or on differences in the value of some other variable such weather on prior days and event days.

When considering what days to choose for the initial reference load calculation, for C&I customers, only business days are typically used. For residential customers, if events only occur on weekdays, weekends would logically be excluded from day selection as usage on weekends tends to be different on average from weekday usage. When it comes to using day matching, one size definitely does not fit all. What works best will vary with customer type, load shape, whether or not the load is weather sensitive, and other factors. On the other hand, an objective is to provide some consistency in the impact estimates across programs to allow for valid program comparisons. Below is a list of methods that have been used or tested in the past. This list is intended to be exemplary, not a complete census of all options:

- Previous 3, 5, 7 or 10 business days or weekdays
- Highest 10 out of 11 prior business days
- Highest 5 of the last 10 business days
- Highest 3 out of 10 prior business days[20]
- 20 days bracketing the event day
- All relevant days in an entire season.

"Same-day" adjustment options include:[21]

---

[20] This reference method is discussed in a recent LBNL report, *Estimating DR Load Impacts: Evaluation of Baseline Load Models for Commercial Buildings in California*, July 2, 2007.

[21] This discussion is based on information in KEMA-XENERGY (Miriam L. Goldberg and G. Kennedy Agnew). *Protocol Development for Demand Response Calculation—Findings and Recommendations.* Prepared for the California Energy Commission, February 2003. p. 2-12. This report uses the term baseline for what we call reference value. Hereafter, we refer to this report as the KEMA/CEC study.

- **Additive Adjustment:** A constant is added to the provisional reference value for each hour of the curtailment period. For simple additive adjustment, the constant is calculated as the difference between the actual load and the provisional reference value load for some period prior to the curtailment. Ad hoc or judgmental adjustments are also possible.

- **Scalar Adjustment:** The provisional reference value load for each hour of the curtailment period is multiplied by a fixed scalar. The scalar multiplier is calculated as the ratio of the actual load to the provisional reference value load for some period prior to the curtailment.

- **Weather-Based Adjustment:** A model of load as a function of some weather parameter is fit to historical load data. The fitted model is used to estimate load (a) for the weather conditions of the days included in the provisional reference value, and (b) for the weather conditions of the curtailment day. The difference or ratio of these two estimates is calculated, and applied to the provisional reference value as an additive or scalar adjustment.

With the additive or scalar adjustment, the two hours prior to an event and the two hours prior to that (e.g., the 3rd and 4th hours prior to the event period) have been tested. There are at least three concerns that must be addressed if the two hours prior to an event period are used to adjust an initial reference value for evaluation purposes.

- **Gaming—**if the two hours prior to the event period are also used as part of the reference value for customer settlement, and this is known by the customer, a customer might intentionally increase energy use in the hours leading up to the event period in order to increase their reference value so as to receive a higher payment.

- **Pre-cooling—**a customer might increase cooling in the hours leading up to the event period in order to retain their comfort level longer if, for example, air conditioning is being controlled during the event.

- **Other pre-event adjustments—**a C&I customer might reduce manufacturing or business operations in anticipation of the event period.

If gaming or pre-cooling occurs, impact estimates based on the two hours prior to the event period will be overstated whereas other types of pre-event adjustments may lead to under estimation. These inaccuracies could still arise when earlier hours in the day are used rather than the two hours prior to the event period, but the bias may be smaller. On the other hand, for weather sensitive loads, using the earlier hours in a day may not be as accurate if temperatures increase significantly as the day progresses.

A variety of research has been done to compare the accuracy and other attributes of various day matching methods. A useful study was completed in 2003 for the California Energy Commission and should be reviewed if a day matching approach is being considered. The KEMA/CEC study examined the relative accuracy, simplicity and other factors associated with a number of day matching methods using data on 646 large C&I customers from utilities scattered throughout the country.

The KEMA/CEC analysis concluded that the reference value calculation method that worked best for a range of load types consists of taking a simple average of the last 10 days of demand data, by hour of the day, and then shifting the resulting profile up or down so that it matches the average observed load for the period 1 to 2 hours prior to curtailment. This method worked well for both weather-sensitive and non-weather sensitive accounts, with both low and high variability, for summer and non-summer curtailments.

The KEMA/CEC study went on to report that, if the default method is problematic either because of the potential for customer gaming or because of a need to curtail more promptly, the next best alternative depends on the weather sensitivity and energy use variability of the account. The default reference value and alternatives that performed reasonably well for different types of accounts and seasons are shown in Table 4-2.

Table 4-2

Methods for Estimating a Baseline Energy Use Profile
By Customer Account Type and Event-Specific Baseline

Season:	Summer				Nonsummer			
Weather Sensitivity:	Weather-Sensitive		Non-Weather-Sensitive		Weather-Sensitive		Non-Weather-Sensitive	
Variability:	Low	High	Low	High	Low	High	Low	High
DEFAULT Baseline Profile : 10 day average baseline load profile (BLP) plus additive adjustment based on actual energy use for 1-2 hours prior to demand response signal.	X	X	X	X	X	X	X	X
ALTERNATIVES:								
10 day average BLP plus additive adjustment based on actual load for three to four hours prior to demand response signal.			X	X			X	X
Weather model to construct BLP based on actual average outside temps when DR signal sent; diagnostics model used to deternine if building loads are suffficiently weather sensitive to use this approach.	X	X	X					
Use only the highest five of the last 10 days to construct BLP, with scalar adjustment based on a Temperature-Humidity Index load model.	X	X				X		
Use only the highest five of the last 10 days to construct BLP, without any adjustment for control day.					X			
Use the highest 10 of the last 11 days to construct BLP without adjustment for control day conditions.			X	X			X	X

Baseline load proflie- The kw demand levels for each hour of a 24 hour day averaged over the previous 10 non curtailment days

Additive adjustment: The adjusted BLP is calculated by adding a fixed amount A to the 10 day average BLP for each hour.

Demand response signal: Price or emergency signal sent to request reduced electric loads for a fixed time period

Scalar adjustment: The adjusted BLP is calculated by multiplying the 10 day average DLP for each hour by a fixed amount S.

Weather model: Hourly load data from non-curtailed days is fit by a regression model using weather and calendar variables. The BLP is the fitted model applied to the observed conditions during and after a DR signal.

Other useful findings from the KEMA/CEC study include:

- An additive adjustment to the load data from two hours before curtailment can often reduce the bias and variability of almost all methods, including weather models, for weather sensitive or non-weather-sensitive load data, high or low variability accounts. Other types of adjustments can improve the performance of averages, but generally with higher bias and variability.

- With this additive adjustment, simple averaging methods in most cases perform essentially as well as complex weather models, even for weather-sensitive accounts.

- Without adjustment, most averages tend to understate the load impacts of a curtailment.

- Bias and variability of weather models tends to be reduced by the use of longer input data series, but not dramatically.

- The decreased variability with longer input series is more noticeable for conditional weather models applied to non-weather-sensitive accounts, particularly high-variability accounts.

- The different average methods performed similarly in terms of bias and variability, except for those that select a subset of days based on high load.

- For summer loads, the "High 5 of 10" average generally reduces the otherwise negative bias. For summer loads using additive adjustment, the "High 5 of 10" days gave the lowest bias measure of any of the averages, for both weather-sensitive and non-weather-sensitive accounts, and comparable variability. The "High 10 of 11" average method gave some bias reduction, but not as much.

- For non-summer loads, however, the "High 5 of 10" average method inflates an already positive bias. The other averages perform better and roughly comparable to each other, in terms of both bias and variability, for both weather-sensitive and non-weather-sensitive accounts. "The High 10 of 11" is somewhat better than the others in terms of the bias and variability measured in this study.

Analysis done by San Diego Gas & Electric (SDG&E) in support of its advanced metering application also found that same-day adjustment improves reference value calculations. SDG&E used data on roughly 340 residential customers from the year 2004 to examine the relative accuracy and bias associated with more than two dozen reference value methodologies. Methodologies using 3, 5 and 7 prior days, with and without various forms of adjustment, were examined. Average error and the sum-of-squared errors (SSE) were calculated for each method. Average error was much closer to 0 for the methods using same-day adjustment and the SSE was also among the lowest for these methods.[22]

Day matching methods are easy to understand and often easier to produce and use than regression methods. With same-day adjustment, day matching methods exist that have very small average errors and that are reasonably precise. If the primary question is, "What was the DR impact for some set of historical event days, or for individual event days?" day matching can be an intuitively appealing and practical approach.

[22] SDG&E AMI proceeding. DRA Exhibit 109.

However, day matching presents challenges in calculating the statistical uncertainty associated with ex post load impact estimates (e.g., the estimates in the right hand columns of Table 4-1). The calculation is mathematically simple for a single customer on a single event day. The reference value of the load for a particular hour on an event day is an average and that average has a standard deviation that describes the uncertainty about the estimated reference load for that hour. If the load is unadjusted, the uncertainty in the load impact can be described by calculating the lower limit of the reference load and adjusting the load impact downward accordingly.

On the other hand, if the load is adjusted for same day load or temperature, then the uncertainty in the adjustment also has to be considered in calculating the uncertainty adjusted load impact. These calculations are not difficult to make, but software would need to be written to produce these estimates in almost all cases because they are not generally provided as output from any standard statistical packages.

The problem of estimating uncertainty adjusted load impacts becomes more complicated when load impacts from more than one customer are aggregated, since sampling weights and the variance of sampling weights must also be included in the uncertainty adjustment. The foregoing calculations can become mathematically cumbersome, even though they are theoretically well defined and within the capabilities of modern desktop computers.

The use of day matching methods to estimate ex ante load impacts and their uncertainty is even more difficult. Indeed, with standard day matching approaches, there is no mathematical function or “bridge” that can be used to relate conditions that are in effect on a given ex ante day type to the load impact that will occur under the prescribed conditions. One can imagine ways of approaching the problem, such as regressing day-matching based impact estimates against explanatory variables such as weather, event characteristics and customer characteristics. However, if regression analysis is needed in order to build a bridge between estimated impacts using day matching and relevant explanatory variables, it is probably better to simply use regression methods directly as the statistical properties will be much easier to calculate and interpret. For this reason, we do not recommend day matching as a suitable approach when the primary focus is on ex ante estimation for day types that differ from those that have occurred historically. On the other hand, if the objective is a simple, straightforward way to develop an ex post estimate, day matching methods can be quite useful as a way of quickly reporting DR results.

4.2.1.1 Day Matching Analysis: An Example

Protocols 1 through 4 require that uncertainty-adjusted impact estimates be developed for event-based programs for each hour of an event day and that the total change in energy on the event day also be estimated. Impacts are to be reported for various day types using a format shown in Table 4-1. Protocol 6 also requires that statistics representing the accuracy of the day-matching method used be reported. In this section, we provide an example of how those protocols can be met using a day-matching methodology.

This example was developed using data from the 2004 and 2005 measurement and evaluation of PG&E, SCE and SDG&E's voluntary critical peak pricing (CPP) and demand bidding (DBP) programs.[23] The requirements for eligibility for the programs vary by utility, but the majority of customers enrolled in the CPP and DBP program are large commercial or industrial customers with demands greater than 200 kW. The 2004 evaluation has more extensive statistics on the accuracy of different day-matching methods and will be used to demonstrate examples of some of the descriptive statistics listed in protocol 6. The 2005 evaluation contains the 24-hour load shapes required by protocols 1 and 4 so data from this evaluation will be used to show how to satisfy protocols 1 and 4.[24] There are a few protocols not satisfied by this study. For example, this study did not estimate the total change in energy use on the event day required in protocol 2 and did not provide confidence intervals required by Protocol 3.

The first step to an evaluation done by day-matching is to select an appropriate day-matching method for the customers of interest. Since the performance of a day-matching method can vary with the type of customer being evaluated, the day-matching method used for evaluation should be verified using the statistics listed in protocol 6. The first step in protocol 6 is to find a set of reasonable proxy days on which to test the reference level. The voluntary demand response programs evaluated were new programs in 2004, so the proxy days chosen to evaluate the day-matching options in this example were 2003 summer weekdays and 2003 summer high demand days. This study did not find much difference between the reference level statistics for high demand days versus all summer weekdays, which is likely because large C&I customers are generally less sensitive to weather than smaller customers.

The statistics in Table 4-3 reflect the results for all summer weekdays. The customer data used to assess reference level accuracy came from the 2004 participants in the voluntary CPP program. Four reference levels were evaluated for CPP customers: the highest 3 of the past 10 days; the average of the previous 10 similar days with no adjustment; the 10-day adjusted by the load the day before the event; and the load on the day prior to the event. For each of the reference levels, three of the statistics listed in protocol 6 were calculated. Protocol 6 calls for these statistics to be calculated on an hourly basis, but only the average error over the event period is presented in this example.

[23] Quantum Consulting and Summit Blue Consulting *Working Group 2 Demand Response Program Evaluation – program year 2004 Final Report* Chapter 6
[24] Quantum Consulting and summit Blue Consulting *Evaluation of 2005 statewide large nonresidential day-ahead and reliability demand response programs Appendix D*

Table 4-3
Selected Statistics for Protocol 6

Reference Level	Average Error (kw)	Median Error (kw)	Relative Average Error (%)	Relative Median Error (%)	Coefficient of Alienation[25]	Median Theil's U
Highest 3 of previous 10 similar days	n/a	n/a	n/a	2.0%	8%	0.10
10-day unadjusted	n/a	n/a	n/a	-2.0%	3%	0.09
10-day adjusted	n/a	n/a	n/a	0.0%	4%	0.08
Prior day	n/a	n/a	n/a	0.0%	6%	0.10

The purpose of the first four statistics in the table is to estimate the bias of the reference level. The bias is the extent to which the reference level systematically overestimates or underestimates the load. The 2% relative median error in the table shows that using the highest 3 out of the previous 10 similar days tends to overestimate the load whereas the 10 day unadjusted median error or -2.0% shows that this reference level systematically underestimates the load impacts. Care must be taken in interpreting the relative median error for two reasons. The first is that median errors are usually less sensitive to changes in reference levels than average errors. It is possible for a reference level with a low median error to have a much higher average error. For this reason, the protocols state that both the average and median errors must be included. The second issue with interpreting the relative median error is that the percentage represents the amount by which the total load is over or underestimated, not how much the demand response impact is overestimated. For example, suppose a reference level overestimates the total load by 5%, and that the actual demand response achieved is a 7% load reduction. An impact analysis using this reference level will report a load reduction of 12%, which is an error of 71% percent relative to the actual load reduction of 7%.

The 5th and 6th statistics in the table, the coefficient of alienation and Theil's U, are intended to measure the variance of the reference level. The prior day reference level and 10-day adjusted by load on the previous day both have relative median errors near zero, indicating that both reference levels are nearly unbiased. However, the Theil's U and the Coefficient of Alienation are both higher for the prior day reference level than for the 10-day adjusted reference level. This indicates that although the prior day reference level does well on average, its

[25] Quantum reported an R-squared rather than a coefficient of alienation. The coefficients of alienation were calculated by subtracting the reported R-squared from 100%

performance varies from day to day by larger amounts than the 10-day adjusted reference level. Another way of stating this is to say that the 10-day adjusted reference level is more consistent than the prior day reference level from one day to the next. The statistics showing the variance of the reference level are important because a reference level that is too low by 50% on one day and too high by 50% on another day has an average error of zero. However, this reference level would have a high Theil's U and a high coefficient of alienation which reveals that it is a poor reference level choice.

Providing the 24 hour load impacts for an event day is straightforward using a day-matching approach. The reference load is provided by the chosen day-matching method. According to Protocol 5, both the average reduction per customer and the total load reduction for the program should be reported. For simplicity, only the average reduction per customer is presented here. Table 4-4 shows the average hourly load reduction from 108 voluntary CPP customers in SDG&E's service territory on August 26th. The reference level used for this example is the 10 day adjusted by the prior day's load. The event period ranged from 11 am to 6 pm, and the higher difference between the reference level and the actual load during this period is evident from the table.

Table 4-4 Average Load Reduction per customer SDG&E Voluntary CPP program 8/26/05							
Hour Ending	**Reference Load (kw)**	**Observed Load (kw)**	**Load Impact (kw)**	**Temperature**	**Uncertainty Adjusted Load Impacts**		
					10%	**50%**	**90%**
1	379	387	-8.29	69	n/a	n/a	n/a
2	365	380	-14.58	67	n/a	n/a	n/a
3	359	371	-12.86	67	n/a	n/a	n/a
4	346	357	-10.86	66	n/a	n/a	n/a
5	348	356	-8.38	67	n/a	n/a	n/a
6	355	357	-2.50	66	n/a	n/a	n/a
7	383	380	2.87	68	n/a	n/a	n/a
8	399	399	0.12	72	n/a	n/a	n/a
9	408	409	-0.45	77	n/a	n/a	n/a
10	419	405	13.85	82	n/a	n/a	n/a
11	419	383	35.74	87	n/a	n/a	n/a
12	410	338	71.58	91	n/a	n/a	n/a
13	377	337	39.14	90	n/a	n/a	n/a
14	359	335	23.64	89	n/a	n/a	n/a
15	359	335	23.74	89	n/a	n/a	n/a
16	380	340	39.45	88	n/a	n/a	n/a
17	390	343	47.45	85	n/a	n/a	n/a
18	387	336	51.47	83	n/a	n/a	n/a
19	391	379	12.14	81	n/a	n/a	n/a
20	399	388	11.07	76	n/a	n/a	n/a
21	401	379	22.32	76	n/a	n/a	n/a
22	393	377	16.23	76	n/a	n/a	n/a
23	389	391	-1.40	73	n/a	n/a	n/a
24	387	395	-8.53	73	n/a	n/a	n/a
					Uncertainty Adjusted Impacts		
	Reference Energy Use	**Observed Energy Use**	**Change in Energy use**	**Degree Hours**	**10%**	**50%**	**90%**
Event day	n/a	n/a	n/a	120	n/a	n/a	n/a

4.2.2 Regression Methodologies

Regression analysis is another commonly used method for estimating the impact of DR programs. Regression methods rely on statistical analysis to develop a mathematical model summarizing the relationship between a variable of interest, known as the dependent variable, and other variables, known as independent or explanatory variables, that influence the dependent variable. When used to determine DR impacts, the dependent variable is typically either energy use[26] or the change in energy use, and the independent variables can include a range of influencing factors such as weather, participant characteristics and, most importantly, variables representing the influence of the DR program. A very simple regression model that relates energy use to temperature and a variable representing the presence or absence of a DR program event is depicted in Equation 4-3.

$$E_i = a + bT_i + c(T_i)(D_i) + e \qquad (4\text{-}3)$$

where E_i = energy use in hour i

T_i = the temperature in hour i

D_i = the program variable, equal to 1 when an event is triggered in hour i, 0 otherwise

e = the regression error term

a = a constant term

b = the change in load given a change in temperature

c = the change in load given a change in temperature when a DR event is triggered.

When the primary interest is ex post impact evaluation, properly specified regression models and day-matching methods often produce similar results. The reader is referred to the KEMA/CEC report for a useful comparison of the relative accuracy and other attributes of a variety of regression models and day-matching methods. However, for ex ante estimates of DR program impacts, regression models are not only recommended, they may be the only feasible approach in many situations.

Regression modeling is complex and requires strong training in statistics and econometrics. There are many different approaches to regression modeling that vary with respect to the general method used (e.g., classical versus Bayesian), estimation algorithms (e.g., Ordinary Least Squares, Generalized Least Squares, Maximum Likelihood Estimation), functional specification (e.g., conditional

[26] Some model specifications use ratios of energy use in different time periods as a dependent variable.

demand analysis, change modeling, etc.), the use of control groups (e.g., participants versus non-participants), and the variables that are explicitly included in the model specification. No single approach will be best in all situations. Indeed, the primary objective of regression–based methods for impact estimation is to choose the method that works best for the application at hand, and to justify that choice. There is both an art and science to regression modeling and there is no substitute for a skilled professional when it comes to the successful application of regression-based methods to DR impact estimation.

4.2.2.1 Overview of Regression Analysis

A very useful overview of regression modeling, including a discussion of the many technical issues that must be considered when developing regression models, is contained in the *The California Evaluation Framework*.[27] This is a good starting point for readers who want a general understanding of some of the options and challenges associated with regression modeling. However, neither that document nor anything said here is intended to be a "how to guide" for using regression analysis for impact estimation.

An important factor to keep in mind when using regression analysis is that the goal is to do the best possible job estimating program impacts, not necessarily to develop the best model for predicting energy usage. This point is expressed well in *The California Evaluation Report* (p. 115), where it states,

> "It is important to recognize that energy savings estimates depend not on the predictive power of the model on energy use, but on the accuracy, stability, and precision of the coefficient that represents energy savings."

A model of energy use as a function of DR program effects and other explanatory variables might have a low R-squared (a measure of the explanatory power of the model), but a very high t-statistic on DR program effect variables, meaning that it may explain the impact of the DR program quite well even if it does not predict overall energy use that well.

There are many different approaches to regression modeling that vary with respect to the estimation method used, functional specification, the use of control groups, pre-treatment data, and the variables that are explicitly included in the model specification. Importantly, most of the work that econometricians do is intended to test whether the key assumptions of the estimator employed are valid, and if not, apply the appropriate corrections or alternative estimation methodologies to acquire accurate, stable, and precise load impacts.

Errors in applying econometric methods can lead to:

[27] TecMarket Works. *The California Evaluation Framework*, June 2004. pp. 105 – 120.

- Biased estimates of the load impacts
- Imprecise estimates of the level of confidence that can be placed on the results
- The inability to mathematically find a solution.

For load impacts, both unbiased estimates and correct portrayals of the uncertainty around those estimates are not only desirable, but necessary.

Table 4-5 identifies potential problems in regression modeling that can influence either the accuracy (lack of bias) or the estimated certainty of the load impacts. It is not intended to be an all inclusive list of potential regression pathologies. Rather, it highlights some of those that can be most damaging to estimating DR impacts using regression methods.

Table 4-5
Issues in Regression Analysis

Problems that potentially bias estimates	Problems that lead to incorrect standard errors
1. Omitted Variable: This is a type of specification error. Omitted variables that are related to the dependent variable are picked up in the error term. If correlated with explanatory variables representing the load impacts, they will bias the parameter estimates.	**1. Serial-Correlation:** Also known as auto-correlation, it occurs when the error term for an observation is correlated with the error term in another observation. It can occur in any study where the order of the observations has some meaning. Although it occurs most frequently with time-series data, it can also be due to spatial factors and clustering (i.e., the error terms of individual customers are correlated).
3. Improper functional form: This occurs when the relationship of an explanatory variable to the dependent variable is incorrectly specified. For example, the function may be treating the variable as linear when, in fact, it is logarithmic. With this error, predictions of load impacts can be wrong	**2. Heteroscedasticity:** This occurs when the variance is not constant but is related to a continuous variable. Depending on the model, if unaccounted for, it can lead to incorrect inferences of the uncertainty of the estimates
4. Simultaneity: Otherwise known as endogeneity, it occurs when the dependent variable influences an explanatory variable. This is unlikely to be a problem in modeling load impacts.	**3. Irrelevant Variables:** When irrelevant variables are introduced into a model, they generally weaken the standard errors of the explanatory variables related to the dependent variable. This leads to overstating the uncertainty associated with the impacts of other explanatory variables.
5. Errors in Variables: When the explanatory variables contain measurement error, it can create bias if the measurement error is correlated with explanatory variables(s).	
6. Influential data: A data point is considered influential if deleting it changes the parameter estimates. Influential variables are typically outliers with leverage. These are more of an issue with Large C&I customers.	

Importantly, a large number of the problems that lead to potential bias are due to model misspecification and the closely related phenomena of correlations between the error terms and the explanatory variables. Despite a large set of diagnostic tools, it is difficult to write down a set of rules that can be use to guide model specification, especially since the best approach for model specification is not a settled question. This is where the art of regression analysis comes into play, making the experience and knowledge base of evaluators and reviewers critical.

Typically, DR load impact analysis involves both a time series and a cross-sectional dimension. This type of data is referred to by a variety of names –

including time series cross-sectional, panel, longitudinal, and repeated measures data. Although the advantages are detailed further in a following section, its ability to lead to more robust findings is noteworthy and deserves an early discussion. With this type of data evaluators are able to account for significant share of omitted variables, including those that are unobservable or not recorded, leading to better specified, more robust regression models.

Panel data can control for omitted and sometimes unobserved factors that vary across individuals but are fixed over the course of the study (fixed effects – e.g. household size, income, appliance holdings, etc.), and for factors that are fixed for all customers but vary over time (time effects -economic conditions). Regression-like models that can be used to analyze panel data include ANOVA, ANCOVA, and MANOVA. These models are similar in that they allow each individual to act as their own control and account for the effects of the fixed, but unmeasured characteristics of each customer.

However, the ability to control for fixed effects comes at a price. By controlling for fixed effects, these models cannot incorporate the impact of explanatory variables that are time-invariant (e.g., A/C) except through interactions with time-variant variables (e.g. temperature). In other words, a fixed effects model only controls for the variation within individual units; it does not control for the variation across individuals units. In many instances impact evaluations will need to take into account how fixed characteristics such as appliance holdings, household size, etc. affect the load response provided, requiring either:

- The use of interactions,

- A two-stage model, where load impacts for each customers are first estimated using individual regressions (or regressions for customer pools defined by criteria such as industry classification) followed by a second stage that regresses load impacts against customer characteristics.

- Using a random effects model which is able to use fixed characteristics as explanatory variables.

Because random effects models can provide biased parameter estimates when the error terms are correlated with the explanatory variables, it important to always start with the more robust fixed effects model and subsequently test whether the resulting coefficients and standard errors are the same. This is typically accomplished via a Hausman test. Interpreting the results of such a test, however, requires the evaluator's judgment. Due to the power of time-series cross sectional load data (which has more time observations than most panel data) and the sensitivity of the Hausman test, even trivial differences in results can be statistically significant when in fact the differences between the two models is virtually nil.

Two additional topics that are particularly pertinent to load impact data deserve mention: auto-correlation and heteroscedasticity. Having both cross-sectional and time-series dimensions there are multiple ways in which the errors can be related. Basic panel data methods generally assume:

- No correlation between the error terms of units in the same time period
- No correlation across units in different time periods
- No auto-correlations within units over time.
- Constant variances over time within a unit (Different variances across units are allowed)

Evaluations will most likely have to account for auto-correlation due to the prevalence of a time dimension in load impact data. However, it is important to distinguish between pure and impure auto-correlation. Impure auto-correlation can arise because of a specification error such as an omitted variable or incorrect functional form. Pure auto-correlation is the correlation that is still present when the model is properly specified. This implies that auto-correlation should be viewed as more than a nuisance to be corrected, but as a signal to further explore the potentially larger problem of misspecification. Correcting the standard errors due to auto-correlation is straightforward and there are a number of options for addressing it, including first differencing, Generalized Least Squares, and the use of Maximum Likelihood estimation that do not assume an error matrix with constant diagonals and zero values in the off-diagonals.

Only heteroscedasticity within individual units is problematic in panel data, although when faced with large variations in customer size and impacts, the evaluator should consider transforming the data to a common metric such as the percent change in load. While heteroscedasticity can typically be corrected for by the use of robust standard errors – also known as Huber-White standard errors and the sandwich standard errors – they do not apply if serial correlation is present[28]. Because of this, the more labor intensive process of testing for heteroscedasticity, determining the specific form of heteroscedasticity, and applying the appropriate data transformation may often be required to identify and correct for heteroscedasticity within units.

Our goal in the remainder of this section is to briefly discuss some topics that are pertinent to the use of regression analysis for ex post DR impact estimation for event-based programs. In the following section, we provide a brief example showing how regression analysis can be used to meet the output requirements specified under Protocol 4.

[28] Page 274-276 of Jeffrey Woolridge's textbook "Econometric Analysis of Cross-section and Panel Data" provides and excellent discussion on serial correlation and the robust variance matrix estimator.

4.2.2.2 The Advantages of Repeated Measures

One of the interesting and useful characteristics of event based programs that differs from the typical situation with both EE evaluation and the evaluation of non-event based DR programs is the fact that you are typically able to observe the impact of the DR program multiple times for the same customer. For an energy efficiency program or for non-event based DR programs, if you have usage data before a customer enrolls in a program, even if you have daily or hourly usage data, you only have two time periods per customer in which the DR program variable(s) differs, one before enrollment and one after. If there is no pretreatment data, you only have one time period for each customer (in which case a suitable control group is needed in order to statistically estimate the impact of the DR program). However, with event-based programs, you get multiple observations for each customer over which the DR program influence either is or is not in effect. For example, if you have twelve days in a year in which a CPP day is called, you have 12 days on which the program influence is in effect, and many more days in which it is not.

The repeated measure effect associated with event-based DR programs has several significant advantages for impact evaluation compared with non-event based programs. One concerns sampling efficiency. As discussed in Section 8, with repeated measures, you may be able to use much smaller sample sizes to achieve the same level of statistical precision. The reduction in sample size is a function of the expected impact size, the coefficient of variation and the number of repeated measures that occur, but a 10-fold decrease may be possible compared with a simple comparison of means using before-and-after data on participants or side-by-side data with participant and control samples.

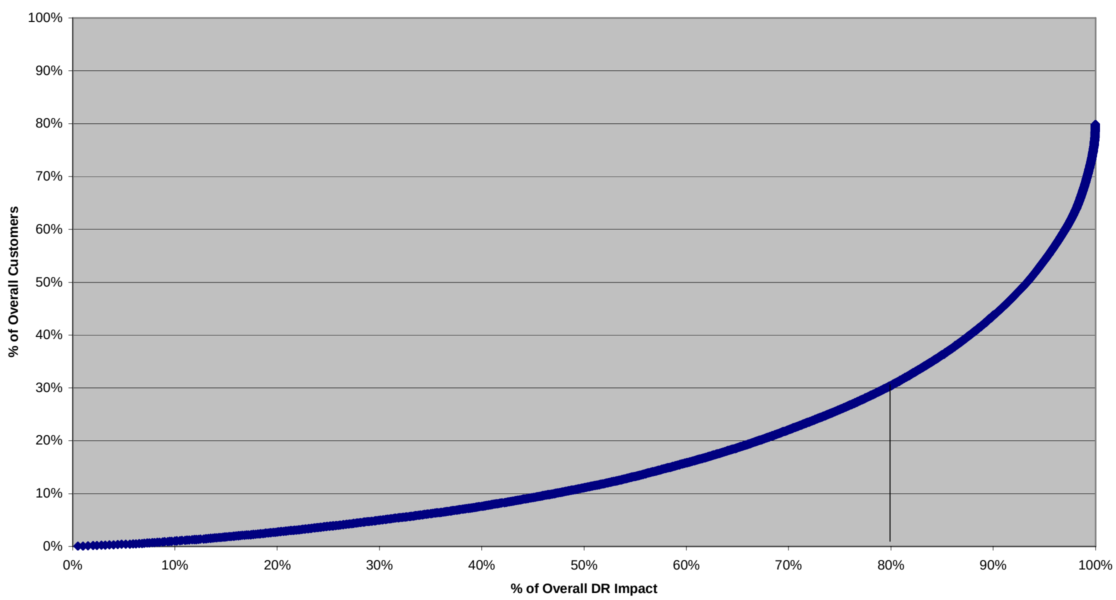
A second advantage of the repeated measure effect associated with event-based programs is that impact estimation typically does not require an external control group.[29] The fact that the DR program influence is in effect on some days and not on others allows you to estimate the influence of variation in factors that change daily, such as weather, along with the influence of the DR program. This, in turn, allows you to estimate the impact of the DR program on any day type that can be characterized in terms of the explanatory variables included in the model without needing a sample of customers who do not participate in the program. This eliminates any concern about internal validity, as there is no opportunity for differences between control and treatment groups to generate biased estimates. This is a significant advantage as long as your primary interest is in estimating

[29] There are situations in which an external control group might still be needed. For example, if an event is only called on the hottest days of the year, and the relationship between energy use on those days is different from what it is on other days, the model may not be able to accurately estimate program impacts on event days. In this instance, it may be necessary to have a control group in order to accurately model the relationship between weather and energy use on the hottest days in order to obtain an unbiased estimate of the impact of the program on those day types.

impacts for a set of volunteers behaviorally similar to those who have participated to date.[30]

A third advantage associated with the repeated measures property of event-based programs is that it allows you to estimate customer-specific regressions. For example, a regression model like the very simple specification shown earlier in Equation 4-1, could be estimated for each individual customer. This would allow you to understand the distribution of impacts across customers, which can be quite useful from a policy perspective, since it allows one to determine if the average impact is more or less typical, or, alternatively, if a relatively small percentage of customers account for the majority of demand response. For example, this type of analysis based on the SPP data produced the distribution of demand response impacts shown in Figure 4-1, indicating that roughly 80 percent of total demand response was provided by roughly 30 percent of participants.

Figure 4-1
Percent Demand Response Impact Relative to Percent Population
California's Statewide Pricing Pilot



A final advantage associated with repeated measures for a cross-section of customers is the ability to better specify regression equations and to produce more robust results.[31] Regressions that have observations over time and across customers can control for omitted variables that vary across customers but are

[30] There may still be some interest in knowing how participants differ from non-participants if there is a need to extrapolate the impact estimates to a population of customers who are unlikely to volunteer (which may differ from those who have not yet volunteered). If so, an external control group may be needed. A more in depth discussion of control groups is contained in Section 5.2.
[31] Peter Kennedy, in *Guide to Econometrics*, provides an excellent discussion of some of the advantages of having repeated measures across a cross-section of customer sin the introduction to Chapter 17.

fixed over the study period, known as fixed effects, and for omitted variables that are fixed across customers but vary over time, know as time effects.

4.2.2.3 Customer Segmentation

The ability to estimate impacts from individual customer regressions can also be quite useful in situations where a few large customers contribute most of the DR impact for a program. In this situation, pooling across customers is likely to violate some of the assumptions that underlie typical regression analysis, such as homogeneity of error terms across customers. Running separate regressions for larger customers can eliminate this type of problem. Furthermore, understanding how these dominant customers behave, which may be quite different from the more numerous smaller customers, may be very useful.

The previous comment about the possibility of estimating separate regressions for large and small C&I customers is just one example of a larger issue—namely, the need to understand how response varies across customer characteristics and the possibility that using different model specifications or even different methodologies may be better than using a single approach. These are important considerations especially when evaluating DR programs targeted at C&I customers, where there is a significant degree of heterogeneity across customers and a greater likelihood that regression methods may not work in all cases.

Regression modeling will not work well under a variety of conditions. For example, it won't work well if there is a large degree of variation in energy use that can't be explained by variation in observable variables and the DR impact is small relative to the total load. This can occur if data on the independent variables that drive this variation is difficult to obtain, as it could be with industrial customers where variation may be caused by industrial process operations that are hard to measure. If the DR impact is small relative to the variation in energy use, and that variation in energy use can't be explained, it will be very difficult for the regression analysis to isolate changes in energy use due to the DR program from the unexplained variation in energy use due to other factors. Given these circumstances, day matching may not work well either since significant variation from day to day means that an average over even a large number of days will not often be a good proxy for what the load would have been on an event day in the absence of the event. If there is less variation in the loads that are contributing to the DR impact than there is in the total customer load, it may be possible to use day matching or regression analysis with sub-metered data for these partial loads.

In contrast to the situation where too much variation creates estimation difficulties is the case where there is too little day-to-day variation in load. For example, with loads that are not at all weather sensitive and, as a result, may not vary much from day-to-day, there may not be much of an advantage in using regression analysis over less complicated and easier to understand methods such as day matching. In these circumstances, regression analysis may be effective for estimating the impact of the DR event, but that impact wouldn't be expected to change from one event to another in response to variation in other observable factors such as weather. As such, one of the primary benefits of regression analysis, the ability to make ex ante estimates for day types or other conditions that differ from the past, is no longer relevant. Given this, if some participants in a DR program have weather sensitive loads, or loads that vary with other observable variables, while other participants have loads that vary very little, using regression modeling to estimate impacts for the variable segment and day-matching to estimate impacts for the non-variable segment may be the best strategy. In these circumstances, using a regression model to estimate the impacts for both types of customers may distort the impacts associated with the market segment with the variable load.[32]

4.2.2.4 Quantifying the Impact of Event Characteristics

Another important consideration when using regression analysis to estimate impacts for event-based programs is how to represent variation in event characteristics in the regression model. Notification lead time and the timing and duration of events may influence demand response for programs in which these factors are allowed to vary across events or across customers (e.g., as in cafeteria style programs). The ability to do this is a function of how much these characteristics vary over the estimation time period or across customers. Given sufficient variation, it is relatively straightforward to include interaction terms in the regression model to determine if impacts vary with these event characteristics. For example, it might be possible to define a set of binary variables representing different event period (e.g., a variable equal to 1 if the event period is less than 3 hours, 0 otherwise). This type of specification would allow you to develop ex ante estimates for specific combinations of event conditions that may be quite useful for operational purposes or for longer term resource planning or program design.

4.2.2.5 Estimating Impacts for Hours Outside the Event Period

As indicated in Protocol 1, impact estimates for event based programs are required for all hours on an event day. This requirement fulfills the need to understand the extent and nature of load shifting that occurs with some types of DR programs, and to estimate the impact of DR programs on overall energy use. Regression modeling can be used to estimate all of these impact types using a variable representing an event day, as distinct from a variable representing an event window, interacted with variables representing individual hours in a

[32] In this instance, separate output tables should be reported for each market segment.

regression analysis that pools all hours in a single regression. The example in Section 4.2.2.11, equation 4-2, illustrates this type of model specification.

4.2.2.6 Weather Effects

Accurately reflecting the influence of weather in load modeling and impact estimation is critical. This is easily done with regression analysis using weather variables and interaction terms as illustrated in the simple model in Equation 4-3 and the example shown in Section 4.2.2.11.

A related factor is heat build up in buildings caused by multiple hot days in a row. This can also be reflected in a regression model, for example, using a variable representing cooling degree hours on days prior to an event day, or cumulative cooling degree hours leading up to the event period (as also illustrated in the example in Section 4.2.2.11).

4.2.2.7 Multi-day Events

Another issue to consider when developing model specifications is variation in impacts across multi-day events. Distinct variables indicating whether an event is the first, second or third day of a multi-day event can be included in a regression specification to determine if impacts vary according to this event feature. Section 4.2 of the *Impact Evaluation of the California Statewide Pricing Pilot*[33] provides an example of this type of specification.

4.2.2.8 Participant Characteristics

Of course, as previously mentioned, a major advantage of regression based analysis is its ability to examine the impact of other influencing variables that might cause DR impacts to vary in the future from what they have been in the past. Primary among these are weather and population characteristics. By incorporating a variable representing the interaction between an event day variable (or variables) and weather (e.g., temperature, degree hours, humidity, etc.), a regression model can estimate impacts under specified weather conditions that differ from those that might have occurred over the historical period. Similarly, by including interaction terms between variables representing event days and customer characteristics, such as air conditioning and/or other equipment ownership, and socio-demographic or firmographic variables such as income, persons per household, business type and others, one can predict how impacts might change as the mix of participant characteristics changes. These topics are discussed in more detail in Section 6. We mention this here because it

[33] Charles River Associates. *Impact Evaluation of the California Statewide Pricing Pilot*. Final Report. March 16, 2005. p. 66.

is important to consider the need for ex ante estimates when developing a model specification designed to do both ex post and ex ante estimation. It might not be necessary to include socio-demographic variables in the model if only ex post estimates are needed, since fixed or variable-effects specifications can control for variation in energy use across customers without explicitly including such variables in the model. However, if ex ante estimation is needed, it will be necessary to explicitly incorporate variables in the specification that are expected to change in the future.

4.2.2.9 Geographic Specificity

Knowing how impacts vary across regions can be very useful for transmission and distribution planning and for operational dispatch decisions by the CAISO, who must balance supply and demand at thousands of points on the grid and who will soon be using locational pricing to help clear markets at numerous transmission nodes. The specific locations for which impacts may be needed in the future are still unclear, and they will vary across utilities and programs. As previously discussed, understanding the extent to which impact estimates are required for specific locations is an important input to evaluation planning.

Whatever the ultimate need, there are two basic approaches to developing location-specific impact estimates. One is to obtain large enough samples at each desired location to develop statistically valid and precise impact estimates based on each geographic sub-population. If the number of geographic regions is large, this could be a very costly approach.

An alternative approach is to incorporate variables in a regression model that explain how impacts vary according to weather and population characteristics that vary regionally. Using survey and climate data to develop estimates of the mean values for each explanatory variable by region, such a model can be used to predict what the impacts will be given the local conditions. This approach can probably be implemented with data on a much smaller sample of customers than the location-specific sampling approach by using stratified sampling methods that ensure sufficient variation in the characteristics of interest to develop the model parameters.

Implementing this approach will be easier and less costly if there is prior knowledge regarding which independent variables drive demand response and if data already exists concerning how relevant variables differ across the regions of interest. California's Residential Appliance Saturation Surveys (RASS) and Commercial End Use Surveys (CEUS) provide a rich database that can be used to inform sample designs and modeling exercises. There is also a growing body of evidence concerning what customer characteristics drive demand response for many program options. As such, there is a greater probability that sufficient prior knowledge exists in California than in many other locations. As such, a model based approach to location specific impact estimates is likely to be less costly than would be developing large enough samples at each location of interest to estimate impacts of comparable validity and precision.

4.2.2.10 Summary

Regression analysis is the leading alternative to day matching methods for estimation of DR impacts. While regression modeling requires more skill and experience to implement, and is not as transparent as most day-matching methods, it offers a number of advantages compared with other methods and should at least be considered whenever there is sufficient data available to support regression analysis. Regression analysis can be used to examine impacts outside the event period and to quantify the influence of event characteristics, heat build up, multi-day events, weather and customer characteristics on demand response. The repetitive nature of event-based programs may allow for regression analysis (or other methods) to be implemented using smaller samples than would be needed for non-event based programs. It also eliminates the need for external samples in most situations, and for customer-specific impact estimates to be developed, thus affording the opportunity to examine the distribution of impacts across the participant population.

4.2.2.11 Regression Analysis: An Example

As indicated in section 4.1, protocols 1 through 4 require that uncertainty-adjusted impact estimates be developed for event-based programs for each hour of an event day as well as the days preceding and following an event day. Impacts are to be reported for various day types using a format shown in Table 4-1. In this section, we provide a simple example of how those protocols can be met using a regression-based methodology.

This example was developed using residential customer data for the CPP rate from California's Statewide Pricing Pilot for the summer of 2004. Only data from climate zone 3 (the hot, climate zone representing California's central valley) was used. This analysis was completed using STATA, a common statistical package. It should be noted that we did not spend a significant amount of time refining the model specification, although this should be a key area of attention for regression-based evaluations. Our focus here is on demonstrating how to use regression techniques to meet the protocol requirements.

The estimated regression model has the following form:

$$E_{iJ} = \beta_0 + \beta_1 CPPday + \sum_{j=2}^{24}\left(\beta_{3j} hour_j \times CPPday\right) + \beta_4 (CPP_j \times CDH_{ij}) + \beta_5 (CAC_i \times CDH_{ij}) +$$

$$\beta_6 (CDHrunup_{ij}) + \beta_7 (CDH24hrlag_{ij}) + \sum_{j=2}^{24} \beta_{8j} hour_j \qquad (4\text{-}4)$$

Where

CPPday = 1 on an event day, 0 otherwise

CPP = 1 during the event period on event days, 0 otherwise

$Hour_i$ = 1 for hour i, 0 otherwise

CAC = 1 for customers with central air conditioning, 0 otherwise

CDH_i = Cooling degree hours to base 75^o F in hour i

$CDHrunup_i$ = cumulative cooling degree hours in the day prior to hour i

$CDH24lag_i$ = cooling degree hours in hour i the day before the event day

The hourly binary variables capture the non-weather dependent load shape on non-critical days whereas the hourly variables interacted with the CPP day binary variable estimate the difference in the load in each hour on CPP days relative to non-critical days. The interaction between the CPP event binary variable and the cooling degree hour variable allows one to estimate the change in the program impact as cooling degree hours change. In order to estimate impacts on the day preceding or following an event day, binary variables representing these days interacted with the hourly binary variables could be included in the specification. For simplicity and ease of interpretation, we did not include these variables in the example.

Figure 4-2 contains the regression output for the model. As seen, the cooling degree hours variable has a strong positive relationship when interacted with central air conditioning, indicating that energy use increases with cooling degree hours for households with air conditioning. The negative sign on the interaction term between degree hours and the CPP variable, which equals 1 during CPP hours, 0 otherwise, indicates that energy use drops more during the event hours when the day is hotter than when it is cooler. This is logical as there is more load to drop on hotter days due to air conditioning use. The positive sign on the interaction term between the hour of the day and the CPP day binary variable for the hours immediately preceding and following the event period indicates a small amount of pre-cooling and a significant snapback effect. Tests of joint significance applied to the results from the event hours and the surrounding hours indicate that the CPP impacts are statistically significant and in the expected direction across the event period hours (2-7 pm), pre-event hours (12-2 pm), and post-event hours (7-9 pm).

Figure 4-2: Regression Output

kWh	Coef.	Std. Err.	t	P>t
CPPday	-0.0097	0.0149	-0.65	0.516
CPPxCDH	-0.0070	0.0013	-5.31	0
CACxCDH	0.0329	0.0006	53.95	0
CDHrunup	0.0017	0.0000	38.64	0
CDH24hrlag	0.0138	0.0006	24.15	0
Hourly weather independent CPP Impacts				
(1:00 am is used as the reference value)				
2:00	-0.0227	0.0150	-1.51	0.1310
3:00	-0.0549	0.0191	-2.88	0.0040
4:00	-0.0644	0.0212	-3.04	0.0020
5:00	-0.0466	0.0225	-2.07	0.0380
6:00	-0.0374	0.0233	-1.61	0.1080
7:00	-0.0240	0.0238	-1.01	0.3150
8:00	-0.0228	0.0242	-0.94	0.3450
9:00	-0.0068	0.0244	-0.28	0.7800
10:00	-0.0311	0.0245	-1.27	0.2050
11:00	-0.0128	0.0247	-0.52	0.6040
12:00	-0.0136	0.0248	-0.55	0.5820
13:00 (pre event hour)	0.0400	0.0249	1.61	0.1080
14:00 (pre event hour)	0.0496	0.0250	1.99	0.0470
15:00	-0.0643	0.0323	-1.99	0.0460
16:00	-0.0334	0.0330	-1.01	0.3110
17:00	-0.0043	0.0330	-0.13	0.8970
18:00	0.0304	0.0321	0.95	0.3430
19:00	0.0479	0.0303	1.58	0.1140
20:00 (snapback period)	0.2314	0.0242	9.54	0.0000
21:00 (snapback period)	0.3037	0.0237	12.8	0.0000
22:00 (snapback period)	0.2595	0.0229	11.35	0.0000
23:00 (snapback period)	0.1953	0.0213	9.18	0.0000
0:00	0.1071	0.0181	5.92	0.0000
Weather independent load profile				
(1:00 am is used as the reference value)				
2:00	-0.1119	0.0044	-25.55	0.0000
3:00	-0.1850	0.0057	-32.49	0.0000
4:00	-0.2152	0.0065	-33.36	0.0000
5:00	-0.2232	0.0069	-32.23	0.0000
6:00	-0.1836	0.0072	-25.36	0.0000
7:00	-0.0949	0.0074	-12.74	0.0000
8:00	-0.0231	0.0076	-3.04	0.0020
9:00	0.0187	0.0077	2.44	0.0150
10:00	0.0579	0.0077	7.48	0.0000
11:00	0.0952	0.0078	12.18	0.0000
12:00	0.1063	0.0079	13.37	0.0000
13:00	0.1243	0.0082	15.21	0.0000
14:00	0.1500	0.0085	17.70	0.0000
15:00	0.1056	0.0087	12.08	0.0000
16:00	0.1744	0.0089	19.55	0.0000
17:00	0.2763	0.0089	31.06	0.0000
18:00	0.3675	0.0087	42.45	0.0000
19:00	0.4090	0.0082	49.74	0.0000
20:00	0.5101	0.0076	66.72	0.0000
21:00	0.5976	0.0071	84.48	0.0000
22:00	0.5572	0.0065	85.47	0.0000
23:00	0.3662	0.0058	63.58	0.0000
0:00	0.1668	0.0045	37.33	0.0000
Constant	0.7133	0.0019	371.94	0.000

FE (within) regression with AR(1) disturbance	Number of obs	=	705421
Group variable (i): custidnum	Number of groups	=	203
R-sq: within = 0.0549	Obs per group: min	=	215
between = 0.0762	avg	=	3475
overall = 0.1497	max	=	3647
	F(33,705185)	=	1241
corr(u_i, Xb) = 0.0249	Prob > F	=	0

Figure 4-3 shows how the predicted values compare with actual values on the average critical event day in 2004. As seen in Figure 4-3, the model does a good job of tracking actual energy use on event days, including the substantial snapback effect that occurs following the end of the event period. The estimated impacts equal the difference in the two lines in Figure 4-3 labeled “predicted energy use without DR” and “predicted energy use with DR.” The figure also illustrates a significant drop in load impacts in the last two hours of the event period. The impact estimates illustrated in Figure 4-3 are shown in Table 4-6, which is in the format required by Protocol 3 for the average event day.

Figure 4-3

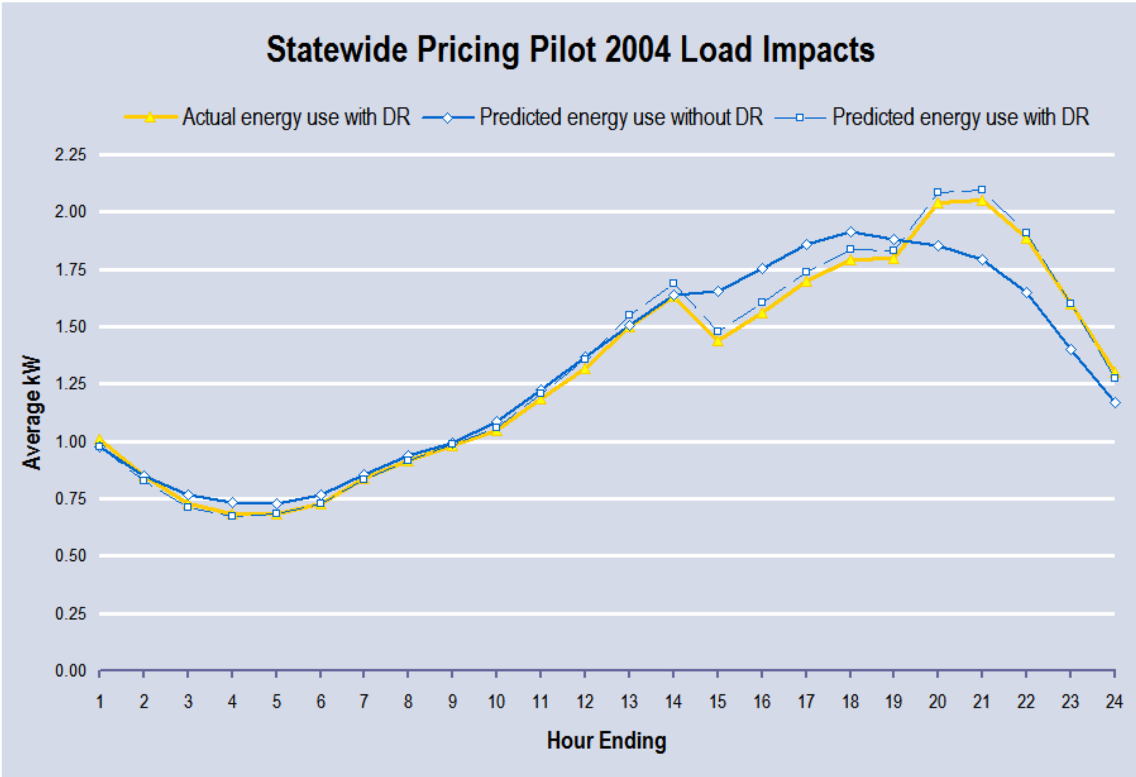


Table 4-6

Day Type: Average Event Day for 2004 SPP Residential - Climate Zone 3

Hour Ending	Temperature	Per participant load impacts			
		Mean (kW)	Percentiles		
			10%	50%	90%
1	71.2	-0.009	0.011	-0.010	-0.030
2	70.0	-0.033	-0.008	-0.032	-0.057
3	68.8	-0.064	-0.039	-0.064	-0.090
4	67.8	-0.074	-0.047	-0.074	-0.101
5	66.9	-0.057	-0.030	-0.057	-0.084
6	66.1	-0.047	-0.019	-0.047	-0.074
7	65.9	-0.034	-0.007	-0.034	-0.061
8	67.2	-0.033	-0.005	-0.032	-0.060
9	70.1	-0.017	0.011	-0.017	-0.044
10	74.4	-0.041	-0.013	-0.041	-0.068
11	78.7	-0.022	0.005	-0.022	-0.049
12	82.9	-0.023	0.004	-0.023	-0.051
13	86.4	0.030	0.058	0.030	0.002
14	89.1	0.040	0.067	0.040	0.013
15	90.8	-0.185	-0.158	-0.185	-0.212
16	91.7	-0.160	-0.132	-0.160	-0.188
17	91.6	-0.131	-0.104	-0.131	-0.159
18	90.5	-0.090	-0.062	-0.090	-0.117
19	88.2	-0.057	-0.030	-0.057	-0.085
20	84.5	0.222	0.249	0.222	0.195
21	80.2	0.294	0.320	0.294	0.267
22	76.7	0.250	0.275	0.250	0.225
23	74.3	0.186	0.210	0.186	0.162
24	72.6	0.097	0.118	0.097	0.077

The uncertainty adjusted load impacts shown in the right-hand columns in Table 4-6 can be generated in two ways.

One approach involves using the regression model to compute the difference in the mean predicted load with and without the DR in effect, and using the standard errors of the predictions to estimate the uncertainty surrounding that difference, i.e., the confidence intervals. The exact equation to use for this calculation will vary depending on whether or not the variances of the predictions are equal and the size of the sample (small samples require adjustments). Assuming the variance of the estimates is equal and the sample size is sufficiently large, the load impact (difference of mean predictions) and the standard error of the difference are given by the following formulas:

$$\text{AVERAGE LOAD IMPACT} = \overline{X}_0 - \overline{X}_1$$

Where $\overline{X}_0$ = Energy use with DR and
$\overline{X}_1$ = Energy use without DR

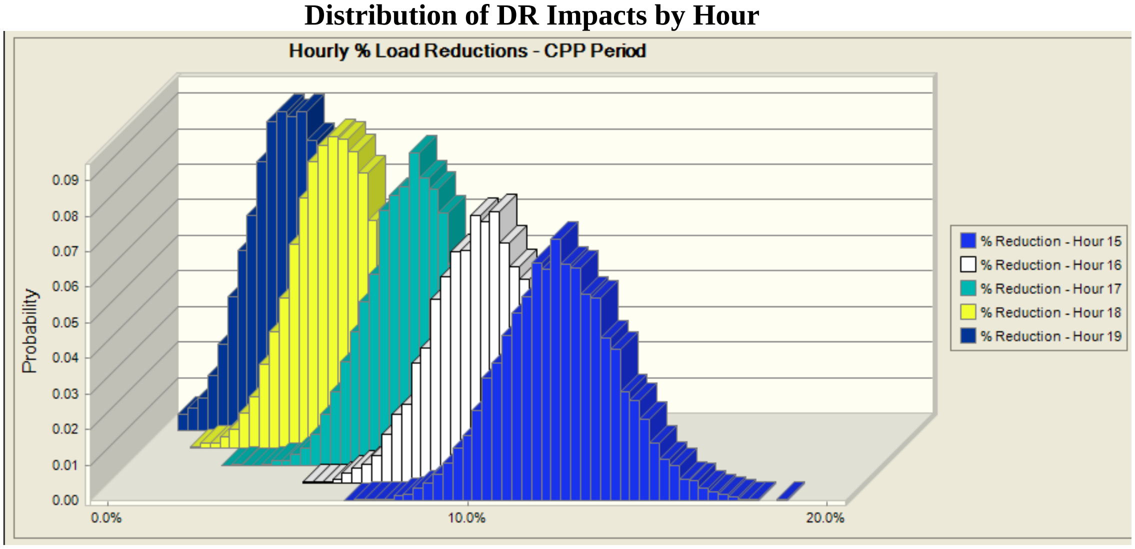
STANDARD ERROR OF THE DIFFERENCE

$$s_{\overline{x}_1 - \overline{x}_0} = \frac{s_0 + s_1}{2}$$

Where s_0 = Standard error of predicted energy use with DR, and
s_1 = Standard error of predicted energy use without DR

An alternative approach to estimating the uncertainty adjusted impacts is Monte Carlo simulation. Monte Carlo simulation will produce the same results provided the simulation includes enough draws (trials). Monte Carlo simulations employ a transparent, brute force approach in which random draws are made from the probability distributions of factors that affect the outcome. The uncertainty adjusted impact estimates in Table 4-6 were produced using Monte Carlo simulation. Figure 4-4 shows how the distribution of DR impacts (in percentage terms) looks in each hour based on the model shown in Figure 4-2.

Figure 4-4
Distribution of DR Impacts by Hour



A significant advantage of Monte Carlo simulation is that it allows you not only to incorporate uncertainty due to forecast error, but to also incorporate uncertainty in the explanatory variables. Given information on the distribution in temperature or degree hours around some mean value in each hour, for example, one can use simulation to produce a probability distribution of impacts that reflects the uncertainty in both weather and the predictive power of the underlying model

The steps outlined below illustrate how to produce probability distributions that reflect both modeling error and uncertainty in the distribution of key explanatory variables such as weather:

1. Predict the load with the DR program in effect by temperature level and hour of day, with a 90% confidence interval for each prediction.

2. Predict the load without DR program in effect by temperature level and hour of day, with the 90% confidence interval for each prediction.

3. Base the weather distribution for each hour of the day by day type from historical data about the same type of day, including a) the mean b) standard deviation c) autocorrelation. A better approach is to fit and compare a range of distributions to the actual weather data, by hour, for that day type and include the hourly auto-correlation of weather. This is preferred because distributions may be skewed, have long tails, or be bimodal, i.e., they may be non-normal.

4. Run a simulation allowing the temperatures to vary (taking into account the actual correlations). The appropriate distribution for the load with and without DR would be created by drawing the mean and standard error of the hour and temperature lookup tables. The draw from these distributions of load with and without the DR program in effect would in turn be used to calculate the load impact for each.

5. Extract the resulting summary statistics from the resulting distributions of the Monte Carlo simulation.

4.2.3 Other Methodologies

Sections 4.2.1 and 4.2.2 provided an overview and discussion of issues associated with the two primary methods of estimating ex post impacts for event-based DR programs, day-matching and regression analysis. Regression analysis is a less transparent but more robust and flexible tool than day-matching. It is the recommended default option whenever ex ante estimation is required unless other considerations, such as erratic consumer behavior, lack of variability, data limitations, budget constraints, or the limited importance of a program due to its small size, suggest that an alternative approach is preferred. This section covers some of the additional options that might be considered if one or more of these conditions is present.

One approach already mentioned is sub-metering. Sub-metering is primarily useful in situations where the load contributing to demand response is relatively easy to isolate without rewiring or other costly procedures. An example is when load response is associated with a single piece of end-use equipment (e.g., an air conditioner, pump or other large motor).

If the isolated equipment is always on except when interrupted for an event, sub-metering will provide a very accurate estimate of load impact by simply comparing load just prior to and after the beginning of an event period. If the equipment has a duty cycle, and one that differs across days due to variation in weather or some other variable, there will still be a need to develop a reference load shape or, alternatively, use regression analysis to predict the "but for" load. However, this task will typically be much simpler when the data being used reflects the only relevant load rather than total premise load. Sub-metering may be necessary if there is significant variability in premise load and the DR impact is small relative to total premise load. In these circumstances, day matching and regression analysis are unlikely to generate statistically significant impact estimates, even if the load reduction is reasonably large in absolute terms (but not relative to the total premise load).

Another method that might be useful in limited situations is engineering analysis. As discussed previously, engineering analysis is much less useful for estimating the impacts associated with most DR programs because impacts are driven much more by consumer behavior than by technology implementation. Even some technology enabled DR programs, such as those using programmable communicating thermostats, have a strong behavioral component since consumers can vary the automated set point and/or override the predetermined setting whenever they wish. For very large loads, there may be situations where the CAISO or utility has direct control over the equipment for emergency purposes, thus eliminating any behavioral influence. Under these circumstances, engineering analysis might produce accurate impact estimates, but these loads are likely to be sub-metered so that impacts can be measured directly.

An example where engineering analysis might be useful would be if a program targeted continuously running pumps and the pumps were remotely controlled during DR events. In this case, one could conduct a survey to gather information on the horsepower associated with each pump and use simple engineering calculations to convert that data into estimates of connected load. DR impacts could then be calculated based on the control strategy that was used for each event. However, this somewhat contrived example may have little practical value as these circumstances are rare.

Another approach is to combine end-use metering with engineering calculations. This approach was employed in the evaluation of SCE's air conditioning cycling program for residential customers, and termed the *Duty Cycle Approach.*[34] The approach is designed to take into account the fact that load cycling impacts vary across program participants by temperature, hour of day, size of the A/C unit, and the share of time the A/C unit is in operation (the duty cycle). The *Duty Cycle Approach* is designed to create a reference value for A/C load by collecting data about:

---

[34] Quantum Consulting Inc. *The Air Conditioner Cycling Summer Discount Program Evaluation Study*. January 2006.

The total connected load for each enrolled participant, and

The share of connected load utilized by hour of day and temperature bin (for non-event days).

The specific load impacts are then calculated by:

- Identifying the average share of connected load utilized during the appropriate temperature and time bins (average duty cycle), and

- Calculating program load impacts by taking into account the average duty cycle, total connected load of each participant, participant cycling selections, and the cycling device failure rate.

Importantly, the approach is able to provide load impact estimates for both ex post and ex ante scenarios as well as information about the uncertainty of those estimates.

Still another approach to impact estimation for event-based programs involves the use of what might be called operational experimentation. By operational experimentation, we mean the selective exercise of a program on a sub-sample of participants with the sole or primary purpose of generating data for impact estimation. This is perhaps best understood with an example constructed once again around an air conditioner cycling program.

Given the typically large number of customers participating in load control programs, there are plenty of customers from which a small sample can be drawn for experimental purposes. One could split this sample into two groups, again using random sampling, and either install an interval meter on the whole house or on the air conditioning unit to obtain the data necessary to determine load impacts. With the metering in place, one could experiment with different load control strategies and event windows across a variety of day types to generate a database that would allow you to estimate impacts under various conditions. The control and treatment groups could be alternated to ensure that there is no correlation between customer characteristics and impacts. Given that this approach provides data on both a control and treatment group on event days, a simple comparison of means on event days would provide a valid estimate of average impacts. However, if ex ante estimates are needed, regression analysis would be required. Operational experimentation would be very cost-effective and straightforward if interval meters were already in place (as they ultimately will be in California), and if incentives are largely fixed (that is, if customer payments are not event-specific). This approach could be quite useful for relatively new DR programs or even for long-standing emergency programs that are not triggered very frequently. In these situations, there may not be sufficient data on event days to estimate impacts using other methods.

4.2.4 Measurement and Verification Activities

Measurement and verification (M&V) refers to data collection, monitoring and analysis activities associated with the calculation of gross energy and peak demand savings from individual customer sites or projects. M&V activities typically focus on measure installation verification, installation quality, manufacturing defects, measure use and operation, equipment maintenance procedures, and in-situ measure efficiency. Such activities can be essential to process evaluations of EE or DR programs and helping to understand ways of improving the program. M&V activities can help determine why estimated impacts might differ from goals or expectations. Put another way, M&V activities aren't often needed to understand what impacts are, but they can be useful for explaining why they are what they are.

There is an extensive literature on M&V protocols and activities in support of EE evaluation. If M&V activities are needed in support of DR impact estimation, evaluators can turn to the following documents to learn more about standard procedures:

- *California Energy Efficiency Evaluation protocols: Technical, Methodological, and Reporting Requirements for Evaluation Professionals*, TecMarket Works. Prepared for the California Public Utilities Commission, April 2006, pp. 49 - 64.

- *The California Evaluation Framework*, TecMarket Works. Prepared for the California Public Utilities Commission, June 2004.

- *The International Performance Measurement and Verification Protocol, Volume 1: Concepts and Options for Determining Savings*, 2002.

As previously discussed, there is probably limited need for M&V activities in support of impact estimation for most DR programs. The prior example discussed in Section 5.2.2.4 where data on motor horsepower is used to develop impact estimates is one case where M&V activities would come into play in support of impact estimation. For technology-based DR programs, such as load control, PCT programs, AutoDR and the like, M&V activities could be useful in helping to understand why impacts are what they are. For example, if impacts for an air conditioning program are not as large as expected, M&V activities could be employed to inspect load control switches to see if faulty installation, equipment deterioration or tampering might explain the result.

5 EX POST EVALUATION fOR NON-EVENT BASED PROGRAMS

This section contains protocols and guidelines for ex post evaluation of non-event based programs. As delineated in Section 2, non-event based programs fall into 3 broad categories:

1. **Non-event based pricing—**This category includes TOU, RTP and related pricing variants that are not based on a called event—that is, they are in place for a season or a year.

Scheduled DR—There are some loads that can be scheduled to be reduced at a regular time period. For example, a group of irrigation customers could be divided into five segments, with each segment agreeing to not irrigate/pump on a different selected weekday.

2. **Permanent load reductions and load shifting—**Permanent load reductions are often associated with energy efficiency activities, but there are some technologies such as demand controllers that can result in permanent load reductions or load shifting. Examples of load shifting technologies include ice storage air conditioning, timers and energy management systems.

The protocols for non-event based programs are similar to those for event-based programs—the primary difference being in the relevant day types and the statistical measures that apply.

Non-event based programs present different analytical challenges compared with event-based programs. For most non-event based programs, the influence of the demand response program is in effect every day of the week (although it may vary across days for some programs). As such, the primary advantages of repeated measures associated with event-based programs are lost, including the ability to use participants as their own control group. For non-event based programs, there is a need for either an external control group, usage data for participants prior to and after joining a program, or both. If an external control group is used, selection bias is a major concern, especially for voluntary programs. In addition, day-matching is not a viable option in most instances since there are no days on which the DR program effect is not present.

Section 5.1 summarizes the protocols for non-event based programs and Section 5.2 discusses technical issues and analysis methods applicable to these program categories.

5.1 Protocols

There are five protocols that apply to non-event based programs. As with event based programs, collectively, these protocols define the time periods, day types, measures of uncertainty, output formats and ancillary data that must be reported when presenting impact estimates for non-event based programs. The day types differ for non-event based programs. The statistical measures protocol associated with day matching methods is typically not relevant for non-event based programs because day matching is not applicable except in rare cases (e.g., scheduled DR). The guidance provided in Section 4 regarding issues that should be addressed whenever regression modeling is used is more relevant to non-event based impact estimation, as regression analysis is more common in this context.

Protocol 8: The mean change in energy use per hour (kWh/hr) for each hour of the day must be estimated for each day type and level of aggregation defined in Protocol 12. The mean change in energy use for the day must also be reported for each day type.

Protocol 9: The mean change in energy use per month and per year must be reported for the average across all participants and for the sum of all participants on a program in each year over which the evaluation is conducted.

Protocol 10: Estimates of the 10th, 50th and 90th percentiles of the change in energy use in each hour, day and year, as described in Protocols 8 and 9, for each day-type and level of aggregation described in Protocol 12, are to be provided.

Protocol 11: Impact estimates shall be reported in the format depicted in Table 4-1 for all required day types, as delineated in Protocol 12.

Protocol 12: The information shown in Table 4-1 must be provided for each of the following day types and levels of aggregation:

- ***For the average across all participants and for the sum of all participants for the average weekday for each month in which the program is in effect***[35]
- ***For the average across all participants and for the sum of all participants for the monthly system peak day for each month in which the program is in effect***

Day type definitions and additional reporting requirements for each day type are summarized below:

---

[35] If a program is seasonal, only the months in which the program is in effect need to be reported.

***Average Week Day for Each Month:* The average across all weekdays in each month during which the program is in effect. In addition to the information contained in Table 4-1, the following information must be provided:**

- ***Average temperature*[36] *for each hour for a typical week day for each month.***

- ***Average degree hours for the typical week day for each month.***

- ***Average number of customers participating in the program each month***

***Monthly System Peak Day for Each Month:* The day with the highest system load in each month. In addition to reporting all of the information shown in Table 4-1, the following information must be provided:**

- ***Temperature for each hour on the system peak day for each month***

- ***Average degree hours on the system peak day for each month.***

- ***Average number of customers participating in the program on the system peak day for each month***

5.2 Guidance and Recommendations

There is a wide range of specific programs that comprise the broad category of non-event programs. The analytical methods that can be used to estimate impacts for these programs are also quite diverse. Impacts associated with TOU and RTP programs are strongly influenced by consumer behavior, suggesting that statistical methods are likely to be best for impact estimation. On the other hand, impacts associated with some of the permanent load shifting options, such as ice storage, might be suitable for engineering methods similar to those used for EE impact estimation. Given that the above options are in effect every day, day matching is not an option, although it might be suitable for estimating impacts associated with scheduled DR, since load on non-scheduled days could be a suitable reference value for load on scheduled days (assuming no free rider ship). Perhaps more so than for event-based programs, impact evaluation for non-event

[36] As noted in Section 4, when reporting temperatures and degree days, it is intended that the temperature be reasonably representative of the population of program participants associated with the impact estimates. If participation in a program is concentrated in a very hot climate zone, for example, reporting population-weighted average temperature across an entire utility service territory may not be very useful if a substantial number of customers are located in cooler climate zones. Some sort of customer or load-weighted average temperature across weather stations close to participant locations would be much more accurate and useful.

based programs must consider a wide range of analytical methods in order to determine which approach is well suited to the specific characteristics of the program being evaluated.

5.2.1 Selection Bias

The primary goal of impact estimation is to develop an unbiased estimate of the change in energy use resulting from a DR program. Impacts can be estimated by comparing energy use before and after participation in a DR program, energy use between participants and non-participants, or both. The primary challenge in impact estimation is ensuring that any observed difference in energy use across time or across consumer groups is attributable to the DR program, not to some other factor—that is, determining a causal relationship between the program and the estimated impact.

One way of ensuring that a causal relationship can be established is through random assignment of treatment and control customers within the context of a controlled experiment. Random assignment helps ensure that any estimated difference in the variable of interest is due to the treatment, not due to any preexisting differences between the treatment and control populations. If participants in an experiment are allowed to self-select into the treatment group, any observed difference between the treatment and control groups could be due to some pre-existing difference between the two groups. A pre-existing difference of this sort will cause selection bias in the estimated impact of the treatment, if measured as the difference between treatment and control customers. Whenever random assignment is not possible, the operating assumption should be that selection bias exists.

Even though ex post estimation of DR program impacts is rarely if ever done in the context of a controlled experiment, the environment of a controlled experiment can be closely approximated if a control group can be selected from among program participants. This requires that data on the variable of interest, in this context, energy use by time period, be available both before and after the DR program influence is in effect. As discussed in Section 4.2.2.5, with event-based programs, it is possible to use participants as their own control group because data is available on days when the program influence is and is not in effect. The primary challenge in this context is controlling for any difference in energy use on event and non-event days that is not due to the program influence.

With non-event based programs, the impact of the DR program is in effect every day. As such, once a consumer joins the program, there are no proxy days that can be used to represent load in the absence of the DR program influence. If preprogram load data is available for each consumer, it may be possible to use this data to develop a reference load, after adjusting for differences in weather or other exogenous factors between the preprogram and program periods. However, until AMI meters are more fully deployed, preprogram data may not be available in most instances. In this case, the only option is to select a suitable control group that can be used to estimate what energy use would have been in the absence of the program influence.

When using an external control group, it is imperative that the control group either has usage characteristics that are quite similar to those of the participant population or that any preexisting differences can be controlled for. For voluntary programs, there are a number of reasons to believe that those who participate might be different from those who do not.

For example, if a TOU rate is revenue neutral compared with energy use for the average consumer in a rate class, customers who use less energy than the average during the peak period relative to the off peak period will see their bills fall even without changing their usage pattern. Preferable load shapes such as these would be associated with consumers who either don't have air conditioning or who have it but typically don't use it during peak periods because, for example, no one is home during that time. These consumers would be considered free riders if they do not change their usage at all, or partial free riders if they adjust their usage in addition to already benefiting by simply going onto the TOU rate. Consumers with this type of preferable load shape may volunteer at a higher rate than those who use more energy during the peak period.

If program participation is driven by the type of selection bias described above, impact estimates based on the difference of loads during the peak period between a control group chosen from the general population of non-participants and a participant group will be comprised of two parts. One part results from any change in behavior that the participant population makes in response to the time varying rate. However, the second part results from any preexisting difference in load shapes between the two groups. In the above example, that preexisting difference would lead to an over estimate of program impacts.

Figure 5-1 helps illustrate how impact estimates can be developed given various scenarios regarding the availability of data for control and participant populations.

Figure 5-1
Impact Estimation Options

	Control Group	Participant Group
Before Participation	C_1	P_1
After Participation	C_2	P_2

In Figure 5-1, P represents the participant population and C represents a control group. Subscripts 1 and 2 represent the time periods before and after program participation, respectively. The ideal situation occurs when usage data is available on participant and control customers for a sufficiently long time period before and after the point at which the group of participants being examined joined the program—that is, during periods 1 and 2 for both groups. In this situation, the program impact can be estimated using the following calculation:

$$\text{Impact} = (P_2 - P_1) - (C_2 - C_1) \qquad (5\text{-}1)$$

That is, the impact equals the difference in energy use in the two time periods for the participant group, adjusted for any difference in energy use between the two time periods for the control group. The second term adjusts for differences in energy use due to exogenous factors, such as weather, economic activity, and the like.

Equation 5-1 above can be rewritten as follows:

$$\text{Impact} = (P_2 - C_2) - (P_1 - C_1) \qquad (5\text{-}2)$$

In this form, the equation can be interpreted as estimating impacts based on the difference in energy use for the participant and control group samples during the program period and the difference between the two groups in the preprogram period. The second term adjusts for any preexisting differences in load shapes between the participant and non-participant population.

If no preprogram period data exists (e.g., there is no data in period 1), impacts could be estimated as $(P_2 - C_2)$, but this estimate will be biased unless $(P_1 - C_1)$ equals 0—that is, unless energy use for the control group is a very good proxy for energy use by the participant population prior to participation. In the future, when AMI is widely deployed, there will be a high likelihood of having pre-program load data in most instances. Today, however, that likelihood is quite low. Without pre-program data, selecting a well matched control group or otherwise controlling for differences between the control and participant population, is essential.

There are a variety of strategies for choosing a good control group or otherwise controlling for relevant differences between the control and participant populations. One approach is to pick a group from the general population that has observable characteristics that match the participant population. For instance, in the above example where it is likely that participants have lower air conditioning saturations than the population as a whole, one could select a control group with the same air conditioning saturation and the same dispersion across climate zones as the participant population. With sufficient survey data on both participants and non-participants, stratification on other characteristics (e.g., pool ownership, size of house, income, etc.) could also be used to decrease the likelihood of any load shape bias influencing the impact estimates.

An alternative to the control matching procedure described above, but one that is conceptually similar, involves incorporating variables representing observable characteristics for the participant and control groups into the impact estimation procedure, and then adjusting the impact estimates to reflect the participant population characteristics. For example, one could estimate a regression model using participant and control group data that would correlate household load (or share of daily energy use) during the peak period with air conditioning ownership. Given this relationship, one could use the saturation of air conditioning for the participant population to produce an unbiased estimate of load and load impact, assuming the difference in air conditioning saturation is the primary determinant of differences in load between the participant and control group, aside from the influence of the DR program itself.

A third approach to addressing selection bias involves developing a two stage model, where the first stage estimates the probability of participating in a DR program, which then becomes a variable that is included in the impact estimation model. A useful discussion of various approaches to modeling and adjusting for self selection in the context of EE evaluation is contained in the *California Evaluation Framework*. [37]

Still another approach to addressing selection bias is to figure out a way of creating a control group from the existing participant population or from future participants. For example, it might be possible to select a sample of current program participants and offer them an incentive to become control customers, thus no longer having them respond to the DR program prices or incentives for some period of time. An alternative would be to focus on future volunteers, asking them to delay their transition onto the program so that they can be used as a control. This approach is comparable to randomly assigning volunteers in an experiment to control and treatment groups once they agree to volunteer.

The problem of selection bias discussed above is equally important for both voluntary opt in and opt out programs. Given the typically high degree of customer inertia, especially among mass market consumers, consumers who opt out of a default tariff or program may be as different from those who stay as are those who volunteer to participate in a program on an opt in basis.

For a mandatory DR program, selection bias is not an issue. Of course, there is also no possibility of selecting a control group from among non-participants. As such, if pre-participation data does not exist, there may be no alternative but to select a sample of consumers and create a control group by removing the influence of the program for some time period.

The approaches outlined above are focused on ensuring the internal validity of the impact estimates. In this context, internal validity refers to establishing a causal relationship between the DR program and the change in energy use for the current program participants. Knowing whether the estimated impacts are also valid for potential future participants is typically also of interest. This is known as external validity. Issues associated with external validity will be discussed in Section 6, as it is a key issue for ex ante estimation.

[37] pp. 142-145.

5.2.2 Demand Modeling

For price-driven DR programs, if there is sufficient variation in prices across time or across consumers, it may be possible to estimate an energy demand model and use the model to estimate program impacts for the day types of interest. A demand model quantifies the relationship between energy demand and price. As prices increase, the amount of energy used decreases and vice versa. Because energy use varies with other factors, such as weather and end use appliance holdings, variables representing these factors are typically also included in the demand equation.

If there is sufficient variation in price to estimate a demand model, the impact of a price-driven DR program can be estimated by predicting energy demand based on the DR program price and what price was prior to a consumer going onto the program. The following equation represents a simple demand model.

$$\ln(E_i) = \beta_0 + \beta_1 \ln(T_i) + \beta_2 \ln(P_i) + \varepsilon \qquad (5\text{-}3)$$

where E_i = energy use in hour i

T_i = the temperature in hour i

P_i = the price in hour i

ε = the regression error term

β_0 = a constant term

β_1 = the change in load given a change in temperature

β_2 = the change in log of energy use given a change in the ln of price

ln = the natural logarithm.

One can use equation 5-3 to estimate energy demand for two price levels, one representing the program price during hour i and the other representing the price in that same hour prior to program participation. The difference between energy use at these two prices is an estimate of the impact of the DR program.

The double-log specification depicted in equation 5-3 is commonly used in empirical estimation of demand models. It is convenient in that the coefficient on the price term, β_2, represents the price elasticity of demand, which equals the percentage change in energy use given a percentage change in price.

Demand modeling works best when there are multiple prices that can be used to estimate the demand function. Real time pricing is an ideal candidate for demand modeling, as prices change hour to hour and day to day. As such, the demand equations can be estimated without using an external control group, thus eliminating the possibility of selection bias due to a mismatch between control and participant populations.

It may also be possible to estimate a demand model for a TOU tariff using only the participant population. This approach will have a higher probability of success if pretreatment data is available and if there is seasonal fluctuation in prices. However, taking advantage of the seasonal fluctuation in prices would require normalizing for variation in weather, seasonal fluctuations in business operations, and other factors. Any omitted variables or misspecification in this regard could easily bias the price parameters and the resulting impact estimates.

5.2.3 Engineering Analysis

Engineering analysis is another approach that might be suitable for some programs such as permanent load shifting programs that are largely technology driven and have much more limited behavioral variation than do pricing programs, for example. Engineering analysis might be useful to estimate impacts for permanent load shifting options such as ice storage technologies or energy management systems.[38]

Engineering methods use basic rules of physics to calculate estimates of energy and demand savings. The technical information required as inputs to engineering models generally come from manufacturers, research studies, and other general references combined with assumed or measured equipment operating characteristics.

In order to estimate savings via engineering methods, one must establish a baseline or reference value from which to compare the energy consumption and demand of facilities included in the evaluation. The program baseline may require a definition of the pre-program equipment or building characteristics and operations, as well as an estimate or measurement of pre-program energy consumption. The program baseline may consist of the following:

- For early equipment replacement (retrofit) programs, the pre-existing and still-functioning equipment replaced during program participation defines the baseline. Pre-program energy consumption may be adjusted to reflect changes in equipment or building operations not related to the program.

- For equipment that is being replaced at the end of its useful life (i.e., in all situations where the customer would have been replacing the equipment in the absence of the program), standard-efficiency new equipment defines the baseline. The program's purpose in these cases is to induce customers to do the replacement with a higher-efficiency alternative than they would have selected in the absence of the program.

Engineering methods can be divided into two basic categories.

---

[38] The remainder of this discussion consists mainly of selected text from *The California Evaluation Framework*, pp. 120 – 129.

- Simple Engineering Model
- Building Energy Simulation Model

Simple engineering models and algorithms are typically straightforward equations for calculating energy and demand impacts of non-weather dependent energy efficiency measures, such as energy efficient lighting, appliances, motors, cooking equipment, etc. Simple engineering models are generally not used for weather dependent measures such as building envelope and HVAC measures; these measures are generally analyzed using building energy simulation models.

Building energy simulation models are computer programs that use mathematical representations of important energy and control processes in an attempt to realistically simulate the thermal and energy systems in a building. Energy calculations are carried out on an hourly or sub-hourly basis for a selected time period or more commonly for an entire year based on typical weather data for the selected building site. The programs are made up of a collection of mathematical models of building components, such as windows, wall sections, and HVAC equipment. The individual component models are linked together to form a complete building simulation. The results predict the performance of the building structure and energy systems under given weather conditions at a selected geographic location.

All building energy simulation programs have limitations that must be well-understood before applying the program to a particular energy estimation problem.[39] For example, most programs are limited to the simulation of common HVAC system types with a predetermined system configuration. Considerable latitude is given to the user with respect to describing system performance parameters, but the basic arrangement of the system component is fixed and defined by common practice in the building design and HVAC industries. This does not present a problem for most buildings and systems, but for complex custom HVAC configurations, the judgment and experience of the user is critical.

It also can be useful to calibrate the simulation models against metered or sub-metered energy usage information in order to ensure that the models are performing well against some empirically based data from the local population.

5.2.4 Day Matching For Scheduled DR

Although day matching is generally not suitable for non-event based programs, one possible exception to this rule may be scheduled DR programs. Scheduled DR programs prearrange with customers that have flexible loads to limit use of certain equipment on

[39] For more information on building energy simulation programs, see *State-of-the-Art Review: Whole Building, Building Envelope and HVAC Component and System Simulation and Design Tools.* (Jacobs and Henderson 2002).

regularly scheduled days. For example, an agricultural customer that does a lot of irrigation pumping might willing to only irrigate on selected days.

With this type of program, it may be possible to use load from non-scheduled days as a reference value for what a customer might have used in the absence of the program on the day that they have agreed not to use electricity. However, there could also be problems with this approach if, for example, the customer uses more electricity on non-scheduled days than they otherwise would have in the absence of the program. In this situation, using other days would overstate the magnitude of the reduction on the scheduled day. Other types of free rider ship might also be present. For example, if a customer agrees not to irrigate or otherwise use load on a day when they typically don't use electricity for those purposes, they would simply be getting paid for doing nothing. Thus, while day matching might work in theory for scheduled DR, it should be used with caution. There may be no real substitute for having pre-participation data on a customer in this situation to determine a suitable baseline.

6 EX ANTE ESTIMATION

This section contains protocols and guidelines for ex ante estimation of both event and non-event based programs. It is useful to distinguish between three types of ex ante estimation:

1. **Existing Program and Participants, Different Conditions:** The objective of this type of ex ante estimation is to determine what impacts would be given weather conditions or event characteristics that differ from what occurred in the past. Examples include determining the impacts for a specific set of weather conditions, such as a 1-in-10 weather year, or specific event characteristics, such as a specific set of prices for an RTP program, or a load control event of specific length on a specific weather day.

2. **Existing Program, Different Participants:** The objective of this type of ex ante estimation is to determine what the average and total impacts would be given growth in program participation and/or a change in the mix of participants over time that might influence the average (and total) impacts. For example, this type of ex ante estimation would forecast what the impacts would be as a program grows in size from say, 100,000 to 200,000 participants, or as the mix of participants changes from primarily large C&I customers to a greater concentration of medium and small C&I customers.

3. **New Program:** The objective of this type of estimation is to develop reasonable forecasts for programs that have no ex post history upon which to base ex ante forecasts. The previous two types of ex ante estimation should take as their starting point ex post impact estimates. With new programs, this is not possible, at least not using ex post estimates from the utility program sponsor. As seen below, it may be possible to borrow ex post impact estimates from other utilities if there are sufficient similarities between the program and population characteristics.

Ex ante estimation involves determining what the impacts are likely to be for a given set of user-defined conditions. It does not include defining what those conditions are. For example, forecasting the size or make up of the participant population at some future point in time is not part of impact estimation. Rather, impact estimation concerns estimating demand response given assumptions about the size and make up of the participant population that are provided to the evaluator by someone else (e.g., regulators, planners or some other stakeholder).

Having said that, the evaluator has an important role in guiding the development of data needed to make such estimates, in that he or she must tell the interested user what information is needed. For example, for a residential critical peak pricing program, it would be important that the evaluator tell the prospective user that air conditioning ownership is a key driver of demand response. As such, it will be necessary for the prospective user to indicate not only that they expect participation in the program to grow from X to Y over the next five years, but also that the percent of participants who own central air conditioning is expected to change from A to B over the same period. With this information, the evaluator can predict how the average impact per customer will change as the air conditioning saturation changes and how total impacts will grow as the number of participants increases.

Ex ante estimation requires development of a model that relates changes in demand response to changes in the exogenous variables that drive demand response. Whenever possible, the model should be based on ex post analysis from existing programs. As such, all of the issues associated with ex post evaluation that have been raised in prior sections apply here as well.

However, there are additional issues that are unique to ex ante estimation. Ex ante estimation requires quantification not only of the uncertainty associated with ex post impact estimates (e.g., due to sample selection, model specification and the like), but also the additional uncertainty associated with the exogenous factors that drive demand response (e.g., uncertainty in weather, participation levels and customer characteristics, etc.). Ex ante estimation also requires determining how demand response might evolve over time as participants become better educated about how to modify behavior in response to demand response stimuli or, alternatively, loose interest in modifying their behavior. The persistence of demand response impacts over time may also be impacted by degradation of or improvement in enabling technology, which may also need to be factored into ex ante estimates. Finally, ex ante estimation must be concerned with difficulties in estimation outside the boundaries of historical experience (e.g., for extremely hot days that might not have occurred over the historical period) where the relationship of demand response and the variable of interest may differ from the relationship that exists within a narrower range of values.

The remainder of this section is organized as follows. Section 6.1 contains the protocols that apply to all ex ante estimation. Section 6.2 discusses technical issues and methods that can be used to address the issues for all ex ante scenarios. Section 6.3 discusses persistence, a key issue associated with ex ante estimation, and Section 6.4 discusses Monte Carlo simulation modeling, which can be used to incorporate into the impact estimates uncertainty associated with the key drivers of demand response.

6.1 Protocols for Ex Ante Estimation

The protocols contained in this section are intended to apply to all three types of ex ante estimation, including estimation for brand new programs. It is expected that, in the vast majority of situations, ex ante estimation for programs that are not new will be based at least in part on ex post evaluation studies involving analysis of historical data. As such, the output requirements and protocols that apply to ex post evaluation should be able to be met for ex ante estimates developed from theses studies, although there are some differences associated with the standard day types and forecast horizon and with factoring in changes in exogenous variables. Meeting the same protocols for brand new programs may be more difficult, as the amount of available data and the statistical rigor that can be applied may be less for new programs than for existing ones. This is not always true, however, as illustrated by the example presented below in Section 6-3. Information on the probability distributions associated with key drivers of demand response, or reasonable assumptions concerning the minimum, maximum and most likely estimates associated with key drivers, can be used along with Monte Carlo simulation modeling to develop uncertainty adjusted impact estimates even for new programs. As such, the same protocols apply for new programs, although it is recognized that even a "best efforts" level of commitment to meeting these requirements may fall short depending upon the nature of the new program and the degree to which data and/or models can be obtained elsewhere.

> ***Protocol 15: Whenever possible, ex ante estimates of DR impacts should be based on ex post empirical evidence from existing or prior programs. Evidence from programs and customer segments most relevant to the ex ante conditions being modeled should be used, regardless of whether they come from the host utility or some other utility. If ex post estimates or models are not used as the basis for ex ante estimation, an explanation as to why this is the case must be provided.***

> ***Protocol 16: The mean change in energy use per hour (kWh/hr) for each hour of the day must be estimated for each day type and level of aggregation defined in Protocol 20. The mean change in energy use for the day must also be estimated for each day type.***

> ***Protocol 17: The mean change in energy use per month must be estimated for non-event based programs and the mean change in energy use per year must be estimated for both event and non-event based programs for the average across all participants and for sum of all participants on a program for each year over the forecast horizon.***

> ***Protocol 18: Estimates of the 10^th^, 50^th^ and 90^th^ percentiles of the change in energy use in each hour, day and year, as described in Protocols 16 and 17, and for each day-type described in Protocol 20, are to be provided.***

> ***Protocol 19: Impact estimates shall be reported in the format depicted in Table 6-1 for all required day types and levels of aggregation, as delineated in Protocol 20.***

Table 6-1
Reporting Format for Ex Ante Estimation

Hour Ending	Reference Load	Load Impact (kWh/hr)	Temperature (degrees F)	Uncertainty Adjusted Energy Impacts		
				10th Percentile	50th Percentile	90th Percentile
1						
2						
3						
4						
5						
6						
7						
8						
9						
10						
11						
12						
13						
14						
15						
16						
17						
18						
19						
20						
21						
22						
23						
24						
				Uncertainty Adjusted Energy Impacts		
	Reference Energy Use	Change in Energy Use	Degree Hours (Base 75)	10th Percentile	50th Percentile	90th Percentile
Day						

It should be noted that there is a difference between Table 4-1, which applies to ex post estimation, and Table 6-1. Table 4-1 contains a column representing the observed load whereas Table 6-1 does not. Obviously, it is not possible to measure observed load in the future. The reference load column is included so that percent impacts can be calculated. Once again, temperature and degree hours are included primarily for comparison purposes across day types and programs. These variables may or may not have been those used in developing the estimates.

Protocol 20: The information shown in Table 6-1 must be provided for each of the following day types and levels of aggregation for each forecast year:

- ***For the average across participants and for the sum of all participants for a typical event day for a 1-in-2 weather year for event-based programs.***

- ***For the average across all participants and for the sum of all participants for non-event based programs for the average weekday for each month in which the program is in effect for a 1-in-2 weather year***[40]

- ***For the average across all participants and for the sum of all participants for the monthly system peak day for each month in which the program is in effect, for a 1-in-2 weather year, for non-event based programs.***

Day type definitions and additional reporting requirements for each day type are summarized below and are intended to be included in this protocol.

Typical Event Day for a 1-in-2 Weather Year: This day type requirement applies only to event-based programs. It is meant to capture both the exogenous factors such as weather and the event characteristics for a day on which an event is likely to be called. The relevant characteristics can be defined by the evaluator. At a minimum, the following information must be provided:

- ***An explanation of how the weather and any other relevant day-type characteristics were chosen***

- ***Detailed information on the timing and duration of the event or any other factors (e.g., notification lead time) that were explicitly factored into the impact estimates (e.g., factors that, if different than those reported, would change the estimated impacts)***

- ***The number of notified consumers included in the aggregate impact estimate***

- ***Any other factors that have been explicitly incorporated into the impact estimate, such as prices for price based programs and population characteristics (e.g., air conditioning saturation, business type, etc.).***

Average Week Day for Each Month In A 1-in-2 Weather Year: This day type applies only to non-event based programs. It is meant to capture the weather conditions and other relevant factors for an average weekday. In addition to the information contained in Table 6-1, the following information must be provided:

- ***An explanation of how the weather and any other relevant day-type characteristics were chosen for the typical weekday in each month***

- ***The number of enrolled consumers included in the aggregate impact estimate***

---

[40] If a program is seasonal, only the months in which the program is in effect must be reported.

- ***Any other factors that have been explicitly incorporated into the impact estimate, such as prices for price based programs and population characteristics (e.g., air conditioning saturation, business type, etc.).***

Monthly System Peak Day For Each Month In a 1-in-2 Weather Year: This day type applies to non-event based programs. It is meant to capture impacts for the day with the highest system load in each month. In addition to reporting all of the information shown in Table 6-1, the following information must be provided:

- ***An explanation of how the weather and any other relevant day-type characteristics were chosen for the typical monthly system peak day***
- ***The number of enrolled consumers included in the aggregate impact estimate***
- ***Any other factors that have been explicitly incorporated into the impact estimate, such as prices for price based programs and population characteristics (e.g., air conditioning saturation, business type, etc.).***

Protocol 21: All ex ante estimates based on regression methodologies shall report the same statistical measures as delineated in Protocols 7 and 14.

It should be noted that the day types described above, and that are incorporated in Protocol 20, are intended to be the minimum set of day types that are required, in part, to allow for comparisons across programs and to support long term resource planning. Additional day types may be of interest to many users. For example, impacts based on weather for a 1-in-10 year or 1-in-10 event day may be a relatively common need.

6.2 Guidance and Recommendations

As described in the beginning of this section, ex ante estimation concerns extrapolating the findings from ex post evaluations (of either the same program or one similar enough so that logical inferences can be drawn) to a set of conditions that differ from those that have occurred in the past. The issues that must be addressed vary depending upon the conditions of interest and how much these conditions differ from those that have occurred in the past. There are four scenarios across which issues and methods vary. There could also be scenarios of interest that combine elements from each of these scenarios.

Scenario 1: The most straightforward scenario is when estimates are needed for a set of conditions that are within the range of those that have occurred in the past. An example would be development of an estimate for a program where the mix of customers is assumed to remain largely the same in the future as it was in the past and the weather conditions of interest, while not exactly the same as any particular day that occurred in the past, can be represented by temperatures that are below the maximum and above the minimum temperatures that occurred during the ex post evaluation period. Another example would be for an RTP pricing program where estimates are needed for a set of prices that are inside the range of prices experienced previously. These examples merely require interpolation between prior extremes. Developing these estimates still requires a model that relates the variables of interest to demand response, but there is every reason to believe that the inferences drawn will be valid.

Scenario 2: A second scenario of potential interest might, once again, involve little change in the participant population. However, in this case, there is interest in knowing what the impacts might be for a day type where the weather or price conditions (or some other variable of interest) are outside the range that has been observed in the past. For example, one might want to know what the impacts would be for a 1-in-10 weather year or weather day, or for highly volatile market conditions where hourly prices exceed any that had previously occurred. These examples are much more challenging, as the functional relationship between the variable of interest and demand response may differ under these extreme conditions from what it was under the observed conditions.

For example, the relationship between the change in energy use associated with air conditioning and a change in temperature is reasonably linear over some range of temperatures, but highly non-linear at both the low and high end of the temperature range. A change in temperature from, say, 65 to 70 degrees will produce very little if any change in energy use because air conditioning typically is not running at either of those temperatures. Similarly, a change in temperature from, say 100 to 105 degrees, may produce little change in air conditioning energy use because most air conditioners are already running flat out at 100 degrees, so higher temperatures don't increase energy use. For the same reasons, demand response may not occur at these extremes, regardless of the magnitude of program stimulus, since thermostat adjustments at these extremes will have little impact on energy use. Consequently, if the model being used for ex ante estimation was developed from data on days that did not include these extreme conditions or, even if such conditions existed, the model assumed a linear relationship across the entire temperature range (e.g., it was misspecified), it will not do a good job of estimating demand response impacts under these extreme conditions.

The same type of problem can arise when using demand models to estimate impacts for prices well outside the range of what has been observed historically. It may be, for example, that customers are not very price responsive at the very low end of the price range, when a change has only a small impact on their bills, or at the very high end of the price range, where they have already made all of the reductions that they are willing or able to make. In between these extremes, customers may be relatively price responsive. Recent evidence from a pricing experiment in New South Wales, Australia, for example, suggests that there is very little incremental effect associated with a change in prices when moving from a peak period price of $1.50/kWh to a price of $2.00/kWh. Implicitly, this evidence suggests that consumers have already made all of the adjustments they are willing to make at the $1.50/kWh price.

Scenario 3: A third scenario concerns estimating the change in demand response associated with a change in participant characteristics that are observable. This could occur, for example, for a demand response program that is targeted at customers with air conditioning but open to all customers. Suppose that the initial marketing effort for this program was quite effective at attracting customers from the target population, perhaps because it was initially only advertised in areas where the saturation of air conditioning was high. However, over time, through word of mouth or because of program expansion into other geographic regions where the saturation of air conditioning is lower, the saturation of air conditioning among participants might decrease. If demand response is tied to air conditioning ownership, this type of shift in the participant population will lead to an overall decrease in average demand response per participant, even as total demand response increases with increased participation.

Producing estimates for this type of scenario requires developing a model that relates the change in demand response to a change in the observable variables that are expected to differ over the forecast horizon. In some instances, this will be relatively straightforward, while, in others, it may be more difficult. In the above example, if the early targeted marketing is so successful that the only customers currently enrolled in the program are those with air conditioning, it will not be possible to establish a relationship between air conditioning ownership and demand response from the historical program data. Under these circumstances, it may be necessary to use information from other utilities with similar programs but a more diverse mix of participants in order to adjust the impact estimates based on current participants so they reflect the future penetration of participants who do not have air conditioning.

Scenario 4: The fourth scenario is the most difficult one of all, as it involves developing estimates when there are reasons to believe that future participants will differ from those in the past in ways that are not easily tied to observable variables. This could be a reasonable expectation for any program that is in the early stages of its lifecycle, as it may have only attracted "early adopters" who may not be terribly representative of the general population. Extrapolation to future participants may be even more challenging in a situation where a program is changing from a voluntary, opt-in marketing approach to a voluntary, opt-out approach or even mandatory participation. If customers who volunteer for a program are more environmentally conscious, more price sensitive, or have life styles or business operations for which any negative aspects of demand response are less impactful then it is for the average customer, for example, extrapolating impacts derived from this group of participants to a much broader population in which those difficult to observe characteristics are much less prevalent will lead to an overestimate of demand response impacts.

6.2.1 Impact Estimation Methods

As indicated above, developing impact estimates for a specific set of weather conditions or population characteristics requires estimation of a model that will predict how demand response impacts change given a change in these conditions. In a regression context, this can be achieved using a model specification that includes interaction terms between exogenous variables of interest and program variables. Equation 6-1 is an example of this type of model. This specification is one of several that were developed to estimate hourly impacts for residential critical peak prices tested in California's Statewide Pricing Pilot.[41]

$$\left[\left(\frac{kWh_i}{\sum_{i=1}^{24} kWh_i}\right)\right] = a_i + b_i\left(\frac{P_p}{P_{op}}\right) + c_i\left[\left(\frac{CDH_i}{\sum_{i=1}^{24} CDH_i}\right)\right] + d_i\left(\frac{P_p}{P_{op}}\right) * CAC + e_i\left(\frac{P_p}{P_{op}}\right) * \left[\left(\frac{CDH_i}{\sum_{i=1}^{24} CDH_i}\right)\right] + \sum_{k=1}^{n} g_k D_k + \varepsilon_i$$

$P_p =$ Peak period price

$P_{op} =$ Offpeak period price

$CDH_i =$ Cooling degree hours for hour i

$D_k =$ Customer k's fixed effect

$\varepsilon_i =$ Error in hour i

CAC = 1 if a household owns a central air conditioner, 0 otherwise.

The above equation estimates the share of daily energy use in each hour as a function of the share of daily cooling degree hours in each hour, the peak-to-off-peak price ratio, air conditioning ownership and binary variables representing each customer (in order to control for cross-sectional differences in energy use). The model coefficients d_i and e_i, respectively, represent the change in price responsiveness given a change in air conditioning ownership and weather. This type of model can be used to produce ex ante impact estimates for any combination of weather conditions and prices that are not too far outside the boundaries of what occurred within the estimating sample, and for a mix of program participants that have any saturation of air conditioning ownership that the evaluator might think is likely to occur over the forecast horizon. Models such as this can be used to produce ex ante forecasts for Scenarios 1 and 3 outlined in the previous section.

This type of model can also be used for ex ante estimation for Scenario 2, estimation outside the boundary of historical experience, but only under certain circumstances. Assuming that data exists for a reasonably wide range of variation in the variables of interest, the first step in model estimation should involve an exploration of different functional forms to assess whether a linear or non-linear relationship fits the data best. If non-linearities are present within the estimating sample and can be captured in the functional form that is fit to the historical data, the model should do a better job of estimating impacts based on input variables that have values outside the historical boundary than if only a linear relationship is exhibited within the estimating sample (and assuming that there are logical reasons to believe that non-linearities exist at the extremes of the distribution even though they aren't detectable from the historical data). The

[41] CRA International. *Residential Hourly Load Response to Critical Peak Pricing in the Statewide Pricing Pilot.* May 18, 2006.

previous examples concerning air conditioning energy use at very low and high temperatures and incremental demand response at very high prices are cases in point. In these examples, logic and/or experience from elsewhere suggest that, at some point, impacts will not change given any incremental change in the exogenous variables.

Another approach to addressing this problem is to incorporate information from other programs or utilities. For example, the highest critical peak price tested in California's SPP was roughly $0.75/kWh for residential customers, which was roughly 5 times the standard price. If there was interest in knowing what the impacts would be for a price closer to $1.25/kWh or even higher, there is a risk that the price elasticities from the SPP would not apply. In this case, one could turn to other pricing experiments, such as the NSW pilot mentioned above, to see if much higher prices and/or price ratios were tested. If they were and the estimated price elasticities were comparable to those found in the SPP, there will be greater confidence in using the SPP model to produce estimates for prices outside the boundary of those tested in the pilot than if a different result were observed elsewhere.

A third approach to addressing Scenario 2 is experimentation. If a program is expected to be large and it is important to understand what happens at the extremes, it may be necessary to plan and conduct an experiment that creates the conditions of interest. For programs such as load control, where an event might only be triggered during system emergencies and such emergencies often, but not always, occur on very hot days, it could be useful to trigger the program when a hot day occurs but an emergency doesn't actually exist. Similarly, for an RTP program, if there is interest in knowing what might happen if prices go really high for a few hours, but such prices have never occurred, it might be possible to get permission to test a very high price signal on a sample of customers under market conditions where prices are typically high, just not as high as they might become at some point down the line.

As previously mentioned, the most difficult challenge occurs in Scenario 4, when a program is expected to undergo a very significant transition from a small group of early volunteers to a much broader level of participation. This might occur due to normal growth over the forecast horizon or a significant shift in marketing approach from a voluntary, opt-in program, for example, to an opt-out or even mandatory program. In these circumstances, the past program history may not be a good guide to the future.

One approach to addressing this scenario is to explore whether or not it is possible to learn enough about the current participants to ascertain how they might differ from potential future participants—that is, to try and turn currently unobservable characteristics into observable characteristics. For example, if one is concerned that early adopters are more environmentally conscious or more budget minded than what future participants would be, it might be possible to conduct a survey to explore whether or not the hypothesis is true. If it is not true, there will be greater confidence in extending the historical findings to future participants. If it is true, the survey data won't necessarily help you solve the problem, but at least it will confirm that you have one.

Another approach is, once again, to look elsewhere for data and information that can be used to gauge whether or not it is proper to extrapolate from the current population and program characteristics to a different set of conditions. It may be that some other utility has a program with the characteristics of interest that can provide guidance into what impacts are likely to be. For example, if there was interest in knowing whether the impact estimates based on large C&I customers participating in a voluntary RTP program are suitable for estimating impacts given a shift to a mandatory RTP tariff, one could examine estimates based on New York's mandatory RTP tariff for large C&I customers and see how they compare to estimates from the current voluntary program. If they are similar, after controlling for differences due to customer mix and price variation, there will be greater confidence in using the current estimates than if they are quite different. There is a growing body of evidence from demand response programs across the country that can and should be used whenever ex ante estimation must be done for a program that is expected to differ significantly from what has occurred in the past.

Another approach to this scenario involves experimentation. This is almost always an option, albeit a potentially expensive and time consuming one, for developing impact estimates where history or information from elsewhere is not a sufficient guide to what might happen over the forecast horizon. In the example discussed above where survey data revealed a difference between current and future participants on attitudes about the environment or cost consciousness, it is likely that demand response impacts will differ for future participants from those estimated from the current participant population. However, it might be impossible to know how impacts are likely to change because the current participant population might not have any people who aren't either environmentally or budget conscious. In this situation, it could be fruitful to conduct a small experiment in which the population of interest is recruited using some form of incentive to secure their participation and assess whether or not there is any difference in demand response between current participants and likely future participants.

6.3 Ex Ante Estimation for New Programs

Ex ante estimation for brand new programs is similar to the situation just discussed for an existing program where the past may not be a very useful guide to the future. The two primary approaches to addressing the problems outlined above, relying on estimates from elsewhere and experimentation, apply here as well. California's SPP is an example of an experimental approach that developed the data necessary for the State's utilities to estimate likely impacts for critical peak pricing programs for residential and small and medium C&I customers in California that did not previously exist. Pilot programs and experiments are important methods to consider when developing ex ante estimates for new programs.

Whenever it is necessary to rely on information from programs or experiments conducted elsewhere, it is important to explore ways of adapting these estimates for differences in the program and population characteristics between the utility from which the information is obtained and the utility for which it is being used. In some situations, the available information may not be robust enough to allow for adjustments to be made, or even to obtain a thorough understanding of whether or not there are differences. Whenever differences are relevant and evident, however, they should be documented, even if they can't be adjusted for. The ideal situation occurs when it is possible to borrow a model from another jurisdiction that allows adjustments to be made.

An example of how data from another utility and a different program type can be used to produce ex ante estimates for a brand new program is documented in testimony filed in conjunction with SDG&E's AMI application.[42] SDG&E required demand response impact estimates for a brand new program, titled the Peak Time Rebate (PTR) program. The proposed PTR program would pay participants a rebate, delineated in cents/kWh, to reduce energy use on critical days. Reductions would be measured based on the difference in energy use during the peak period on critical days and a reference load value based on a day-matching method.

SDG&E had not tested this program concept but another utility, Anaheim Public Utilities (APU), had completed a pilot program that was quite similar to the program of interest. The APU pilot paid an incentive equal to $0.35/kWh for all energy reduced during the peak period on critical peak days during the summer of 2005. For the purpose of determining the incentive payment amount, reductions were calculated relative to a reference value equal to energy use during the peak period on the three highest, non-critical days during the summer period for each customer. The incentive was paid as a bill credit at the end of the summer.

The peak period in the APU program was from noon to 6 pm and there were 12 events called during the summer period, which ran from June 1st through October 31st. Approximately 120 customers participated in the pilot. Approximately 71 treatment customers and 52 control customers participated in the pilot.

There were three concerns about using the APU impact estimates for predicting demand response for SDG&E's proposed program. The primary one concerned a difference in the rebate amounts. APU paid a rebate equal to $0.35/kWh whereas SDG&E was considering a rebate of $0.65/kWh or even higher. Since APU only tested a single price, it was not possible to estimate price elasticities using the APU data that could be used to estimate impacts for the very different incentive value of interest to SDG&E. Another issue concerned differences in the saturation of air conditioning between APU and SDG&E customers. The SPP indicated that demand response varied significantly with air conditioning saturation, so it was important to adjust any estimates based on the APU pilot for differences in the saturation of air conditioning for the APU and SDG&E customer populations. Still another factor included differences in climate.

A key issue in approaching this analysis concerned whether the demand models estimated from the SPP could be used to predict impacts associated with the implicit price signals

---

[42] Amended testimony of Dr. Stephen S. George, Application of San Diego Gas & Electric Company (U-902-E) for Adoption of An Advanced Metering Infrastructure Deployment Scenario and Associated Cost Recovery and Rate Design. Chapter 6, Demand Response Benefits. July 14, 2006, Revised September 19, 2006.

inherent in the PTR rebate. The SPP price elasticities were developed from information on a CPP tariff, which is a traditional "carrot-and-stick" price signal in which customers can reduce their bills if they respond to the higher critical peak prices but will pay more than they otherwise would have on the standard rate if they did not adjust their energy use. The PTR program, on the other hand, represents a "carrot-only" incentive in which customers can reduce their bills if they reduce energy use during the critical peak period, but their bills would not change compared to what they would be under a standard rate if they did not. On the other hand, the marginal price signal, that is, what customers would save for each kWh of avoided energy use, is the same in both instances. If it could be established that customers are likely to respond in the same manner to both types of price signals, then the SPP model could be used to predict impacts for the PTR program based on SDG&E's price incentive and population characteristics.

In order to assess whether the SPP model could be used, SDG&E acquired data from APU and determined that the average reduction across the 12 critical days for the average customer equaled 11.9 percent. Next, the demand models from the SPP were used to predict what the reduction would be using the APU implicit price signal and air conditioning saturation and weather statistics representing the APU population. The SPP model predicted a reduction of 11.4 percent. Given the high degree of similarity between the estimated impact from the APU pilot and the predicted impact using the SPP models, SDG&E concluded that customers are likely to respond quite similarly to the two different types of price signals and felt comfortable using the SPP models to predict what impacts would be for their proposed PTR program. This is a good example of how information based on a similar program from another utility, combined with a model based on a different kind of tariff, can be used to predict impacts for a brand new program.

6.4 Impact Persistence

Impact persistence refers to the period of time over which program-induced impacts are expected to last. There are two key questions that influence how estimation of impact persistence might be approached:

- Do impacts persist beyond the life of the program?
- Do average impacts per customer change over time due to changes in consumer behavior and/or technology degradation?

For most demand response programs, the answer to the first question is no. In most instances, demand response can only be expected to occur for as long as the incentives being paid to induce response continue. For example, for customers that are on time-varying rates or interruptible rates, or for customers who are paid an incentive to participate in a load control program, if the tariff or incentive stops, impacts will also stop even if some technology was installed to enable the impacts to occur in the first place (e.g., like a load control switch).

An exception to this rule might be for some permanent load reduction programs, such as ice storage. If a utility program is implemented that subsidizes ice storage, for example, the overall load shifting associated with the technology will probably persist as long as the technology remains operational, which could extend well beyond the termination of the program. Estimating persistence in this case requires estimating the effective useful life of the technology. Persistence may not extend beyond the program life for all permanent demand response programs, however. For example, demand response associated with energy management systems or time switches may dissipate once a program is terminated, as consumers might disable the time switch or adjust their energy management system so they can operate end use equipment at times that are more convenient once an incentive is no longer provided.

For technology enabled demand response programs, such as direct load control, programmable communicating thermostats and autoDR, the average impact per participant may change over time even when incentives are continuing to be paid because of technology degradation. Unless there is a proactive effort to maintain and/or replace the technology to ensure that it remains operational, technology will eventually fail and the impacts associated with the technology will no longer exist. Persistence estimates for technology enabled programs must account for technology degradation.

For both technology and non-technology enabled programs, changes in human behavior must be considered. For some programs, such as price-driven demand response, the average impact might increase over time as consumers become better educated and learn better ways to reduce energy use during peak periods or as they invest in equipment on their own, such as time switches or programmable thermostats in order to increase their demand response. On the other hand, responsiveness may fall over time if the savings associated with participation are not large enough to sustain the behavior initially observed while customer inertia or some other factor (e.g., mandatory participation) keeps participants in the program even though they are no longer providing the same level of demand response.

The *EE Protocols* contain an extensive discussion of methods and protocols for estimating the effective useful life of various kinds of energy efficiency equipment. For programs that have been in place for an extended period of time and that have undergone multiple evaluations, surveys and on-site inspections of equipment can build a database over time that will allow for estimation of logistic curves and other functional forms that can be used to estimate the effective useful life of equipment. Given that most demand response programs are new and few evaluations have been done, these kinds of methods may not be an option currently, although there may be exceptions to this fact. For example, traditional load control of air conditioners has been used in the US by many utilities for many years and it may be possible to obtain data from some of these other programs that can be used to estimate annual failure rates for this type of technology.

The *EE Protocols* define a basic rigor level for degradation studies as follows:

> *Literature review for technical degradation studies cross a range of engineering-based literature, to include but not limited to manufacturer's studies, ASHRAE*

studies, and laboratory studies. Review of technology assessments. Assessments using simple engineering models for technology components and which examine key input variables and uncertainty factors affecting technical degradation.

These methods should also be considered for demand response impact persistence estimation for programs where technology is a key component.

A potentially much more difficult aspect of persistence estimation concerns predicting how consumer behavior may change over time. The extent to which this is a concern will vary significantly across programs. Programs involving the establishment of firm service levels and substantial penalties for violation of agreements are unlikely to see much degradation in demand response over time. On the other hand, price based programs such as critical peak pricing or RTP, or even incentive programs such as PCT programs where overrides are possible, might experience either an increase or decrease in average response depending upon how much consumers value the financial benefits that are actually received relative to the discomfort, inconvenience or other disbenefits that might be associated with behavior modification. Most dynamic rate programs have not been in place long enough anywhere in the US to obtain good information on which direction these behavioral changes might go, or whether there is likely to be any change at all compared with the response that was estimated over a relatively short program history.

6.5 Uncertainty in Key Drivers of Demand Response

As mentioned at the beginning of this section, with ex ante estimation, not only is it important to reflect the degree of uncertainty associated with the ex post evaluation parameters, which is largely tied to the accuracy and statistical precision of model parameters, but also the uncertainty associated with any significant drivers that underlie the ex ante estimates. Everything is uncertain in the future, and providing point estimates based on specific values for key variables can significantly overstate the true confidence that underlies the estimates.

Incorporating uncertainty in input values into estimates of demand response is straightforward using Monte Carlo simulation methods or similar approaches, but requires caution. With Monte Carlo analysis, each variable that drives demand response would be represented by a probability distribution defined by an explicit set of characteristics. Standard software packages, such as Crystal Ball, can also accommodate intercorrelations among exogenous variables (e.g., the fact that both price elasticities and reference values may increase with weather). The analysis software will pick a value from each input distribution and predict the demand response associated with that set of input values. This process will be repeated many times (1,000 draws from each distribution is relatively common) in order to produce the distribution of impact estimates that reflects the uncertainty associated with the driving variables as well as the model parameters.

The challenge in employing this (or any) method to represent the uncertainty in ex ante forecasts is developing probability distributions for the input values and incorporating the interdependencies in the relationships. In some cases, data exists that will allow for empirical estimation of the distributions. This is often the case for weather variables, and it might be true for things like market prices (in a competitive wholesale market with a reasonably long history, for example). For many other important variables, such as program participation, it might be possible to develop reasonable estimates of minimum, maximum and most likely values. If so, the information can be used to create a triangular or beta distribution to represent the uncertainty, as these types of distribution can be fully defined with just these three values. Regardless of how the distributions for key drivers or DR are arrived at, the shape of those underlying distributions should be clearly described.

6.5.1 Steps for Defining the Uncertainty of Ex ante Estimates

In the case of a regression based impact evaluations, incorporating uncertainty in the regression parameters and in the input values for ex ante estimates is relatively straightforward, and involves the following steps:

1. Obtain the regression output, recording the parameters and their respective standard errors

2. Obtain the variance-covariance matrix of the parameters

3. Convert the co-variances between parameters into correlations

4. Create a Monte Carlo model that replicates the parameters, their distributions, and inter-correlations

5. Incorporate the uncertainty associated with key drivers (e.g. temperature, participant characteristics) of the ex ante estimates and the intercorrelations among these drivers.

6. Run the simulation many times and obtain the confidence intervals.

An accurate estimate of the uncertainty associated with the model precision requires obtaining the regressions full variance-covariance matrix and incorporating any intercorrelations among the parameters. Simulating each parameter independently provides an inaccurate estimate of the confidence intervals. Likewise, correlations among the load impact drivers should be incorporated; otherwise the uncertainty estimates will be inaccurate.[43]

Nearly all statistical packages provide the full variance-covariance matrix of parameters if requested explicitly, and many easily provide it in the form of a correlation matrix, which is the format required by most Monte Carlo simulation software packages.

[43] Section 7.2.3 provides a detailed example of how failure to account for correlations can distort uncertainty estimates.

In cases where statistical packages do not translate the parameter co-variance matrix to correlations, the correlations can be obtained by the following equation:

$$\mathrm{corr}(\beta_i, \beta_j) = \frac{cov(\beta_i, \beta_j)}{\sigma_i \sigma_j} \mathrm{T}$$

where

σ_i and σ_j are the standard error for coefficients β_i and β_j

6.5.2 Defining the Uncertainty of Ex Ante Estimates: Example

To illustrate how incorporating the uncertainty associated with any significant drivers affects the load impact estimates, the example from Section 4.2.2.11 is extended here.

The example was developed using residential customer data for the CPP rate from California's Statewide Pricing Pilot for the summer of 2004. Only data from climate zone 3 (the hot, climate zone representing California's central valley) treatment group was used. As such the impacts reflect the incremental impact of a CPP rate layered on top of a TOU rate.

The key difference between ex post and ex ante uncertainty adjusted load impact estimates in the example is the fact that weather is unknown ex ante. In an ex ante setting the historical weather for the defined scenario, a typical event day, can be used to create the distributions by hour. For this example, the process was oversimplified and, instead of looking through historical data, a hypothetical distribution of weather across event days was developed. Importantly, the hour to hour correlation for weather had to be incorporated in order for the uncertainty adjusted load impacts to be accurate.

Table 6-2 presents the uncertainty adjusted load impacts for both a fixed scenario and one that incorporates the uncertainty in weather.

Table 6-2

Day Type: Average Event day for 2004 - SPP Climate Zone 3

Hour Ending	Temperature	Fixed scenario: Mean Impact (kW)	Fixed scenario: Percentiles 10%	Fixed scenario: Percentiles 50%	Fixed scenario: Percentiles 90%	Stochastic Scenario (incorporates uncertainty associated with weather): Mean (MW)	Stochastic: Percentiles 10%	Stochastic: Percentiles 50%	Stochastic: Percentiles 90%
1	71.1	-0.01	-0.03	-0.01	0.01	-0.01	-0.04	-0.01	0.01
2	69.9	-0.01	-0.03	-0.01	0.01	-0.01	-0.03	-0.01	0.01
3	68.8	-0.01	-0.03	-0.01	0.01	-0.01	-0.03	-0.01	0.01
4	67.7	-0.01	-0.03	-0.01	0.01	-0.01	-0.03	-0.01	0.01
5	66.8	-0.01	-0.03	-0.01	0.01	-0.01	-0.03	-0.01	0.01
6	66.1	-0.01	-0.03	-0.01	0.01	-0.01	-0.03	-0.01	0.01
7	65.9	-0.01	-0.03	-0.01	0.01	-0.01	-0.03	-0.01	0.01
8	66.9	-0.01	-0.03	-0.01	0.01	-0.01	-0.03	-0.01	0.01
9	69.7	-0.01	-0.03	-0.01	0.01	-0.01	-0.03	-0.01	0.01
10	74.0	-0.01	-0.03	-0.01	0.01	-0.02	-0.06	-0.02	0.01
11	78.3	-0.03	-0.05	-0.03	-0.01	-0.04	-0.09	-0.03	0.00
12	82.5	-0.06	-0.08	-0.06	-0.04	-0.06	-0.12	-0.06	-0.01
13	86.0	-0.09	-0.11	-0.09	-0.06	-0.09	-0.15	-0.08	-0.03
14	88.8	-0.11	-0.14	-0.11	-0.08	-0.11	-0.18	-0.10	-0.04
15	90.5	-0.12	-0.15	-0.12	-0.09	-0.12	-0.19	-0.11	-0.05
16	91.3	-0.12	-0.16	-0.12	-0.09	-0.12	-0.20	-0.12	-0.05
17	91.3	-0.12	-0.16	-0.12	-0.09	-0.12	-0.21	-0.12	-0.05
18	90.2	-0.12	-0.15	-0.12	-0.08	-0.12	-0.20	-0.11	-0.04
19	87.9	-0.10	-0.13	-0.10	-0.07	-0.10	-0.18	-0.10	-0.02
20	84.2	-0.07	-0.10	-0.07	-0.05	-0.08	-0.15	-0.07	-0.01
21	79.8	-0.04	-0.06	-0.04	-0.02	-0.05	-0.11	-0.04	0.00
22	76.4	-0.02	-0.04	-0.02	0.00	-0.03	-0.08	-0.03	0.00
23	74.1	-0.01	-0.03	-0.01	0.01	-0.02	-0.06	-0.02	0.01
24	72.4	-0.01	-0.03	-0.01	0.01	-0.02	-0.05	-0.01	0.01

Figures 6-1 reflects the uncertainty associated with the load reduction, presented as percent change in energy use, during the peak period hours for the fixed scenario. Figure 6-2 reflects the uncertainty adjusted load impacts that incorporate the uncertainty of weather. Both figures employ the same horizontal scale in order to allow for easy comparisons.

As seen in the figures, the difference is not trivial. The ex ante estimate under a fixed scenario presents substantially narrower distributions. If, for example, a planner was interested in the load impacts that could be obtained with 90% confidence (i.e., the 10^{th} percentile) the fixed scenario produces an estimate of a reduction of 0.09 kW per customer. The stochastic scenario, on the other hand, produces a 10^{th} percentile estimate of .05 kW per customer. This is a difference of roughly 80%.

Figure 6-1

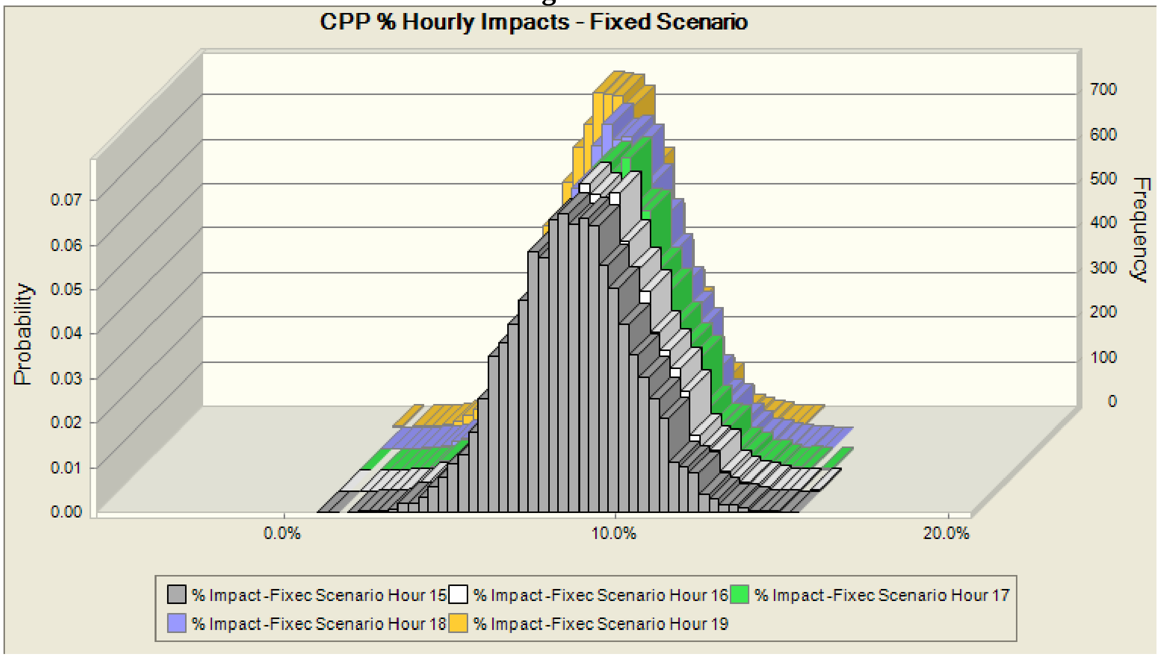
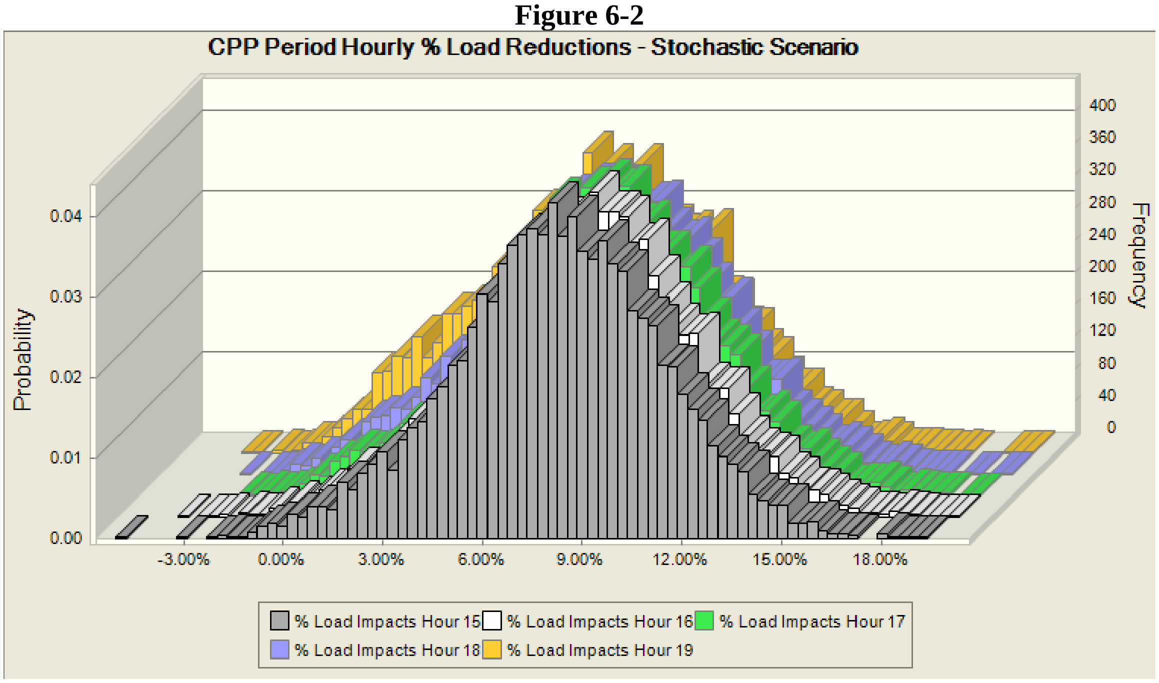


Figure 6-2



7 Estimating Impacts for Demand Response Portfolios

7.1 Introduction

The application of DR programs to operations and planning problems may require estimation of the load impact, and the uncertainty associated with that estimate, for a given day type and hour for all (or a subset) of DR programs taken together. Estimation of simultaneous load impacts across programs for a given day and/or time requires analysis of a portfolio of DR programs. This section discusses issues related to estimating the load impacts for DR portfolios and quantifying the uncertainty associated with using DR portfolios for operations and planning purposes.

DR portfolios include multiple programs across different customer segments. Each program will have advantages, constraints, and uncertainties with respect to fulfilling the resource requirements for a specific period of time (e.g., event day, season, etc.). Portfolios of DR programs can be designed with several objectives in mind including:

- capturing the benefits of synergies across programs,
- minimizing the impact on risk and uncertainty,
- maximizing overall yield, and
- investing in resources that are more economically efficient.

Because of its diversity, a DR portfolio provides many decision-making options in the short and long-run. From an operations perspective, a key question is: What is the optimal subset of DR and other resources to use in order to meet demand given the cost of resources and the tolerance for risk. From a long term planning perspective, a key question is: What resources should be expanded given long term demand forecasts and the uncertainties associated with fuel costs and the delivery of resources. Rational planning for system operations and expansion requires the selection of resources to meet electricity demand based not just on the benefits and costs of the individual resources but also on the risk and uncertainty associated with the combination of resources available. This is true for the portfolio of all resources needed to meet system loads as well as for DR portfolios.

To estimate portfolio level load impacts and uncertainty, it will be necessary to combine the estimated load impacts and uncertainties from multiple independent evaluation studies. Because calculating DR portfolio load impacts is an exercise of deliberate aggregation, a common lexicon for the load impacts from individual programs and the uncertainty associated with them is necessary.

This section proceeds by:

- Discussing basic issues in aggregating load impact estimates from individual programs to analyze the portfolio properties of programs taken together;
- Discussing the similarities and difference between energy efficiency and demand response portfolios
- Presenting a framework for aggregating load impacts across DR portfolios and quantifying the portfolio risk

7.2 Issues in Portfolio Aggregation

Aggregating DR program impacts and uncertainties across programs can increase or decrease the level of uncertainty about combined DR resource performance depending on the strategies used to deploy programs and the relationships among the various programs.

Key considerations in calculating a portfolio's uncertainty are:

- Ensuring proper aggregation of individual program load impacts,
- Correctly modeling the form of statistical distribution of the load impact (e.g., normal, beta, gamma, etc.), and
- Correctly taking account of correlations among the effects of the various programs.

The challenge in aggregating load impact estimates from individual programs to arrive at portfolio level load impact estimates is not in calculating the expected value (mean) of the load impact. Rather, it is in describing the level of certainty associated with the aggregated estimates. Estimating risk requires more than knowledge of the mean, it requires an accurate description of the uncertainty in the estimates.

With Energy Efficiency portfolios, the portfolio analysis framework calls for aggregating program impacts assuming they are independent and normally distributed. This simplifies the calculations necessary for aggregation. In the case of DR portfolios, these simplifying assumptions are probably not appropriate or at least should not be assumed to be appropriate without question.

To simultaneously take account of the complexities involved in convolving load impacts from probability distributions with different shapes whose performance may be correlated with weather and other factors will probably require the use of Monte Carlo simulation – a transparent brute force approach – grounded on real (not hypothetical) distributions and correlations whenever possible.[44]

Incorporating the uncertainty and relationship between programs for a DR portfolio is straightforward using Monte Carlo simulation methods or similar approaches, but it does require caution. With Monte Carlo analysis, the load impacts from each program in the

[44] In theory, the convolutions of the underlying distributions of load impacts from different DR programs could be accomplished with calculus, but it is much easier to do so with Monte Carlo simulation.

portfolio should be represented by a probability distribution defined by an explicit set of characteristics. Standard software packages, such as Crystal Ball, can accommodate correlations among the load impacts in distributions of load impacts for individual programs.

Throughout the following discussion of the basic issues involved in portfolio aggregation, a simplified, hypothetical example of a portfolio consisting of four programs is used to describe how load impact estimates should be aggregated; and the consequences of not aggregating loads correctly -- principally by making incorrect assumptions about the form of the probability distribution of load impacts or by failing to include correlations among the impacts produced by the programs in the portfolio.

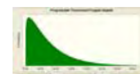
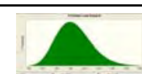
The example portfolio involves an Interruptable/Curtailable program, a Critical Peak Pricing program for small and medium commercial customers, a residential A/C cycling program, and two-way programmable thermostat program. Keep in mind that the examples used here are strictly hypothetical.

Some of the programs have load impacts that are not normally distributed and some of the load impacts from the programs are correlated. The distributions and correlations for the example are intentionally exaggerated to better illustrate the three basic complexities in portfolio aggregation. For each of the three key considerations, the example proceeds by presenting the expected load impacts, the standard deviation, and the distribution shape for the DR portfolio using both the correct and incorrect approach to aggregation, as explained below. A visual depiction of the distributions and the correct and incorrect portfolio aggregation estimates is shown in the right hand column of each table below.

7.2.1 Ensuring Proper Aggregation of Individual Program Load Impacts

A common mistake made in DR planning is to de-rate individual programs and then sum the de-rated values for a de-rated portfolio value. For example, for a planning application, a utility may be interested in only counting the DR load impacts that are 90% certain to be delivered when called upon. The incorrect but intuitive approach would be to de-rate individual programs, and sum the 90th percentile values for each program to reach an estimate of the 90th percentile value for the DR portfolio. This approach incorrectly calculates the uncertainty in the portfolio and undercounts DR when 90% certainty is required over the portfolio. Table 7-1 contains estimates using the correct and incorrect methods.

Table 7-1
Comparison of DR Portfolio Aggregation:
Correct Approach versus Summing Across All Program Percentiles

PROGRAM	Expected Load Impact Capacity	σ	Percentiles 10%	50%	90%	Distribution Visual
Interruptible/Curtailable	401.2	98.7	279.3	395.0	531.0	
CPP - Small and Med C&I	120.7	31.6	88.5	113.1	162.9	
A/C Cycling	320.5	33.5	284.8	313.5	365.4	
Programmable Thermostat	85.0	14.0	72.0	81.0	103.4	
PORTFOLIO	927.4	129.6	770.5	916.6	1,098.3	
PORTFOLIO - INCORRECT AGGREGATION	927.4		724.6	902.6	1,162.6	

The difference between this approach and the correct value may not be trivial. The incorrect approach undercounts the load impacts for the DR portfolio with 90% certainty by roughly 46 MW. The DR program can provide 6.3% more resources than calculated by the incorrect method. The disparity associated with the inaccurate method would be even greater if more of the distribution were normal or skewed to the right.

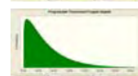
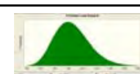
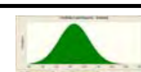
Importantly, the improper aggregation method also incorrectly describes the magnitude and distribution of the uncertainty. This can be seen by looking at the distribution depicted at the right hand column of Table 7-1 for the rows associated with the correct and incorrect approaches, both of which are on the same scale. The example highlights the importance of properly aggregating individual load impacts and risk.

7.2.2 Non-normal Distributions

Another error is to assume that all distributions are normal when in reality they are not. Using the above hypothetical example, assume that we are interested in aggregating the portfolio for a stochastic, i.e., probabilistic, event day scenario. Again, the example utility is interested in calculating the DR load impacts that are 90% certain to be delivered when called upon at the portfolio level. In this example the certainty estimates need to account for both the statistical precision and the stochastic component of the scenario. Generally, the uncertainty of statistical estimates is normally distributed and non-normal distributions will arise because non-fixed variables that drive demand response, such as weather, may not be normally distributed. In the hypothetical example, the A/C cycling and programmable thermostat programs are most skewed since their impacts can be expected to increase with temperature, the driver of load impact variation that is not normally distributed.

Under this example, shown in Table 7-2, the difference between the correct and the incorrect approach has a smaller impact, but, clearly, the shape and amount of uncertainty around the estimate is not properly represented if normal distributions are assumed

Table 7-2
Comparison of DR Portfolio Aggregation:
Correct Approach versus One That Assumes Normal Distributions

			Percentiles			
PROGRAM	Expected Load Impact Capacity	σ	10%	50%	90%	Distribution Visual
Interruptible/Curtailable	401.2	98.7	279.3	395.0	531.0	
CPP - Small and Med C&I	120.7	31.6	88.5	113.1	162.9	
A/C Cycling	320.5	33.5	284.8	313.5	365.4	
Programmable Thermostat	85.0	14.0	72.0	81.0	103.4	
PORTFOLIO	927.4	129.6	770.5	916.6	1,098.3	
PORTFOLIO - INCORRECT	927.4	109.8	786.7	927.4	1,068.2	

7.2.3 Correlations

Correlations are particularly important to incorporate in the process of estimating the portfolio impacts from the individual programs. Their impact can be mixed, depending on the relationship between the load impacts of programs. If the load impacts across programs are positively correlated, it has the effect of increasing the variation in the portfolio level load impacts. If the load impacts are negatively correlated, the correlation has the opposite effect – that of narrowing the variance for the portfolio impacts.

In practice, portfolios may have a mix of positive and negative correlations between individual program impacts. As will be detailed later, the preferred approach is to incorporate the interdependencies when estimating the certainty of the individual program load estimates, a bottom-up approach. By accounting for the drivers of load response for each program, which may be similar across programs, the correlation due to the common drivers is accounted for, as long the drivers are identified. However, it is also possible that impact evaluations do not account for all of the factors that affect the amount of load response provided, particularly as some of these factors become apparent with program experience.

In practice, key drivers that could lead to correlations across DR programs are temperature, industry type, day of week. If factors correlated with load impacts of multiple programs are not accounted for in the estimates for the DR portfolio aggregation scenario, it could potentially lead to correlation between the individual program impact estimates.

In the hypothetical example, shown in Table 7-3, the individual program estimates already factor in a key driver of demand response, temperature. Hence the correlations could be due to factors that were not incorporated in producing the program level impact estimates. In our example, geographical distribution could be a common factor unaccounted for in the load impacts if, say, customers in some regions are more or less pre-disposed to provide larger amounts of load impacts entirely separate from the temperature effect. Admittedly, the correlations employed in this example are unrealistic, but it illustrates a key point. Figure 7-1 provides a visual display of the assumed correlations.

For the example, the analysis that did not incorporate the correlations incorrectly stated the amount and shape of the uncertainty for the portfolio's DR resources.

Table 7-3
Comparison of DR Portfolio Aggregation:
Correct Approach versus One that Does Not Correct for Correlations

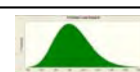
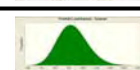
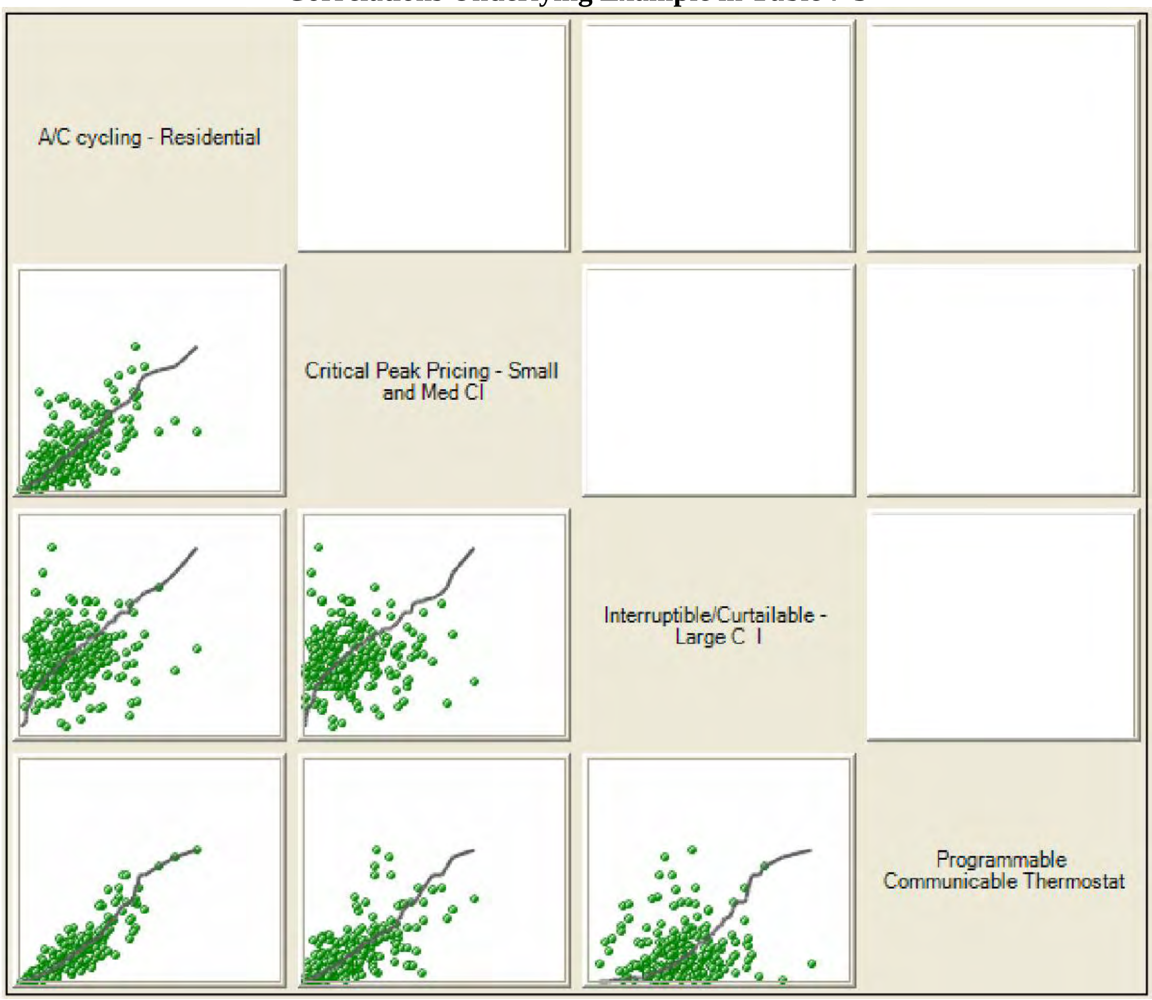
			Percentiles			
PROGRAM	Expected Load Impact Capacity	σ	10%	50%	90%	Distribution Visual
Interruptible/Curtailable	401.2	98.7	279.3	395.0	531.0	
CPP - Small and Med C&I	120.7	31.6	88.5	113.1	162.9	
A/C Cycling	320.5	33.5	284.8	313.5	365.4	
Programmable Thermostat	85.0	14.0	72.0	81.0	103.4	
PORTFOLIO	927.4	129.6	770.5	916.6	1,098.3	
PORTFOLIO - INCORRECT	927.4	109.2	792.2	922.4	1,072.6	

Figure 7-1
Correlations Underlying Example in Table 7-3



7.3 Similarities and Differences Between EE and DR Portfolios

There are several similarities between how DR and EE portfolios are created and analyzed. They both require the ability to aggregate across individual studies. They both depend on unbiased estimates of the actual impact of the program. They both require that the evaluations provide a clear measure of the uncertainty associated with the impact estimate, whether it be the actual distribution or an error bound.

However, there are differences between aggregating individual DR program impacts and the aggregation conducted for an EE portfolio.

- The quantity of demand response that can occur varies within programs and across participants based on conditions that vary systematically with weather, day-of-week, etc.

- Interactions between DR programs need to be explicitly considered, and in turn affect the aggregated results. In practice, participants can enroll in multiple programs and programs are often triggered under similar conditions. For example, a customer may be enrolled in both a demand bidding program and a curtailable program with a firm load level. The customer can submit bids at any point, but the load impacts for demand bidding should take into account whether or not curtailment notices were sent.

- Individual DR programs may or may not be deployed at the same time and, even if deployed at the same time, may not be deployed to full potential.

The value of the portfolio is not simply represented by instances in which participants reduce load, it also includes the option value of having the DR as a resource, making it important to obtain an accurate assessment of the total DR resources and the uncertainty/confidence surrounding that estimate.

7.4 Steps in Estimating Impacts of DR Portfolios

Assuming unbiased estimates of the impacts, there are five steps required to properly analyze the DR resources in a portfolio:

- Define the event or day-type scenarios
- Assess which DR resources are available during event days or day-type scenarios
- Calculate the load impact for the average customer for each program, hour, and the certainty around those estimates while explicitly accounting for interdependencies and scenario uncertainties (if the scenario is not fixed)
- Aggregate the load impacts for individual programs given the participant population enrolled or notified
- Aggregate the load impacts across DR programs.

If there is interest in the actual load reductions instead of DR portfolio capabilities, the aggregation of load impacts should be done for the customers who were sent a notification.

7.4.1 Defining Event Day Scenarios

Since demand response load impacts can vary by temperature, day type and other factors, using the common event or day type definition for all portfolio elements is a necessary first step in calculating DR portfolio load impacts. Day types can be defined in a variety of ways, including based on temperature, system load, and ISO system-wide or zonal emergencies.

System load levels are useful in defining event days and scenarios because the need for DR tends to coincide with high-load days, although the relationship is not perfect. Generation outages, transmission outages, level of imports, wind generation, and the accuracy of the load forecast are also key factors affecting whether or not a resource shortage occurs on a particular day.

Weather patterns are useful in helping define suitable event days or scenarios for DR portfolio aggregation because, in many cases, they are directly related to the amount of demand response that individual programs may deliver. For example, an A/C cycling program may deliver more load reduction on a hotter day than on a cooler one, while a demand bidding program may deliver less if events had been called for the days leading up to that day.

ISO called emergencies may also serve to define the common DR portfolio event days or scenario. Importantly, they will reflect the conditions under which DR resources are most needed and, by default, factor in the other drivers of resource shortages. The one drawback to this approach is that a larger set of data may be desirable in helping define the common event day or scenario for purposes of DR portfolio aggregation.

The definition of event days or scenarios preferably should be grounded in historical data. If the level of load response for a program is affected by temperature, it will be necessary to compute a weighted temperature for the scenario that reflects the geographical distribution of the participant population. While the scenario remains common, the temperature used to get estimates from individual programs may differ from program to program because of different participant characteristics and geographical distribution.

7.4.2 Assessing DR Resource Availability

Individual DR programs have different triggers, event durations, notification periods, restrictions on operations and hourly impacts. At times they directly interact, for example, when a participant is enrolled in two demand response programs. In other instances, the load impacts from individual programs are correlated, which affects the certainty of the portfolio load impacts.

The first step in analyzing the DR portfolio’s resources based on the common scenario is to determine the likelihood that individual programs could operate simultaneously given the scenario characteristics and the program trigger and notification requirements.

The second step is to assess whether the programs share the same participants and, if so, whether the load reduction in question is sufficiently large so as to require attention and resources for untangling those load impacts.

The third step is to assess whether or not the load impacts are correlated, and, if so, in what way. This step requires some attention and caution, as correlations in load impacts across programs affect the uncertainty associated with the portfolio load impacts.

Specifically, there are two types of correlation that must be accounted for, correlations among the load response drivers (e.g. temperature and compliance), and, in the case of regressions, correlations between model parameters.

Correlations become more important when the day type/scenario definition involves probabilities (i.e., it is stochastic). They are less of an issue for cases where the day type/scenario definition is fixed. With a fixed scenario, say a weekday with fixed ISO weighted temperatures by hour, if the correlation across program load impacts is due to a common factor, such as weather, and that factor is fully accounted for in the estimates of individual program load impacts. So, it is not an issue. With a stochastic scenario, temperature becomes a much larger issue if it affects load impacts for a variety of programs. In the latter case, if the correlations are unaccounted for, the resulting confidence bands will likely be inaccurate, as demonstrated in the example in section 7.2.3.

7.4.3 Develop Average Load Impact Estimates per Customer

With the scenarios defined and the participants properly allocated, the next step is to estimate the individual impacts and the level of certainty around those estimates for each program. Preferably, this is where any relationships between programs must be accounted for. In cases where the common scenario is not fixed, calculating the confidence intervals for the DR portfolio will require a Monte Carlo simulation approach.

The first task is to identify the common factors that drive load response for multiple programs, e.g., weather and day of week. These drivers should already be identified and accounted for in the individual evaluation studies, since factors that affect load are of particular interest.

The second task is to model the load impacts for individual programs taking into account 1) the uncertainty in the parameters and the stochastic components of the scenario, and 2) inter-correlations between model parameters and scenario characteristics. By including the drivers of demand response in individual programs, which may be common across programs, the relationships across programs are accounted for in the certainty bands. It is critical for evaluations to model and account for factors that influence the customer load response. If this is not done, it will not only provide less accurate load impact estimates for the program, but the correlations with other demand response programs cannot be easily controlled for regardless of the attentiveness paid in aggregating the portfolio.

7.4.4 Aggregate Load Impacts Across Customers

Given unbiased estimates of the average hourly load impacts, estimating an individual program's load impacts is straightforward.

To obtain the load impact estimates of a program for any given hour of a day type, the average hourly load impacts and confidence bands are multiplied by the number of participants enrolled or the number notified. Multiplying the average by the number of enrolled customers produces an estimate of program potential. Multiplying by the number of individuals notified provides an assessment of what the load impacts would be, ex ante, in an operations setting.

$$\text{Load Impact Capacity}_{dt} = \text{average load impact}_{dt} \times \text{total enrollees}_{d}$$

Where d = day type
t = hour of day

7.4.5 Aggregate Load Impacts Across DR Programs

The final step involves aggregating the load impacts, by hour, across DR programs. This requires both adding up the mean values and calculating the uncertainty or confidence bands around the portfolio level estimate.

Aggregating the mean load impact of a DR portfolio is straightforward. The sum of the average expected load impacts of individual programs produces the average load impacts for the entire portfolio, provided that the individual programs operate simultaneously. If the programs do not operate simultaneously under the defined scenario, it is necessary to weight the individual program load impact capacity by the likelihood that the programs would operate under the scenario.

The more complex task is best calculating the certainty of the DR portfolio's load impacts. How best to approach this depends on whether or not the distribution for all program load impact estimates are normally distributed. The methods for aggregating load impacts under the two different circumstances is discussed below, but first it is important to clarify the difference between the two potential cases, so that the correct method is identified and employed in the aggregation.

The individual program impacts are likely to be normally distributed when the underlying scenario is fixed and the relationships of load impact to, for example, weather and day of week are already accounted for in the hourly individual program predictions. This is because the certainty is largely tied to the accuracy and statistical precision of model parameters.

The individual program impacts are less likely to be normally distributed when the underlying scenario is stochastic. For example, in the stochastic scenario, the expected ISO weighted temperature for hour 1700 may be between 95-104 degrees with a median of 98.5. The implications are twofold. The confidence bands for the DR portfolio load impacts must incorporate uncertainty in weather since the scenario is not fixed. Second, the estimates and confidence bands will vary, by hour, for programs with weather sensitive load impacts and those impacts will be correlated. With non-normal distributions, the DR portfolio uncertainty can be computed via either calculus or Monte Carlo methods.

The impact of correlations and stochastic scenarios on the certainty of the estimates are preferably accounted for in step 3, as described above. If, however, the evaluations do not calibrate the load impact estimates for factors that drive load impacts for multiple programs, it might be necessary to incorporate the correlations at the aggregation stage, although it will likely be difficult to obtain empirically based, accurate estimates of correlations across programs at this stage.

When the certainty around impact estimates for all DR programs are normally distributed, they can be accurately described by the standard errors, which can be aggregated and used to produce the certainty around the DR portfolio estimates. This approach mirrors the discussion on integrating the results from multiple evaluation studies presented in Chapter 12 of *The California Evaluation Framework.*

The standard error from multiple programs can obtained be by the following equation:

$$\sigma_{\text{Portfolio}} = \sqrt{\sigma_1^2 + \sigma_2^2 + \sigma_3^2 + \ldots\sigma_n^2}$$

The standard error of the portfolio, along with the mean, can then be used to recreate the distribution and compute confidence intervals as described in any basic statistics textbook. An alternative method is to create a joint probability surface, incorporating the load impact uncertainties for individual programs via Monte Carlo simulation.

In cases where the common scenario is not fixed, the certainty around individual program estimates may or may not be normally distributed. Calculating the confidence intervals for the DR portfolio requires either calculus or a Monte Carlo method approach. If done properly, both will produce the same results, but Monte Carlo methods are less prone to error and more transparent to the reviewers.

As previously discussed, Monte Carlo simulations are a brute force approach, in which a value is drawn from the distribution of each program (taking into account correlations, if specified) and the portfolio estimate is computed. The process is repeated thousands of times, creating a joint probability surface that incorporates the uncertainty of individual programs and the relationship among those uncertainties.

7.5 Conclusions

Despite several similarities, the inherent differences between the aggregation of DR and energy-efficiency portfolios requires a different process for analyzing DR portfolios, their impact, and level of certainty. A pre-requisite for portfolio analysis is the existence of evaluations that produce the information that not only allows reviewers to assess the quality of the evaluation, but provide the correct metrics and statistics so as to allow ex ante predictions and the analysis of portfolio risk. A common lexicon for the load impacts from individual programs and the uncertainty associated with them is necessary. A second key prerequisite is the use of a common event or day type definition in the portfolio aggregation and analysis. However, the scenario in question will likely vary depending on the application of the portfolio and the utility itself.

8 SAMPLING

Sampling is a useful procedure in estimating DR load impacts because information needed for load impact estimation (i.e., interval load measurements) often is not available for the customers who are participating in programs; and installing interval meters is often too costly to consider for an entire program. Even in the future when most customers have interval meters, sampling may be useful as a means to reduce analysis costs when the volume of data available for describing load impacts is large. Despite these obvious advantages, relying on sampling for estimating load impacts increases uncertainty about the accuracy and precision of load impact estimates.

If interval load data is available for the entire population of DR program participants, evaluators should strongly consider using all available information to estimate DR program load impacts. Analyzing the data from the entire population of program participants eliminates the need for sampling and the attendant concerns about potential sampling bias and sampling precision discussed in this section.

Sampling adds three potential sources of uncertainty about the magnitude of load impact estimates:

- The potential for bias or inaccuracy resulting from the processes used to select and observe load impacts (i.e., sampling bias);

- Increased imprecision in the load impact estimates arising from sampling error (i.e., error arising from the inherent sample-to-sample variation that will occur when samples are used to estimate load impacts from the population); and

- Concern about the reliability of load impact estimates obtained from samples (i.e., concern that the results obtained from the sample may accidentally over or understate load impacts obtained from DR programs).

These issues should be directly addressed whenever sampling is used to estimate load impacts. Recommended approaches and resources for dealing with these issues are discussed below.

8.1 Sampling Bias

By far the most dangerous source of uncertainty arising from sampling is sampling bias. This is because, when sampling bias occurs, what is true of the sample is <u>not</u> necessarily true of the population – no matter how large the sample is.

There are two important sources of sampling bias:

- Under-coverage bias – the situation in which the sample frame from which the study participants are selected does not represent significant elements of the population. (At present, under-coverage bias is not a problem with samples chosen for DR program impact estimation because the population of participants in DR programs is known); and

- Selection bias – the situation in which elements in the sample are selected in such a way that they are not representative of the population of interest.

The best way to control sampling bias is to eliminate it by sampling observations for study at random from the populations under study. This practice will ensure that the initial sample is "representative" of the population of interest. Whenever possible, this approach to sampling should be employed. Unfortunately, it is virtually impossible to completely enumerate (i.e., observe all sampled members) a random sample when humans are involved; and this opens up the possibility of sampling bias even when random sampling has been undertaken.

There are many ways in which randomly selected observations can be systematically "selected out" of a given study before they can be observed. Examples of potential sources of selection bias include:

- Technical constraints associated with telecommunications, meter installation or other physical constraints may limit the installation of interval meters to a subset of sampled customers;

- Participants may refuse to supply information that is necessary for impact estimation (i.e., non-response to survey elements that may correspond with load impact measurements; and

- Participants may migrate out of the study while it is in progress.

Because selection bias is a strong possibility in any DR load impact evaluation, evaluators should use the following procedures to ensure that it is minimized and that its existence is detected and documented:

1. the population(s) under study should be clearly identified and described – this should be done for both participants and control groups to the extent that these are used;

2. the sample frame(s) (i.e., the list(s) from which samples are drawn) used to identify the population(s) under study should be carefully and accurately described and if the sample frame(s) do not perfectly overlap with the population(s) under study, then the evaluator should describe the measures they have taken to adjust the results for the sample frame so that it reflects the characteristics in the population of interest – this would include the use of matching or regression analysis;

3. the sample design used in the study should be described in detail including the distributions of population and sample points across sampling strata (if any);

4. a digital snapshot of the population and initial sample from the sample frame should be preserved – this involves making a digital copy of the sample frame at the time at which the sample was drawn as well as a clean digital copy of the sample that was drawn including any descriptors needed to determine the sampling cells into which the sampled observations fall;

5. effort should be expended to ensure that a large percentage of originally sampled observations are included in the study;

6. the “fate” of all sampled observations should be tracked and documented throughout the data collection process (from initial recruitment to study conclusion) so that it is possible to describe the extent to which the distribution of the sample(s) may depart from the distribution of the population(s) of interest throughout the course of the study;

7. If significant sample attrition is found to exist at any stage of the research process (i.e., recruitment, installation, operation), a study of its impact should be undertaken. This study should focus on discovering and describing any sampling bias that may have occurred as a result of selection. Generally this will be done by comparing the known characteristics of the observed sample with the known characteristics of the population. Known characteristics would include such variables as historical energy use, time in residence, geographical location, reason for attrition from sample and any other information that may be available for the population and sample.

8. If selection bias is suspected, the evaluator should describe it as well as any efforts made to control for it either by weighting the sample to correctly reflect the population; by regression analysis designed to adjust for selection effects; or by selection modeling (ala. Heckman’s technique[45]).

It is important to keep in mind that the mere fact that some randomly sampled observations are not completely observed (i.e., have been selected out of the sample at some point) does not necessarily mean that the resulting sample has been <u>biased</u> in some significant way. Whether bias is induced by selection depends on whether the selection is somehow related to the magnitude of the impact of the DR program. This can only be determined by carrying out the work outlined above.

[45] The problem of controlling for selection bias has been discussed at great length in the literature on econometrics. The seminal articles on this topic are by James Heckman “The common structure of statistical models of truncation, sample selection and limited dependent variables and a simple estimator for such models”, in *The Annals of Economic and Social Measurement* 5: 475-492 1976; and Sample selection bias as a specification error” in *Econometrica*, 47: 153-161

8.2 Sampling Precision

A sample is a subset of the population of interest and as such will not in general have exactly the same statistical measurements as the population as a whole. Correspondingly, sample estimators such as means, standard deviations, frequency counts etc. will vary from random sample to random sample. Thus, whenever sampling is used to describe the characteristics of a population, there is some uncertainty about the estimates from the sample that comes from random variation in the sampling process. While we sometimes find it convenient to talk about the results obtained from a sample as though they were "point estimates" of the measures of the population of interest, it is generally inappropriate to interpret the results of sampling without considering the sample-to-sample variation that is likely to have occurred. This is the problem of sampling precision.

The extent of sample-to-sample variation in measurements generally depends on the inherent variation in the factor of interest in the population (in this case hourly loads) and the number of observations that are sampled. In general, the more homogeneous is the population of interest with respect to the variable of interest, the lower the sample-to-sample variation in measurements that can occur. If everything in the population is the same or nearly the same, then there will be little sample to sample variation obtained through random sampling.

It is also true that the larger the sample size, the lower the sample-to-sample variation in measurements. This is true because the standard error of the mean (average distance of the sampled mean from the true population mean) decreases with the square root of the sample size. This can be seen in the formula for the standard error of the mean shown in equation 8-1.

$$\sigma_m^2 = \frac{\sigma^2}{n} \qquad (8\text{-}1)$$

Where σ_m^2 is the standard error of the mean, σ^2 is the variance of the population, and n is the sample size.

Both determinants of sampling precision can be manipulated to establish desired levels of sampling precision.

The standard error or average distance of sampled means from the center of the sampling distribution is a useful measure of the sampling precision because it tells us how far on the average the sample can be expected to stray from the mean of the population given its variance and the sample size. However, an even more useful measure of sampling precision can be derived from the standard error of the mean by computing the interval within which the true population estimate is likely to be found. This is called the confidence interval. The confidence interval for a sample estimator is the interval in which the true population value is likely to be found with a certain probability. So, for example, you often see sample estimators described in terms of upper and lower confidence limits expressed in terms of percentages. The confidence interval for a given estimator is obtained by multiplying the standard error of the mean times the area under the sampling distribution for the mean associated with the observation of a given extreme value (i.e., 90%, 95% or 99%). This can be seen in the formula for the confidence interval of the mean shown in formula 8-2.

$$\bar{x} - z\sigma_m^2 \leq \mu \leq \bar{x} + z\sigma_m^2 \qquad (8\text{-}2)$$

Where $\bar{x}$ is the sample mean, z is the value of z associated with the appropriate alpha, and σ_m^2 is the standard error of the mean.

The confidence interval is a useful statistic because is reflects the upper and lower limits within which the true population value will be found with a given level of certainty. It is particularly useful in operations and resource planning where users will generally want to incorporate the maximum amount of load impact they can confidently expect to occur in their decision making and planning. Whenever load impacts are calculated based on sampling, the upper and lower confidence limits should be reported. The confidence levels or probabilities employed in the calculation should be determined in consultation with the users of the information.

It is important to keep in mind the fact that sampling precision and sampling accuracy are two very different things. One cannot overcome inaccuracy or bias in load impact measurements induced by inaccurate reference load measurements or sample selection by increasing sampling precision as this will simply result in a more precise wrong answer.

8.2.1 Establishing Sampling Precision Levels

Samples can be made to be nearly perfectly precise for all intents and purposes. However, sampling precision is not inherently valuable and it comes at a cost in terms of meter installation, maintenance and database management. In essence, the reduction in uncertainty associated with sampling error has to be balanced against the increased cost of obtaining more precise estimates in sampling.

An important step in designing a DR load impact evaluation is to identify the extent of sampling precision required to support decision making. There are no hard and fast rules concerning how much sampling precision is enough. It depends on how the information is intended to be used. Establishing an appropriate level of sampling precision is best done by consulting with the intended users of the information and asking them to agree to an acceptable sampling error rate.

There are really two related issues that must be decided in this conversation – identification of an acceptable level of sampling precision (e.g., plus or minus 5% or 10% or whatever) and identification of the desired reliability of the estimate (e.g., 95% reliable, 90% reliable, etc.). In the end, it is important to agree with intended users about both the precision and reliability of the estimators coming from the sample – since these two issues can be traded off against one another. Once the desired level of sampling precision has been determined, an appropriate sample design can be identified.

8.2.2 Overview of Sampling Methodology

Sampling is a well developed scientific discipline and there are well known textbooks that outline technical approaches to sample design that are appropriate for designing samples to be used in DR load impact estimation. These include classics such as Cochran, Kish and Deming.[46] While an in depth treatment of sample design is well beyond the scope of these protocols, there are certain sample design options that are more appropriate for DR load impact estimation than others and the remainder of this section discusses issues that favor using some designs over others under certain conditions.

Sample design is a highly technical art that requires training and experience in statistics and survey sampling. If the expected level of investment in metering and data collection is significant for a given program, then it is recommended that evaluators consult with an expert survey statistician in identifying an efficient sample design for DR impact evaluation.

8.2.2.1 Simple Random Sampling

Any discussion of sampling and sample design has to proceed from a review of simple random sampling because it is the basis of most sampling procedures that are appropriate for DR load impact estimation. However, for reasons that will be discussed below, simple random sampling will seldom be appropriate in studies of DR load impacts.

In simple random sampling, population units are selected for observation with probability 1/N. That is, all of the elements in the population have an equal chance of being selected for study. Statistical estimators obtained from such simple random samples are unbiased and consistent.

Equation 8-3 identifies the formula for determining the sample size required to obtain a given level of precision under simple random sampling.[47] :

$$n = \frac{z^2\sigma^2}{r^2\bar{x}^2} \qquad (8\text{-}3)$$

[46] Classic textbooks useful in survey sampling include:
Sampling Techniques: third edition, by William Cochran, John Wiley and Sons. 1977
Survey Sampling, by Leslie Kish, , John Wiley and Sons, 1965
Sample Design in Business Research, by William Deming, John Wiley and Sons 1960

[47] The actual equation for calculating sample size includes a correction for the size of the population called the finite population correction. This adjustment has been left off of the equation for ease of exposition. In general its effect on the sample size calculation is deminimous when the population of interest is large. – say more than a few thousand.

Where n is the sample size, z is the value in the z distribution associated with alpha (probability of Type II error), σ^2 is the population variance, r^2 is the relative error (error as a percentage of the mean), and $\bar{x}$ is the population mean.

Notice that this formula requires just two pieces of information: a desired level of sampling error; and an estimate of the standard deviation of the variable of interest in the population. In most cases, the standard deviation of the variable of interest in the population is <u>unknown</u> and must be estimated by *proxy* from the distribution of some variable for which these values are known. It is also possible to substitute an estimate of the coefficient of variation (CV) for the standard deviation in the above equation and solve for sample size. The CV is equal to the ratio of the standard deviation to the mean.

Load research has been underway for many years in the utility industry and in most cases it is possible to identify a reasonable proxy for the standard deviation of an electric load in the population of interest or, in the absence of that, a reasonable estimate of the coefficient of variation. Using the above information, the sample size required to obtain a given level of statistical precision is easy to calculate.

Simple random sampling is easy to do, and the results obtained from it can be directly used to estimate population parameters from sample values by multiplying the sample estimates times the sampling fractions (e.g., population weights). So, what's not to like about simple random sampling?

While simple random samples are easy to create and use they have certain limitations in practice. First, because sample elements in simple random samples are selected exactly in proportion to the prevalence of conditions in the population, they may produce relatively small numbers of "interesting" population members that occur relatively rarely. For example, commercial office buildings comprise only a small fraction of all commercial accounts. Too few of these buildings may be selected in a simple random sample of commercial accounts to meaningfully describe the impacts of DR programs on loads in these buildings. To the extent that it is useful to describe the DR load impacts of important subsets of the population, a simple random sample may not be a practical approach to sampling because the sample size required to select them at random from the population is extremely large.

A second limitation in the usefulness of simple random sampling in DR load impact estimation arises from the fact that customer loads vary widely within populations of DR programs with known customer characteristics (i.e., geographic location, customer type, connected load, etc.). It is not unusual to observe coefficients of variation for energy use and hourly loads ranging from 1 to 4 for these populations. Left unchecked, this variation can lead to greatly inflated requirements for sample size.

These problems are common to most scientific research and many sample design alternatives have evolved to solve them. Consequently, in many applications, more complicated sample designs are often preferred over simple random samples.

8.2.2.2 Stratified Random Sampling

In stratified random sampling, each and every element of the population of interest is pre-sorted into one and only one category for purposes of sampling. Then samples are drawn at random from each category. The sample sizes obtained from each category are generally not proportional to the distribution of the population across the strata, so the sample per se is not representative of the population of interest (i.e., it is biased). This distortion, however, can be used to good effect if properly constructed.

Stratification is very useful in load impact estimation because it allows the researcher to exactly control the distribution of the sample across meaningful categories. Examples of useful stratification variables include: weather zones, usage categories, utility service territories, business types, occupancy patterns and a host of other variables that can have an effect on customer loads. Stratified random samples can be constructed in such a way as to supply known levels of sampling precision within strata and for the population as a whole. In this way they can be used to develop statistically precise estimates of load impacts within weather zones, usage categories and so on. They can also be useful for developing sample designs that are statistically more efficient (i.e., have higher statistical precision at given sample sizes) than simple random samples.

The sample estimators (i.e., means, standard deviations, etc.) for the sampling strata are unbiased estimators of the parameters of interest for the population within each stratum. However, to estimate total population parameters using estimators from stratified random samples, it is necessary to properly weight the estimates obtained from each of the sample strata so that the effects of the measurements from the strata (e.g., mean, standard deviation, proportion, etc.) are proportional to the sizes of the populations in the strata. All statistical estimators obtained through stratified random sampling must be corrected in this manner to produce unbiased total population estimates.

Identification of appropriate sample sizes for stratified random samples is somewhat more complicated than it is in the case of simple random samples. If the purpose of stratification is to obtain designated levels of sampling precision within the strata, then the sample sizes within each stratum are obtained using the formula for simple random sampling – using the estimated standard deviation and desired sampling precision for the stratum. It is not unusual for decision makers to specify that they require a given level of sampling precision for each utility, or by weather zone. In such cases, the sampling precision within the strata will determine the overall sampling precision obtained for the population. The sampling precision for the combined sample (i.e., with all the strata taken together) is obtained by calculating the weighted standard error of the estimate.[48] The sampling precision for the entire population should be substantially higher than it is for any of the strata taken alone.

On the other hand, or in addition to the above consideration, stratification can be used to enhance sampling efficiency. In this case, the sample is distributed among the strata in

[48] Ibid

such a way as to minimize the weighted standard error of the total population estimate. Procedures for identifying optimal stratum boundaries and for calculating sample sizes within strata to achieve desired levels of statistical precision in stratified random sampling have been developed by Delanius and Hodges and Neyman respectively.

Stratified random sampling will almost always be required in assessing DR program impacts – particularly for programs where it is important to develop reasonably precise measurements within geographic locations or for different customer types. It may also be useful for improving the efficiency of sample designs – though in the case of many programs, the improvements in sampling efficiency obtained from repeated measures designs (discussed below) will overshadow any improvements that may be obtained by pre-stratifying on the basis of customer size.

Whenever stratified random samples are used to estimate DR load impacts, researchers should carefully describe the sample design. Oft-reported measures include;

1. the distribution of the population across sampling strata;

2. the distribution of the sample across sampling strata;

3. any procedures used to identify optimal stratum boundaries used in pre-stratification and the impacts of pre-stratification on sampling efficiency (i.e., if Delanius-Hodges and/or Neyman allocation are used, the researcher should provide a rationale for their choice of the number of strata and stratum boundaries used in the design and their respective impacts on sampling precision);

4. the expected statistical precision for estimators within each strata (including a discussion of any use of proxy measures of the standard deviation used in this calculation); and

5. the expected statistical precision for estimators in the population overall.

8.2.2.3 Sample Designs Using Alternative Estimators

Beyond stratification there are several other important ways of enhancing the statistical precision of sample estimates. These are used in conjunction with the basic sample designs outlined above. They involve using alternative estimators to those used in designing conventional samples. The conventional sample designs discussed above are focused on identifying sampling procedures that will achieve a certain level of statistical precision in estimating well known parameters of statistical distributions such as the mean and standard deviation. In the case of DR load impacts, these sample designs can be used to achieve a certain level of precision in estimating the average load impact, its standard deviation and confidence intervals.

It is possible and in many cases desirable to create samples designed to measure other parameters in the population that can be used to develop more precise estimates of load impacts than the elementary sample means and standard deviations. Two important alternative estimators that should be considered in developing DR load impact evaluations are ratio estimators and regression estimators. Under certain circumstances, these estimators can be used to greatly enhance the precision of statistical estimates obtained from sampling and thereby significantly lower the cost of impact evaluation.

Ratio Estimation

Sampling to observe ratio estimators improves efficiency by sampling to observe the relationship in the population between an unknown variable (e.g., the actual load observed during a DR event) and a property that is known for all population members (e.g., the contractual firm service level for subscribers to the program). To the extent that the actual load observed during the DR event is correlated with the firm service level, the ratio of the two variables will have inherently lower variation than the metric value of the loads involved in the numerator or denominator; and the estimated load impact can be measured with substantially greater precision than the metric loads underlying it. Correspondingly, significantly smaller numbers of sample points are required to observe the ratio of the two variables in the population than would be required to estimate the value of either the numerator or denominator. This is called ratio estimation. Designing samples for ratio estimation follows the same basic logic as for conventional sample designs – except the variable of interest in establishing sampling precision is the ratio, not the metric value of the loads of interest.

The EE protocols devote considerable attention to the technical details of developing samples for ratio estimators and these protocols should be consulted if the use of ratio estimators is being considered in DR load impact estimation. Ratio estimators are very useful in EE program evaluation because it is relatively easy to conceive of the impact of an EE program as a ratio of achieved savings to estimated savings for measures that were supposed to have been adopted. DR programs that are excellent candidates for sampling based on ratio estimation are those where participants agree to reduce loads to firm service levels on command and those where participants are demand bidding – both cases where the program impact is easily defined as a ratio.

Regression Estimation

An extension of the logic of ratio estimation is regression estimation. In regression estimation, sampling efficiency is improved by sampling to observe the relationship in the population between the regression adjusted mean (in this case of hourly load) and variables that influence the value of the regression adjusted mean (e.g., time of day, program participation, ambient temperature, household size, load in hours prior to the event, etc.). To the extent that hourly loads are correlated with factors that vary systematically in the population, it is possible to define a regression function that will predict those loads more or less precisely.

An interesting property of the regression adjusted mean is that its standard error decreases with $(1-R^2)$. This means that if R^2 (e.g., the proportion of the variation in the load explained by the regression function) is .9, the standard error of the regression adjusted mean is 10% of the standard error of the population mean. Thus, substantial improvement in sampling precision can be obtained if the regression adjusted mean and standard error are estimated instead of the population mean. Of course, the smaller the R^2 for the regression equation, the smaller will be the improvement in sampling precision.

While the potential for improvement in sampling efficiency from regression estimation is tantalizing, researchers have to bear in mind that the extent of improvement in sampling efficiency depends entirely on the predictive power of the regression function that is specified. Practically speaking, this means that the researcher must have some a priori knowledge that the predictors to be included in the regression function actually have substantial predictive power before developing a sample design based on regression estimation. Fortunately, there is ample evidence in prior research concerning customer loads that information about type of customer, time of day, temperature, day of week, and other variables are highly predictive of hourly customer loads. Examples of load impact estimates that have been made using regression estimation are provided in Appendix A.

If the relationships between predictor variables and hourly loads have been studied in prior research, sample sizes for estimating regression functions including variables from the prior research can be calculated directly. This is done by observing the R^2 of the prediction equation (applied to past data) and making a reasonable guess about the incremental increase in R^2 that will result from addition of the effect variable (a new predictor).

Most statistical packages provide algorithms for estimating sample sizes for estimation of effects using multiple regressions. These require making assumptions about R^2 of the model without the effect predictor, the incremental improvement in R^2 that will result from the inclusion of the predictor variable, desired statistical power and alpha (probability of Type II error). For examples of these algorithms see STATA and SPSS.

In the case where no prior information is available concerning the predictive power of the regression function, sample sizes can be estimated using various rules of thumb involving assumptions about desired statistical power, Type II error (alpha) and the number of predictors in the regression equation. See Tabachanick and Fidell (2001)[49] for a discussion of the various rules of thumb that have been applied historically to estimating sample sizes required to estimate regression parameters. Various rules have been suggested. For example, one rule suggests that the minimum sample size for estimating regression coefficients should not be less than 104 plus the number of predictors in the regression equation. Another rule suggests that the sample size should be at least 40 times the number of independent variables in the regression equation. Still another rule says that the minimum sample size should depend both on the effect size that is to be detected and the number of variables in the equation. This rule calculates the minimum sample size as [8/(effect size)] plus the number of independent variables minus 1. All of

[49] *Using Multivariate Statistics (3rd ed.),* Tabachnick, B. G., & Fidell, L. S. New York: Harper Collins (1996).

these rules have some basis logic and experience, but none can be said to be robust and capable of producing efficient sample size decisions.

Given the uncertainty that may exist about the predictive power of regression models, if circumstances permit, it is advisable to set sample sizes for estimating regression functions using double sampling. In double sampling, an initial sample is drawn that is thought to be sufficient and the parameters in the distribution of interest (in this case regression parameters) are calculated. The initial sample might be drawn according to the first rule of thumb described above which would yield less than 120 observations in most cases. If the initial sample is insufficient to precisely estimate the parameters of interest, sufficient additional samples are then drawn to supplement the first sample.

Regression estimation can be used to good effect in estimating load impacts for most DR programs.

8.2.2.4 Repeated Measures Designs

For event based programs it is possible to employ repeated measures designs. The availability of repeated measures of the outcome variable (i.e., hourly loads) is an interesting complication (and great advantage) in load impact estimation. When multiple events occur over a given period of time (e.g., critical peak days, interruptions, calls for curtailment) each conventionally sampled “point” (i.e., customer) actually produces multiple observations of the program impacts (hourly loads). In effect, the study design that is being undertaken is a panel in which repeated measurements are taken over some number of time periods.

To talk about this sort of study design one needs to distinguish between two kinds of measurements – cross-sectional measurements and time series measurements. Repeated measures study designs typically have both kinds of measurements. The cross-sectional measurements are those that vary over customers but not over time – things like location, customer type and income. Time series measurements are those that vary over time within a given member of the cross-section. These are variables like energy use, cooling degree hours, day of week, season and whether a DR event has been called.

Variation in customer loads arises out of variation in factors in the cross-section and out of variation in factors in the time series. For example, in a given hour, one customer in the cross section might use 2kWh of energy while another might use 4kWh. Such a difference could be because one of the customers has twice the air conditioner capacity of the other or it might be because one of the customers has a chest freezer in the garage and is charging the battery on their electric car during the time the energy use is observed. The sources of variation among customers that account for these differences are numerous and some are very difficult to measure. From hour to hour for any given customer, the loads also vary as a result of factors that are changing with time – factors such as season, day of week, temperature, occupancy patterns, and whether or not a DR event is called, etc. Some of these are also difficult to measure.

Because observations are being made across the variables in the cross-section and over time, it is possible with repeated measures designs to isolate the effects of cross-sectional and time series variables. In particular, it is possible to observe the main effect of a DR program in isolation from the cross sectional variation and to observe the interaction between the DR program and the cross sectional variables of interest. These can be used to produce a very powerful predictive model of the load impacts of event based DR programs.

As explained in Section 4, repeated measures designs offer several powerful advantages.

- These designs are statistically much more powerful than conventional designs in which a single observation is taken per sampled point. That is, much smaller cross-sectional samples can be used to estimate average load impacts than would otherwise be necessary.

- There is typically no need for a control group in estimating load impacts because load impacts for sampled units (e.g., households, firms, etc.) can be estimated as the difference between loads for "event" days and "non-event" days for each sampled unit. This eliminates the attendant risks of selection bias in comparing volunteers in the DR program with those who have not volunteered in the general population of interest;

- The potential for estimation bias arising from fixed omitted variables in the estimation equation can be completely eliminated; and

- Variation in load measurements arising from factors in the cross-section can be isolated and accurately described.

The conventional sample design techniques discussed under simple random sampling and stratified random sampling provide no basis for selecting an appropriate sample size for this sort of study because they are based on the notion that the sampled observations are independent of one another. The observations within the time series are not.

The sampling precision in a repeated measures design is a function of the size of the cross-section, the number of repeated measurements that occur and the correlation between the measurements. All other things being equal, sampling precision and statistical power increase significantly as the number of measurements increases. For DR programs involving six to ten events per season, sampling precision can be increased very dramatically – making it possible to detect relatively small effects (i.e., load reductions in the range of 5-10%) with only a few hundred observations. A good example of the analysis of repeated measures to observe relatively small load impacts is the SPP.

It is possible to calculate the sample size required to detect effects of a given size with repeated measurements in time given the:

- mean of the variable of interest;

- standard deviation of the variable of interest;
- number of repeated measurements by type (event and non-event);
- the number of groups in the analysis;
- acceptable probability of Type II error (alpha);
- desired power of the statistical test;
- correlation between measurements in the time series (rho);
- type of model used to estimate impact (e.g., Pre/Post, Change, ANOVA or ANCOVA); and
- minimum effect size that is to be detected.

A procedure for making this calculation is available in STATA's sampsi program.[50]

It is possible to use information from load research samples to estimate the parameters that are required to calculate the sample sizes necessary to undertake a repeated measures study. In general, this will be the minimum sample size required to estimate the load impacts of the DR program.

Sample sizes calculated in this way do not include any provision for estimating the effects of the interactions of cross-sectional variables with the treatment effect. Accounting for the effects of the cross-sectional variables on the load impact will in most cases require additional samples. There are two reasons for this. First, the effect size specified in the sample design calculation must be reduced substantially if the effect sizes for the interactions are to be observed because interacting the cross sectional variables with the treatment will, in effect, decompose the treatment effect into smaller pieces (effects). Second, to observe the effects of the cross sectional variables it will be necessary to ensure that these variables have sufficient variation to permit regression type estimation.

If the effects of cross-sectional variables are to be included in repeated measures calculations it is probably more appropriate to employ sample sizes that would be required to estimate cross sectional effects in regression models (i.e., stratified random sampling).

8.3 Conclusion

Sampling adds uncertainty about the accuracy, precision and reliability of load impact estimates. When interval load data is available for the entire population(s) under study

---

[50] See Frison and Pocock (1992) "Repeated measures in clinical trials: An analysis using mean summary statistics and its implications for design", in Statistics in Medicine 11: 1685-1704 for a technical discussion of the method used to estimate the impacts of repeated measures on sampling precision and sample size.

evaluators should consider using it to avoid these sources of uncertainty. However, there are certain instances where using data for the entire population might be impractical and sampling will be the appropriate method for observing DR load impacts. This will be true for mass market programs where interval metered data is not available for all population members. The use of sampling may be desirable even when information is available for a large mass market program because a more focused effort on a properly designed sample can produce more accurate information than may be available through an attempt to analyze the information for the entire population.

When sampling is used care must be taken to ensure that it is representative of the population of interest and that it is sufficiently precise to meet the needs of the various stakeholders. There are well accepted sampling techniques that should be used whenever sampling is employed. These include: random sampling from the populations of interest and stratifying the random sample to achieve an acceptable level of statistical precision.

In most cases, stratified random sampling will be required for DR program evaluations because it will be necessary to precisely estimate load impacts for important subsets of the populations under study (e.g., by utility service territories, weather zones and customer types defined in various ways). It may also be necessary to stratify samples by usage or other variables representing customer size in order to achieve acceptable sampling precision within budget limitations. Whenever stratified random samples are used, care must be taken to consider the impacts that sample weighting will have on subsequent analyses and to make sure that sampling weights are appropriately applied when summary measures for the population are calculated.

Efficiency gains arising from regression based estimators and repeated measures designs will generally favor the use of these analysis techniques in DR load impact estimation. Sampling to support the use of these techniques is not straightforward. It is possible in both cases to use either simple random sampling or stratified random sampling to establish appropriate sample sizes for DR load impact evaluations. Sample sizes established using these procedures will be conservative since the effects of the covariates and repeated measures will only serve to make the measurements more precise.

The most robust approach to estimating the sample size required for regression modeling presupposes an understanding of the variation in the customer loads in the population of customers under study; and the relationship between those loads and the factors that are being considered for use as control variables. In some cases, this information is available from prior studies (e.g., SSP) or from load research samples. Whenever such information is available, it should be used to identify an appropriate sample size required to support the analysis. If this information is not available, the sample design should be developed using conventional stratified random sampling techniques (i.e., those that only require information about the population mean and standard deviation within strata).

There are well developed procedures for establishing sample sizes for repeated measures studies used in experiments and clinical trials. An important determinant of the sample size required in a repeated measures design is whether interactions between cross-sectional variables and the effect of the program have to be estimated. If this is not required, then the sample can be designed using the simple procedures that are appropriate for establishing sample sizes for clinical trials and experiments. On the other hand, if the -interactions of the cross-sectional variables -sectional variables are to be described it is probably more appropriate to employ sample sizes that would be required to estimate cross-sectional effects in regression models. The resulting sample size will be larger than what is possible with a repeated measures design, but will ensure that the cross section is large enough and diverse enough to estimate the cross-sectional effects.

Given the complexity of the analysis procedures used in DR load impact estimation, evaluators are advised to consult with a qualified and experienced survey statistician in developing sample designs to be used in DR load impact estimation. This is particularly true if significant resources will be expended installing meters and surveying customers.

9 REPORTING PROTOCOLS

Evaluation reporting has a variety of objectives, including:

- Presenting the detailed impact estimates developed as part of the evaluation
- Comparing these findings with program goals and the impacts that have been used to report progress toward goals, and explain any differences
- Thoroughly documenting the methodologies used in sufficient detail so that, given access to the same data and information, a trained evaluator would be able to reproduce the impact estimates that are reported
- Reporting any deviations from the requirements of these protocols and the reasons why it was not possible to meet them
- Providing recommendations regarding program modifications and modifications to the impact estimates used for program progress reports
- Providing recommendations concerning future evaluation activities.

The evaluation reports should generally be written for a wide range of individuals, including people who are not familiar with evaluation approaches or the field's specialized terminology. Technical information associated with the evaluation methodologies, research design, sampling, M&V efforts, regression analysis, bias detection, bias correction and other technical areas must be reported and should not be avoided to ensure readability by a wider range of audience. While a summary of the methodology, findings and decisions covering these issues should be written for a wider audience, the more technical details relating to these reporting categories must also be provided.

Evaluation reports should include, at a minimum, the following sections:

1. Cover
2. Title Page
3. Table of Contents
4. Executive Summary - this section should very briefly present an overview of the evaluation findings and the study's recommendations for program change
5. Introduction and Purpose of the Study - this section should briefly summarize the program or programs being evaluated and provide an overview of the evaluation objectives and plan, including the research issues that are addressed. It should also provide a summary of the report organization.

6. Description of Programs Covered in the Study - this section should provide a detailed description of the program being evaluated in enough detail that readers can understand the program and program components that delivered the evaluation identified effects. The description should include a program history, summary of program goals (both in terms of enrollment and demand impacts), tables showing reported progress toward goals, projections of future goals and known program changes and other information deemed necessary for the reader to obtain a thorough understanding of how the program has evolved over time and what changes lie ahead.

7. Study Methodology - this section should describe the evaluation approach in enough detail to allow a repetition of the study in a way that would produce identical or similar findings. (See additional content information below.)

8. Reliability Assessment of the Study Findings – this section should include a discussion of the threats to validity and sources of bias and the approaches used to reduce threats, reduce bias and increase the reliability of the findings, and a discussion of precision levels. (See additional content information below.)

9. Detailed Study Findings - this section presents the study findings in detail. (See additional content information below.)

10. Recommendations - this section should contain a detailed discussion of any recommended changes to the program as well as recommendations for future evaluation efforts.

The **Study Methodology** section must include the following:

1. Overview of the approach and evaluation plan;

2. Questions addressed in the evaluation;

3. Description of the study methodology, including not just the methodology used and the functional specification that produced the impact estimates, but also methodologies considered and rejected and interim analytical results that led to the final model specification. The intent of this section is to provide sufficient detail so that a trained reviewer will be able to assess the quality of the analysis and thoroughly understand the logic behind the methodology and final models that were used to produce the impact estimates;

4. How the study meets or exceeds Protocol requirements or, if any protocols were not able to be met, an explanation of why and recommendations for what it will take to meet these protocols in future evaluations;

5. How the study addresses the technical issues presented in these Protocols;

6. Sampling methodology and sample descriptions (including all frequency distributions for population characteristics from any surveys done in conjunction with the analysis).

The **Reliability Assessment** section of the report should focus on the targeted and achieved precision levels for the key findings presented, the sources of uncertainty in the approaches used and in the key findings presented, and a discussion of how the evaluation was structured and managed to reduce or control for the sources of uncertainty. All potential threats to validity given the methodology used must be assessed and discussed. This section should also discuss the evaluator's opinion of how the types and levels of uncertainty affect the study findings. Findings also must include information for estimation of required sample sizes for future evaluations and recommendations on evaluation method improvements to increase reliability, reduce or test for potential bias and increase cost efficiency in the evaluation study(ies).

The **Detailed Study Findings** section should include the following:

1. A thorough discussion of key findings, including insights obtained regarding why the results are what they are.

2. All output requirements and accompanying information shown in protocols 1 through 7 for ex post evaluation of event based programs, protocols 8 through 14 for non-event based programs, and protocols 15 through 20 for ex ante estimation. If the number of data tables is large, the main body of the report should include some exemplary tables and explanatory text with the remaining required tables provided in appendices. Detailed data tables should also be provided in electronic format.

3. For ex post evaluations of event-based programs, a table summarizing the relevant characteristics associated with each event and the date of each event over the historical evaluation period. At a minimum, the table should include for each event: date, weather conditions (for weather sensitive loads), event trigger (e.g., emergency, temperature, etc), start and stop times for the event, event duration in hours, notification lead time, number of customers notified, and number of customers enrolled.

4. For ex ante forecasts, detailed descriptions of the event and day type assumptions underlying the estimates

5. For ex ante forecasts, assumptions and projections for all exogenous variables that underlie the estimates for each forecast year, including but not necessarily limited to, the number of customers enrolled and notified (for event based programs), participant characteristics, weather conditions (if relevant), prices and price elasticities (if relevant), other changes in demand response over time due to persistence related issues and the reasons underlying the changes for the average customer.

6. A comparison of impact estimates derived from the analysis and those previously obtained in other studies and those previously used for reporting of impacts toward program goals, and a detailed explanation of any significant differences in the new impacts and those previously found or used.

APPENDIX A: ANNOTATED BIBLIOGRAPHY

This appendix contains summaries of selected studies that are good examples of the application of various evaluation methods. Due to time constraints, the summaries contained in this draft are limited. We may include additional studies in a subsequent draft.

2005 Smart Thermostat Program Impact Evaluation

Sponsoring Entity:	San Diego Gas & Electric	Sector:	Residential with A/C
Produced By:	KEMA	DR Type:	Direct Load Control

Available At:

PROGRAM OVERVIEW

SDG&E implemented the Smart Thermostat Pilot beginning in 2002 in response to a California Public Utilities Commission mandate. The report evaluated the impacts of the program during 2005, when the program was deployed for 12 events. The equipment installed in the roughly 4,000 enrolled sites allows SDG&E to remotely raise or reset the thermostat cooling set points for participating customers by 4 degrees for a 2 hour period. Resetting the cooling point upwards produces a drop in energy use for participating customers. Importantly, customers have the option to override the thermostat reset.

EVALUATION OBJECTIVES:

- Calculate the impacts for each reset period for potential contributors only.
- Calculate the fraction of units potentially contributing to savings for each event.
- Adjust the impact for potential contributors with the actual fraction of potential contributors to derive a per unit impact estimate for all program participants.

KEY FINDINGS

- The estimated impact across all events averaged 0.35 kW per unit with lower and upper bounds (90% confidence level) of 0.14 and 0.57 kW, respectively. A review of the more detailed estimates by hour and temperature are highly recommended.
- Seventeen percent (17%) of A/C units were not used during summer weekdays. The estimate may need to be adjusted for meters providing false zero readings.
- Across all events, an average of 8 percent of the participants did not receive reset signals.
- The override rate across the 12 events averaged 22 percent and ranged from a low of 9 percent to a high of 39 percent.

OBSERVATIONS ON METHODOLOGY

The study was selected primarily because of the strength of the research design and the measurement. It used an alternated assignment approach, meaning that participants switched between being a control and a treatment group. During each event, roughly half of the sampled participants acted as a control group while the other half acted as a treatment group, with their roles reversing for the next event. The sampled participants were randomly assigned to one group or the other. This design not only accounts for selection bias, it also provides powerful information about individual participant response patterns. It also efficiently makes use of the fact that there are multiple events. Importantly, the strength of the research design allows for multiple types of analysis ranging in complexity from direct comparisons to regression methods.

The study metered both the whole premise and the A/C unit. The A/C meters provided accurate measurement of the impact from the primary end-use affected by the program, as well as smaller variances, leading to more precision in the estimates. The whole building meter data accounted for the potential of altered end-use consumption such as increased fan use due to the thermostat reset. Having the two data sources allows the researcher to parse out the direct impacts of the thermostat reset on the A/C use from any behavioral adjustments due to the higher thermostat setting. The A/C metering also helps to identify the share of participants using A/C and to check whether the thermostat did indeed reset, an important fact since active demand response relies on reliable communication of an event or price signal.

The study also deliberately modeled estimates of A/C use and override rates across a range of weather scenarios, on an hourly basis, providing inputs needed for long term planning and grid operations. While the comprehensiveness of the analysis deserves accolades, the modeling of the override rate and its robustness was not entirely transparent in the report. The study employed a survival analysis approach, but did not provide details about the covariates included in the analysis, the model specifics, the fit, or whether or not the proportional hazards assumption of the model held true.

Evaluation of California's Real Time Energy Metering (RTEM) Program

Sponsoring Entity:	California Energy Commission	Sector:	Commercial and Industrial
Produced By:	Christensen Associates	DR Type:	Non-Event-Based Pricing

Available at: http://www.energy.ca.gov/2005publications/CEC-400-2005-021/CEC-400-2005-021.PDF

OVERVIEW

In March 2001, the California Assembly (in AB29X) authorized $35 million to install advanced automatic meter reading (AMR) devices for all customer accounts with peak demands greater than 200 kilowatt s(kW) in the state. In addition, customers who were not already on Time-of-Use rates (TOU) had their flat energy prices converted into TOU energy prices. The overall objective of the study was to measure changes in the consumer load patterns that could be directly attributed to the switch to TOU pricing or to providing customer access to information about their energy usage patterns.

EVALUATION OBJECTIVES:

- Calculate the impact of TOU rates on customer's energy use patterns.
- Calculate the impact of providing customers access to information about their energy usage patterns.
- Assess how customers make use of meter data and whether or not they self-reported taking load changing actions.

KEY FINDINGS

When taking into account demand charges, customers already faced higher summer peak rates (i.e., an implicit TOU rate). The conversion to TOU rates had only a modest incremental effect on peak period prices since those customers already faced effective TOU price signals due to the summer demand charge.

Customers responded to strong TOU price signals and adjusted their loads. Customers with weaker TOU price signals lowered peak demand, but by a lesser amount.

For customers who switched from the flat rates with demand charges to TOU rates, it is not possible to definitively attribute changes in energy to the switch. This is partly because they already faced an effective TOU rate due to the demand charges.

- In nearly every Standard Industrial Code (SIC) group, approximately 10 percent of the customer accounts showed evidence of strong price responsiveness, suggesting summer peak load reductions of 20 percent or more, with the larger customer accounts in SIC 2 through 4 showing the largest load reductions.

OBSERVATIONS ON METHODOLOGY

The study was selected because of a) the unusual set of challenges faced by the evaluator, b) the innovative methods of working around them, and c) using multiple types and levels of analysis to ensure robust results.

The evaluation had to be conducted without two fundamental evaluator tools: control groups and pre-treatment data (where the treatment is the change in prices). Because the meters and rate switch were mandated in the midst of the California power crisis, the evaluators: lacked a control group, since all customers had switched to mandatory TOU prices; lacked pre-TOU interval data for customers who recently had meters installed and switched to TOU prices; lacked comparable pre AB29X data for the customers who had been on TOU prior to AB29X. While the data existed, it reflected customer behavior during a crisis period during which customers were strongly encouraged to reduce consumption, particularly during peak periods.

To further complicate the analysis, the switch to TOU pricing provided a weak signal. Prior to the switch, PG&E and SCE customers faced summer demand charges that effectively sent a TOU-like price signal. Moreover, the new TOU energy prices provide a relatively low Peak/Off-peak price ratio and relatively small incremental price signal.

To overcome these challenges, the evaluation relied largely on the fact that the SCE and PG&E TOU rates differentiated strongly between the summer and non-summer months. The analysis approach was designed to measure differences in energy usage during summer peak periods relative to the same time period in non-summer months, after controlling for the effect of weather and other variables expected to impact energy consumption. The approach relied on three separate regression equations at two different levels of analysis. Regression functions were develop to detect changes in 1) the energy use by time period, 2) the share of energy consumed by time period, and 3) daily energy use. The study included pooled analyses of customers based on their Standard Industry Code (SIC) and individual customer-level analyses. The analyses provided estimates of the TOU impacts for both customers who were on TOU pricing before AB29X and those who switched to TOU pricing as a result of the legislation.

Overall, the study quickly identified its unique evaluation challenges and presented an approach that provided sufficiently robust and insightful results despite constraints. Importantly, the robustness of the study is the direct result of the multiple types and levels of analysis and their internal consistency. Without pre-treatment data and control groups, regressions are more prone to model specification errors (which potentially bias impact estimates), making it particularly important to ensure the results are robust and credible. While, ideally, evaluators would have access to pre-treatment data and/or well-matched control groups, these options may become less available over time, particularly due to the possibility of default or mandatory DR tariffs and/or programs.

The 2003/2004/2005 Energy-Smart Pricing Plan^{SM} Evaluations

Sponsoring Entity:	Community Energy Cooperative, Chicago, IL	Sector:	Residential
Produced By:	Summit Blue Consulting	DR Type:	Non-event based pricing (RTP)
Available at:	http://www.energycooperative.org/reports.php		

PROGRAM OVERVIEW

The Energy-Smart Pricing Plan (ESPP) was initiated in January 2003 by the Community Energy Cooperative, in cooperation with Commonwealth Edison (ComEd). The program sets hourly rates that are made available to customers one day in advance. A price cap of $0.50 per kWh is in place, and consumers receive a phone and/or email alert whenever prices are expected to exceed $0.10 per kWh. In 2003, 650 customers were enrolled in the pilot program, and an additional 100 were placed in a control group that received all of the program materials and had an interval meter installed, but were not put on the real-time pricing plan. The pilot had more than 1,000 residential customers enrolled in 2004, including 57 who had switches installed on their A/C units that could cycle the unit at 50% during high price periods. Through 2004, participants had realized an 11% cumulative savings on their bills.

EVALUATION OBJECTIVES:

- Evaluate the responsiveness of customers to hourly market-based electricity price information.
- Evaluate the impacts of day-ahead hourly RTP information on customer electricity use.
- Evaluate the impacts of A/C cycling on customer loads.
- Evaluate how customer satisfaction and behavior patterns changed over time, particularly given that the summers of 2003-2004 were significantly cooler (with correspondingly lower prices) than the summer of 2005.
- Determine whether any differences exist between the types of customers who joined the program in the first year versus the second year.

KEY FINDINGS

- Customers responded to hourly rates, and impacts observed in the program's first year persisted in its second and third years.
- Customers' average price elasticity to pricing information was calculated to be -.080 in the 2004 report.
- There was little behavioral response to "high price hour" notifications (that prices would exceed $0.10 per kWh) in 2004, in part because that summer was quite cool in the survey area and little A/C was used.

- In 2005 (a hotter summer), the overall price elasticity was calculated to be -.047. However, customers were also reported to be more responsive to the high-price period notifications in 2005 (doubling their average response when they successfully received a notification).
- In 2005, customers' response to high-price period notifications was observed to decline somewhat as the number of consecutive notification days increased, and as the length of the high-price period increased.
- Customers in multi-family homes had greater price elasticities than those in single-family homes. Also, customers with no A/C were the most price elastic, followed by those with window A/C units. Customers with central A/C were least price elastic. These effects are likely due to income effects.
- Customers are generally pleased with the program, and 98% believe that participating in the program has lowered their energy bills.
- Compared to those who enrolled in the program at its outset, participants who enrolled during the second year, on average, had higher incomes, larger households, greater internet access, and greater use of central A/C.

OBSERVATIONS ON METHODOLOGY

The study was selected as an example because a) it is one of few existing residential RTP programs, b) it carefully collected and analyzed data about differences between participants and non-participants, c) it included a detailed analysis of high versus low responders, and d) it sought to model the response as the duration of a high price period progressed.

The study used a fixed-effects panel data model to relate hourly electricity consumption (based on fifteen minute load data) to weather and pricing information. The model regressed the change in hourly electricity consumption on the changes in hourly electricity price (price elasticities), and weather variables.

The study also modeled customers' decisions about whether or not to enroll in the program and used that information to correct for selection bias. This was accomplished by collecting data about the characteristics of participants and non-participants that were well aware of the program. The participation model was used to address the fact that those who chose to be in the pilot were self-selected due to unique characteristics. In this study, including a self-selection variable in the energy use model did not change the results. In other words, while there are differences between participants and non-participants, these differences are not correlated with customers' responses to hourly energy prices. While, technically, differences between participants and non-participants do not affect the impacts of the program, accounting for the selection decision is important if the results are to be extrapolated to non-participants. In other words, the selection decision becomes more of an issue when it comes to the external validity of the results.

Finally, the study segmented the participants into high and low responder groups and used the information about household characteristics to better define how different characteristics correlated with the overall demand response provided by households. An alternative more insightful and more labor intensive approach would be to calculate price elasticities for individual customers and then assess how the household characteristics correlated with price responsiveness. This would allow for ex ante estimates of response if the program were expanded and/or targeted at a different participant mix.

Importantly, the study made use of a phased implementation to create a randomly assigned control group of customers that self-selected into the program. The control group customers experienced all the non-price incentives and benefits of the Energy Smart Pricing Plan in the first summer of the study and were subsequently enrolled into the program. This not only helped ensure more robust results for the initial year of the study, it also provided a benchmark against which later analyses without a control group could be compared with, both in terms of changes to the results and the accuracy of the model specification.

The 2004 study was also notable in that it employed A/C cycling for a small subset of the treatment group in order to determine its effects on load. Unfortunately the study provides insufficient detail about how these participants were selected or how the impact evaluation was made.

NYISO Price Responsive Load Program Evaluations (2001-2004)

Sponsoring Entity:	NYISO and NYSERDA	Sector:	Commercial, Industrial, and allowed load aggregation in residential sector. Minimum bids of 100kW.
Produced By:	Neenan and Associates, LLP	DR Type:	Event-Based Pricing and Callable DR

Available at: http://www.nyiso.com/public/products/demand_response/index.jsp

PROGRAM OVERVIEW

The NYISO implemented two price-responsive load (PRL) programs, the Day-Ahead Demand Response Program and the Emergency Demand Response Program, during the summer of 2001 and has continued them until the present. The Day-Ahead Demand Response Program (DADRP), allows industrial, commercial, and aggregations of residential customers to offer demand reduction bids into New York's day-ahead electricity market to help reduce system demand and receive market prices for any load reduction. The Emergency Demand Response Program (EDRP) targets the same set of potential participants but has the primary goal of bolstering reserves during times of system emergency. Participants in EDRP are notified at least two hours in advance of when emergency system conditions are imminent, and are guaranteed a minimum price for any load curtailment during the emergency period.

Customers that qualify their load curtailment capability also can sell their capacity as a special case load resource, allowing Load Serving Entities to meet their installed capacity requirements. Roughly 40% of customers in the price-responsive load programs also offer their load into the NYISO installed capacity market (ICAP). Participation in the capacity market offers up-front payments that the price-responsive load programs do not. There is a tradeoff, however. If the NYISO calls an event due to reserve shortfall, customers who serve as special case resources must provide the load curtailment or face penalties assessed for noncompliance. Importantly, curtailments called under the Emergency Demand Response Program have been mostly coincident with the calls for ICAP special case resources.

EVALUATION OBJECTIVES:

- Calculate the impact of load reductions on the locational based marginal prices (LMBPs) for both the day ahead and the real time markets.
- Characterize the load response of existing program participants to higher electricity prices or other program payments for load curtailments.
- Identify the shape of the short-run electricity supply curve for the energy markets and the impact of DR on them.
- Identify the factors that lead customers to participate in the current programs.
- Estimate the amount of load reduction that enrolled customers will contribute to the Emergency Demand Response Program.

- Estimate the amount of load reduction that will be bid into the Day Ahead Demand Response program at different prices.
- Understand how changes in current program designs are likely to affect customer participation rates, particularly for recruiting customers in future years.

KEY FINDINGS

Institutional and industrial customers had the highest price sensitivity and office buildings had the lowest price sensitivity.

There is substantial variation in customer's price responsiveness. The firm-level variation reflects differences in the ability of customers to respond on certain days, and it also reflects differences in the Customer Base Loads (CBLs) against which performance is measured.

- The implicit elasticities of response varied considerably with the size of the firm, as defined by their average electricity usage. Participants with low elasticities were, in general, equally distributed amongst the firm size bins. However, as firm size increased, so did the percentage of participants with high elasticities of response (elasticities less than -0.20).
- Most participants without on-site generation remained steadfast in their load curtailment contribution once they committed to the EDRP event.

OBSERVATIONS ON METHODOLOGY

The NYISO price responsive load evaluations were selected because the evaluations performed work on a) estimating the load reduction bid into markets (as opposed to the amount enrolled), b) quantifying the factors that drive load response and participation, and c) examining the effect of customer participation in multiple programs. In addition, the reports are notable for their work on using load impacts to quantify the impacts of DR on energy markets.

One of the distinguishing factors of the evaluation of the NYISO programs was the lack of access to hourly data, and the innovative methods that were employed to assess the price responsiveness (the decision to participate and demand reduction) of customers enrolled in the program. The NYISO settlements relied on a day-matching approach that on their own provided limited information on price responsiveness and/or how different customer, event, and market characteristics/conditions affected overall responsiveness. To better make use of the information, implicit price elasticities were estimated for all the customers. The price elasticities where then sliced to reveal price elasticities by customer (unpublished), zone, customer type, customer size, etc. To calculate the implicit price elasticities, the study used information about the tradeoffs faced by customers when deciding whether to participate and how to bid.

The evaluation also carefully distinguished between subscribed load reduction and actual load reduction, creating two performance measures used to characterize and compare program participants: load reduction provided relative to subscribed load reduction, and the share of potential load reduction deployed.

Because one of the evaluation goals was to better understand how changes in program designs would affect customer recruitment and responsiveness, it included the analysis of the revealed preferences and stated preferences. The revealed preferences method reflected the relative impact of various factors on actual decisions about whether or not to participate in the program. The revealed preferences analysis relied on a survey that posed hypothetical program variants, allowing for a level of variation in program designs that could not normally be introduced in practice, and computed the relative impacts of each of the program design attributes.

Overall, the evaluation provided a comprehensive picture of the amount and reliability of demand response, as well as the ability to assess how changing the program design would affect both program enrollment and responsiveness.

Time-of-use Prices and Electricity Demand: Allowing for Selection Bias in Experimental Data (1997)

Utility:	Ontario Hydro	Sector:	Small C&I <50kW
Produced By:	John C. Ham, Dean C. Mountain, and M.W. Mountain	DR Type:	Non-event based pricing

PROGRAM OVERVIEW

Ontario Hydro implemented a TOU pricing experiment to assess the price responsiveness of small and medium commercial and industrial customers. The experiment involved 120 customers and their interval data was collected from October 1985 to October 1987. It tested three different TOU rate structure, each with different peak hours and peak to off-peak price differentials

EVALUATION OBJECTIVES:

- Estimate load shifting and load reductions by hour
- Account for selection bias
- Assess how customers respond to different rate structures and price differentials

KEY FINDINGS

- Small and medium commercial customers showed significant price responsiveness given a sufficiently short peak period (5 hrs) and sufficiently large peak/off-peak price differential

OBSERVATIONS ON METHODOLOGY

The study was selected primarily because of the informative structure of the research design and the innovative work on selection bias. Prior to attempting to enroll customers in the experiment or exposing them to time-varying prices, Ontario Hydro placed meters at 120 customers sites and collected a full year of data. The 120 customers were randomly assigned to four groups of thirty – one control and three treatment groups. Each of the treatment groups were offered TOU pricing with different peak period lengths and peak to off-peak differentials. While all customers were invited to participate in the TOU pricing trial, customers selected into the treatment (the time-varying prices), introducing the potential for selection bias. The research performed in noteworthy on two areas.

The structure of the research provided an opportunity to clearly assess both the share of customer accepted the offer as well as the drivers for acceptance rates. The roll out of the experiment and information requested provided the opportunity to create discrete choice model and, importantly, assess how peak-period length and pricing differential affected participation.

The study performed innovative work in controlling for selection bias. While the data collected allowed the researchers to control for selection bias using Heckman's selection bias procedure (citation), the innovation was in creating a selection bias procedure that could be tested and did not depend on a well defined participation equation, distributional assumptions, or exclusion restrictions.

While Heckman's approach and its successors have been widely praised and criticized, a difficulty in applying Heckman's approach is in specifying the selection rule and, in particular, the variables in the participation equation that are not included in the regression. The procedure is also highly sensitive to violation in the assumptions.

The authors took a different selection bias correction approach that relies on a realistic and testable assumption: that establishments make participation decisions on the basis of permanent, not transitory, factors. The procedure, however, does require the use of pre-treatment usage data. For full details on the selection bias correction used and how to test the key assumption please refer to the full paper.

Switches or Stats: Who Wants What? A Comparison of Switches and Web-enabled Programmable Thermostats for Mass Market Demand Response

Utility:	Wisconsin Public Service Corporation	Sector:	Residential
Produced By:	Mary Klos	DR Type:	Direct load control & smart thermostats

PROGRAM OVERVIEW

WPSC conducted a residential thermostat pilot during the summer of 2005, installing web-enabled programmable thermostats in 86 residential homes in northern Wisconsin. Concurrently, a larger direct load control study was also taking place, and hourly data for 1,170 of these customers was also collected for this study. Impact evaluations were done to compare the effects of different control strategies for the two technologies, and data were also collected on customers' notice and comfort levels during the load control tests.

EVALUATION OBJECTIVES:

- Estimate load shifting and load reductions by hour for various load control strategies.
- Estimate customer comfort level and percentage of customers who noticed control protocols.
- Compare load impacts and customer comfort levels associated with direct load control and smart thermostat control strategies.

KEY FINDINGS

- The average load impacts of a 50% cycling strategy with regular switches are approximately equivalent to those of a 4-degree rise with thermostat control. Thus similar load reductions can be induced using smart thermostats or switches (which are far less expensive).
- Customers reported relatively high comfort rates, and relatively low rates of noticing the control program for both strategies, although comfort and notice levels were better for the switching strategy than the smart thermostat strategy.
- Relatively low opt-out rates (3%-9%) were reported for the thermostat pilot program.

OBSERVATIONS ON METHODOLOGY

The study was selected primarily because of it is one of the only comparisons of smart thermostat and direct load control load impacts, as well as the fact that it measures customer comfort and notice levels under both strategies. WPSC collected hourly load data, as well as Composite Temperature Humidity Index (CTHI) data for the 1,170 customers in the direct load control pilot program and the 86 customers in the smart thermostat program throughout the summer of 2005.

These customers were placed in one of six different treatment groups – there were three direct load control groups and three smart thermostat groups, each with its own control, notification, and rate incentive protocol (and sometimes a variety of different protocols were used within a single group over the course of the summer).

A primary weakness of this paper, however, is that it completely ignored the issue of selection bias – it never addressed how customers were selected to participate in the experimental program in general, or how they were assigned to the various treatment groups. In fact, toward the end of the paper it is suggested that one of the thermostat groups is significantly different from the population at large, with greater energy and A/C use, implying that selection bias was a problem but that it was never addressed. However, hourly load data was also collected on a control group of 1,500 residential customers with A/C. Comparing the average load curves of the treatment groups with this control group indicated that only one of the treatment groups was significantly different from the control group, although this difference was not directly addressed.

The load impact estimation technique was based on a regression technique. An hourly reference level load curve was generated for each customer by regressing usage data on two variables: hour of the day, and weather (CTHI). The impact is then calculated by taking the difference between the customer's actual usage on an event day and his reference level usage for that day. Impacts for each treatment group were then averaged over all event days in the summer to arrive at an average load impact value for each treatment.

Customers participating in the experiment were sent monthly surveys in which they were to record their daily observations about comfort level and whether they noticed the external control of their load. Customers reported relatively low levels of discomfort and notice in both the direct load group and the thermostat group, although discomfort and notice levels were significantly lower in the direct load group. However, a direct comparison is not fair, because the thermostat group experienced more extended periods of 100% load shed and also had an indoor control indicator (instead of an outside indicator on their switch box) which they were much more likely to notice.

Joint IOU Framework
Framework for Evaluating Cost Effectiveness of DR Resources
July 16, 2007

1 Executive Summary

This Chapter summarizes the main recommendations of this Framework for Evaluating the Cost Effectiveness of Demand Response (DR) programs. The reader should refer to the chapters of this document to understand the complete details of that framework and the rationale for each of these recommendations.

1.1 Summary of Recommendations

1.1.1 Recommendation 1

Evaluation of the cost effectiveness of Demand Response (DR) programs should be consistent with the Standard Practice Manual (SPM) where possible. However, to provide the proper guidance for evaluating the cost effectiveness of a DR program, the SPM must also be modified and extended due to the significant inherent differences between Energy Efficiency (EE) programs and Demand Response (DR) programs, and between different types of DR programs.

1.1.2 Recommendation 2

The focus of this DR cost effectiveness evaluation framework is on providing guidance for forecasting DR impacts for long-term resource planning.[1] The methods described in this framework should be used for *ex ante* evaluation of DR cost effectiveness. *Ex ante* cost-effectiveness evaluations should be based on the most recent expected values of energy and capacity market prices. Using outdated earlier expectations of those market prices would produce unreliable results.

Ex post cost evaluations of the cost effectiveness of DR would not be appropriate, because in general demand response resources provide "insurance" against low probability events that have severe consequences when they occur. If those extreme events did not occur during a given time period, it does not necessarily mean that those demand response resources were less valuable or less cost effective ex post.

This framework is applicable both to individual DR programs and to a DR portfolio,[2] and includes methodologies to evaluate the cost effectiveness of both event-based DR and non-event based DR (*e.g.,* time differentiated rates and dynamic pricing options).

1.1.3 Recommendation 3

Unlike energy efficiency programs, most of the benefits DR programs provide are related to avoiding relatively low probability future events (*e.g.,* unusually high peak demand and/or energy prices) in relatively few hours, whose occurrence could have significant economic consequences. Furthermore, event-based DR programs provide options whose value is related to the degree of uncertainty about exactly when those events will occur, and the actual magnitude of the resulting consequence.

---

[1] Staff Guidance, pp. 9 – 10.
[2] Staff Guidance, p. 35.

Therefore, the cost effectiveness of most DR programs usually should be evaluated on an hourly basis over relevant time periods by using analytic methods that take into account the uncertainties about when and how often these types of circumstances will occur, the magnitude of the resulting economic consequences, and the impact of behavior on the load changes that occur under those DR programs.

1.1.4 Recommendation 4

This framework recommends that the (All Customers) Total Resource Cost (TRC) test and/or Societal TRC test should be used to evaluate the overall cost effectiveness of DR programs. These tests net out the impact of transfers between participating and non-participating customers. As a result, all of the remaining costs and benefits are actual resource costs or resource cost savings.

The other SPM tests should be used to evaluate the distributional impact that DR programs, and should be investigated as part of an overall DR program evaluation.

1.1.5 Recommendation 5

Staff Guidance directs that DR cost effectiveness methods should assume that DR programs ought to be viewed as an alternative to supply side resources from a long-term planning perspective. Therefore, the cost effectiveness of DR programs should be evaluated by comparing the costs and benefits of DR programs, to the costs and benefits of the alternative supply-side resources.

1.1.6 Recommendation 6

Evaluations of cost effectiveness of DR programs should distinguish between avoided capacity costs, and avoided energy costs, rather than using only “all-in” avoided costs.

The use of an “all-in” avoided cost approach would preclude the use of the types of stochastic techniques that are necessary to value the avoided capacity costs and avoided energy cost benefits that can be provided by event-based DR.

1.1.7 Recommendation 7

DR programs, by their design, avoid the need for generation capacity since their function is to reduce customer usage during periods of peak demand. The amount of DR capacity and supply-side capacity that the Commission requires each LSE to maintain, from a resource planning perspective, is determined by the Resource Adequacy (RA) requirements established by the Commission.

As a result, the extent to which a DR programs enables an LSE to avoid generation capacity costs, depends upon the extent to which the Commission’s RA “counting rules” allow that LSE to count the capacity of that DR program in complying with its RA requirement.

DR cost effectiveness evaluations should not attribute any avoided capacity cost to a DR program that an LSE cannot count in meeting its RA requirement.

1.1.8 Recommendation 8

The RA requirements for each LSE currently require that LSE to maintain a 15 to 17% capacity reserve margin over that LSE's share of a "1 in 2 temperature peak demand" in the control area (as well as local RA capacity requirements in certain transmission constrained load pockets).

A DR program that an LSE cannot count in meeting its RA requirement, or which would cause the LSE to have more capacity than it needs to comply with that RA requirement, may nonetheless provide a higher level of reliability than that reflected in the planning reserve margin adopted by the Commission.

Therefore, this framework recommends that the methods used to evaluate the cost effectiveness of DR capacity should not incorporate the implicit assumption that a higher planning reserve margin is needed, by assigning avoided generation capacity value to DR capacity that would cause LSE to exceed its RA requirement.

1.1.9 Recommendation 9

The generation capacity costs avoided by a DR program that a Load Serving Entity (LSE) can count toward its RA requirement is the market value of the additional generation capacity that LSE would otherwise need to build or procure in order to comply with that requirement.

DR usually avoids the need for relatively high heat rate generating capacity that is only dispatched during peak demand periods. In the long run, the market value of that type of generating capacity in California will be based on a new natural gas-fired combustion turbine (or CT). Therefore, new natural gas-fired CT capacity should be used as a proxy to derive the market value of the generation capacity avoided by demand response programs (*i.e.,* generation capacity costs).

1.1.10 Recommendation 10

Although market prices for energy are available, market prices for capacity are not. Therefore, the market value of the capacity of new CT capacity (*i.e.,* the generation capacity cost avoided by DR) must be estimated. In the long run, additional peaking generation capacity will be built only if prices for energy and capacity are expected to be high enough to recover the variable and fixed operating costs of that capacity, the amounts invested in the construction of that generating capacity. The gross margins that new CT capacity is expected to achieve (*i.e.,* revenues from selling energy at wholesale market prices, less the variable fuel and O&M costs incurred in generating that energy) will cover part of the fixed costs of that capacity.

Therefore, estimates of the market value of CT generation capacity costs should be derived as follows. First, the lifecycle market value of the capacity of a new CT should be estimated by subtracting:

(1) the present value of the gross margins the new CT capacity is expected to earn from selling energy at those times when revenue from wholesale electricity market sales exceeds variable operating cost; from,:

(2) the present value of the total fixed costs of that new CT.

The resulting lifecycle "net" capacity cost should then be converted to an annual real economic carrying charge, and treated as an estimate of the annual market value of new CT capacity in each year.[3]

Eventually, California may have a liquid, robust, public market for generation capacity. At that time, the Commission should consider requiring LSEs to utilize the market prices of long-term generation capacity from that source to evaluate the cost effectiveness of DR programs.

1.1.11 Recommendation 11

The resulting estimates of the generation capacity costs avoided by DR program should also be adjusted upward, to reflect the T&D line losses avoided by that DR program capacity, and the capacity planning reserve margin avoided by that DR program on a case-by-case basis.

1.1.12 Recommendation 12

In each year, the portion of the annual market value of the new CT capacity that will be avoided by RA-qualified DR program capacity depends upon the hours in which that DR capacity will be available in that year. Therefore, in order to estimate the generation capacity costs that will be avoided in each year by that DR capacity, it is necessary to estimate the hourly market value of new CT generation capacity (in $/kW-hour) in each hour of that year.

Estimates of hourly market prices for generation capacity in each utility area can be derived from the estimated annual market price of generation capacity ($/kW-year) in that year, by allocating that annual market price of capacity among the hours in that year, in proportion to the relative need for generation capacity in each hour.

That need, in turn, is proportional to the probability that some of the load in that hour in that region will not be served (*i.e.,* unserved energy) due to a lack of sufficient generation capacity (*i.e.,* a loss of load). The likelihood of that occurring in a given hour is the "Loss of Load Expectation" (LOLE) for that hour.

Therefore, this framework recommends allocating the annual market value of new CT capacity among the individual hours in each year, in proportion to the loss of load expectation (LOLE) in each hour.

1.1.13 Recommendation 13

The capacity benefits of a DR program should be adjusted for differences between the DR program and the capacity value of a new combustion turbine.

DR programs may contain a number of constraints that affect the capacity value of the program relative to the capacity value of a new combustion turbine including seasonal operation, limits on the number of DR program event calls, the available hours of each program event call, the number of consecutive days duration of each program event, and the total number of program event days available. Further, there may be differences in program event notification periods that impact value.

These differences should be accounted for, and Section 5 and Appendix A describe the methods that PG&E and SCE have utilized for this purpose (call option valuation models and Monte Carlo runs of dispatch simulations).

1.1.14 Recommendation 14

DR also enables LSEs to avoid energy costs. An LSE can always buy or sell energy in the wholesale energy market. Therefore, the expected wholesale market price of energy in each future time period is the relevant opportunity cost for ex ante estimates of the value of the energy that will be avoided by a demand response resource (or provided by a supply side resource).[4]

In the case of the energy avoided by a demand response resource, the avoided energy costs at those wholesale market prices should also be adjusted upward, to reflect the distribution line losses that demand response load reductions would avoid in those hours.

1.1.15 Recommendation 15

Non-event DR program load reductions in any given hour are fairly predictable. Therefore, future energy costs avoided by non-event DR programs should be based on expected hourly prices for the hours in which the DR program is expected to reduce load.

In contrast, DR program events are most likely to be called in hours when prices are higher than normal (*i.e.*, higher than expected). Therefore, expected hourly prices tend to underestimate energy prices in the hours in which an event-based DR program will actually reduce loads, if there is a correlation between the occurrence of a condition that would trigger the program, and the actual price of energy in that hour.
The higher than expected energy prices that are likely to occur at those times when a DR program event is triggered, can be estimated by using methods. Therefore, those methods should be used to estimate the energy costs avoided by event-based DR programs.

1.1.16 Recommendation 16

Although important, avoided energy costs usually account for only a small share of the total costs avoided by a demand response program, but their estimation is complex. Estimates of avoided energy costs for event-based DR programs should reflect constraints on their availability and exercise. The methods and criteria they require to estimate avoided those avoided energy costs also are technically challenging. Given that avoided energy costs account for a small fraction of the value of most demand response programs, the additional analytic complexity associated with estimating avoided energy costs more precisely may not be warranted if more precise methods do not have a material impact on estimates of cost-effectiveness.

4 Unless noted otherwise, throughout this straw proposal chapter the terms “energy price” and “electricity price” refer to wholesale market electricity prices.

1.1.17 Recommendation 17

If a demand response program can provide highly predictable load reductions on very short notice in very specifically defined locations on the grid, that program may enable an LSE to avoid or defer investments in transmission and/or distribution capacity. However, making that determination for specific demand response programs requires very detailed, geographically specific T&D studies, and analyses of substantial data on the historical load reductions that DR program provided in the past.

Therefore, whether or not it would be feasible to estimate potential avoided or deferred T&D capacity investments should be determined on a case-by-case basis.

1.1.18 Recommendation 18

Environmental benefits that DR resources may provide should not be calculated separately, except for greenhouse gas benefits.

<u>Reductions in Emissions of Criteria Pollutants</u>

Capital and operating costs of utility area specific CT capacity proxy resources used for avoided generation capacity costs estimates should include capital costs incurred to comply with existing environmental regulations including acquisition of offsets for criteria pollutants (NOx, PM 10, VOCs, SOx). In that case, the criteria emission pollutant-related costs that can be avoided by DR programs are already reflected in estimates of the capacity costs avoided by that DR program.

<u>GHG Emissions</u>

Implementation of AB32's requirement to reduce GHG emissions starting in 2012 is still in formative stages. Federal action on GHG emissions reductions may occur. There is considerable uncertainty about future GHG emission compliance costs that a DR program might avoid. At present, this framework recommends that the value of the GHG emissions that a DR program would avoid, starting in 2012, be estimated in a manner that is consistent with Commission direction in D.05-04-024. In that Decision, the Commission suggested that a reasonable estimate of the avoided benefits associated with avoided GHG emissions would be $8 per ton in 2004, escalated in later years. The estimates of the volume of GHG emissions avoided by a DR resource would be based on the operating and emission rate characteristics of utility-specific new CT capacity.

This approach to estimating the value of the GHG emissions avoided by a DR program should be re-evaluated when sufficient additional information becomes available on federal and state programs to limit GHG emissions, including AB32.

<u>Land and Water Quality Impacts</u>

There are several other environmental impacts that might be avoided depending on the specific type(s) of capacity– generation, transmission, or distribution -- that the DR program is expected to defer or avoid. These potential environmental impacts include environmental justice (particularly for supplying electricity in urban areas), biological impacts, impacts on cultural resources, diminishing visual resources (*e.g.,* due to power plant stacks or transmission towers), water quality/consumption, and noise pollution.

As with criteria pollutants, the preferred approach is to incorporate these benefits in cost-effectiveness evaluation of DR programs by incorporating the compliance costs in avoided energy, avoided generation capacity, and avoided T&D capacity costs.

However, there may be residual benefits in addition to existing compliance costs. In most instances, these residual benefits are extremely difficult to value in a reliable manner, and therefore should not be included in evaluating the cost-effectiveness of DR programs. In specific situations where those additional environmental impacts clearly cause regulatory agencies to impose significantly higher control costs or fines, the value of the additional environmental impacts can be based on the additional control costs or penalties. Those values can then be included in cost effectiveness analysis of DR programs that would avoid those costs.

1.1.19 Recommendation 19

The reliability of electricity service can be increased by adding either DR capacity or supply-side generation capacity. To the extent that the reliability benefits provided by DR programs and by supply-side alternatives are the same, cost effectiveness evaluations of DR programs should not consider DR reliability benefits that would instead be attributable to adding additional capacity due to higher RA requirements.

Assertions that DR can provide a larger physical hedge against extreme events, lower “insurance” costs of events, more opportunity to manage risk, more valuable ancillary services, and larger physical hedging benefits are also incorrect. All of these benefits also could be achieved by increasing peaking generation relative to expected peak load. Therefore, the substitution of DR resources for supply-side capacity with the same operating characteristics would not provide any more of these benefits than could be provided by that supply-side alternative.

More importantly, these asserted benefits should not be considered because they implicitly assume a change in the level of reliability achieved by compliance with RA capacity requirements that are based on the capacity planning reserve margin the Commission adopted and is reviewing in another proceeding.

1.1.20 Recommendation 20

Some assert that cost effectiveness evaluations of DR programs should take into account the fact that adding DR resources can provide improve market performance by reducing market power, reducing energy price volatility, provide financial hedges against energy procurement costs, and/or lower the market prices paid for energy provided to customers that do not participate in DR programs.

Even if adding more DR resources would achieve those benefits, those benefits also could be obtained by adding more supply-side capacity. Therefore, these asserted benefits are not additional benefits provided by a DR program, but rather reflections of the fact that adding additional DR and/or supply side capacity would increase capacity reserve margins.

In other words, the Commission could obtain these benefits by requiring LSEs to maintain higher capacity reserve margins, and therefore comply with higher RA requirements by incurring additional, not lower, capacity costs

Cost effectiveness evaluations of a DR program should not include benefits of additional DR capacity that could also be obtained from additional supply-side capacity, such as increased efficiency in price signals, market power mitigation, reductions in volatility of market energy prices, physical hedging benefits, physical hedges against procurement costs,

1.1.21 Recommendation 21

The suggestion has been made that DR programs provide system operators with "more flexible resources" to meet contingencies than supply-side alternatives. This view ignores the challenges with maintaining customer enrollment and participation if DR programs go on hold. Also, quick start supply-side resources also may provide the flexibility to meet contingencies. Where the option of substituting a new CT in time to avoid potential rotating outages in the near-term is not available, the modular nature of DR may a benefit, but a specific analysis is recommended. The value of the lost load associated with these benefits may be quantified but considerable judgment is required both in selecting and applying information on value of lost load in these instances.

1.1.22 Recommendation 22

Other potential benefits associated with DR programs (*e.g.,* portfolio benefits, future potential reductions in DR technology costs, and spillover energy efficiency benefits) may exist, but are difficult to quantify and highly speculative for reasons discussed herein. Therefore, it would is inappropriate to include these asserted benefits in evaluating the cost effectiveness of DR resources.

1.1.23 Recommendation 23

The replacement of "flat rates" by time-differentiated or dynamic prices provides customers with the information required to more closely match their use with the costs of supplying the electricity. Without time-differentiated rates, customers "overpay" for electricity during off-peak hours and "underpay" for electricity during on-peak hours. In addition, customer utilization is below the optimal usage levels during off-peak hours, and above the optimal usage levels during on-peak hours. The economic result of flat rates is what has been termed a "dead-weight loss" or resource loss.

Although time-differentiated or dynamic prices pricing can reduce those dead weight losses, developing reliable estimates of the magnitude of that potential benefit is usually difficult without using confidential commercially sensitive information (*e.g.*, the expected net short or long positions of an IOU during the hours in which DR resources are most likely to be utilized) that has been given confidential treated in Commission decision D.06-06-066.

1.1.24 Recommendation 24

DR programs require customers to incur some degree of physical discomfort, some change in business operations, or some loss of business profits.[5] Because the customer cost is unknown, and are likely to be highly variable among customers, a conservative assumption is that customer costs are equal to the incentive

[5] In contrast, energy efficiency programs encourage more efficient usage of electricity while providing the same benefits to customers.

payment under a voluntary DR program. If customer costs were higher than the incentive, the customer would not participate.

Because a TRC (or all customers) test and a Societal TRC test only take into account the costs and benefits to all customers or society as whole, incentive payments are ignored because they do not change the total benefits and costs to all customers combined, or to society,

Therefore, in evaluating the cost effectiveness of DR programs, incentives paid to participants in DR programs should be treated as transfers from non-participating to participating customers, except in the case of voluntary DR programs.

1.1.25 Recommendation 25

Utilities currently evaluate supply-side resources using different methods for dealing with future uncertainty. It is more important to allow diversity of methods for DR cost effectiveness to be consistent with supply-side evaluations than to require a single prescriptive method.

1.1.26 Recommendation 26

Although evaluations of the cost effectiveness of DR programs should eventually take into account geographic differences into capacity and energy costs avoided by those programs, it is not yet possible to do so. Therefore, this framework recommends that the initial round of DR program design that takes place for the 2009-2011 program cycle should not include any geographic granularity.

1.1.27 Recommendation 27

Most DR program incentive payments are paid upon actual load reductions delivered during a DR program event. The straw proposal recommends that these costs be included in the cost effectiveness analysis of a voluntary DR program on a probability basis in the exact same fashion as avoided energy prices. The incentive costs will be based on probabilistic estimates of the number of DR program event calls and the associated hours that are expected based on the methods described in Chapter 6 to calculate avoided energy payments.

1.1.28 Recommendation 28

The DR cost effectiveness evaluation methods presented under in this framework promote transparency by using published data and public data sources where practicable. While these DR cost effectiveness methods attempt to promote transparency wherever practical, confidential, sensitive or proprietary data and analyses underlying utility cost effectiveness are entitled to the confidentiality protections recognized in Commission decisions.[6] .

[6] *See* Section 454.5(g) of the California Public Utilities Code and D.06-06-066.

2 Cost Effectiveness Evaluation Objectives and Scope

2.1 Background

California regulatory agencies have a long history of attempting to ensure that demand-side management (DSM) activities are cost effective. In 1983, the California Public Utilities Commission (Commission), in conjunction with the California Energy Commission (CEC), established the Standard Practice Manual (SPM) to create greater uniformity in the assessment of DSM programs.[7] However, DSM practitioners have encountered difficulties in developing the inputs for and applying SPM tests to the DR programs of the California investor-owned utilities ("IOUs").[8] [9]

Noting that an industry-accepted methodology for evaluating cost effectiveness of DR programs had not yet been established, in 2005 the Commission directed the Energy Division to recommend whether to open a proceeding to develop measurement and evaluation protocols and cost effectiveness tests for DR.[10] After a March 2006 workshop devoted to examining DR measurement and evaluation and cost effectiveness issues, Energy Division recommended that a Commission rulemaking be opened to further examine these issues.

Rulemaking (R.)07-01-041 was opened on January 25, 2007 to (i) establish a comprehensive set of protocols for estimating the load impacts of DR programs; (ii) establish methodologies to determine the cost-effectiveness of DR programs; (iii) set DR goals for 2008 and beyond, and develop rules on goal attainment;, and (iv) consider modifications to DR programs needed to support the California Independent System Operator's (CAISO) efforts to incorporate DR into market design protocols.[11] The DR cost effectiveness methodologies were to "address the broad variety of DR approaches, including current and anticipated future activities," and "identify all relevant quantitative and qualitative inputs (other than load impacts) that are important for determining cost-effectiveness of DR."[12] In addition, the cost effectiveness methodologies should "recommend values for the inputs, or at least recommend methodologies for determining the value of the inputs."[13]

Energy Division staff guidance issued on May 25, 2007 ("Staff Guidance") made clear that the focus of the DR cost effectiveness methodologies should be on determining a useable overall framework and methodology for evaluating the cost effectiveness of each of the different types of DR activities, with the key task of suggesting the relevant perspectives and cost effectiveness tests.[14]

---

7 D.03-06-032, p. 54.

8 *Id.*

9 "DR" (DR) is defined as changes in electricity consumption by customers in response to signals in the form of electricity prices, incentives, or alerts during periods when the electricity system is vulnerable to extremely high prices or compromises to reliability. *See* April 18, 2007 Ruling of ALJ Hecht. The term is used generally in this framework to encompass all DR resources (*e.g.,* programs, purchase agreements, portfolios) and all types of DR, including dispatchable (*e.g.,* curtailable programs) and non-dispatchable (*e.g.,* permanent load shifting), as well as reliability (*e.g.,* interruptible programs) and price-responsive, (*e.g.,* dynamic rates).

10 D.05-11-049, p. Ordering Paragraph 11.

11 Order Instituting Rulemaking 07-01-041 at Section I.

12 Assigned Commissioner and Administrative Law Judge's Scoping Memo and Ruling, issued April 18, 2007, p. 5-6.

13 *Id*.

14 Administrative Law Judge (ALJ) Hecht's Ruling Distributing Staff Guidance for Straw Proposals, 5/25/07, p. 2 - 3.

Consistent with the Staff Guidance, this framework builds on existing work on DR cost effectiveness evaluation where practical[15] and leverages cost effectiveness methodologies developed in other regulatory proceedings. More specifically, this framework uses the SPM as a starting point, and draws from the cost effectiveness methodologies developed for the Self-Generation Incentive Program (SGIP) and from cost effectiveness analysis performed as part of Working Group 2 in R.02-06-001 to refine the SPM framework to more completely and accurately reflect the benefits and costs of DR programs.

2.2 Objectives

The objectives of this DR cost effectiveness evaluation framework are to:[16]

- Determine the perspectives from which the cost effectiveness of DR should be evaluated in formulating policy decisions relating to DR;
- Provide a formal framework for assessing the cost effectiveness of DR programs; and,
- Propose workable methods for valuing the benefits and costs of DR under that framework.

This DR cost effectiveness evaluation framework maintains the philosophy of the SPM by providing "rules" that should be viewed as appropriate guidelines for developing the primary inputs for the cost effectiveness equations contained in this framework, but not requiring excessive and unnecessary rigidity in the application of the methodologies. This philosophy is described succinctly in Appendix A of the SPM:

> A comprehensive review of procedures and sources for developing inputs is beyond the scope of this manual. It would also be inappropriate to attempt a complete standardization of techniques and procedures for developing inputs for such parameters as load impacts, marginal costs, or average rates. Nevertheless, a series of guidelines can help to establish acceptable procedures and improve the chances of obtaining reasonable levels of consistent and meaningful cost effectiveness results.[17]

2.2.1 Specific Uses

The focus of this DR cost effectiveness evaluation framework is on providing guidance for forecasting DR impacts for long-term resource planning.[18] Thus, it is intended that the methodologies developed in this framework will be used for *ex ante* evaluation of DR cost effectiveness. *Ex ante* cost-effectiveness evaluations should be based on the most recent expected values of energy and capacity market prices. Using outdated earlier expectations of those market prices would produce unreliable results.

Further, this framework is applicable both to individual DR programs and to a DR portfolio,[19] and includes methodologies for both non-event (*e.g.,* time differentiated rates and dynamic pricing options) and event-based DR.

---

15 Staff Guidance, p. 24.

16 These objectives draw from those developed by Itron, Inc. for the Self-Generation Incentive Program. *See* Self-Generation Incentive Program Framework for Assessing the Cost effectiveness of the Self-Generation Incentive Program, Itron, Inc., March 2005, p. 2-2.

17 California Standard Practice Manual, p. 26

18 Staff Guidance, pp. 9 – 10.

19 Staff Guidance, p. 35.

2.2.2 Feasible/Workable Methodologies

The DR methodologies in this framework attempt to balance theoretical purity, analytical rigor, and computational complexity. These methodologies deliver accuracy levels adequate for program assessment and policy-making while avoiding methods that greatly increase computational complexity but provide little or unknown increases in accuracy. This approach supports the goal "to improve the cost effectiveness processes used for DR assessment without setting objectives that are beyond the reach of this rulemaking, given its scope and timeline."[20]

2.2.3 Promote Transparency Consistent with Preserving Confidentiality of Commercially Sensitive Information of Each IOU

The cost effectiveness evaluation methods under this framework promote transparency by using published data and public data sources where practicable. This transparency combined with the use of feasible/workable methodologies described above will yield results and supporting analyses more readily understandable to DR stakeholders.

While these cost effectiveness methods attempt to promote transparency wherever practical, it must be recognized that some of the data and analyses underlying utility cost effectiveness are market sensitive and subject to confidentiality protections recognized in Commission decisions.[21] Utility estimates of the volatility of future power and natural gas prices are examples of commercially sensitive information that require confidentiality.

2.3 Scope

The scope of the cost effectiveness methodologies in this framework includes:

- identifying all relevant quantitative and qualitative inputs (other than load impacts) that are important for determining cost effectiveness of DR;
- recommending methodologies for determining the value of the inputs.

This framework attempts to address all of the expectations in the Staff Guidance, recognizing that some of the processes may be viewed as interim and may be refined or reassessed in the future by:

1. Listing material factors and attributes of DR activities that a comprehensive cost effectiveness framework should be able to address. To the extent practicable, this framework addresses the broad variety of DR approaches, including current and anticipated future activities.

2. Addressing the identified material factors from item 1 (above), which can be used in upcoming DR assessments and applications, with the knowledge that future work may continue to address and refine important components of this framework.

3. Listing relevant quantitative and qualitative inputs (other than load impacts) that are important for determining cost effectiveness of DR.

---

20 Staff Guidance, p. 25.

21 *See* Section 454.5(g) of the California Public Utilities Code and D.06-06-066.

4. Recommending methodologies for determining the value of the inputs. Further work needed to develop satisfactory inputs is identified, and interim methodologies (*e.g.,* estimates used in other proceedings) may be recommended for use until additional work can be undertaken. In some cases, the practical solution may be to use ranges for the values of some inputs, or the proposal of a research agenda needed to produce values.

5. The broad variety of DR approaches, including current and anticipated future activities, may require that there be different cost effectiveness methodologies that are appropriate for different types of DR activities.[22]

2.4 Methodological Consistency: Demand-Side vs. Supply-Side Resources

2.4.1 Define methodology

This framework uses multiple perspectives consistent with the SPM. These perspectives are described in Chapter 4 herein. Like the SPM, the tests in this framework are not intended to be used individually or in isolation. Rather, the tests are to be compared to each other, and tradeoffs between the tests considered.[23] Also like the SPM, the results of each perspective are based on the net present value of program impacts over the lifecycle of those impacts.[24]

While the SPM is the starting point for the cost effectiveness methodologies in this framework, modifications have been made to selected elements of the SPM methodologies to better adapt them for use with specific types of DR.[25]

2.4.2 Complete listing of DR benefits

This framework attempts to capture all DR benefits, including:

- Avoided Generation Capacity Costs
- Avoided Energy Costs
- Avoided T&D Capacity Costs
- Other Economic Benefits.[26]

Detailed descriptions of the methodologies used to value these DR benefits are provided in chapters 5 – 8 of this framework.

2.4.3 Highlight direct comparison with other resource options

Staff Guidance makes clear that the DR cost effectiveness methodologies should allow DR activities to be compared to other alternatives in developing a forward-looking resource plan[27] and that the cost-

---

22 Staff Guidance, p. 24.
23 Standard Practice Manual, p. 6.
24 Standard Practice Manual, p. 4.
25 Staff Guidance, p. 25.
26 These costs may include environmental externalities and other benefits listed on page 32 of the Staff Guidance.

effectiveness framework should balance the benefits and costs of DR activities, both individually and in a DR portfolio, with other resource investments.[28] Accordingly, the methodologies in this framework are consistent with the methods used to evaluate the cost effectiveness of available supply-side alternatives. The valuation and counting of demand-side and other resources are consistent with Resource Adequacy Requirements.

The consistent analytical treatment of other available resources and DR also extends to the treatment of uncertainty. This framework includes suggested methods to address various types of uncertainty associated with the calculation of cost effectiveness.

2.5 Modifications to DR Programs Needed to Support MRTU

Staff Guidance states that the cost effectiveness evaluation framework will likely need to fit with the planned CAISO markets and coordinate with the CAISO market planning and zonal requirements.[29] As MRTU is implemented, the IOU DR programs are anticipated to compete directly with conventional generation resources in the market on a level playing field. However, Staff Guidance acknowledges that specific product definitions and dates for MRTU have not yet been set.[30]

IOU DR programs are expected to interact with the CAISO through measured load impacts in defining and meeting emergency conditions. In the future, IOU DR programs are expected to be more integrated into CAISO planning.

One of the strengths of this framework is its flexibility. Even though the specifics of MRTU are still defined, the framework is sufficiently flexible that it can be modified as needed to incorporate the specific requirements of MRTU as it evolves.

This framework is based on an integrated resource planning perspective in which LSEs are responsible for meeting resource adequacy requirements by acquiring resources, including DR resources. Therefore, this framework does not include evaluating the cost effectiveness of arrangements under which IOU customers choose to participate in CAISO energy, capacity or ancillary markets in response to prices in those markets in non-emergency conditions.

27 Staff Guidance, p. 24.
28 Staff Guidance, p. 35.
29 Staff Guidance, p. 35.
30 Staff Guidance, p. 5.

3 Nature of Existing Programs

3.1 Background

Demand side management is the process of reducing customer demand for electricity as an alternative to increasing supply. It requires customer participation in the balancing of demand relative to supply. The objective is to optimize the efficient production and use of electric energy and includes a number of different strategies. Alternatively, the primary strategy to manage demand for electricity in recent years has been energy efficiency. This approach relies on efficient use of energy for the same level of customer value and is typically characterized as a reduction in overall energy consumption. However, in addition to reducing total energy consumption, there are significant benefits that can be achieved by taking actions that will reduce the variability or spiky attributes of customer demand for electricity.

The DR strategy is aimed at reducing the high customer driven demands that cause the need for increased capacity of supply and result in inefficient generation. Specifically, these high demands are referred to as peak demands and efforts to alter customer demand at specific times of the day or days of the year can increase the efficiency of the generation of electricity.

A variety of options are available to achieve customer demand reduction. This section will describe characteristics of DR, for current and anticipated future programs, that are designed to alter customer demand for electricity. The categorization of program characteristics are focused at designating attributes that are significant for differentiating how to measure the demand reduction, or impact, of the program. Additional categorizations relevant to valuing the demand reduction are also discussed.

3.2 Event vs. Non-Event DR programs:

After reviewing existing programs, one differentiating characteristic identified for DR programs was whether the program was designed to be dispatched at the discretion of the utility or CAISO. This type of program is identified as an event based program. In this situation a program can be dispatched for either the cost of energy or a constraint of the supply or delivery of energy. Event based programs are triggered by system conditions and the type of trigger can be a factor in differentiating the value of the resulting DR load reduction. As a result they are dispatchable resources within the limits set up for the program. Non-event based programs are designed to provide a customer incentive to alter demand at predetermined times. They are primarily based on historic or forecasted price conditions. Both event and non-event programs can be behavioral or technology enabled responses to incentives for DR load reductions. Incentives can be in the form of tariff prices, payments, penalties, or societal good will. Examples of current programs categorized by event and non-event programs are described below.

3.2.1 Event based program examples:

Critical Peak Pricing – after a notification of activation call on a limited number of days for a limited set of hours a customer pays a higher price for energy in return for a lower price for energy use other hours and days;

Demand Bidding Program – in response to a call for bids a customer can indicate a load reduction amount for specific hours for an offered level of financial compensation;

Curtailable/Interruptible Program – in response to a call for load reduction a customer will reduce demand to a predetermined firm service level or pay a penalty in return for a lower cost for energy all other times;

Air-conditioning cycling program – in response to a remote signal from the utility the customer's air-conditioner turns off for a period of time.

3.2.2 Non-event based program examples:

Time-of-Use (TOU) – customers are provided with prices of energy that vary by time-of-day based on historic cost of service to reduce demand during on peak periods or shift on peak demand to off-peak periods;

Real Time Pricing (RTP) – customers are provided with prices that vary on an hourly basis based on actual, or short term forecasts of, cost of service to reduce or shift energy use from high cost of service peak periods;

Scheduled Load Reduction Program (SLRP) – a customer chooses a day where they agree to a reduced demand level during historic peak periods or pay a penalty;

Permanent Load Reduction Program – a customer agrees to alter their demand pattern, often with an enabling technology, to shift load from peak periods to off-peak periods.

3.3 Price Responsive Programs

At times pricing programs are differentiated from programs that are considered reliability or capacity bidding programs. However this distinction is not relevant for cost effectiveness; each program is viewed to have a price incentive involved for participation. Whether it is a penalty for non-compliance or a value of lost service, an economic decision is made. The primary concern for evaluating the cost effectiveness of a DR program remains whether it is an event or non-event based program. The trigger for an event based program, becomes a critical factor in determining the value of the load reduction. The hours of operation and seasonality are also significant factors in determining the value of the load reduction.

Non-event based programs will have fixed times to use in valuing the estimated load reduction while event based programs will have a more complicated method of estimating the times in which an event is likely to occur given the trigger and operating characteristics of the program.

3.4 Technology Enabled Programs

Technology can be used to produce or enhance the DR potential. Technology, like price, can be used for both event and non-event based programs. An AC cycling program for example is an event based program. The trigger can be either price based, event based, or both depending on the program design. For a non-event program, technology can be used to shift load off of peak periods on a regular basis. Technology can be used with or without a pricing program that provides an incentive for DR.

4 Framework for Cost Effectiveness Analysis

A cost effectiveness analysis compares the likely benefits and/or costs of a project or program with its other alternatives, to understand the relative effectiveness of the project or program in achieving cost and performance objectives, including meeting the Commission's DR goals and providing the amount of capacity required to achieve the level of reliability implied by the capacity planning reserve margins required by the Commission for resource adequacy.[31]

This framework recommends the use of the Standard Practice Manual (SPM) as the basis for evaluating DR programs, with adjustments to reflect the differences between energy efficiency and DR programs. The structure of the SPM approach is well known to the Commission and to DR stakeholders.

The critical element of the SPM is a multi-perspective approach wherein all applicable benefits and costs are expressed with separate cost-benefit tests for participating customers, non-participating customers, all customers and for society overall, which can address the distributional impacts of programs (across customers) and the beneficial impact of programs on achieving broad societal goals.

The Commission oversees IOU administration of energy efficiency programs through the use of an adopted protocol (policy manual) and guidance on various evaluation inputs, such as D.06-06-063, which adopted the interim use of an avoided cost forecasting methodology. This framework for evaluating the cost effectiveness of DR programs recommends not using the same protocols and forecasting methods that are used for energy efficiency. There are significant differences between energy efficiency and event-based DR that require different techniques for implementing and administering programs.[32] The avoided cost forecasting methodology adopted in D.06-06-063, designed specifically for energy efficiency programs, is not appropriate for event-based DR program evaluation. This framework recommends a more appropriate cost-effectiveness evaluation for DR.

4.1 Standard Practice Manual Evaluation Framework

Table 4-1 below summarizes the various perspectives specified in the SPM from which tests evaluates a demand-side resource, and the costs and benefits that are considered in applying the perspective.

---

31 Under current resource adequacy requirements, the amount of qualified capacity a load serving entity (LSE) must have includes a 15 – 17% planning reserve margin over that LSE's share of a "1 in 2 temperature peak demand" in the CAISO control area.

32 For example, DR programs require close coordination with the California Independent System Operator (CAISO) and involve elements of tariff design and implementation not generally a part of energy efficiency program implementation. Also, the load impacts are less well developed for DR programs.

Table 4-1
SPM Perspectives

	Participating Customer Perspective	Non-Participating Customer Perspective	Program Administrator Perspective	All Customer or Societal Perspective
Benefits	• Bill Savings • Incentives Received	• Capacity Cost Savings • Energy Cost Savings • Delivery Savings • Market Effects	• Capacity Cost Savings • Energy Cost Savings • Delivery Savings	• Capacity Cost Savings • Energy Cost Savings • Delivery Savings • Market Effects • Externalities (Societal)
Costs	• Impact of Lower Reliability • Device Costs	• Incentives Paid • Revenue Reduction from Bill Savings • Device Costs • Admin Costs	• Admin Costs • Incentive Payments • Device Costs	• Impact of Lower Reliability on Participating Customers • Device Costs • Admin Costs

4.1.1 Participating Customer Perspective

The participating customer perspective considers the viewpoint of a customer choosing to participate in a DR program (or required to participate in the case of a mandatory program). The participating customer benefits as a result of any incentive payments received for participation in a program and any bill savings which result from a decrease in usage. For programs which involve curtailing electricity consumption during peak periods, such an interruptible programs or air conditioning cycling, the customer loses the beneficial use of electricity during the curtailment period. For a residential customer, this may result in discomfort or inconvenience (*e.g.,* coping with a higher house temperature due to restricted air conditioner operation). For a business customer, this may result in lost revenues or lost production capacity, which reduces business profits. There are some DR programs, such as thermal energy storage systems, which may shift when electricity is consumed without any reduction to the value of service received by the customer. Finally, some DR programs require the participating customer to pay for the equipment necessary to participate in the program.

Mandatory dynamic pricing programs may require a more detailed analysis of participating customer impacts than is typical for energy efficiency program evaluations because of differences between the ways that dynamic pricing programs affect differently situated customers are affected. In general, this framework recommends that in instances where there is a potential for significant distributional impact among participating customers (*i.e.,* winners and losers), that the participating customer group be subdivided to separately identify these impacts. This approach can also be used in the case of voluntary dynamic pricing programs, where self-selection by participating customers increases rates paid by the remaining customers within the affected group.

4.1.2 Non-Participating Customer Perspective

The non-participating customer perspective considers the perspective of the customer that does not participate (or is ineligible to participate) in a DR program, yet bears the impact of the overall effect of the program on rates or service reliability. In general, the non-participating customer perspective captures all program impacts that are spread broadly across all the utility's customers, and thus includes a proportionate share of these overall impacts which are associated with customers that participate in the specific program.

Non-participating customers benefit from the power cost savings (reduced energy and capacity requirements) and any delivery cost savings (reduced transmission and distribution system infrastructure). Environmental costs which are internalized in the cost of power are included in this perspective. If the DR program has overall market impacts, such as an influence on market competitiveness or an effect on overall system reliability, these effects would be included in this perspective.

Non-participating customers bear the cost of encouraging customer participation in voluntary DR programs, including any incentive payments and the revenue loss associated with any bill savings received by the participating customers. Non-participating customers also bear any device costs and administrative costs which are spread across all utility customers.

For energy efficiency programs, it is common to reduce the power and delivery cost savings and revenue losses due to bill savings to account for "free riders" who would have participated in the program without the incentive. This adjustment is made using a "net to gross" ratio applied to the non-participating customer perspective. In most instances, this framework does not anticipate using net to gross ratios for evaluating DR programs. To the extent a specific program is susceptible to free riders (such as a customer who participates in an air conditioning cycling program who is not at home during peak period hours and does not run their air conditioner at these times), this impact is expected to be incorporated into the load impact protocols.

4.1.3 Program Administrator Perspective

The program administrator perspective is essentially a measure of the degree to which a program administrator is able to leverage direct program expenditures to achieve overall resource benefits. Power and delivery system cost savings and any market effects are considered benefits. Costs borne by the program administrator, including administrative cost, incentive payments, and device costs are included, but revenue losses from participant bill savings are excluded.

4.1.4 All Customer (Total Resource Cost) Perspective

The all customer perspective is essentially the sum of the participating and non-participating customer perspectives. Since any transfers between non-participating and participating customers are netted out in this perspective, the remaining costs are typically related to actual resource costs or resource cost savings. Thus, this perspective is commonly described as the Total Resource Cost (TRC) test.

4.1.5 Overall Societal (Total Societal Resource Cost) Perspective

The societal perspective is a variant of the all-customer perspective that evaluates the overall costs and benefits from a societal perspective that is broader than a "California ratepayer" perspective. Tax-related transfers (such as sales taxes and federal income taxes and credits) are sometimes omitted from this perspective. The cost of externalities, such as the societal impact of air pollutants or other emissions (such as green house gases) that are not internalized in the financial costs incurred by customers or the program administrator, can be included in this perspective.

4.2 Total Resource Cost (TRC) and/or Societal TRC Test Should be Used to Evaluate DR Cost Effectiveness

This framework recommends that the TRC and/or Societal TRC test should be used for DR cost effectiveness evaluation. The TRC test nets out the impact of transfers between participating and non-participating customers. As a result, all of the remaining costs and benefits are actual resource costs or resource cost savings.[33]

The other SPM tests should be used to evaluate the impact that DR programs have on a specific group – participating customers, non-participating customers, or the program administrator – due to the transfers that occur between groups because of a demand-side resource. This may require the investigation of bill savings or other ratemaking impacts.

[33] The TRC test does include outside-of-California transfer payments - such as federal tax credits or tax liabilities - as benefits or costs, and also treats incentive payments paid to "free riders" who would have participated in the demand-side resource as a cost, rather than being netted out as a transfer.

4.3 Purpose of the Cost Effectiveness Measurement

4.3.1 DR Programs Should Contribute to Resource Adequacy

All DR programs should be designed to contribute to resource adequacy either directly or indirectly through load reductions. If the effective capacity available from a DR program cannot be used to satisfy an LSE's RA requirement, that program will not enable that LSE to avoid generation capacity costs, even if the availability of that DR capacity makes the system somewhat more reliable.[34]

The additional reliability provided by the availability of the effective capacity of a DR resource or a supply-side generation resource that cannot be counted for resource adequacy can make electricity service more reliable by reducing the probability of system outages. However, the value is much smaller, as in adding planning reserves beyond 17 percent. Therefore, DR programs should be required to impact resource adequacy to be counted as RA.[35]

The DR load impact protocols should address the statistical requirements to be counted as equivalent to a supply-side resource.

4.4 Scope and Structure of the Analysis

There are a number of analytical issues associated with the scope and structure of the cost effectiveness analysis that should be addressed as part of the Commission's adopted methods.

---

34 As noted in Section **Error! Reference source not found.**, the additional reliability provided by the availability of the effective capacity of a DR resource or a supply-side generation resource that cannot be counted for resource adequacy can make electricity service more reliable by reducing the probability of system outages. The RA requirement is based on a capacity planning reserve margin that, in principle, aims to balance the marginal cost of acquiring additional capacity, with the marginal value that customers place (on average) on avoiding the potential outages avoided by those additional resources

It is sometimes suggested that DR programs that cannot be used to satisfy an LSE's RA requirement should be assigned an avoided capacity value, if they cost less than the value that customers place on avoiding the additional potential outages those DR programs would avoid.

That suggestion is not correct. If the planning reserve margin equates the marginal cost of acquiring additional capacity, with the marginal value that customers place (on average) on avoiding those potential additional outages, the capacity value of DR programs that cannot be used to satisfy RA requirements will be less than the marginal cost of additional generation capacity.

The question of whether the Commission should adopt a higher capacity planning reserve margin, and therefore higher RA requirements, is currently being addressed in another CPUC proceeding. The issue does not appear to be within the scope of this DR cost effectiveness.

35 It is sometime suggested that DR programs (and some supply side resources) both avoid RA-related generation capacity costs and provide "insurance" or "hedge" value. Obviously, this is not the case, since DR capacity can only be used for one purpose or the other.

4.4.1 Portfolio vs. Program Evaluation

Many administrative and marketing costs are incurred as part of implementing an overall DR program portfolio, and are not readily attributed to individual programs. For the purpose of program design and approval, it would not make sense to allocate a share of these common costs to individual programs. The result might be to reject an individual program as not being cost effective, when the program would actually be net beneficial but for the allocation of these common costs. Nevertheless, administrative and marketing costs cannot be completely ignored.

This framework recommends that the cost effectiveness analysis that accompanies applications for a three-year program funding cycle be performed on both an individual basis and on an aggregate portfolio basis. Only administrative and marketing costs attributable to an individual program would be included in the individual program cost effectiveness analyses, but all administrative and marketing costs would be included in the portfolio analysis.

4.4.2 Length of Time Period

In general, the time horizon for an analysis should cover the economic life of the most significant component of program investment. For instance, if a DR program requires installation of a device with an expected 10-year life, then the analysis period should be for 10 years. Ongoing costs for equipment repair/replacement or to solicit new customers to replace those who leave the program should be included in the ongoing cost of the program. This approach recognizes that maintaining an existing program for which significant sunk costs have already been incurred is likely to be quite cost effective.

For ongoing programs that have significant mid-stream costs which need to be incurred in a subsequent program cycle, it may be appropriate to perform a cost effectiveness analysis (treating the initial expenditure to initiate the program as a fixed unavoidable cost) to determine whether the program should be continued.

For programs that have few upfront costs, it will usually be appropriate to perform the analysis over the three-year program cycle. However, if it is expected that the program will continue in future years, then an analysis can be performed for a longer period.

4.4.3 Temporal Granularity

This framework recommends that analysis results be presented on a calendar year or monthly basis. Within each year, analysis should be performed on at least a time of use period basis (*e.g.,* summer on peak) and on an hourly forecast basis where the nature of the program justifies this level of granularity, and where sufficient information exists to perform an hourly calculation.

4.4.4 Geographic Granularity

Currently, visible market prices in the CAISO service area are reported in broad geographic zones, generally reflecting PG&E's service area (north of path 15, or NP-15) and a combined SCE and SDG&E service area (south of path 15, or SP-15). Within these zones, congestion is managed by means of tariff rules that allocate the cost of resolving congestion to market participants without using price signals. The CAISO has proposed broad market changes that are expected to go into effect in early 2008 (the Market Redesign and Technology Upgrade or MRTU). Of significance to the evaluation of DR, separate generation pricing nodes will be established at substations and other facilities throughout the existing zones. As a result, intra-zonal congestion and intra-zonal line losses will be implicitly reflected in the prices for each node. Once an understanding of how prices vary by node emerges, it may be possible to incorporate locational information in the cost effectiveness evaluation of DR.

However, this framework recommends that the initial round of program design that takes place for the 2009-2011 program cycle should not include any geographic granularity. There will be very limited information available (if any) regarding locational prices in time for preparing the mid-2008 submissions for this program cycle.

5 Economic Benefits: Avoided Generation Capacity Costs

5.1 Definition of Avoided Generation Capacity Costs

DR programs, by their design, avoid the need for generation capacity since their function is to reduce customer usage during periods of peak demand. Thus, avoided generation capacity costs will usually account for most of the economic benefits provided by a DR program.[36]

A DR program can avoid generation capacity costs if an LSE can utilize that program in meeting its Resource Adequacy (RA) requirement.

The generation capacity cost avoided by a DR program that a Load Serving Entity (LSE) can count toward its RA requirement is the market value of the additional generation capacity that would otherwise be needed to comply with that requirement.

Because DR programs are used to reduce demand during periods of relatively high demand, DR usually avoids the need for relatively high heat rate generating capacity that is only dispatched during peak demand periods. In the long run, the market value of that type of generating capacity in California will be based on a new natural gas-fired new combustion turbine (or new CT). Therefore, this framework recommends that new natural gas-fired new CT capacity be used as a proxy to derive the market value of the generation capacity avoided by DR programs.

That market value would be the same, regardless of whether that capacity is owned by the LSE that administers the DR program or procured from a third party (*e.g.,* under a power purchase agreement). Although market prices for energy are available, market prices for capacity are not. Most of the generating capacity that California IOUs procure from third parties is purchased through bilateral transactions in which energy is also purchased as well. The prices and terms of those bilateral transactions are usually confidential, and if not would often still have to be analyzed to separate capacity value from energy value. Therefore, the market value of the capacity of a new CT must be estimated.

As the remainder of this chapter explains, the lifecycle market value of the capacity of a new CT should be estimated by deducting:

(3) the present value of the gross margins the new CT capacity is expected to earn from selling energy (when wholesale electricity market prices exceed variable costs); from'

(4) the present value of the total fixed costs of that new CT.

The resulting lifecycle "net" capacity cost should then be annualized, to estimate the annual market value of new CT capacity in each year.

---

36 Some DR programs may also be operated to reduce transmission or distribution facility loading; the potential for avoiding transmission and distribution costs is discussed in Chapter 7.

That annual market value of new CT capacity should then be allocated among the time periods within that year, and then adjusted for any additional value that DR programs provide in each period, including the avoidance of capacity planning reserves and reductions in transmission and distribution line losses. Commission

5.2 Resource Adequacy Capacity Costs Avoided by DR Programs

DR capacity only enables a Load Serving Entity (LSE) to avoid generation capacity costs if the LSE can use that capacity to meet its Resource Adequacy (RA) requirement. Under current Commission rules, an LSE can count load reductions achieved by DR as reductions in the peak loads used in determining Resource Adequacy (RA) requirement.[37] Therefore, the extent to which a DR program enables an LSE to avoid generation capacity costs depends upon how much of the load reduction capacity available from that program ("enrolled MWs") has been deemed to "qualify" for system RA ("qualifying MWs"), based on "RA counting rules" established by the Commission and implemented in cooperation with the CEC.

Evaluations of the generation capacity costs that a DR program can avoid should be consistent with the Commission's RA "counting" rules for those programs.

The criteria that DR capacity must meet under current RA counting rules in order to qualify for RA includes a requirement for that capacity to be available for at least 48 hours in each summer season, and for at least four consecutive hours.

The overall RA requirement for the CAISO control area is set in a manner that attempts to balance the cost of acquiring additional resources with the value that customers place (on average) on avoiding the potential for outages that would be avoided by those additional resources. Thus, if the CAISO control area (or some transmission constrained portion of the control area) is over-resourced, the avoided capacity value of additional DR is less than the net capacity cost of a new CT.

As noted above, if a DR program cannot be used to satisfy an LSE's RA requirement, that program will not enable that LSE to avoid generation capacity costs, and therefore should not be ascribed any capacity value. This recommendation is based on a presumption that the Commission's long-term procurement planning reviews (for utilities) and RA requirements (for all LSEs) will be adequate to maintain a balanced portfolio of demand and supply resources in the CAISO control area. This topic is addressed further below.

Aspects of the Commission's resource adequacy program are currently under review in R.05-12-013. One party, the California Independent System Operator (CAISO), has suggested that certain categories of DR should not count for meeting RA. In addition, the Commission has included the subject of coordination of DR programs within the scope of this proceeding, and the CAISO has also established a stakeholder process which will address various DR topics.

The eventual resolution of these issues may affect the quantification of DR benefits and costs in ways that are not immediately apparent. Thus, any DR cost effectiveness framework adopted by the Commission will necessarily be a "living document" which will be subject to ongoing modification in response to changes in the institutional setting in which resource adequacy is determined.

---

[37] D.04-10-035, p.54, Conclusion of Law 19.

5.2.1 Using DR to Meet System-Wide RA Requirements

Under the Commission's current rules, the amount of qualified generation and DR capacity an LSE must have to meet its "system-wide" RA requirement is determined by adding a 15% to 17% planning reserve margin to that LSE's share of a "1 in 2 temperature peak demand" within the CAISO control area.[38]

Due to transmission constraints in certain parts of the CAISO control area, LSE's also must demonstrate that they will have enough additional capacity in certain geographically-specific load packets to satisfy their "local" RA requirements.

As of now, by September 30th of each year each LSE must demonstrate that it has procured and/or owns enough capacity to meet at least 90% of its system-wide RA requirement in each month of the following year, and 100% of its local resource adequacy requirement in each month of that year.

At the end of each month in that following year, each LSE also must also demonstrate that it has enough additional generation and DR capacity to meet the remainder of its system RA requirement for the month after next.

5.2.2 Using DR to Meet Local RA Requirements

Due to the existence of transmission and distribution constraints, LSE's must also demonstrate that they will have enough additional capacity in certain geographically-specific load pockets.

In a recent decision[39], the Commission established guidelines for counting dispatchable DR for local RA starting 2008. The Commission also affirmed its earlier decision that emergency and interruptible DR could be counted for local RA requirements.

Therefore, this framework recommends developing estimates of the geographically-specific local generation capacity cost avoided by DR that qualifies local RA, after enough historical data on MRTU nodal prices has been accumulated.

A related issue is the possibility that DR in an area with local constraints might defer or avoid transmission and distribution investments that would otherwise have to be made. That issue is discussed in Chapter 7.

---

[38] PG&E has recently recommended in R.05-12-013 that the Commission adopt a higher planning reserve margin requirement that provides a 16% reserve margin over a "1 in 10 temperature" peak demand. Some parties have supported PG&E's proposal, while others have opposed it.

[39] D.07-06-029 issued on June 21, 2007, pp. 37-40.

5.2.3 Value of DR In Excess of the RA Requirement

The Commission establishes the RA requirement for the CAISO control area by adopting a capacity planning reserve margin that aims to balance the marginal cost of acquiring additional capacity, with the marginal value that customers place (on average) on avoiding the economic losses associated with the potential outages avoided by those additional resources.[40]

The additional reliability provided by the availability of the effective capacity of a DR program that exceeds RA requirements can make electricity service more reliable by reducing the probability of system outages. The same is true for additional supply-side generation resources.

It is sometimes suggested that DR programs procured in excess of RA requirements should therefore be assigned value.[41] Based on this reasoning, DR programs that do not avoid generation capacity costs would be cost effective if they cost less per kW than the value that customers place on avoiding outage.

However, that suggestion is incorrect if the planning reserve margin chosen by the Commission balances the marginal cost of acquiring additional capacity with the marginal value that customers place (on average) on avoiding the economic losses associated with the potential outages avoided by those additional resources.

In this case, the "insurance" or "hedge value" of generation and/or DR capacity in excess of RA requirements will be less than the marginal cost of acquiring additional capacity. Therefore, assigning an "insurance" or "hedge value" to DR capacity in excess of RA requirements would place a value on that additional DR capacity higher than that the equilibrium which would be otherwise normally be obtained by assuming that DR capacity is procured (along with any other supply side resource) to exactly meet RA requirements.[42]

The question of whether the Commission should adopt a higher capacity planning reserve margin, and therefore higher RA requirements, is currently being addressed in another Commission proceeding. The issue does not appear to be within the scope of this DR cost effectiveness.

Therefore, this framework recommends that the methods used to evaluate the cost effectiveness of DR capacity should not incorporate the implicit assumption that a higher planning reserve margin is needed, by assigning avoided generation capacity value to DR capacity that would cause LSE to exceed its RA requirement.

If the Commission does adopt a higher capacity planning reserve margin, that would increase the RA requirements of LSEs. The capacity benefit provided by a DR program would then still depend upon the

---

40 The marginal value that customers place on avoiding economic losses associated with outages is termed Value of Lost Load (VOLL).

41 The Commission recently cited that rationale when it required the IOUs to obtain more DR after the July 2006 heat wave, and when the Commission approved SCE's contract for the Long Beach project, despite capacity reserve margins that exceeded the Commission's current "15% to 17" capacity planning reserve margin standard.

42 It is sometime suggested that DR programs (and some supply side resources) both avoid RA-related generation capacity costs and provide insurance or hedge value. Obviously, this is not the case, since DR capacity can only be used for one purpose or the other.

extent to which an LSE could utilize that program to avoid generation capacity costs in meeting that revised RA requirement.

5.3 Estimating Avoided Generation Capacity Costs

Different types of generation capacity, for instance a natural gas-fired new combustion turbine (new CT), a natural gas fired combined cycle unit (CCGT), or a coal plant, have different costs, heat rates, and operating characteristics (*e.g.,* start up time and dispatchability).

An efficiently designed generation system will have a mix of different resource types. Baseload generating units generally cost more to build, but have lower variable costs. These units will run whenever they are available, except during extremely low load "off-peak" periods.

At the other extreme, peaking generating units are relatively inexpensive to build, but have high variable operating costs. These units will run only during high load periods, when all available baseload and intermediate load generation is already in operation. Intermediate units fall in between these extremes.

Therefore, the avoided capacity cost associated with a change in demand in a high load period is generally considered to be the capital and other fixed costs of a peaking unit that are not covered by the gross margins (*i.e.,* revenues minus variable operating costs) earned from selling energy at prices that cover variable operating costs The avoided capacity cost associated with a change in demand in a moderate or low load period is generally considered to be the capital and other fixed costs of an intermediate or base load unit, respectively, that are not covered by gross margins from selling energy.

Because DR programs are used to reduce demand during periods of relatively high demand, the capacity available from those programs usually avoids the need for relatively high heat rate generating capacity that is only dispatched during peak demand periods.

However, an intermediate or base load unit runs during higher load periods at an operating cost below the cost of the unit on the margin, and produces a gross margin that compensates for the higher capital cost (the so-called energy related capital cost).

Therefore, the avoided capacity cost associated with changes in demand in a high load periods is generally considered to be the net capacity cost of a peaking unit (*i.e.,* the amount by which the unit's fixed costs exceed the gross margins achieved by selling energy when revenue at wholesale market prices exceed variable operating costs).

Thus, it is common to express avoided capacity costs as the annualized net capacity cost of a peaker (allocated to time periods within the year based on the relative need in each time period for the reliability contribution made by the peaker).

5.3.1 Natural Gas-Fired new CT as the Proxy Resource

Because DR programs are used to reduce demand during peak usage periods, the capacity available from those programs usually avoids the need for relatively high heat rate generating capacity that is only dispatched during those peak usage periods. In California, natural gas-fired new combustion turbines are usually considered to be the marginal resources during peak usage periods, and are widely used in cost effectiveness evaluations.

For this reason this framework recommends using a new CT as the proxy resource for evaluating DR program cost effectiveness.

The market value of generation capacity for any given time period is, in principle, determined by the market price at which the amount of capacity supplied equals the amount of capacity that is demanded during that period. The market price of generation capacity for each time period therefore depends upon what type of generation capacity is most likely to be on the margin during that period.

Issues are sometime raised as to whether the net capacity costs of older high-heat rate units which typically run only during peak periods would be an appropriate alternative to a new CT as a proxy resource.

In the long run, additional peaking generation capacity will be built only if prices for energy and capacity are expected to be high enough to recover the variable and fixed operating costs of that capacity. This includes both the amounts invested in the construction of that generating capacity, and the rates of return required by the lenders and shareholders that provided the capital which financed the construction of that new capacity.

In contrast, an existing plant can be expected to remain in operation as long as its revenues from selling energy plus its revenues from selling capacity are expected to cover both its going forward (incremental) fixed and variable costs. In a period of growing peak demand, existing peaking plants will be retired when the going forward fixed and variable costs incurred to continue operation exceed the cost of a new CT.

Therefore, the annualized net capacity cost of a new CT provides a reasonable estimate of the annual market value of additional generating capacity.

5.3.2 Adjustment for new CT "Energy Benefits"

The owner of a new CT obtains both "energy benefits" (*i.e.,* the gross margins earned by generating and selling electricity whenever wholesale market prices exceed the variable costs incurred in producing that power), and "capacity benefits" (*i.e.,* the revenues generated by selling the right to dispatch the capacity of the new CT in order to maintain the reliability of electricity service)

As noted above, those energy benefits (*i.e.,* gross margins) must be subtracted from the fixed costs of the new CT capacity to determine the net capacity cost of the new CT capacity.[43]

A new CT's expected energy benefits are calculated by determining the gross margins which the proxy unit could achieve over its future operating life by selling energy in those time periods in which revenues at wholesale market prices exceed the variable fuel and O&M costs incurred in generating that energy. Those expected gross margins depend upon:

[43] A new CT may also be able to sell ancillary services to the CAISO. Since a DR program may also provide ancillary service value it may be appropriate to reflect differences between the ability of a new CT and the DR program in the cost effectiveness evaluation. This issue is discussed further in Section 8.3.

(1) the heat rate of the new CT (Btu/kWh);
(2) variable O&M costs ($/kWh) of that new CT;
(3) natural gas prices ($/MMBtu);
(4) wholesale market energy prices ($/MWh); and,
(5) the expected operating life of the new CT.

There are several ways to estimate these expected gross margins.

5.3.2.1 Production Cost Simulation Modeling

New CT energy benefits have traditionally been estimated by using a production cost simulation model to develop a long-term forecast of wholesale market energy prices, and determining the net present value of the operating profits the new CT would achieve by selling its output at the forecasted prices.

Natural gas prices and, in particular, wholesale electricity prices are volatile. Wholesale market electricity prices vary enormously from hour to hour, day to day, and month to month, and are far more volatile than the prices of almost any other commodity.

Therefore, it is common for such efforts to distinguish between a new CT's "intrinsic value" relative to a deterministic forecast of future market prices, and the new CT's "extrinsic value" relative to a stochastic simulation which accounts for a range of possible future market prices.

Deterministic models tend to significantly underestimate the expected energy benefits of a new CT, because the models use point (*i.e.,* deterministic) forecasts of fuel and market prices, that do not take into account the fact that a new CT is likely to operate more hours when wholesale electricity prices (relative to fuel prices) are high, and fewer hours when wholesale electricity prices are low. That is, deterministic forecast-based production cost simulations do not take into account the correlation between:

(1) the difference ("spark spread") between wholesale electricity sales revenues and variable fuel costs, which is determined by the market prices for electricity and the unit's heat rate; and,

(2) how much power the new CT will be used to generate.

In contrast, stochastic models use techniques to simulate multiple forecasts of future fuel and market prices (*e.g.,* Monte Carlo methods), and calculate an overall expected energy benefit by averaging these various scenarios.

5.3.2.2 Hourly "Spark-Spread" Call Option Estimates of Expected Gross Margins

An alternative to using a production cost simulation model is to forecast a new CT's extrinsic energy benefits by using "spark-spread" call option valuation model. Option valuation models are widely used for financial valuation purposes in competitive and bilateral power procurement transactions. Under this approach, the expected gross margins are estimated by treating the ability to dispatch the new CT as a series of (hourly) "spark spread" call options.

Hourly "spark-spread" call option valuation models can provide an accurate estimate of the expected future gross margins of new CT capacity, because those models explicitly take into account:

(1) the expected volatility of hourly wholesale electricity prices in each period;
(2) the expected volatility of future natural gas prices in each period; and,
(3) the expected correlation in each period between those future electricity and natural gas prices.

As a result, hourly "spark-spread" call option models reflect both the "intrinsic" and the "extrinsic" values of the future options to dispatch new CT capacity to generate energy whenever it would be profitable to do so.

Whether or not the estimates of expected gross margins obtained from hourly "spark-spread" call option valuation models are more accurate than those obtained from deterministic and stochastic production costing models depends upon the reliability of the methods used to estimate:[44]

(1) expected future natural gas prices;
(2) expected future wholesale market prices for energy;
(3) the expected volatility in each period of future hourly wholesale electricity prices;
(4) the expected volatility in each period of future natural gas prices; and,
(5) the expected correlation in each period between those future electricity and natural gas prices.

5.3.2.3 Recommended Methods for Estimating new CT Energy Benefits

This framework recommends that estimates of the market value of new CT capacity that are used to evaluate the cost effectiveness of DR resources, should be based on the net capacity costs obtained by subtracting the gross margins that new CT capacity is expected to achieve from selling energy, from the expected total fixed costs of that new CT capacity.

We recommend that those gross margins be estimated by using methods that capture both the "intrinsic" and "extrinsic" values of those expected gross margins.

We do not, however, recommend a particular method for deriving those estimates.

5.3.3 Other Adjustments to the new CT Capacity Benefits

5.3.3.1 Adjusting for Avoided T&D Line Losses

Due to transmission and distribution voltage level line losses, more than one MW of capacity is needed to meet one MW of customer-meter level demand from customers that receive electricity service at distribution voltage levels.

44 The forward power prices, volatilities and correlations that are used in applying hourly "spark-spread" call option models to estimate expected gross margins from selling energy should be consistent with those used to estimate the energy costs avoided by utilizing the DR program to actually reduce the demand for energy.

Due to transmission voltage level line losses, more than one MW of capacity is also needed to meet one MW of demand from customers that receive electricity service at transmission voltage level.

Marginal transmission and distribution voltage level line loss rates vary from hour to hour, due to variations in load and ambient temperatures.

Therefore, before generation capacity costs avoided by available DR capacity can be estimated, the estimated annual and hourly market prices of new generation capacity should be adjusted upward for transmission and distribution line losses that would be avoided by customer meter-level demand reductions associated with the RA-qualified DR capacity available in those periods.

5.3.3.2 Adjusting for Avoided RA Reserve Margin Capacity

As noted above, under current Commission rules, an LSE can count reductions in peak loads available from DR programs as reductions in the loads used in determining its system-wide Resource Adequacy (RA) requirement. That RA requirement includes a 15% to 17% capacity reserve margin.

Therefore, estimates of the future (avoided T&D line-loss adjusted) annual and hourly market prices of new generation capacity also must be adjusted upward to take into account the RA reserve margin new CT generation capacity that will be avoided by the RA-qualified DR capacity that is available in those periods.

5.3.4 Annualizing the Market Value of new CT Capacity

The lifecycle net capacity of a new CT coming on line in a given year is the present value of the stream of future amounts the owner of new CT capacity would have to charge for the right to dispatch that capacity in order to cover the present value of its future net capacity costs,[45] including an appropriate after tax rate of return on the amount invested in building or acquiring that new CT capacity.[46]

The annual market price of new CT capacity in that year ($/kW-year) is then derived from that lifecycle net capacity cost by determining the annual real economic carrying charge (denominated in constant dollars/kW-year) that would recover the future net capacity costs of a new CT that began operating in that year and is expected to remain in operation over the remainder of its useful life.

---

45 The future net capacity costs of the CT are the amounts by which its future fixed costs are expected to exceed the future gross margins that will earned by selling energy when revenues at wholesale market prices exceed the variable fuel and O&M costs incurred in generating that energy.

Those future variable costs would include costs associated with acquiring CO2 emission allowances after 2012 under a “cap and trade” emission allowance market that California may establish pursuant to AB 32.

Those future fixed costs include fixed O&M costs, property taxes, insurance expenses, state and local income taxes, the recovery of the amounts invested in constructing or acquiring the plant. Those future costs also include any other fixed “internalized” monetary costs that will be incurred due to the environmental impacts of that new CT capacity (*e.g.,* the cost of mandatory emission control systems and permits).

46 That return reflects the after tax cost of the debt and equity capital that financed the construction or acquisition of that CT capacity

That annual real economic carrying charge is an amount which, when escalated at the rate of inflation, results in a stream of annual cash flows that has the same present value as the stream of that new CT's future annual net capacity costs. The rate of inflation used in this calculation should reflect the expected increase over time in the cost to build new CT capacity taking into account potential productivity improvements in plant construction.

5.3.5 Hourly Market Price of New Capacity

In each year, the portion of the annual market value of the new CT capacity that will be avoided by RA-qualified DR capacity depends upon the hours in which that DR capacity will be available in that year.

Therefore, in order to estimate the generation capacity costs that will be avoided in each year by that DR capacity, it is necessary to estimate the hourly market price of new CT generation capacity (in \$/kW-hour) in each of the hours of that year during which that DR capacity will be available.

Estimates of hourly market prices for generation capacity in each utility area can be derived from the estimated annual market price of generation capacity (\$/kW-year) in that year, by allocating that annual market price of capacity among the hours in that year in proportion to the relative expected need for generation in each hour.

The amount of generation capacity required in any given hour in each region is proportional to the probability that some of the load in that hour in that region will not be served (*i.e.,* unserved energy) due to a lack of sufficient generation capacity (*i.e.,* a loss of load). The likelihood of that occurring in a given hour is the "Loss of Load Expectation" (LOLE) for that hour.

Therefore, this framework recommends using loss of load expectation (LOLE) as the basis for assigning the annual new CT capacity cost to individual hours during each year.

A Monte Carlo simulation can be used to estimate hourly LOLEs for each utility area. Each of the many runs included in a Monte Carlo simulation compares stochastic estimates of the load that would occur in each of the 8,760 hours of the year within that region, to the capacity of the resources that would be physically available in each hour to serve that load.

The LOLE for each hour is the percentage of all of those runs in which the load in that hour exceeded the capacity of physically available resources.

5.4 Estimating Generation Capacity Costs Avoided by DR Programs with Different Availability and Dispatch Constraints

The capacity that will be available from different dispatchable DR resources depends upon limitations on the time periods during which each resource will be available, and on the duration, frequency, and number of times that each resource can be dispatched.

Therefore, in order to estimate the generation capacity costs that would be avoided by the capacity of a DR resource, it is necessary to model the impact of those restrictions on the sum of the hourly generation capacity costs which that resource is likely to avoid.

A DR resource that is not available in as many hours as a new CT, or that cannot be dispatched for the same hours as a new CT, will not have the same capacity value as a new CT. However, most DR programs will be called during a relatively small number of critical peak periods during the year. Therefore, the generation capacity costs avoided by most limited DR programs will be only somewhat lower than the market value of the capacity of a new CT.

Currently, each of the IOUs uses a different method to adjust the capacity value of DR programs for limited use restrictions. At present, we do not have a unified recommendation on methodological approach. This is an area that will benefit from further discussion and evaluation of the different approaches.

Conceptually, as noted earlier, the annual capacity value of a new CT ($/kW-year) can be assigned to individual hours ($/kW-hour) in that year, in proportion to an estimate of the LOLE in each hour of that year.

Within the time periods for which a DR program can be triggered, each of the IOUs makes analytical assumptions about how DR would be utilized to maximize its value over the hours with the highest LOLEs.

SDG&E assumes that it has sufficient use of DR to fully capture reliability benefits.

SCE uses a simulated optimal dispatch to identify a set of hours (up to the available number of times the program can be used) which in combination would enable the dispatch of a DR program to provide the highest total expected value, with an adjustment to account for the inability to predict critical events with perfect foresight. Appendix A contains a brief description of this approach.

PG&E uses formal explicit optimization methods, which assume that a DR resource will be used in a manner that achieves the highest possible avoided capacity cost, after explicitly taking into each of the limits on the availability and use of that resource. Appendix A contains brief descriptions of two examples of those methods.

5.5 Present Value of Avoided Capacity Costs

Under the TRC test or the Societal TRC test, the cost effectiveness of a DR resource is evaluated by comparing the present value of the expected future costs of that resource, to the present value of the expected future benefits of that resource. Therefore, the future generation capacity costs that will be avoided by the capacity available from a DR resource should be discounted and converted to a present value. Chapter 10 discusses the determination of that discount rate.

The present values of the economic benefits and economic costs used in applying the TRC and societal TRC test to determine the ex ante cost effectiveness of a DR resource should be based on expected cash flows, not expected book expenses determined under financial accounting rules. Therefore, whether specific types of generation capacity costs would be treated as expenses, or capitalized and amortized over time, for accounting purposes is irrelevant.

6 Economic Benefits: Avoided Energy Costs

As noted in Chapter 4, the cost effectiveness of a DR program should be evaluated on an *ex ante* (*i.e.,* prospective) basis. Therefore, estimating avoided energy costs due to DR requires projections of future energy prices.

To estimate the energy costs that an LSE will avoid in the future by using a DR resource, we need to estimate future energy prices, and model the timing and magnitude of the reductions in energy consumption that will occur, in the periods in which the DR resources are dispatched.

Developing appropriate forecasts of future wholesale market energy prices is comparatively straightforward. Modeling the magnitude and timing of DR can be somewhat more complicated.

Given that avoided energy costs account for a small fraction of the value of most demand response programs, the additional analytic complexity associated with estimating avoided energy costs more precisely may not be warranted if more precise methods do not have a material impact on estimates of cost-effectiveness

6.1 Avoided Energy Costs Should Be Based on Market Prices

Whether avoided energy costs are viewed as avoided market procurement costs or avoided production costs should not matter. Given that an LSE can always buy energy from or sell energy to the market, the wholesale market price of energy is the relevant opportunity cost for valuing the energy avoided by a DR resource (or provided by a supply side resource).[47]

6.1.1 Energy Price Projection Methods

Forward prices for energy represent the prices today for energy that will be delivered in the future. Forward prices are established through trading on organized exchanges such as NYMEX and ICE as well as through bilateral transactions in over-the-counter markets. To the extent that forward energy prices are determined in liquid, well-functioning markets with the temporal and geographic granularity needed to value DR, they can be used to estimate the energy costs that will be avoided by DR.

Unfortunately, forward prices are often not available for all of the specific times and locations at which a DR resource would reduce electricity consumption. In the absence of sufficient forward prices, there are a variety of acceptable approaches to developing estimates of future energy prices, as discussed below.

[47] Unless noted otherwise, throughout this chapter the terms "energy price" and "electricity price" refer to wholesale market electricity prices.

6.1.1.1 Using Production Cost Models to Forecast Electricity Prices

Detailed production cost simulations can be used to estimate future energy prices, based on the projected values of fundamental determinants of power prices, such as fuel prices, heat rates, loads, and available resources. These simulations can provide estimates of future energy prices with the degree of temporal granularity (*e.g.,* hourly prices) needed to estimate the energy costs that will be avoided by DR resources.

Even for periods for which forward prices are available, production cost simulations can be calibrated against existing forward prices: in other words, the inputs needed to perform these simulations can be adjusted so that they forecast prices that are equal, on average, to available forward prices. If the calibrated production cost simulations forecast energy prices with the necessary greater degree of temporal granularity, those prices can used to estimate energy costs that will be avoided by a DR resource.

6.1.1.2 Using Forward Curves to Forecast Future Electricity Prices

If forward electricity prices are not available for delivery dates sufficiently far in the future to value a DR program, another method must be used to forecast those prices. One approach is to use the statistical estimates of the relationships between forward power prices and other variables to extrapolate the forward electricity price curve. For example, the ratio of power prices to natural gas prices might be relatively stable over a time period for which forward prices for both power and natural gas prices are available. That ratio could then be used to extrapolate future power prices, based on available projections of natural gas prices, such as those prepared by the U.S. Energy Information Administration (EIA).

Advantages of one method of forecasting electricity prices over another have not been determined. Both can be considered acceptable methods.

6.1.1.3 Temporally Disaggregated Prices

Electricity prices are more volatile than those of almost any other commodity. As a result, there are large differences between wholesale market electricity prices at different times (*i.e.,* wholesale market prices for electricity are highly time-differentiated). Because DR programs are typically used for only a few hours per year, estimating electricity prices that reflect that time-differentiated volatility is particularly important in estimating the energy costs that a DR will avoid.

DR programs are typically used for only a few hours per year, therefore the method used to model the timing and magnitude of the demand reductions that will take place under a DR program should take into account those characteristics of the program that are most likely to determine those specific hours.

Unfortunately, forward prices often are not available at the hourly granularity level that would be needed to estimate the energy costs that a DR program will avoid.

There are two general approaches to developing appropriately temporally disaggregated prices.

The first approach involves applying "price shapes" to available or projected forward prices. A "price shape" reflects the relative prices, rather than the actual price levels, at different times over an extended period of time (*e.g.,* relative hourly prices over the course of a year). Price shapes might be estimated from the historical variation of temporally disaggregated prices around long-term average values.[48] The risk of this approach is that historical price shapes may be quite different from future price shapes. For example, the magnitude of price spikes observed during California's energy crisis is unlikely to be repeated for the foreseeable future. Also, as older generating capacity is replaced by newer resources, there may be greater differences between the costs of using different units to generate electricity. Given that in a well-functioning, competitive market, market prices are closely related to the marginal cost of generation, as the generating capacity supply stack changes, hourly price shapes will change as well.

The second approach to developing hourly prices involves production cost modeling as described above in section 6.1.1.1.

Recognizing the limitations of historical price shapes, this framework does not have a preference between the use of an appropriate price shape or production cost modeling for developing time differentiated energy prices.

6.1.2 Projecting Energy Prices at Different Locations

DR capacity may avoid more energy costs in a load pocket, such as San Diego, than in other locations, which will not be reflected in market prices for electricity delivered to broader geographic areas such as SP-15.

However, there are a variety of problems associated with assigning an avoided energy cost to a DR program that fully reflects the value of the location of that program. First, it may be difficult to determine the precise location of the resources involved in a DR program. The IOUs have large and geographically diverse service territories. For example, the market value of energy in a coastal climate zone may be lower than the value of energy in the Central Valley. Similarly, the market prices for energy delivered to densely populated areas where it is more difficult to build conventional resources may have higher than market prices for energy delivered to less densely populated areas.

Under MRTU, separate generation pricing nodes will be established at substations and other facilities throughout the existing zones. Prices will be determined at each of those nodes. As a result, intra-zonal congestion and intra-zonal line losses will be reflected in the energy prices at each of those nodes. After enough data has been collected to develop a better understanding of the relationship over time between electricity prices at different nodes it may be possible to incorporate locational differences in energy prices into evaluations of the cost effectiveness of DR resources.

---

48 For example, some parties use historical price shapes derived from average hourly prices from the California PX market. That is somewhat problematic for the framework here, in part because PX market prices were those sellers charged buyers for products that reflected capacity and energy, rather than energy alone, *i.e.,* when the PX existed, there was no explicit RA requirement.

However, this framework recommends that evaluations of the cost-effectiveness of the initial round of program design which takes place for the 2009-2011 cycle should not reflect the geographic granularity in post-MRTU electricity prices, because there will be very limited information available (if any) regarding locational prices in time to prepare the mid-2008 submissions for this program cycle.

6.1.3 Adjustments for Avoided Line Losses

Energy costs that are expected to be avoided by the use of DR resources must be adjusted for the line losses that will be avoided by the resulting reductions in energy consumption.

Most energy prices are quoted at the transmission voltage level (*e.g.,* for firm energy delivered to NP-15 or SP-15). On the other hand, DR resources reduce customer meter-level energy consumption. Therefore, those prices already reflect transmission voltage level line losses

Reductions in customer meter-level electricity consumption due to the use of DR resources not only avoid wholesale energy costs but also distribution voltage level line losses as well.

Distribution line loss rates vary over time. In particular, they tend to be higher when the ambient temperature and/or loads are higher, which is when DR program events are most likely to be called. Therefore, if the necessary data are available, market price-based estimates of the energy costs avoided by the use of DR resources should be adjusted based on the marginal transmission and distribution voltage level line loss rates for the periods in which DR is likely to reduce energy consumption, rather than average or typical line loss rates.

This framework recommends that avoided energy costs include value for avoided transmission and distribution losses.

6.2 Valuation

There are several different approaches to estimating the energy costs that will be avoided by event based and non-event based DR programs. It is recommended that the methods for estimation of market prices for event based programs and non-event based programs be different. Several methods are described below.

6.2.1 Event-based programs

Estimating the energy costs that will be avoided by event-based DR programs is usually more complicated because on an *ex ante* basis, there is a substantial amount of uncertainty about when and whether the events that might lead to a specific DR program to be dispatched are going to occur. In addition, energy prices are likely to be positively correlated with types of conditions under which a DR program might be dispatched. Ideally, valuation methodologies for event-based DR should capture the range of uncertainty about whether and when DR programs will be dispatched, as well as the extent of the correlation between energy prices and the occurrence of conditions under which DR programs are likely to be dispatched. There are a variety of acceptable methods to capture this uncertainty and correlation. However, this is also an area in which more complicated models may not provide significantly different or more accurate estimates of avoided energy costs.

Event-based DR programs typically have constraints on at least some of the following:

(1) the number of times that they can be called (events),
(2) how long they can be called when they are called (event duration),
(3) the number of days (consecutive or non-consecutive) on which they can be called, and,
(4) the aggregate number of hours during which they can be called over a period (*e.g.,* a summer season).

Fully modeling these types of constraints requires dynamic programming techniques. Although reflecting these constraints in a cost-effectiveness valuation model can be somewhat complicated, it is not impossible.

Incorporating these types of constraints in estimating energy costs that will be avoided by a DR program can be complicated. Therefore, it is usually appropriate only if incorporating those constraints is likely to provide significantly different or more accurate estimates of avoided cots. If not, reasonable approximations are acceptable.

One simple approach involves making reasonable assumptions about when and how often a program will be dispatched. For example, it can be assumed that a DR program will be dispatched only during peak price periods. Peak prices forecast by methods discussed in section 6.1 **Error! Reference source not found.** can then be used to estimate the energy costs that will be avoided by that DR program..

In estimating how often a program might be called, it is important to bear in mind that most DR programs usually are not dispatched as frequently as their limits allow. Therefore, it is reasonable to estimate a range of avoided energy costs, by using a range of assumptions about how often DR program events will be called (*e.g.,* no dispatch, dispatch up to the limits of a program, and perhaps some intermediate level of dispatch). Alternatively, an assumption can be made about the proportion of the time the program will be dispatched during the hours in which that program will be available.

In cases in which making reasonable assumptions has the potential to significantly impact the accuracy of estimates of the cost-effectiveness of a DR program, *i.e.,* when avoided energy costs are a significant fraction of the total value of a program, more complicated methods may be warranted. The appendix to this chapter (Appendix B) provides detailed descriptions of more complicated methods for calculating avoided energy costs for event-based DR.

6.2.2 Non-event based programs

It is comparatively easy to estimate avoided energy costs for non-event based programs because they produce load changes in well-defined, pre-determined time periods.

Non-event based DR resource resources such as Permanent Load Shifting. TOU pricing, and Scheduled DR programs produce regular and predictable changes in load. In some cases, all or a portion of that load will be shifted from higher to lower price periods.

Because these programs produce regular and predictable changes in load that are (usually) not related to wholesale market prices or other exogenous factors, the energy costs avoided by a DR resource can be calculated in the same way as avoided energy costs are calculated for energy efficiency. For example, the net energy costs that a Permanent Load Shifting program will avoid is determined by the sum of the amounts obtained by multiplying the amount by which load is expected to be reduced (or increased) in each time period due to that program, by the electricity prices in each of those time periods, after making appropriate adjustments for line losses, location, *etc.*.

Real-time pricing is sometimes classified as non-event-based DR. To the extent that the utilities introduce real-time pricing programs, more complicated avoided energy cost valuation methods will be warranted. Estimating the avoided energy costs associated with real-time pricing requires not only projections of electricity prices or ranges of electricity prices, but also detailed models of how consumers are likely to respond to these prices. Developing these models is closely related to the techniques discussed in the load impact protocols.

6.2.3 Projecting Energy Prices at Different Locations

DR capacity may have additional value in a load pocket, such as San Diego, that is not reflected in market prices for electricity delivered to broader geographic areas such as SP15.

However, there are a variety of problems associated with assigning an avoided energy cost to DR that fully reflects the value of the location of that DR. First, it may be difficult to determine the precise location of the resources involved in a DR program. The IOUs have large and geographically diverse service territories. For example, the market value of energy in a coastal climate zone may be lower than the value of energy in the Central Valley. Similarly, energy delivered to densely populated areas where it is more difficult to build conventional resources may have higher value than energy delivered to less densely populated areas.

Under MRTU, separate generation pricing nodes will be established at substations and other facilities throughout the existing zones. Prices will be determined at each of those nodes. As a result, intra-zonal congestion and intra-zonal line losses will be implicitly reflected in the energy prices for each node. Once an understanding of how prices vary by node emerges, it may be possible to incorporate locational information in the cost effectiveness evaluation of DR resources.

However, this framework recommends that evaluations of the cost-effectiveness of the initial round of program design which takes place for the 2009-2011 cycle should not include any geographic granularity. There will be very limited information available (if any) regarding locational prices in time for preparing the mid-2008 submissions for this program cycle.

7 Economic Benefits: Avoided T&D Capacity Costs

7.1 Overview

By providing for peak load reductions, DR programs have the potential to defer and possibly avoid transmission and distribution ("T&D") capacity investments. There peak load reductions may be for a transmission system or for specific sub-areas targeted by a DR program.

The framework described below is designed to be flexible and to acknowledge possible differences in how DR is integrated into the planning of T&D systems by the three IOUs. This flexibility is achieved by proposing a series of guidelines for the valuation of avoided T&D capacity costs. These guidelines are designed to recognize differences in the attributes of specific DR programs as well as differences in the T&D systems of the three IOUs.

7.2 Transmission and Distribution Systems

The transmission and distribution systems built, maintained and to a large extent operated by the IOUs are comprised of three key elements: 1) interties, 2) local network transmission, and 3) local distribution systems.

7.3 Capital Investment Decisions

The IOUs' T&D systems have and will continue to require significant capital investments both to meet increases in regional load growth and replace components that have reached their useful life. In many instances, these investments are projected to provide benefits (*i.e.,* the economic life) and to require some maintenance for multiple decades.

Decisions to make investments in the T&D systems are guided by the obligation to provide reliable service at reasonable cost. To meet this obligation, T&D systems are developed by planning for various contingencies such as equipment failures and other difficult to predict but expected events (*e.g.,* high winds, fires in remote locations, car hits pole). For major transmission systems, multiple redundancies are built into the system to avoid customer interruptions when contingencies do occur. To provide guidance to those responsible for planning the T&D systems, these redundancies generally are incorporated into various planning criteria for various elements of the system. The development and application of these criteria have occurred over time and reflect the experience of both operators of the T&D systems and policy makers.

Implementation of these planning criteria is based on peak load forecasts for the relevant geographical area. For distribution systems, these forecasts are for specific geographic areas. The objective is to determine the maximum peak load that a specific line or system may experience over a specific time period (*e.g.,* 1 in 5 years) both with all equipment operating as well as with one or more contingencies. The specific time period considered as well as the number of contingencies considered will depend on the number of customers likely to be affected if the projected maximum peak load is exceeded requiring customer load to be interrupted. The timing of the maximum peak period for a sub-area of the utility's system will more than likely not be coincident with the system peak.

In making these projections, the characteristics of the customer load are a key concern. Whether the load is attributable to a few large industrial customers with distributed generation facilities or many residential customers will affect the projections of peak load requirements. For areas dominated by industrial customers, the peak load requirements for the distribution system may occur when the distributed generation equipment of one or more customers is out of service for maintenance. For areas dominated by residential customers, the peak load requirements for the distribution system will frequently occur late in the afternoon or early evening of a day with unusually high temperatures.

7.4 Deferral of Capital Investments

Given the importance of peak load requirements for a number of T&D capital investment decisions, DR programs have the potential to defer and, in some instances, to avoid capital investments. The extent to which DR programs may defer or avoid specific capital investments depends on a number of factors, including:

1. The characteristics of the existing IOU system, including the extent to which DR programs are being used to manage both system load and load in geographically targeted areas;

2. The specific T&D investment proposed (*i.e.,* the base case);

3. The characteristic of the customer load to be served by the proposed T&D investment;

4. The attributes of the proposed DR program;

5. The level of uncertainty associated with the projected load impacts of the DR program.

For example, T&D capacity deferral benefits of a specific DR program will depend on the extent to which the load projections for the system and for specific geographical areas materialized. A decision to not go forward with the planned building of a large housing development in the Inland Empire or in the Central Valley would have an effect on the capital investments in distribution and possibly transmission systems.

The T&D capacity deferral benefits of event-based DR programs will also depend on the extent to which other event-based DR programs have been deployed in the same geographical areas. Greater diversity (*i.e.,* more participating customers) in the DR programs will reduce the implications associated with equipment failure, such as problems communicating with the customers.

7.5 Quantification Guidelines and Methodology

This section provides general guidelines and recommendations for incorporating the benefits of DR in deferring capital investments associated with T&D systems.

1. Capacity (not energy) benefits: Quantification of capital deferral benefits for T&D should be in terms of $/kW (and not $/kWh). As described above, capital investments in T&D systems are dependent on peak load projections and not on energy projections.

2. Separation of Distribution from Transmission: Differences across T&D systems are such that estimates of the value of deferring distribution investments should be distinguished from estimates of the value of deferring sub-transmission and transmission investments.

3. Flexible Methodology: Methodology should be flexible in recognition of the differences in the various factors influencing the benefits of deferring capital investments in T&D systems over the long term. These factors include the characteristics of the T&D systems, the extent to which event-based DR programs have been deployed, the attributes of the proposed DR program, and the certainty of the DR program load impact estimates.

4. Identification of T&D Projects: The T&D projects that may be deferred or avoided for each IOU system need to be identified for each IOU system and each DR program. Given the differences of T&D systems and DR programs, consideration should not be given to developing a generic adder to be applied to all IOU systems In some instances, the share (or slice) of T&D capital investments that can be deferred by relatively certain reductions in peak load requirements either for a utility system or a specific geographical area would be determined. In other instances, the specific T&D project and its associated capital investments to be deferred or avoided by a given DR program would be identified. The extent and timing of the assumed deferral of the T&D capital investments will depend on the certainty of the projected peak load projections from the DR program. Experience with the specific DR program may be required prior to including the deferral of specific T&D capital investments in the cost-effectiveness for the specific DR program (*e.g.,* no avoided T&D capital costs are assumed for the first two years following program implementation). In other instances, a customer commitment of multiple years may be called for in order to provide sufficient certainty for the deferral of distribution capital investments.

5. Estimation of Benefits of Deferred Capacity: The methodology for quantifying the avoided (or deferred) capacity benefits for T&D should be consistent with the methodology used to calculate avoided capacity costs. Values should be expressed in $/kW-year. All assumptions regarding escalation rates and carrying costs should be explicit.

6. DR Attributes: Attributes of specific DR programs should be clearly described along with the basis for projected load impacts and the duration of those impacts. The uncertainty associated with these load impacts needs to be presented in such a way to allow for a determination of the probability of reducing peak load reductions on specific sub-systems of the transmission and distribution system consistent with individual utility's existing planning criteria.

7.6 Present Value of Avoided T&D Capacity Costs

The cost effectiveness of a DR resource is evaluated by comparing the present value of the expected future costs of that resource to the present value of the expected future benefits of that resource. Therefore, the future T&D capacity costs that a DR resource is expected to avoid should be discounted using an appropriate discount rate to allow for the calculation of the present value of the stream of future benefits. Section 10.2 discusses the determination of that discount rate.

The present values of the economic benefits and economic costs used in applying the TRC test to evaluate the cost effectiveness of a DR program (or portfolio of programs) should be based on expected cash flows, not expected book expenses determined under financial accounting rules. Therefore, whether specific types of T&D capacity costs should be treated as expenses or capitalized and amortized over time for accounting purposes is irrelevant.

8 Other Possible Economic Benefits

8.1 Overview

Other potential economic benefits of DR programs have been identified in addition to the financial benefits associated with avoided capacity, avoided energy, and avoided T&D.[49] For purposes of the discussion below, those potential economic benefits are grouped into four major categories: (1) environmental, (2) reliability, (3) market performance; and (4) energy efficiency. Those benefits that do not logically fit in these four categories are put into a fifth category entitled "other".

To determine whether and how to incorporate each of these other potential economic benefits into the evaluation of the cost-effectiveness of DR programs, the following questions should be addressed:

1. Is that potential economic benefit already included in the financial benefits attributable to avoided capacity costs, avoided energy costs, and avoided T&D capacity costs?

2. If that potential benefit is in addition to these financial benefits, are methods available to develop reliable and reasonably accurate estimates of the value of that potential benefit?

3. If it is possible to obtain reliable and reasonably accurate estimates of the value of that potential benefit, how significant is the relative value of that potential benefit for the DR program (or portfolio of DR programs) that is being evaluated?

4. If the value of that potential benefit is likely to be significant, can the method selected to value the benefit be used in a manner that is consistent with the objectives of transparency and accuracy at a reasonable and prudent administrative cost, while not disclosing commercially sensitive information that IOUs are allowed to keep confidential under relevant statutes and Commission decisions?[50]

The following sections discuss each of these other potential benefits attributed to DR, summarize the recommendations on whether or not those potential benefits should be included in analyzing the cost effectiveness of DR resources and describe the methods that should be used to evaluate each of the potential benefits that should be included.

8.2 Environmental Benefits

8.2.1 Overview

Environmental benefits are attributed to DR programs based both on the extent these programs avoid or defer development of supply-side resources and on the extent to which these programs avoid the environmental effects of generating electricity. In a number of instances, the costs of environmental

---

49 U.S. Department of Energy (2006) "Benefits of DR in Electricity Markets and Recommendations for Achieving them: A Report to the United States Congress Pursuant to Section 1252 of the Energy Policy Act of 2005," February 2006; California Energy Commission, Public Interest Energy Research Program (PIER) (2006) "Development of a Comprehensive/Integrated DR Value Framework," prepared by Summit Blue Consulting, Contract No. 500-03-026, January 2006; California Energy Commission (2007) "The State of DR in California: Draft Consultant Report," CEC-200-2007-003-D, April 2007.

50 *See e.g.,* Section 454.5(g) of the California Public Utilities Code and D.06-06-066.

compliance are already included in the financial benefits associated with the avoided capacity, avoided energy, and avoided T&D capacity attributed to DR programs. To the extent this is the case, these environmental benefits should not be separately valued to avoid possible "double counting".

To the extent a DR program causes certain customers to shift load from one time period to another, rather than simply reduce load in one time period, the increase in energy generation as a result of the load shift will influence the environmental benefits associated with the DR program. Shifting load from one period to another could have a positive environmental impact if the shifted load is served by a mix of generation resources with a lower emission rate (*i.e.,* pollutants emitted at lower rates) than that of the mix of generation resources that would otherwise have been used in the periods from which that load was shifted.[51]

8.2.2 Air Quality Criteria Pollutants (NOx, PM-10, Sox and VOCs)

Criteria pollutants are those pollutants that are presently regulated by local air quality management districts consistent with state and federal air quality regulations. The criteria pollutants of most importance are NOx, PM-10, VOCs and SOx. The potential damage associated with these pollutants depends on the concentration of these pollutants as well as ambient air conditions and is not a simple linear relationship (*i.e.,* each additional unit added to the air does not equate to the same level of damages and the level of damage depends on a number of factors, including temperature). At present, there are a number of differences among the regulations governing criteria pollutants in the various local air quality management districts.

In evaluating the cost-effectiveness of DR programs, the criteria pollutant emission control and compliance costs avoided by reductions in criteria pollutants are appropriately included as a benefit. These avoided emission control and compliance costs may be included in the capacity and energy costs avoided by that DR resource.

Estimates of avoided capacity costs should be based on utility-specific new CT capacity (see Chapter 5). If the costs of complying with the requirements of the local air quality management district have been included in the cost of that new CT capacity, those compliance costs should not be double counted because the benefits of avoiding those compliance costs are already included in the avoided capacity costs.

---

51 DR load reductions frequently occur during summer daytime hours in which temperatures are very high. As a general rule, these high temperature hours often correspond to times of poorer air quality and "Spare the Air" days when the air quality can be harmful to human health. These also will be days when ambient air quality standards established by local air quality management districts may potentially be exceeded.

If the DR resource reduces generation by peaking plants on these days with poor and unhealthy air quality, there may be an additional benefit to public heath associated with these DR programs. Additionally, there may be an additional benefit by DR associated with having avoided violations of an ambient air quality standard. The violation of such standards for a certain number of days each year can lead to higher costs, if local air quality management districts take actions that impose costs on the communities to reduce these violations of air quality standards.

DR programs can also help avoid the environmental implications associated with CAISO Stage 3 events (rolling blackouts). During Stage 3 events, some customers are expected to turn on their emergency back up generators. These generators are typically allowed to operate only during such emergencies because of their high pollutant emissions rates and minimal emission controls. These back-up generators typically burn diesel fuel, which emit air toxics at higher rates than conventional power plants, and have lower stack heights, which do less to dilute those emissions. This may also lead to violations of air quality standards and any resulting increased costs for those using back-up generators or for the community as a whole.

Estimates of avoided energy costs should be based on market prices for electricity (see Chapter 6). These market prices are likely to include any variable costs associated with compliance with emission restrictions. Therefore, as a general rule, there is no need to include adders to account for the benefits of avoiding emission compliance costs that are already reflected in avoided energy costs.

There may be residual benefits associated with reductions in emissions of criteria pollutant emissions and air toxics over and above those mandated by existing compliance requirements. These residual benefits might include avoiding violations of ambient air quality standards during days of poor air quality and avoiding the use of emergency diesel generators during CAISO Stage 3 events. At present, the valuation of these residual benefits is likely to be quite speculative. One possible exception is the case in which these residual emissions would have caused customers or IOUs to incur significantly higher emission control costs, fines or other costs as a result of actions on the part of the air quality regulator. In this situation, a specific analysis of these benefits (*i.e.,* the potential compliance costs avoided by DR programs) could be incorporated in the cost-effectiveness analysis.

8.2.3 Greenhouse Gases (GHG)

At present, there is widespread recognition of the benefits associated with reducing greenhouse gases. This past year, AB32 was passed and signed into law creating a framework for a GHG emission reduction to be implemented in 2012.

Implementation of this legislative directive designed to limit GHG emissions will result in compliance costs for electric ratepayers. As with criteria pollutants, DR programs would avoid these compliance costs and incorporation of these avoided compliance costs in the cost-effectiveness analysis for DR programs would be appropriate beginning with 2012.

At present, there is considerable uncertainty associated with these future compliance costs. The approach suggested is to value avoided CO2 emissions starting in 2012 in a manner that is consistent with Commission direction in D.05-04-024 (*i.e.*, only for evaluation purposes). In that Decision, the Commission suggested that a reasonable estimate of the avoided benefits associated with avoided CO2 emissions would be $8 per ton (annual levelized costs for 2004) with a 5 percent nominal annual escalation rate for an initial period. Under this framework, the estimates of the volume of CO2 emissions avoided by a DR resource would be based on the operating and emission rate characteristics of utility-specific new CT capacity (see Chapter 5).

The proposed approach described above should be re-evaluated as additional information becomes available on federal and regional programs to limit GHG emissions, as well as the including the AB 32 cap and trade framework.

8.2.4 Land Use, Water Quality and Other Environmental Impacts

DR programs may have the potential to defer or avoid the use of land for large scale generation capacity projects and for T&D installations. By avoiding some energy consumption, these programs may also contribute to improving water quality by reducing water consumption, and avoiding the impacts of heated water discharged by water cooled power plants.

There are several other environmental impacts that might be avoided depending on the specific type(s) of capacity– generation, transmission, or distribution -- that the DR program is expected to defer or avoid.

These potential environmental impacts include environmental justice (particularly for supplying electricity in urban areas), biological impacts, impacts on cultural resources, diminishing visual resources (*e.g.,* due to power plant stacks or transmission towers), and noise pollution.

As with criteria pollutants, the preferred approach is to incorporate these benefits in cost-effectiveness evaluation of DR programs by incorporating the compliance costs in the avoided energy, avoided generation capacity, and avoided T&D capacity costs. However, as with criteria pollutants, there are potentially residual benefits in addition to existing compliance costs that should be addressed. In most instances, these residual benefits are extremely difficult to value in a reliable manner, and should not be quantified for inclusion in the cost-effectiveness analysis of DR programs. However, there may be specific situations in which it can be demonstrated that those additional environmental impacts would have caused regulatory agencies to require IOUs or communities to incur significantly higher control costs or fines. In those specific cases, the value of avoiding those additional environmental impacts could be based on the additional control costs or penalties that IOUs or customers would have incurred due to those regulatory actions. The benefits attributed to DR in avoiding these costs could be included in the cost-effectiveness analysis of DR programs.

8.3 Reliability Benefits

Reliability benefits have been attributed to DR programs, including reductions in energy procurement hedging costs and reductions in the volatility of market prices. In a report prepared for the U.S. Department of Energy ("DOE"), reliability benefits have been defined as "the operational security and adequacy savings that result because DR lowers the likelihood and consequences of forced outages that impose financial costs and inconvenience customers."[52]

In a report prepared for the CEC, reliability benefits have been grouped with risk management benefits in a category titled "Risk Management and Reliability." In describing this category of benefits, five specific benefits are identified. These include: provision of a physical hedge against extreme events, lower cost of "insurance" against extreme events, real option flexibility/portfolio resource diversity, improved ancillary services, and possible opportunities to manage financial and/or outage risks. Both physical and financial hedging benefits are addressed under the following section entitled "Market Performance."

All of these benefits are potentially associated with DR programs. However, in most instances, the same benefits may also be provided by new combustion turbines and other supply-side resources. Whether or not these benefits should be attributed to DR programs depends on what is assumed to be the supply-side alternative to the DR programs.

At present, IOUs are responsible for providing electricity service to customers at a level of reliability that is consistent with the capacity planning reserve margin established by the Commission. IOUs accomplish that objective by using a mix of supply-side and demand-side resources that complies with the level of reliability specified by the Commission and the loading order preference for cost-effective DR.

[52] U.S. Department of Energy (2006) "Benefits of DR in Electricity Markets and Recommendations for Achieving them: A Report to the United States Congress Pursuant to Section 1252 of the Energy Policy Act of 2005," February 2006, p. vi.

As a general rule, the underlying assumption in cost-effectiveness evaluations is that DR programs can be substituted for an alternative supply-side resource (generation capacity or T&D capacity). In those instances, the reliability benefits attributed to DR programs can also be obtained by using that alternative supply-side resource. Including these benefits for DR resources would be inappropriate unless these benefits are also attributed to the supply-side resource. Because the DR resource and the alternative supply-side resource provide the same amount of reliability in most instances, the DR resource does not provide more reliability than the supply-side resource.

As a general rule, the recommended approach is to specify the level of reliability to be maintained (*e.g.,* the capacity planning reserve margin adopted by the Commission and any equivalent requirements for T&D capacity), and to include the costs of maintaining this specified level of reliability in the costs of the alternative supply-side resource as well as in the costs of the DR resource.[53] This should be accomplished by properly defining and measuring avoided generation capacity costs and avoided T&D capacity costs, not by attributing additional reliability benefits to a DR resource that could avoid the need for that generation capacity or T&D capacity.

8.4 Market Performance Benefits

8.4.1 Overview

DR programs provide customers with improved price and event-triggered signals, thereby reducing energy use during high-priced periods and increasing energy use during low-priced periods. These changes are expected to provide system and overall societal benefits by increasing the efficiency of energy use. In addition, the reduction in energy use is also expected to reduce on-peak energy prices and to potentially reduce the incentives for generators to exercise market power.

In considering the potential benefits associated with improved market performance, one needs to specify whether DR programs are assumed to replace a proposed supply-side resource or are assumed to be in addition to proposed supply-side resources. In evaluating the cost effectiveness of DR programs, the underlying assumption is that DR programs usually avoid supply-side resources, rather than increasing the total amount of resources needed to provide electricity service at a given level of reliability. Due to differences between DR programs and supply-side resources, such as the ability to quickly expand and geographically target existing DR programs, there are some instances where DR programs to provide benefits not provided by supply-side alternatives. These benefits are addressed under the section entitled "Modular Nature of DR."

8.4.2 Improved Efficiencies from Improved Price Signals

As has been widely recognized, the replacement of "flat rates" by time-differentiated or dynamic prices provides customers with the information required to more closely match their use with the costs of

[53] It should be noted that in addition to providing capacity and energy value, a new CT recommended in this framework as the alternative supply-side resource can provide ancillary services by bidding into CAISO ancillary service markets. Currently, DR programs in California are not being used to supply ancillary services to the CAISO, although such opportunities are being investigated. For example, the Federal Energy Regulatory Commission (FERC) has recently requested comments on this topic. If a DR program provides a different level of ancillary service value than the CT proxy, either higher or lower, then it may be appropriate to make an adjustment to reflect the DR program's relative potential for ancillary service revenues. For simplicity, the ancillary services value is embedded as part of the capacity value described in detail in Chapter 5.

supplying the electricity. Without time-differentiated rates, customers "overpay" for electricity during off-peak hours and "underpay" for electricity during on-peak hours. In addition, customer utilization is below the optimal usage levels during off-peak hours, and above the optimal usage levels during on-peak hours. The economic result of flat rates is what has been termed a "dead-weight loss" or resource loss.[54]

Estimates of the value of this "dead-weight loss" have been developed using illustrative information on customer demand curves at different time periods with and without time-differentiated prices and on supply curves for generation resources.[55]

As suggested, there are potential benefits associated with reducing the dead weight loss that could be attributed to improved price signals. There is a question whether estimates of the dead-weight loss could be developed that would be sufficiently robust without requiring the use of confidential information (*e.g.*, the expected net short or long positions of an IOU during the hours in which DR resources are most likely to be utilized).

8.4.3 Price Elasticity and Market Power Mitigation Benefits

As noted above, several observers have suggested that DR programs can mitigate the exercise of market power on the part of supply-side providers, because DR resources can reduce market prices during high-priced hours either through event-triggered programs or high price signals. The underlying assumption is that in those hours there are no remaining supply-side resources available either due to physical limitations or the withholding of resources from the market, and, as a consequence, the supply-side curve in those hours is relatively inelastic (*i.e.*, small reductions in supply provided can lead to significant price increases). By reducing demand, the DR program is assumed to reduce market prices and therefore the benefits generators would receive by withholding supply from the market in those hours.

The proposed framework for evaluating DR assumes that, in most instances, DR programs are substitutes for supply-side resources and are not in addition to supply-side resources. Therefore, DR resources do not simply shift the demand curve to the left with the implied reduction in market prices. Instead, the assumption is that while DR programs shift the demand curve to the left (*i.e.*, a higher demand at any given price), the substitution of the DR program for the supply-side resource means that the supply curve also shifts to the left (*i.e.*, a larger supply at any given price). As a consequence, if DR resources replace supply-side resources, there are no benefits in terms of reduced prices and market power mitigation.

While reducing market power is a desirable objective, any benefits of mitigating market power that may be provided by DR programs could also be provided by the alternative supply-side resources. Treating market power mitigation as a benefit provided by a DR resource is not appropriate unless market power mitigation benefits are also attributed to the avoided supply-side resources.

---

54 U.S. Department of Energy (2006) "Benefits of DR in Electricity Markets and Recommendations for Achieving them: A Report to the United States Congress Pursuant to Section 1252 of the Energy Policy Act of 2005," February 2006, pp. 69-72.

55 California Energy Commission (2007) "The State of DR in California: Draft Consultant Report," CEC-200-2007-003-D, April 2007 "Brattle Report", pp. 62-63.

8.4.4 Reduced Price Volatility

The same arguments made in Section 8.4.3 about market power mitigation benefits can be applied to the potential benefits of reduced price volatility attributed to DR programs. Supply-side resources can also reduce price volatility. As such, the same reduced price volatility benefits are also available from supply-side resources. Increasing the planning reserves would also provide the same benefit. Attributing these benefits to DR resources would not be appropriate because the same benefit can be provided by the alternative supply-side resources. Therefore, potential reductions in the volatility of market prices should not play any role in evaluating the cost-effectiveness of DR resources.

8.4.5 Physical Hedging Benefits

Some have asserted that DR resources provide an additional benefit by serving as physical hedges against unusually high levels of demand. These benefits are classified as a market performance benefit, but could also be classified as a reliability benefit. Both DR resources and the alternative supply-side resources can provide physical hedges against unusually high levels of demand that have a very low probability of occurring. As a result, the physical hedging benefit available from a DR resource is, all other things being equal, the same as the physical hedging benefit that is available from the alternative supply side resource.

To the extent that the physical hedging benefits of a new combustion turbine and the DR resources are the same, those benefits should not play a role in evaluating the cost effectiveness of that DR resource.

8.4.6 Financial Hedges Against Procurement Costs

Some have asserted that DR resources provide an additional benefit by serving as a financial hedge against unusually high supply-side procurement prices.

Supply-side procurement costs include the cost of the generation capacity ($/kW), the cost of the energy ($/kWh), and any administration costs. Although administrative costs may be a function of the number of transactions, administrative costs are not directly related to the quantity of capacity or energy that is procured.

The cost of procuring a DR resource also includes the equivalent costs to allow for avoiding capacity and energy costs. As such, the extent to which a DR resource actually provides a financial hedge against unusually high supply-side procurement prices depends upon the level, volatility and correlation between the cost of a DR resource and the cost of the alternative supply side resource.

The method recommended in this framework for evaluating the cost-effectiveness of a DR resource already takes into account the levels, volatilities and correlations between the cost of a DR resource and the cost of the alternative supply-side resource (*see* Chapter 5 and Chapter 6). Thus, there is no need to attribute a financial hedge benefit to the DR resource.

8.5 Other Suggested Benefits

8.5.1 Modular Nature of DR

The suggestion has been made that DR programs provide system operators with "more flexible resources" to meet contingencies.[56] For example, dynamic pricing can be changed easily year to year. Incentive programs such as peak time rebates can also be easily modified. This flexibility will also increase in the future as smart thermostats are widely installed. After wide-scale installation of smart thermostats, the incentives paid to customers can be adjusted from year to year from zero to large payments depending on the need for demand reduction.

While there are potential modular benefits associated with DR programs, there are also costs associated with these benefits. First, while DR may be a more flexible resource to meet contingencies, quick start supply-side resources may also be quite flexible in meeting contingencies. Second, discussions of the ability to put DR programs on hold tend to ignore the issue of maintaining customer enrollment and participation.

There are instances where existing DR programs have been expanded quickly to handle anticipated capacity shortfalls. Without these programs, customers might have been subject to rotating outages and the associated loss in value. The option of substituting a new CT in time to avoid these potential rotating outages was not available. For these instances where the modular nature of DR provides a benefit, a specific analysis is recommended. The value of the lost load associated with these benefits may be quantified but considerable judgment is required both in selecting and applying information on value of lost load in these instances.[57]

8.5.2 Development of Appropriate Technology and Behavioral Incentives

There are potential benefits attributable to technology development and to experience with the design of various program incentives that can be attributed to some DR programs, particularly those that are particularly innovative. Some would term this benefit as "learning by doing." Developing estimates of these benefits is extremely speculative and is not recommended at this time.

8.5.3 Portfolio benefits of an expanded set of options for meeting peak and high-cost loads.

There are potential benefits from including DR resources in the portfolio of resources for meeting peak and high cost loads, just as there are potential benefits from having a diverse mix of fuels and supply-side technologies for dealing with unanticipated events. In most instances, these portfolio benefits can be provided by both DR resources and supply-side resources. However, given differences in the attributes of DR programs compared to supply-side alternatives, there may be additional portfolio benefits for DR programs. These potential benefits would be considered on a detailed, case-by-case basis with attention given to avoiding attributing benefits that have been previously accounted for in other areas.

---

[56] U.S. Department of Energy (2006) "Benefits of DR in Electricity Markets and Recommendations for Achieving them: A Report to the United States Congress Pursuant to Section 1252 of the Energy Policy Act of 2005," February 2006, p.xiv.

[57] Reference to Chapter 5.

8.5.4 Customer options for better managing their electric bills

DR programs do give customers more choices for managing their electric bills. By participating in programs and by shifting load from high-priced to low-priced periods, customers are able to reduce their electricity bills. This benefit aligns with the Vision Statement of Working Group 1 in which DR programs provide customers greater control by enabling them to choose technology options appropriate for their situation and control strategies consistent with their preferences to control their costs.[58] As discussed in Chapter 9, there are some costs to customers from participating in DR programs, but they also receive the benefit of being able to better manage their electric bills.

Further work is required to better understand the value customers perceive receiving from DR programs prior to developing methods to quantify these benefits. At present, these benefits are already incorporated in the assumption described in Chapter 9 that the level of incentive is equal to the customer costs. Including a reduction of the customer cost below the incentive level of a DR resource is not recommended at this time.

8.6 Energy Efficiency and Program Costs

As noted above, DR programs are primarily focused on reducing peak load requirements. However, there is evidence that some DR programs have served to reduce overall usage and to improve energy efficiency.

The improvements in energy efficiency are attributable to a number of factors and do depend on the specific program characteristics. Some DR programs may result in a significant decrease in energy use during on-peak periods and no increase in off-peak periods. One obvious example would be a peak load reduction program that focuses on reducing commercial lighting during peak periods; there is no obvious reason to believe that customers would increase their lighting loads over prior levels during off-peak periods. In addition, DR programs may increase energy awareness and lead to behavioral changes and purchases of energy efficient equipment. There is also the possibility that the costs of implementing a combined DR and energy efficiency program could reduce program costs by this "holistic" approach to program implementation.[59]

The avoided costs of energy addressed in Chapter 6 already do incorporate these energy efficiency benefits to the extent these effects are reflected in the load impact estimates for DR programs. In most instances, those load impact estimates are based on actual program implementation and would not take into account changes in energy usage resulting from the impact on future behavior of increased energy awareness. These potential benefits of DR programs through encouraging energy efficiency and reducing energy efficiency program costs need to be acknowledged either in the evaluation of DR programs or energy efficiency programs.

Where evidence is available to support the relationship between DR programs and future reductions in energy usage, these reductions could be incorporated in future load projections. Similarly, where evidence exists to support the relationship between DR program participation and future program costs, these cost reductions could also be incorporated in the cost-effectiveness analysis.

---

[58] D.03-06-032, Appendix A (Vision Statement of Working Group 1).

[59] Nemtsow, David, Dan Delurey, and Chris King (2007) "The Green Effect: How DR Programs Contribute to Energy Efficiency and Environmental Quality," Public Utilities Fortnightly, March 2007, pp. 40-45.

In both areas, however, there is limited program specific information to justify incorporating these specific benefits at this time. Further work is needed to better understand the relationship between DR program participation, energy awareness, and future reductions in energy usage and program costs.

8.7 Recommended Approach for Issues Addressed on Page 32 of Staff Guidance Document

The Staff Guidance at page 32 contains a list of ten potential additional benefits attributable to DR.[60] This objective of this section is to explain how these benefits are addressed in this framework. These benefits are listed in Table 8-2 with a reference to the section in this framework that describes how these benefits are addressed.

[60] Staff Guidance at p. 32.

Table 8-2: Page 32 Issues

"Page 32 Issue"	Where Addressed in Framework
1. Reduced market power by allowing demand to adjust in response to higher prices. Increasing the price elasticity of demand decreases market power	Section 8.4.3
2. Lowered market prices for all customers, not just DR participants	Chapter 5 (avoided capacity costs); Chapter 6 (avoided energy costs); Chapter 7 (avoided T&D costs); Section 8.4 (market performance)
3. Lowered long-term trend in rising electricity costs	Chapter 5 (avoided capacity costs); Chapter 6 (avoided energy costs); Chapter 7 (avoided T&D costs); Section 8.4 (market performance)
4. Appropriate technology and behavioral incentives for development of methods to reduce peak loads and shift peak loads to off-peak periods	Section 8.5 (other suggested benefits)
5. Portfolio benefits of an expanded set of options for meeting peak and high-cost loads (can impact reliability)	Section 8.5 (other suggested benefits)
6. Hedge values due to mitigating the impacts of adverse energy market outcomes	Section 8.3 (market performance)
7. Ability to locationally target DR to address geographic market constraints in capacity and/or transmission	Chapter 5 (local resource adequacy); Chapter 7 (avoided T&D capacity costs)
8. Modular nature of DR in that it is built up in increments over a period of time, and can be maintained when there is adequate capacity, *i.e.,* no new customers are added	Section 8.6
9. For customers, DR can provide them with options for better managing their electric bills, either through time-differentiated/dynamic pricing or by participating in a curtailable load program	Section 8.5 (other suggested benefits)
10. Environmental benefits in terms of reduced emissions, land use and water use. This also reduces the risks associated with uncertain environmental compliance costs in the future	Section 8.2

8.8 Present Value of Other Economic Benefits

Under the TRC test or the Societal TRC test, the cost effectiveness of a DR resource is evaluated by comparing the present value of the expected future costs of that resource, to the present value of the expected future benefits of that resource. Therefore, the other economic benefits/costs that will be avoided by a DR resource should be discounted and converted to a present value. Chapter 10 discusses the determination of that discount rate.

The present values of the expected economic benefits and economic costs used in applying the TRC and societal TRC test to determine the ex ante cost effectiveness of a DR resource should be based on expected cash flows, not expected book expenses determined under financial accounting rules. Therefore, whether other economic benefits would be treated as expenses, or capitalized and amortized over time, for accounting purposes is irrelevant.

9 Discussion of Specific Cost Issues

9.1 Present Value of DR Costs

As noted in earlier sections, under the TRC test or the Societal TRC test, the cost effectiveness of a DR resource is evaluated by comparing the present value of the expected future costs of that resource, to the present value of the expected future benefits of that resource.

Therefore, the future costs that are expected to be incurred under a DR or dynamic pricing program, should be discounted and converted to a present value. Section 10.2 discusses the determination of that discount rate.

9.2 Costs Based on Cash Flows, Not Book Expenses

The present values of the economic benefits and economic costs used in applying the TRC and societal TRC test to determine the ex ante cost effectiveness of a DR resource should be based on expected cash flows, not expected book expenses determined under financial accounting rules.

Therefore, whether specific costs would be treated as expenses, or capitalized and amortized over time, for accounting purposes is irrelevant.

9.3 Customer Costs

Customer costs are defined as those costs incurred by the customer to participate in a DR program. As discussed in Chapter 3, a DR program is designed to reduce customer demand during specific times of high demand as an alternative to increased supply. The customer's cost can include the incremental capital costs of equipment needed to shift or reduce demand, and participation costs such as employee time to implement a demand reduction strategy.

In addition, unlike energy efficiency programs that assume the customer will experience the same value with less energy consumption, DR participants may experience reduced value from participation in a DR program. This can include such items as lost service or reduced product production. Finally, participation can result in increased costs due to increased electricity prices or penalties.

9.3.1 Differences Based Eligibility Requirements

One critical factor in determining customer costs is the eligibility requirements for customer participation. A DR program can be 1) voluntary opt-in, 2) default opt-out, or 3) mandatory. In the case of a voluntary opt-in program it is reasonable to assume that the participant costs are less than or equal to the incentives offered by the program; otherwise the customer would not voluntarily choose to participate. In the case of a program that is defaulted on a customer with the option to opt-out, participation costs may also be assumed to be less than or equal to the incentive. However, this second situation is likely to include a subset of participants that experience a participation cost but because of inertia or lack of awareness of alternatives fails to opt-out of the program. Finally, the mandatory DR program will most likely include a group of participants that will experience participation costs.

9.3.2 Value of Service Loss

An additional factor in determining customer costs will result from the customer segment participating in a DR program. The value of service differs by customer segment. Residential customers tend to have less lost value from load loss than business customers in value of service studies. Business customers are more likely to be able to quantify the value of lost or reduced production. Agricultural customers tend to be seasonal and the participation cost will vary significantly based on the time of year.

9.3.3 Impact of Program Characteristics

Finally, customer participation costs will be affected by DR program characteristics. For example, inclusion of enabling technology in the program design, such as auto Critical Peak Pricing (CPP), could affect participant costs if the participant is expected to contribute any of the cost of the technology. Program characteristics that require monitoring prices or reacting to notifications could cause participants to incur costs associated with employee training and participation in initiating demand responsive actions.

As a result, a forecast of customer costs will depend on the program characteristics, customer segment, and eligibility requirements. For example, if the program is a voluntary residential air-conditioning load management program where the capital costs are included in program costs and the customer has over-ride options if they perceive a reduced value of service, the customer costs can be assumed to be less than or equal to the incentive. Alternatively, if a similar program was offered without an over-ride alternative, a loss of value could be experienced by some participants. And finally, if the air-conditioning program was a voluntary program where the customer was required to invest in the load management technology to participate, the cost of the technology would be estimated for all participants and included in customer costs.

9.4 Incentive Payment Costs

Incentives paid to customers that participate in DR programs can be divided into two categories:

(1) "capacity" incentives paid for the availability of DR capacity, which are paid regardless of whether or not participants were actually required to reduce demand in a given period; and,
(2) "energy" incentives paid in proportion to the amounts by which participants actually reduced demand during a given period.

Total expected capacity incentive payments can be estimated based on:

(1) the number of customers that are expected to participate in the program in each period,
(2) the amount of DR capacity that will be available from those participants (because that usually determines the total amount of capacity incentives they will be receive in each period);
(3) the incentive amount that a participant will be paid for each unit of that capacity in that period.

Total expected "energy" incentive costs can be estimated based on:

(1) the amounts by which customers are expected to reduce their demand in each period (*i.e.,* an ex ante load impact forecast); and,
(2) the energy incentive that the program is expected to pay per unit of demand reduction in that period

As a result, ex ante estimates of energy incentive payments in each period will be proportional to the amount of energy that the DR program is expected to avoid in each period.

The amount of energy that the DR program is expected to avoid in each period should, in turn, be derived from the energy costs that program is expected to avoid in each period, divided by the expected price of energy in that hour. (The methods that should be used to estimate those avoided energy costs are described in Chapter 6).

9.5 Treatment of Incentive Payment Costs

The purpose and appropriate treatment of incentive payments is somewhat different in the case of DR programs than in the case of energy efficiency programs. Incentives are paid to participants in DR encourage them to begin and to continue participating in that program.

Frequently, participants in DR programs are required to accept some degree of discomfort (such as a higher household temperature during an air conditioning cycling program event) or reduction in business profits (such as revenue losses or increased costs when a business shuts down operations when an interruption is called). Therefore, incentives compensate customers for this reduction in the value they would normally receive from the beneficial use of electricity, and the additional "out of pocket" costs that they incur due to their participation. Currently these costs are not well known and can be difficult for a customer to quantify without program experience.

In contrast, energy efficiency programs encourage more efficient usage of electricity without requiring participating customers to change the benefits they obtain by using the products and services that require electricity. For example, more efficient appliances provide the same service as less efficient appliances, but use less electricity. Therefore, incentives paid to participants in energy efficiency programs primarily to overcoming customer inertia, market barriers such as cost of the appliances, or as a marketing tool, although there are some exceptions to these generalizations.

Incentives for some permanent load shifting programs, such as thermal energy storage systems, are more like typical energy efficiency incentives. Also, utilities have been exploring ways to use incentives as a marketing technique to increase DR program participation, such as SCE's give-away of a coffee vendor's gift card to new enrollees in its air conditioning cycling program.

Nevertheless, the differences between the typical application of incentives in DR and energy efficiency programs needs to be recognized in cost effectiveness analysis.

9.5.1 Incentive Payments Treated As Transfer Payments

Because a TRC (or all customers) test and a Societal TRC test only take into account the costs and benefits to all customers or society as whole, incentive payments are ignored because they do not change the total benefits and costs to all customers combined, or to society,

Therefore, in evaluating the cost effectiveness of DR resources, it is common to treat incentives paid to participants in energy efficiency programs as transfers from non-participating to participating customers, because the cost of paying those incentives are recovered from non-participating customers.

This would be appropriate in evaluating the cost effectiveness of DR programs, except in those cases where it would also be necessary to quantify the loss of service value which the participating customer incurs as a result of program participation.[61]

The next section discusses the circumstances under which it would not be necessary to quantify the loss of service value which the participating customer incurs as a result of program participation.

9.5.2 Incentive Payments as a Proxy for "Out of Pocket" and "Loss of Service" Costs Participants Expect to Incur

It is difficult to develop reliable and accurate estimates of the costs that customers would incur under a voluntary DR program. Absent current information on customer out of pocket and loss of service costs, it is recommended that the incentive payment serve as a proxy for those customer costs.

If participation in a DR program is voluntary, it is reasonable to treat incentive payments[62] as compensation for the "loss of service" and "out of pocket" costs that participating customers expect to incur under that program.

That approach is based on the reasonable assumption that a customer will only chose to participate in a voluntary DR program if the sum of:

(1) the present value of the loss of services that customer expects to experience due to participation in that program; plus,

(2) the present value of any equipment and other "out of pocket" costs that customer expects to incur due to participation in that program is less than or equal to the present value of the incentive payments that customer expects to receive due to participation in that program.

---

[61] The treatment of incentive payments as a transfer payment introduces complexity in the application of SPM tests to those energy efficiency programs which involve a net-to-gross ratio. Since net-to-gross ratios typically do not need to be considered in DR programs, this will usually not be an issue.

[62] The discount prices that participants in TOU pricing programs pay in off-peak hours, and the higher prices that those participants pay in peak hours simply transfer costs between participants and non-participants. Because such transfers do not change total costs to customers as whole or to society, these transfer are irrelevant under a TRC test or a Societal TRC test.

As a result, the present value of the net benefits that a customer expects to obtain under a voluntary DR program can be assumed to be positive because the present value of the (out of pocket and "loss of service") costs a participating customer expects to incur under a voluntary DR program, can be assumed to be no higher than the present value of the incentives that customer expects to receive under that program.

In applying the TRC test or the Societal TRC test, the incentives that are expected to be paid to participating customers can be used as a proxy for any costs that participating customers expect to incur due to participation in that program.

Furthermore, treating incentive payments under a voluntary DR program as a proxy for the costs that participants expect to incur is conservative because, as noted above, it is logical to assume that for some customers the present value of the incentive payments participants expect to receive will be higher than the present value of the costs that those participants expect to incur. As a result, treating expected incentive payments as a proxy for expected customer costs results in an understated benefit/cost ratio tends to understate, not overstate, the benefit/cost ratio.

Therefore, *ex ante* estimates of the cost effectiveness of voluntary DR programs should assume that the present value of the total costs participants expect to incur will be equal to the present value of the total incentives they are expected to be paid. If there is documented information available to indicate that total customer costs are lower than the incentives paid, the customer costs can be adjusted downward.

In the case of mandatory DR programs that are mandatory, however, the incentives that customers expect to receive may not fully offset the "out of pocket" and "loss of service" costs they expect to incur due to the program.

Therefore, in order to evaluate the cost effectiveness of mandatory DR programs, it may be necessary to develop estimates of the "out of pocket" and "loss of service" costs that would be incurred by customers that are expected to participate in the program. As noted earlier, these costs can be difficult for customers to estimate without program experience. Current research provides little insight. For example, an estimate of bill impacts of a mandatory pricing program could be used to estimate some customer costs. However, in some cases the higher bills could be mitigated with an appropriate customer education program. This topic should be included in further research on customer costs.

9.5.3 Applications Involving Third-Party Aggregators

Some utility DR programs use third-party aggregators to solicit customer participation and to implement load reductions when these programs are called upon to reduce usage.

Given the nature of the contracting process, the utility that administers the program may not know how much of the payments it expects to make to these third party aggregators will fund the incentive payments those aggregators expect to pay to the customers who participate in the third-party programs, as opposed to the aggregators' expected administrative and equipment costs (plus a reasonable expected rate of return).

If the DR program operated through aggregators is voluntary, and if the aggregators were awarded contracts through a thorough and competitive bidding process, for the reasons summarized above incentive payments can be treated as a proxy for:

a) the resource costs that aggregators expect to incur, plus,
b) the amounts by which the “out of pocket” and “loss of service” costs participants expect to incur exceed the incentives that participants expect to receive from those aggregators.

If, however, the utility contracts aggregators to operate a mandatory DR program, the incentives that aggregators expect to pay to participants may not fully offset the “out of pocket” and “loss of service” costs that those participants expect to incur.

Therefore, in order to evaluate the cost effectiveness of mandatory DR programs operated through aggregators, it would be necessary to:

(1) develop estimates of the “out of pocket” and “loss of service” costs that would be incurred by customers that are expected to participate in the program; and,

(2) obtain data on the costs that aggregators expect to incur, in addition to the incentives those aggregators expect to pay to participants.

However, third party aggregators may be less likely to bid for the right, or would charge higher amounts, to operate such programs if they are required to reveal confidential data on their costs and profits.

Therefore, in using the TRC or Societal TRC test to evaluate the ex ante cost effectiveness of a mandatory DR program, it should be assumed that the present value of the total costs that the participating customers expect to incur due to that program, equals the present value of the amounts aggregators expect to be paid in exchange for soliciting customer participation and implement load reductions when these programs are called upon to reduce energy consumption.

9.6 Program Costs

In order to start up a new or modified DR program and to provide for its on-going operations and administration, the program administrator will incur costs. The activities to which these costs are related are fairly typical (e.g. program management, supporting system enhancements, marketing, customer education, on-site hardware, etc.).[63]

Only incremental program costs should be included in an ex ante evaluation of the cost effectiveness of a DR resource. In other words, the evaluation should take into account only the costs that would be directly attributable to the resource, and that would not have been incurred in the absence of that resource.

Furthermore, an ex ante estimate of a DR resource usually should not take account of “sunk costs” incurred before the resource existed (or was being designed or developed).

[63] Whether these costs are treated as expenses, or capitalized and amortized over time is irrelevant, because the present values of benefits and costs used in applying the TRC and societal TRC test should be based on cash flows, not accounting numbers.

Some examples of sunk costs are the carrying and operating costs of existing communications infrastructure, the value of real estate housing program staff, salaries and benefits for employees who support the program but for whom costs would been incurred anyway (rate case funded). These are sunk costs and should not be included in a cost-effectiveness analysis.

Incremental costs attributable to DR resource would include payments to third parties for:

- Billing system and customer event notification system enhancements
- Customer notification, e.g. charge per call or page
- Web site development and maintenance
- Design and printing of promotional material
- Production of TV, cable or radio ads
- Market research activities
- Postage
- Purchase or leasing of dedicated server for program operations
- Consulting support
- Temporary staffing for customer enrollment and tracking activities
- Enabling technology installed at the customer's location
- Metering equipment provided free of charge to a DR participant that is not otherwise available to the participant under a utility tariff.

Program management costs may be considered incremental, if the costs are incurred to acquire resources or personnel specifically for that program, or represent personnel or resources that are transferred to the program after being replaced.

9.6.1 Matching of Costs and Benefits

In the start up phase of a DR resource, there may be significant costs incurred in the first year that provide benefits over a number of years. Some examples are participant acquisition costs through marketing campaigns and enabling technology provided by the administrator to the participant.[64] In order to achieve efficiencies in the scale of operations, it may be prudent to ramp up enrollment in a DR program as quickly as possible to reach an optimal level of DR. The resulting DR may persist for a number of years. In the case of enabling technology, it may have a service life of 10 years or more, even though the cost of acquiring and/or installing it is incurred in the year in which it is installed.

Under these circumstances, the period covered by the cost effectiveness evaluation should cover all the years in which benefits will be provided due to these costs.

As noted in Section 9.2, cost effectiveness evaluations should be based on cash flows. Therefore, whether specific costs would be treated as expenses, or capitalized and amortized over time, for accounting purposes is irrelevant.

[64] This example assumes that these costs are classified as expense, not capital.

9.6.2 Capital Costs

From a cash flow perspective there is no difference between costs that are accounted for as expenses and those that are accounted for as capital investments.

An asset whose acquisition cost is accounted for as a capital investment may have an extended useful service life. A common example of a capital item is automated metering equipment provided to the participant at no charge.

Under these circumstances, the period covered by the cost effectiveness evaluation should cover all the years in which benefits will be provided due to the cash flow cost incurred in acquiring that asset, rather than over the period during which that capital investment is amortized.

9.6.3 Incentive Payments

DR program incentive payments often are paid upon actual load reductions delivered during a DR program event. This framework recommends that for voluntary DR programs, these incentive costs will be based on the number of DR program event calls and the associated hours described in Chapter 6 to calculate avoided energy payments. The costs used in the cost effectiveness test may be lower than the total DR program budget, which is based on budgeted payment of the incentive for the maximum available hours that each DR program can be called.

10 Present Value and Discount Rate

As noted earlier in Section , under the TRC test (or the Societal TRC test), a DR resource is "cost effective" for utility customers as a whole (or for society as whole), if the economic benefits that DR resource is expected to provide, are higher than the expected costs of that resource.

If the costs and the benefits of a DR program occur in a single year, present value considerations are not a major concern.

However, if the benefits and/or costs of the DR resource will occur over a multi-year period, the DR resource will be cost effective, only if the present value of all the economic benefits that DR resource is expected to provide over that period, is higher than the present value of all the expected costs of that resource during that period.

The need to use present values to evaluate benefits and/or costs that will occur over an extended time period is a standard basic economic and financial principle that is reflected in the evaluation formulas presented in the California Standard Practice Manual provides (SPM).[65]

10.1 Time Horizon

For a two or three year program, the impact of discounting future years to present value is will be small. However, utilizing present values has a much greater impact if the program's benefits will occur over a long period of time, or if the timing of the costs is significantly different from the timing of the benefits that are due to those costs.

For example, the equipment installed under an air conditioning direct load control program may be expected to operate for fifteen or twenty years, even if the bulk of the costs (*i.e.,* equipment acquisition and installation costs) are incurred in the first few years.

In general, the longer the time period over which benefits and/or costs are likely to occur, and/or the greater the differences in timing between those costs and benefits, the greater the impact of discounting future nominal dollar-denominated costs and benefits to the valuation date used for the analysis.

Using the expected life of the benefits provided by the DR measure is consistent with the approach of the SPM.[66]

In order to use the expected life of those benefits, there is a need to make assumptions about what benefits will occur after the end of the period for which the DR program has been authorized.

---

[65] California Standard Practice Manual: Economic Analysis of Demand-side Programs and Projects, July, 2002, Table 1, page 5.

[66] California Standard Practice Manual: Economic Analysis of Demand-side Programs and Projects, July, 2002, Basic Methods section, page 3.

For example, for the purposes of ex ante cost effectiveness evaluations of energy efficiency programs, there is a rebuttable presumption that the projected benefits from measures installed in the last year of the authorized program (*e.g.,* 2011 in the 2009-2011 period) can be assumed to continue at the same level after the end of the period for which the program was authorized, until the end of the remaining life of the measures installed under that program.

Unlike energy efficiency programs, the benefits provided by event-based DR resources would probably decline or cease if the programs were discontinued. However, it seems more than reasonable to assume that California's long-term commitment to DR will not change. Therefore, in evaluating the cost effectiveness of DR resources, it should be assumed that technology-based DR programs will continue long enough to achieve benefits over the useful life of that technology

10.2 Discount Rate

The annual discount rate used to determine the present value of future costs and benefits is the rate at which present and future costs and benefits can be traded off.

The extent to which a DR resource is or is not cost effective under the TRC test will determine whether that resource will increase or decrease the revenue requirements of the utility that administers that resource.

In the long run, revenue requirements will reflect utility costs. Therefore, the expected benefits from DR resources (*i.e.,* the avoided future costs) represent expected future reductions in that utility's revenue requirements. The expected costs of a DR resource, on the other hand, represent expected future increases in that utility's revenue requirements.

Therefore, the appropriate discount rate for determining the present values of those future costs and benefits is the utility's weighted average cost of capital (WACC) that reflects the rate of return that utility is authorized to earn.

The WACC is the rate at which revenue requirements in different years are equalized since the WACC represents the cost that ratepayers incur to compensate utility investors for dedicating their assets to public utility use. It is appropriate to use the WACC as a discount rate since DR programs involve a variety of risks (such as equipment non-performance, technological obsolescence, and changing regulatory requirements). The utility's WACC is the discount rate that utility usually uses when making decisions about investments in or the procurement of most supply-side resource resources which promotes comparability.

A utility also uses its WACC as the discount rate in evaluating the cost effectiveness of energy efficiency programs. The latest Energy Efficiency Policy manual (Version 3) states that the utility's WACC is the appropriate discount rate for the Total Resource Cost test:

> This Commission relies on the Total Resource Cost Test (TRC) as the primary indicator of energy efficiency program cost effectiveness, consistent with our view that ratepayer-funded energy efficiency should focus on programs that serve as resource alternatives to supply-side options. The TRC test measures the net resource benefits from the perspective of all ratepayers by combining the net benefits of the program to participants and non-

participants. The benefits are the avoided costs of the supply-side resources avoided or deferred. The TRC costs encompass the cost of the measures/equipment installed and the costs incurred by the program administrator. The TRC should be calculated utilizing a discount rate that reflects the utilities' weighted average cost of capital, as adopted by the Commission.

In evaluating the cost effectiveness of DR resources under the TRC test, use of a discount rate that is consistent with what is used to evaluate the cost effectiveness of supply-side resources would be appropriate.

Therefore, a utility's WACC is the appropriate basis for the discount rate that should be used in ex ante evaluations of the expected cost effectiveness of that utility's DR resources.

11 Confidentiality

There are several important confidentiality issues that are likely to arise with regards to the Commission's review of future DR program proposals, and the adopted DR cost effectiveness framework may have implications for the extent to which such issues arise.

11.1 Confidentiality of Procurement-Related Information

One viewpoint is that maintaining simplicity and transparency of the evaluation process justifies reliance on public data or estimates from public sources, even if this sacrifices some accuracy in the valuation process and undermines consistency with how supply-side resources are chosen. Another viewpoint favors establishing close alignment between the valuation process used for supply-side resources, which necessarily requires the use of confidential information, and would result in application of the IOU "confidentiality matrix" to the underlying data inputs used to assess value.[67]

It should be noted that any confidentiality restrictions would apply to market participants, and would not preclude Commission staff and various consumer representatives from access to such data. Confidentiality restrictions would, however, introduce a degree of administrative complexity in review proceedings and limit the public release of some information on which the Commission may have relied upon for its decisions.

11.2 Confidentiality of Third-Party Solicitation Information

Using any confidential information from prior solicitations is not recommended, either for supply-side or demand response third party proposals. Third-party bidders had expectations that their bids will be confidential, since their prices and pricing strategy have proprietary value to the extent they will compete in subsequent solicitations.

[67] D.06-066, Appendix 1.

Appendix A

METHODS PG&E USES TO EVALUATE IMPACT OF LIMITS ON AVAILABILITY AND EXERCISE OF DR CAPACITY

Method One

Under the first method, the generation capacity costs that will be avoided in a given year by the RA-qualified capacity available from a DR in that year are estimated by:

1) Multiplying the estimated annual market price of new CT capacity in that year (denominated in $/kW-year), by

2) The ratio of:

   a) the LOLE in each hour of the year in which that RA-qualified DR capacity will be available; by,
   b) the sum of the LOLEs in all 8760 hours of that year;

3) Adjusting each of the resulting hourly avoided generation capacity costs (denominated in $/kW-hour) upwards for:

   a) the T&D line losses that would be avoided in that hour by a reduction in demand, based on the marginal line loss rates in that hour at the voltage levels at which those customer meter-level demand reductions would occur; and;
   b) the RA reserve margin that would be avoided by those reductions in customer demand;

4) Multiplying each of the resulting hourly (adjusted) avoided capacity costs (denominated in $/kW-hour) for each of the hours in that year, by the MWs of RA qualified DR capacity that will be available in that hour; and,

5) Adding up all the resulting hourly (line loss- and reserve margin-adjusted) avoided generation capacity costs for that year.

This method cannot be used if the total number of hours in which the program will be available in that year, is less than the sum of the hourly LOLEs for that year (because that would imply that the total number of hours I which the program will be available is less than the expected number of hours in which additional capacity would be needed to avoid an outage).

Method Two

Some DR programs allow potential participants to choose between several different constraints on when their DR capacity will be available and/or the time periods in which they actually will reduce their demands.

In these situations, it may be necessary to use LOLEs in a different way to estimate to estimate the generation capacity costs that will be avoided in a given year by the RA-qualified DR capacity that will be available in that year.

In order to apply that method, the expected participants would first be divided into different groups, in which each group included all the participants that chose the same limitations on the availability of their DR capacity.

For example, each group would contain only participants that chose:

(1) the same specific days and hours in which their load reductions would be available from that group of customers (*e.g.,* 24 hours per day for 7 days per week vs. noon through 7 PM on weekdays);

(2) the same maximum number of DR events per season in which their load reductions would be available;

(3) the same maximum number of hours per event in which they would reduce their demand;

(4) the same maximum number of event hours per year in which they would reduce their demand;

(5) the same maximum number of consecutive days in which they would reduce their demand.

The generation capacity costs that will be avoided in a given year by the DR capacity available from each group under that program can then be estimated by:

2) Multiplying the (avoided T&D line loss- and avoided RA reserve margin-) adjusted annual market value of new CT capacity in that year (denominated in $/kW-year), by

3) The percentage of all the hours in that year with a positive (*i.e.,* non-zero) LOLE in which the DR program's capacity would be available, after taking into account:

    a) the restrictions on load reduction availability ("event window") chosen by that group of participants; and,
    b) the following restrictions on the events in which that group of participants agreed to reduce their demand:[68]

        i) the maximum number of events;
        ii) the duration of each event;
        iii) the days on which an event could be called; and

[68] Depending upon the number and complexity of these restrictions, it may be necessary to use a linear program or mixed integer program to determine the maximum percentage of positive LOLE hours in which the DR capacity of that group of customers would be available.

iv) the maximum number of consecutive events; and,

4) Multiplying the resulting amount by the MW of RA-qualified DR capacity that will be available in that year from that group of participants

The total of the generation capacity costs that will be avoided in that year by the RA-qualified DR capacity available from that program is then determined by adding together the generation capacity costs that will be avoided in that year by the RA qualified DR capacity available from each group of participants

METHOD SCE USES INCALCULATING CAPACITY VALUE
Generally, the capacity value (*Vc*) of a supply- or demand-side resource is equal to the expected deliverable capacity (*EDC*) of that resource multiplied by the avoided cost of capacity (*AC*). The expected deliverable capacity should be quantified and reported in accordance with the appropriate DR protocols.

For the purposes of analyzing DR cost effectiveness, the *AC* (measured in $/kW-yr) is based on the estimate of capacity value (a Combustion Turbine or CT Proxy) adjusted for its value across time (time differentiation).

DR programs are limited energy resources, meaning, they can only be exercised for a limited number of hours per year. Each program has a specific number of callable events available for a maximum duration. Capacity only has value if it can be called upon for energy or defers the need for energy. The dispatch limitations of DR programs will reduce their value of capacity relative to a CT proxy, which is available year-round. To account for this reduction in value, we should apply an adjustment factor (*A*) to the dispatchable *EDC* portion:

$$Vc = (A \times EDC \times AC)$$

Where the *A*-factor is less than or equal to 1.

The avoided cost of capacity (*AC*) is assumed to be based on a CT proxy, which is a day-of call option[69] for power. Certain DR programs are designed to be (one) day-ahead options. Generally speaking, a day-of call option has more intrinsic value than a day-ahead call option by virtue of the former having greater flexibility in time of need. To credit the full value of capacity as defined by a CT proxy to a day-ahead program would not be a fair evaluation and would overstate its value. Therefore, the equation should be modified to reflect this adjustment in capacity value with a factor (*B*).

For a DR program that can be dispatched day-of, the B-factor by default equals 1.

Time Differentiating Capacity Value

Both marginal energy and capacity values are time differentiated. Energy costs vary according to daily gas prices and hourly system incremental heat rates. Gas prices and heat rates are typically higher during peak demand periods, therefore differentiating energy value across time. Likewise, capacity value varies according to need and the relative risk of low reserve margin events. Periods of supply shortages during system contingencies (unanticipated or anticipated) tend to increase the value of capacity, therefore differentiating capacity value across time.

The marginal capacity value of the CT proxy is an annualized value and not yet differentiated by time. Thus, we must "spread" or allocate the annual marginal capacity value using relative loss of load expectation (LOLE) values to indicate time differentiated

---

[69] The capacity from a CT proxy can be used for energy with one hour notice.

values based on peak period usage.[70] LOLE is a measure of system reliability that indicates the ability (or inability) to deliver energy to the load. A more detailed description is provided in the next section.

Loss of Load Expectation

There is always some probability, however small, that the electricity system will be unable to serve demand. The risk of a generation shortage can be reduced by over supplying generation, but over investment and high operating costs would significantly increase customer bills. Determining the optimum supply and demand balance requires the study of expected system operations using a probabilistic risk assessment approach. Analysis of a system's LOLE is an appropriate risk assessment approach – it is a measure of system reliability that indicates the ability (or inability) to deliver energy to the load.

The LOLE metric provides a method for allocating annualized capacity value across time-of-use periods in proportion to when the loss of load is likely to occur.[71] For example, if the LOLE is greatest in the summer period primarily due to load conditions, particularly during the on-peak, then most of the value we would attribute to capacity will be assigned to those periods. On the other hand, if the probability for loss-of-load is essentially zero during winter off-peak periods, we would assign very little capacity value at those times. LOLE makes it possible to evaluate the relative reliability contribution of different resources across a range of time-of-use periods. This method of analysis employs a Monte Carlo approach by way of two-factor mean reversion sampling of loads and resources. The analysis performed 250 simulations of the entire Western Electricity Coordinating Council (WECC), each unique with regard to hourly supply and demand. From the Monte Carlo analysis, we are able to extract hourly resource availability and loads from each of the 250 simulations. An LOLE event occurs in hour *h* when the load (*L*) exceeds available resources (*R*).

$$L_h - R_h > 0$$

For each simulation, the load in a particular hour can be compared to each of the 250 Monte Carlo outcomes of resource availability in that same hour. In other words, the load in hour *h* is assumed to have the same likelihood of occurring in any of the 250 resource outcomes in hour *h*. The same is true from another viewpoint: the resource availability in hour *h* is assumed to have the same likelihood of occurring in any of the 250 load outcomes in hour *h*. Effectively, this approach yields 250 × 250 or 62,500 possible combinations of load and resources in hour *h*. The above equation can be modified to illustrate this method.

$$L_{h,i} - R_{h,j} > 0$$

Where *i* and *j* are from the respective simulations for load and resources.

70 This approach is a standard utility practice and has been used in prior proceedings.

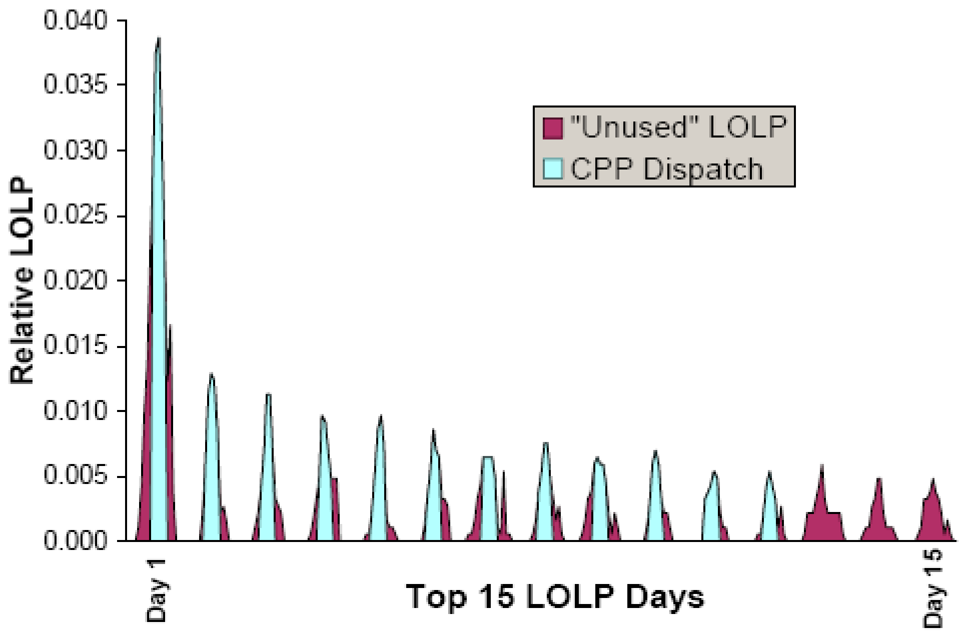
71 The purpose of this LOLE analysis is not to forecast the precise timing of future low-reserve margin events, nor is it to forecast the absolute magnitude of any single loss-of-load event. Rather, it is intended to be a relative distribution of risk used to allocate capacity value across time.

The range of loads and resources is determined by stochastic parameters tied to historical performance. Each load and resource combination is given equal probability of occurring assuming short-term variations in loads (*i.e.,* weather) and available resources (*i.e.* forced outages) are purely random. Combinations in which available resources are unable to meet the load (hence, loss-of-load) contribute to the LOLE for that hour. For example, if 125 out of the 62,500 combinations resulted in loads exceeding available resources, then the LOLE for that hour is 0.2 percent (125 divided by 62,500), or a probability of 1 in 500. The hourly LOLEs, or stochastic LOLEs, are normalized over all hours of the year such that the sum of the normalized LOLEs equals 1. This effectively creates a relative relationship of the hourly LOLE across time. The stochastic LOLE approach takes into account as much uncertainty as can reasonably be captured within the limitations of the model. These are the same uncertainties facing today's system operators (load forecast, supply availability, and hydro conditions). We believe this approach provides a reasonable representation of estimating the relative risk of not serving the load in any given hour, realizing that not all of the market's inefficiencies can be captured in any single model.

The *A*-Factor

The *A*-factor is determined by simulating an optimal dispatch of a sample DR program[72] against an LOLE forecast, and calculating the percentage of time the program is able to "displace" LOLE events, subject to the program's dispatch limitations. As discussed earlier, the LOLE forecast is a method of allocating capacity value across time. To the extent the DR program can be available during times of need (as defined by the LOLE forecast), it will be credited capacity value during those times. In the optimal dispatch simulation, the DR program is assumed to be called upon as often as allowed during periods of greatest LOLE. The following figure illustrates the highest LOLE hours over the top 15 days. Each daily LOLE extends for several hours within the day, ranging between 11 AM and 9 PM. Although the sample DR program is optimally dispatched, the five-hour window is not enough to capture all LOLE hours in each day. Furthermore, since the sample program is limited to 12 calls per year, it does not capture LOLE events beyond the 15th day.

[72] The sample program has 12 callable events with a maximum duration of 5 hours each



This analysis results in an *A*-factor of 50.2 percent. The *A*-factor can be increased in three ways: 1) increase the number of allowable events per year beyond twelve; 2) extend the duration of each event to more than five hours; or 3) allow the program to be called during a greater span of hours.

The *B*-Factor

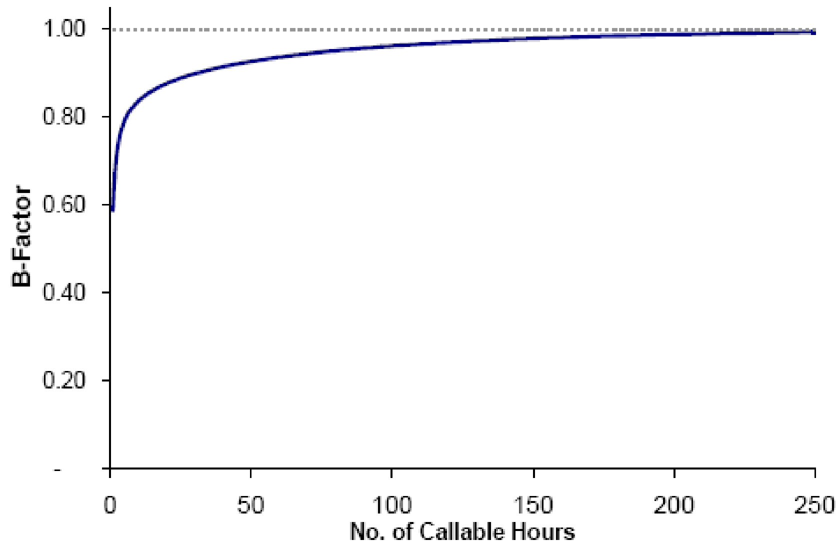
The *B*-factor is based on the difference in value between a day-ahead and a day-of call option for power. A CT is essentially a day-of call option with a strike price equal to the variable operating cost of a CT proxy. The CT proxy value should be adjusted downward for DR programs that are callable on a day-ahead basis. The CPP program, for instance, is a day-ahead call option resource. For a DR program that can be dispatched on a day-of basis, the *B*-factor equals 1 by default[73]. One approximate method to capture the difference in value between a day-ahead and a day-of program is to compare the value of a day-ahead and day-of call option resulting from a Black-Scholes option model[74]. The Black-Scholes model is a standard tool for valuing energy options, but can be used to estimate the relative "payoff" of DR resource options with differing times to expiration (time horizon). Inputs to the model are the forward view of LOLE, day-ahead and day-of volatility of LOLE, and time to expire. The output of the model is a relative value of each call option. Comparing model outputs for day-ahead volatility inputs versus day-of volatility inputs provides a relationship that can be used to approximate the relative value of a day-ahead versus a day-of DR program.

---

73 If the notification time for a day-of CPP program is greater than the time between dispatching a CT and receiving energy, then the value of the B-factor is less than 1.

74 John C. Hull, 3rd Edition, p. 393.

For a day-ahead DR program[75], the *B*-factor is represented in the Figure below.



The value of a day-ahead call option approaches the same value of a day-of call option if the day-ahead option has sufficient callable events. A DR program that is dispatchable for 300+ hours will likely capture all of the LOLE events in a year, regardless of whether the program has a day-ahead or day-of dispatch requirement.

75 The *B*-factor only applies to DR programs with a zero strike price.

Appendix B

PG&E'S METHOD FOR DETERMINING AVOIDED ENERGY COSTS

Option Valuation Method for Valuing Event Based DR

The dispatch of event based DR programs can be modeled explicitly. The manner in which dispatch is modeled depends on the specific characteristics of the DR program. In particular, the dispatch of "reliability" programs should be modeled in a different manner than the dispatch of "economic" programs.

Reliability programs

"Reliability" programs are programs whose dispatch is tied to exogenous events such as CAISO system emergencies. Because many of the triggers for programs in this category are related to reliability, these programs are sometimes referred to as Reliability DR. Direct Load Control programs are examples of reliability programs.

The probability of a reliability program being dispatched is essentially the probability of occurrence of a relevant triggering. For example, for a program that is only dispatched under emergency conditions, the LOLE for an hour might be a good estimate of the probability that a program will be dispatched in that hour.

When estimates of the probability of dispatch are available, avoided energy costs can be estimated in two different ways, one that ignores any correlation between an event occurring and energy prices, which the second takes into account that correlation.

Estimates Based on Hourly Forward Prices and Hourly Probabilities of DR Events

One simple approach to estimating the energy costs that will be avoided in each year by a Reliability DR resource is to multiply the probability of an event occurring in each hour of that year, by the expected market price of energy in that specific hour, and then adding together the resulting products for all hours in that year.

However, this approach tends to understate the energy cost that will be avoided by a Reliability DR resource, because it fails to take into account the positive correlation between the hourly probability that an event will occur in a given hour, and the market price of energy in that hour.

This point is illustrated by the following example. Suppose the expected price is \$100/MWh and the probability of an event occurring is 50%. If the event occurs, the price is \$150/MWh. If the event does not occur, the price is \$50/MWh. Taking the product of the event probability and the expected price yields an estimate of avoided energy cost of 50%*\$100/MWh or \$50/MWh. In reality, if the event occurs and the DR program is triggered, the energy price will be $150/MWh. Therefore, the expected avoided energy cost is 50%*$150/MWh or $75/MWh.

Estimates Based on Conditional Prices

It is also possible to improve estimates of the energy costs that will be avoided by a reliability DR resource by using the appropriate conditional market prices (*i.e.,* estimates of the energy market prices that would occur when the programs are actually triggered).

The underlying correlation between event triggers and market prices has to be estimated in order to estimate those conditional market prices. Estimates of cross-commodity market price correlations, such as the correlation between the market price of electricity and the market price of natural gas can be estimated statistically based on data on historical prices for the different commodities, and are frequently used in option valuation.

However, it is more difficult to estimate the correlation between electricity prices and the probability of different DR events, largely because those events are relatively infrequent. For example, what is the correlation between the hourly probability of a Stage 2 Emergency and hourly market prices for electricity? In theory, electricity prices might be expected to reach the CAISO's price cap that type of an emergency. In reality, as the July 2006 heat storm demonstrated, market prices for electricity are frequently below the CAISO price cap even under emergency conditions.

Production cost simulations could be used to model the relationship between events such as emergencies and prices provided that the prices produced by the model reflect real-world constraints such as automatic mitigation procedures.

Economic programs

Economic DR programs differ from reliability programs in that their dispatch is tied to trigger based market prices rather than physical events such as system emergencies.

In explaining how the net capacity costs of a CT should be calculated in order to estimate the market price of generation capacity, Section 5.3 describes how option valuation models and/or production cost simulation modeling methods could be used to model the "economic" dispatch of peaking generation capacity. It also includes methods to calculate the gross margins that would be achieved by generating and selling energy in those time periods in which revenue at wholesale market prices is expected to exceed variable fuel and O&M costs. These option valuation and simulation can also be used to simulate and estimate the energy costs avoided by "economic" dispatch of DR resources. For example, a DR program that costs $500/MWh to dispatch generally will be dispatched when the market price exceeds $500/MWh. Option valuation techniques, whether they are based on closed-form option-pricing formulas or simulation techniques, provide estimates of the frequency with which prices are likely to exceed $500/MWh. In other words, how frequently the DR program is likely to be dispatched and the extent to which prices are likely to exceed $500/MWh in the event that the program is dispatched, *i.e.,* the avoided energy cost when the program is dispatched.

Using Option Models to Estimate Avoided Energy Costs

A DR resource that is dispatchable is essentially a series of call options. Each call option gives the owner of the option, namely the LSE, the right but not the obligation to obtain reductions in demand in a specific time period by paying a pre-specified strike price to the party that "sold" the option (usually, a customer who enrolled in that DR program).

Simple call option pricing models, such as the Black's formula, can be used to value such options.

The value of a call option depends on the price of an underlying asset, which in this case is the price of electricity. The Black's formula derives an estimate of the value of an option from the forward price of the underlying asset, the volatility of the underlying asset, the strike price of the option, and the time to maturity of the option. The Black's formula produces an option value that is net of the strike price. Therefore, to use the value produced by the Black formula to calculate avoided energy costs, it is necessary to add back the strike price to the value of the option.

In the case of a dispatchable DR resource, the relevant forward price is the price for delivery of electricity during the time period in which the DR program might be dispatched. The volatility of that forward price (*i.e.,* the volatility of the market price of electricity in each period) is a statistic that reflects the degree of uncertainty associated with what the actual market price of electricity will be at the time of delivery (*i.e.,* how far above or below the forward price the actual market price will be at the time the DR option call option could be exercised). The estimate of the conditional price for that period is therefore a function of the volatility estimate, which in turn reflects current market information. The strike price is the amount that the LSE which administers the dispatchable DR resource would pay a participating customer in the event that the DR program is dispatched.

Given that most DR programs can be dispatched with hourly or finer granularity, under the call option pricing approach, DR programs should be valued as bundles of hourly options. Hourly inputs are needed to value hourly call options. Black's formula requires two market price inputs: a forward price and a statistic that reflects the expected volatility of that price. As discussed above, because forward hourly prices do not generally exist, they must be interpolated from forward prices for more aggregated (*i.e.*, longer) time periods.

Market information needed to estimate the volatility of forward prices is less readily available than for forward prices. Quotes for "implied volatility" are available going out a few years from brokers such as Amex. These implied volatility quotes are implied in the sense that they are calculated based on options prices. Given the price of an option and information about every input to Black's formula besides the volatility is available, implied volatility can be calculated as the volatility that would have resulted in the observed option price of that observed price equals the option value by applying Black's formula.

Volatility quotes are generally available for monthly and daily products. It is common practice to use daily volatilities in hourly option valuations (*e.g.,* there is no reason to believe that the uncertainty (in percentage terms) of prices in hour ending 17 a year from now is any greater than the uncertainty associated with prices for the entire on-peak period).

In the absence of market information, volatility must be interpolated and extrapolated from available market information.[76]

Simulation approach

The simulation approach to estimating avoided energy costs involves modeling the dispatch of DR by using portfolio simulation tools, such as Global Energy's Marketsym. These tools are essentially production cost models in which a set of resources are dispatched against a given set of prices. Simulations differ with respect to the amount of foresight they assume in dispatching resources, and with respect to how accurately they reflect actual constraints on the dispatch of different resources. One approach is to assume perfect foresight so that, for example, a DR program would be dispatched in only the set of hours that yield the highest value ex post, even though it may be difficult in real world operations to determine on an ex ante basis when those periods are likely to occur.

Reliability/Economic Hybrids

Some hybrid DR programs have elements of both reliability and economic programs. Estimating avoided energy costs for these hybrid programs requires modeling both their reliability and economic aspects. If the option pricing approach is applied rigorously, this can require resolving significant technical challenges. For example, option value (gross of the strike price) typically is estimated by multiplying the probability that an option is in the money in each time period, by the price of the underlying asset in that period conditional on the option being "in the money." If that option value can only be realized when a triggering event that is unrelated to price also occurs, then the option value needs be adjusted downwards to reflect the joint probability of the option being in the money and an event trigger being satisfied. Developing estimates of these joint probabilities and incorporating them into a valuation can be complicated.

---

[76] Certain models of forward prices impose constraints on the relationship between implied volatilities and instantaneous volatilities (*i.e.,* the standard deviation in the return of a specific asset at a point in time). These models can be used to develop estimates of implied volatility in cases in which market data for implied volatility do not exist but there are data on the instantaneous volatilities of related assets.

Constraints

Estimates of avoided energy costs for event-based DR programs should reflect constraints on their availability and exercise. Although reflecting these constraints in a cost-effectiveness valuation model can be somewhat complicated, it is not impossible. For example, PG&E has incorporated these types of constraints into spreadsheet-based analytic option pricing models that it has used to evaluate the cost-effectiveness of a number of dispatchable DR resources.

For example, an option that can be called a limited number of times during a fixed period is a swing option. There generally are no simple formulas that can be used to value swing options. There are shortcuts, however, that may be able to capture the salient elements of some of the constraints.

Consider a program that is only available for 100 hours in a summer. One conservative approach is to assign avoided energy costs to the program only for the 100 hours with the highest expected avoided energy costs. Given that the probability that an event will be triggered in any of those hundred hours is less than one (1.00), this approach guarantees that the program is valued in a manner which ensures that no value is ascribed to exercising the option in more than 100 hours.

On the other hand, if the probabilities of event triggers in the 100 hours included in the valuation are significantly below one, this approach may be overly conservative. For example, suppose the probability of an event in each hour included in a valuation is 0.1. The probability that the 100 hour constraint would be satisfied in those hours is only 0.1^{100} (*i.e.,* roughly zero for practical purposes).

CERTIFICATE OF SERVICE BY ELECTRONIC MAIL OR U.S. MAIL

I, the undersigned, state that I am a citizen of the United States and am employed in the City and County of San Francisco; that I am over the age of eighteen (18) years and not a party to the within cause; and that my business address is Pacific Gas and Electric Company, Law Department B30A, 77 Beale Street, San Francisco, California 94105.

I am readily familiar with the business practice of Pacific Gas and Electric Company for collection and processing of correspondence for mailing with the United States Postal Service. In the ordinary course of business, correspondence is deposited with the United States Postal Service the same day it is submitted for mailing.

On July 16, 2007, I caused to be served a true copy of:

STRAW PROPOSALS FOR LOAD IMPACT AND COST EFFECTIVENESS OF SOUTHERN CALIFORNIA EDISON COMPANY, SAN DIEGO GAS & ELECTRIC COMPANY, AND PACIFIC GAS AND ELECTRIC COMPANY PURSUANT TO THE ASSIGNED COMMISSIONER AND ADMINISTRATIVE LAW JUDGE'S SCOPING MEMO AND RULING DATED APRIL 18, 2007 – R. 07-01-041

[XX] By Electronic Mail – serving the enclosed via e-mail transmission to each of the parties listed on the official service list for R.07-01-041 with an e-mail address.

[XX] By U.S. Mail – by placing it for collection and mailing, in the course of ordinary business practice, with other correspondence of Pacific Gas and Electric Company, enclosed in a sealed envelope, with postage fully prepaid, addressed to all parties of record on the service list for R.07-01-041 who do not have an email address.

I certify and declare under penalty of perjury under the laws of the State of California that the foregoing is true and correct.

Executed this 16th July, 2007 at San Francisco, California.

/s/

PAMELA J. DAWSON-SMITH

THE PUBLIC UTILITIES COMMISSION OF THE STATE OF CALIFORNIA
EMAIL SERVICE LIST

Downloaded July 16, 2007; last updated on July 13, 2007

Commissioner Assigned: Rachelle B. Chong on January 31, 2007; **ALJ Assigned:** Jessica T. Hecht on January 31, 2007

CPUC DOCKET NO. R0701041

Order Instituting Rulemaking Regarding Policies and Protocols for Demand Response Load Impact Estimates, Cost-Effectiveness Methodologies, Megawatt Goals and Alignment with California Independent System Operator Market Design Protocols.

Rulemaking 07-01-041

(Filed January 25, 2007)

e-recipient@caiso.com;mrw@mrwassoc.com;kea3@pge.com;case.admin@sce.com;philha@astound.net;gayres@energycoalition.org;stephen.baker@constellation.com;bdicapo@caiso.com;GLBarbose@LBL.gov;dbarker@semprautilities.com;brbarkovich@earthlink.net;barrettlarry@comcast.net;cbaskette@enernoc.com;bernardo@braunlegal.com;clark.bernier@rlw.com;bboice02@yahoo.com;abonds@thelen.com;wbooth@booth-law.com;ja_boothe@yahoo.com;adamb@greenlining.org;l_brown369@yahoo.com;marian.brown@sce.com;abb@eslawfirm.com;agc@cpuc.ca.gov;tcarlson@reliant.com;CentralFiles@semprautilities.com;janet.combs@sce.com;joc@cpuc.ca.gov;kcooney@summitblue.com;larry.cope@sce.com;rcounihan@enernoc.com;sdebroff@sasllp.com;bdicapo@caiso.com;douglass@energyattorney.com;zap@cpuc.ca.gov;jellis@energyconnectinc.com;dcengel@fscgroup.com;ahmad.faruqui@brattle.com;cpuccases@pge.com;garwacrd@sce.com;stephengeorge@fccgroup.com;skg@cpuc.ca.gov;jgoodin@caiso.com;jeffgray@dwt.com;alhj@pge.com;marcel@turn.org;jhe@cpuc.ca.gov; ;andrea.horwatt@sce.com;dhungerf@energy.state.ca.us;hvidstenj@kindermorgan.com;imungi@energycoalition.org;bsk@cpuc.ca.gov;tomk@mid.org;chris@emeter.com;klatt@energyattorney.com;evk1@pge.com;dnl@cpuc.ca.gov;claufenb@energy.state.ca.us;jlaun@apogee.net;joyce.leung@sce.com;liddell@energyattorney.com;karen@klindh.com;jody_london_consulting@earthlink.net;mark.s.martinez@sce.com;rmccann@umich.edu;keith.mccrea@sablaw.com;sem4@pge.com;mike.messenger@powerauthority.on.ca;crmd@pge.com;kmills@cfbf.com;wmitchellrunner@socal.rr.com;jym@cpuc.ca.gov;demorse@omsoft.com;jeff@jbsenergy.com;stacia.okura@rlw.com;pxo2@pge.com;roger.pelote@williams.com;cfpena@sempra.com; ;clark.pierce@us.landisgyr.com;ka-wing.poon@sce.com;ted@energy-solution.com;snuller@ethree.com;cpjoe@gepllc.com;david.reed@sce.com;janreid@coastecon.com;Service@spurr.org;jk1@cpuc.ca.gov;lms@cpuc.ca.gov;gayatri@jbsenergy.com;rschmidt@bartlewells.com;msherida@energy.state.ca.us;sherifl@calpine.com;nes@a-klaw.com;jshields@ssjid.com;carl.silsbee@sce.com; ;ksmith2@semprautilities.com;sas@a-klaw.com;latd@pge.com;sharon@emeter.com;filings@a-klaw.com;pthompson@summitblue.com;vthompson@sempra.com;LATc@pge.com;wtr@cpuc.ca.gov;aulmer@water.ca.gov;rogerv@mid.org;crv@cpuc.ca.gov;elvine@lbl.gov;dviolette@summitblue.com;rwalther@pacbell.net;joyw@mid.org;jweil@aglet.org;lwilloughby@semprautilities.com;saw0@pge.com;dwood8@cox.net;vwood@smud.org;ewoychik@strategyi.com;eric@strategyi.com;jwwd@pge.com;dwylie@aswengineering.com;jyamagata@semprautilities.com;

THE PUBLIC UTILITIES COMMISSION OF THE STATE OF CALIFORNIA SERVICE LIST
Downloaded July 16, 2007; last updated on July 13, 2007
Commissioner Assigned: Rachelle B. Chong on January 31, 2007; **ALJ Assigned:** Jessica T. Hecht on January 31, 2007
CPUC DOCKET NO. R0701041 CPUC REV 07-13-07
Total number of addressees: 125

LEGAL AND REGULATORY DEPARTMENT
CALIFORNIA ISO
151 BLUE RAVINE ROAD
FOLSOM CA 95630
Status: INFORMATION

LEGAL AND REGULATORY DEPARTMENT
CALIFORNIA ISO
151 BLUE RAVINE ROAD
FOLSON CA 95630
Email: e-recipient@caiso.com
Status: INFORMATION

MRW & ASSOCIATES, INC.
1814 FRANKLIN ST, STE 720
OAKLAND CA 94612
Email: mrw@mrwassoc.com
Status: INFORMATION

KEN ABREN
245 MARKET ST
SAN FRANCISCO CA 94105
Email: kea3@pge.com
Status: INFORMATION

CASE ADMINISTRATION
SOUTHERN CALIFORNIA EDISON COMPANY
LAW DEPARTMENT
2244 WALNUT GROVE AVE
ROSEMEAD CA 91770
Email: case.admin@sce.com
Status: INFORMATION

PHILIPPE AUCLAIR
11 RUSSELL COURT
WALNUT CREEK CA 94598
Email: philha@astound.net
Status: INFORMATION

GEOFF AYRES
THE ENERGY COALITION
15615 ALTON PARKWAY, STE 245
IRVINE CA 92618
Email: gayres@energycoalition.org
Status: INFORMATION

STEPHEN D. BAKER SENIOR REGULATORY ANALYST
FELLON-MCCORD AND ASSOCIATES
CONSTELLATION NEW ENERGY-GAS DIVISION
9960 CORPORATE CAMPUS DRIVE, STE. 2000
LOUISVILLE KY 40223
Email: stephen.baker@constellation.com
Status: INFORMATION

DI CAPO BALDASSARO
CALIFORNIA ISO
151 BLUE RAVINE ROAD
FOLSOM CA 95630
Email: bdicapo@caiso.com
Status: INFORMATION

GALEN BARBOSE
LAWRENCE BERKELEY NATIONAL LAB
MS 90-4000
1 CYCLOTRON RD.
BERKELEY CA 94720
Email: GLBarbose@LBL.gov
Status: INFORMATION

DAVID BARKER
SAN DIEGO GAS & ELECTRIC COMPANY
8306 CENTURY PARK COURT
SAN DIEGO CA 92123
Email: dbarker@semprautilities.com
Status: INFORMATION

BARBARA R. BARKOVICH
BARKOVICH & YAP, INC.
44810 ROSEWOOD TERRACE
MENDOCINO CA 95460
Email: brbarkovich@earthlink.net
Status: INFORMATION

LARRY B. BARRETT
CONSULTING ASSOCIATES, INC.
PO BOX 60429
COLORADO SPRINGS CO 80960
Email: barrettlarry@comcast.net
Status: INFORMATION

CARMEN BASKETTE
ENERNOC, INC.
594 HOWARD ST, STE 400
SAN FRANCISCO CA 94105
FOR: EnerNoc, Inc.
Email: cbaskette@enernoc.com
Status: APPEARANCE

RYAN BERNARDO
BRAUN & BLAISING, P.C.
915 L ST, STE 1270
SACRAMENTO CA 95814
Email: bernardo@braunlegal.com
Status: INFORMATION

CLARK BERNIER
RLW ANALYTICS
1055 BROADWAY, STE G
SONOMA CA 95476
Email: clark.bernier@rlw.com
Status: INFORMATION

BARB BOICE
4309 NORWOOD AVE, APT. 160
SACRAMENTO CA 95838
Email: bboice02@yahoo.com
Status: INFORMATION

ASHLEE M. BONDS
THELEN REID BROWN RAYSMAN&STEINER LLP
SUITE 1800
101 SECOND ST
SAN FRANCISCO CA 94105
Email: abonds@thelen.com
Status: INFORMATION

WILLIAM H. BOOTH ATTORNEY
LAW OFFICES OF WILLIAM H. BOOTH
1500 NEWELL AVE, 5TH FLR
WALNUT CREEK CA 94596
FOR: California Large Energy Consumers Association
Email: wbooth@booth-law.com
Status: APPEARANCE

JAMES BOOTHE
THE ENERGY COALITION
9 REBELO LANE
NOVATO CA 94947
FOR: The Energy Coalition
Email: ja_boothe@yahoo.com
Status: APPEARANCE

ADAM BRIONES
THE GREENLINING INSTITUTE
1918 UNIVERSITY AVE, 2ND FLR
BERKELEY CA 94704
Email: adamb@greenlining.org
Status: INFORMATION

LYNNE BROWN VICE PRESIDENT
CALIFORNIANS FOR RENEWABLE ENERGY, INC.
24 HARBOR ROAD
SAN FRANCISCO CA 94124
FOR: California for Renewable Energy, Inc.
Email: l_brown369@yahoo.com
Status: APPEARANCE

MARIAN BROWN
SOUTHERN CALIFORNIA EDISON
6040A IRWINDALE AVE.
IRWINDALE CA 91702
Email: marian.brown@sce.com
Status: INFORMATION

ANDREW B. BROWN ATTORNEY
ELLISON, SCHNEIDER & HARRIS, LLP
2015 H ST
SACRAMENTO CA 95814
Email: abb@eslawfirm.com
Status: INFORMATION

Andrew Campbell
CALIF PUBLIC UTILITIES COMMISSION
EXECUTIVE DIVISION
505 VAN NESS AVE RM 5203
SAN FRANCISCO CA 94102-3214
Email: agc@cpuc.ca.gov
Status: STATE-SERVICE

TRENT A. CARLSON
RELIANT ENERGY
1000 MAIN ST
HOUSTON TX 77001
Email: tcarlson@reliant.com
Status: INFORMATION

CENTRAL FILES REGULATORY AFFAIRS
SAN DIEGO GAS & ELECTRIC CO.
8330 CENTURY PARK COURT-CP31E
SAN DIEGO CA 92123-1530
Email: CentralFiles@semprautilities.com
Status: INFORMATION

JANET COMBS ATTORNEY
SOUTHERN CALIFORNIA EDISON COMPANY
2244 WALNUT GROVE AVE
ROSEMEAD CA 91770
FOR: Southern California Edison Company
Email: janet.combs@sce.com
Status: APPEARANCE

Joe Como
CALIF PUBLIC UTILITIES COMMISSION
LEGAL DIVISION
505 VAN NESS AVE RM 5033
SAN FRANCISCO CA 94102-3214
Email: joc@cpuc.ca.gov
Status: STATE-SERVICE

KEVIN COONEY PRINCIPAL/CEO
SUMMIT BLUE CORPORATION
1722 14TH ST
BOULDER CO 80302
Email: kcooney@summitblue.com
Status: INFORMATION

LARRY R. COPE ATTORNEY
SOUTHERN CALIFORNIA EDISON
2244 WALNUT GROVE AVE
ROSEMEAD CA 91770
FOR: Southern California Edison
Email: larry.cope@sce.com
Status: INFORMATION

RICHARD H. COUNIHAN
ENERNOC, INC.
45 FREMONT ST, STE 1400
SAN FRANCISCO CA 94105
FOR: EnerNoc, Inc.
Email: rcounihan@enernoc.com
Status: APPEARANCE

SCOTT H. DEBROFF
SMIGEL, ANDERSON & SACKS
RIVER CHASE OFFICE CENTER
4431 NORTH FRONT ST
HARRISBURG PA 17110
FOR: Elster Integrated Solutions & CellNet
Email: sdebroff@sasllp.com
Status: APPEARANCE

BALDASSARO DI CAPO, ESQ.
CALIFORNIA ISO
LEGAL AND REGULATORY DEPARTMENT
151 BLUE RAVINE ROAD
FOLSOM CA 95630
FOR: California Independent System Operator
Email: bdicapo@caiso.com
Status: APPEARANCE

DANIEL W. DOUGLASS ATTORNEY
DOUGLASS & LIDDELL
21700 OXNARD ST, STE 1030
WOODLAND HILLS CA 91367
FOR: Alliance for Retail Energy Markets/Western Power Trading Forum
Email: douglass@energyattorney.com
Status: APPEARANCE

Tim G. Drew
CALIF PUBLIC UTILITIES COMMISSION
ENERGY RESOURCES BRANCH
505 VAN NESS AVE AREA 4-A
SAN FRANCISCO CA 94102-3214
Email: zap@cpuc.ca.gov
Status: STATE-SERVICE

JACK ELLIS
ENERGY CONNECT, INC.
490 RAQUEL COURT
LOS ALTOS CA 94022
FOR: Energy Connect, Inc.
Email: jellis@energyconnectinc.com
Status: APPEARANCE

DANIEL C. ENGEL SENIOR CONSULTANT
FREEMAN, SULLIVAN & CO.
101 MONTGOMERY ST, 15TH FLR
SAN FRANCISCO CA 94104
Email: dcengel@fscgroup.com
Status: INFORMATION

AHMAD FARUQUI
THE BATTLE GROUP
353 SACRAMENTO ST, STE 1140
SAN FRANCISCO CA 94111
Email: ahmad.faruqui@brattle.com
Status: INFORMATION

LAW DEPARTMENT FILE ROOM
PACIFIC GAS AND ELECTRIC COMPANY
PO BOX 7442
SAN FRANCISCO CA 94120-7442
Email: cpuccases@pge.com
Status: INFORMATION

RUSS GARWACRD
SOUTHERN CALIFORNIA EDISON COMPANY
2244 WALNUT GROVE
ROSEMEAD CA 91770
Email: garwacrd@sce.com
Status: INFORMATION

STEVE GEORGE
GSC GROUP
101 MONTGOMERY ST, 15TH FLR
SAN FRANCISCO CA 94104
Email: stephengeorge@fccgroup.com
Status: INFORMATION

Sudheer Gokhale
CALIF PUBLIC UTILITIES COMMISSION
ELECTRICITY RESOURCES & PRICING BRANCH
505 VAN NESS AVE RM 4209
SAN FRANCISCO CA 94102-3214
FOR: DRA
Email: skg@cpuc.ca.gov
Status: STATE-SERVICE

JOHN GOODIN
CALIFORNIA ISO
MARKET & PRODUCT DEVELOPMENT
151 BLUE RAVINE RD.
FOLSOM CA 95630
Email: jgoodin@caiso.com
Status: INFORMATION

JEFFREY P. GRAY ATTORNEY
DAVIS WRIGHT TREMAINE, LLP
505 MONTGOMERY ST, STE 800
SAN FRANCISCO CA 94111-6533
Email: jeffgray@dwt.com
Status: INFORMATION

ARTHUR HAUBENSTOCK ATTORNEY
PACIFIC GAS AND ELECTRIC COMPANY
77 BEALE ST, B30A
SAN FRANCISCO CA 94105
FOR: Pacific Gas and Electric Company
Email: alhj@pge.com
Status: INFORMATION

MARCEL HAWIGER ATTORNEY
THE UTILITY REFORM NETWORK
711 VAN NESS AVE, STE 350
SAN FRANCISCO CA 94102
FOR: TURN
Email: marcel@turn.org
Status: APPEARANCE

Jessica T. Hecht
CALIF PUBLIC UTILITIES COMMISSION
DIVISION OF ADMINISTRATIVE LAW JUDGES
505 VAN NESS AVE RM 5113
SAN FRANCISCO CA 94102-3214
Email: jhe@cpuc.ca.gov
Status: STATE-SERVICE

GRAYSON HEFFNER
15525 AMBIANCE DRIVE
N. POTOMAC MD 20878
Status: INFORMATION

ANDREA HORWATT
SOUTHERN CALIFORNIA EDISON COMPANY
2244 WALNUT GROVE AVE
ROSEMEAD CA 91770
Email: andrea.horwatt@sce.com
Status: INFORMATION

DAVID HUNGERFORD
CALIFORNIA ENERGY COMMISSION
DEMAND ANALYSIS OFFICE
1516 NINTH ST, MS-22
SACRAMENTO CA 95814
Email: dhungerf@energy.state.ca.us
Status: INFORMATION

JOEL M. HVIDSTEN
KINDER MORGAN ENERGY PARTNERS
1100 TOWN & COUNTRY ROAD, STE 700
ORANGE CA 92868
Email: hvidstenj@kindermorgan.com
Status: INFORMATION

MWIRIGI IMUNGI
THE ENERGY COALITION
15615 ALTON PARKWAY, STE 245
IRVINE CA 92618
Email: imungi@energycoalition.org
Status: INFORMATION

Bruce Kaneshiro
CALIF PUBLIC UTILITIES COMMISSION
ENERGY RESOURCES BRANCH
505 VAN NESS AVE AREA 4-A
SAN FRANCISCO CA 94102-3214
Email: bsk@cpuc.ca.gov
Status: STATE-SERVICE

TOM KIMBALL
MODESTO IRRIGATION DISTRICT
1231 11TH ST
MODESTO CA 95354
Email: tomk@mid.org
Status: INFORMATION

CHRIS KING
EMETER CORPORATION
ONE TWIN DOLPHIN DRIVE
REDWOOD CITY CA 94065
Email: chris@emeter.com
Status: INFORMATION

GREGORY KLATT ATTORNEY
DOUGLASS & LIDDELL
411 E. HUNTINGTON DRIVE, STE. 107-356
ARCADIA CA 91006
FOR: Direct Access Customer Coalition
Email: klatt@energyattorney.com
Status: APPEARANCE

Dorris Lam
CALIF PUBLIC UTILITIES COMMISSION
ENERGY RESOURCES BRANCH
505 VAN NESS AVE AREA 4-A
SAN FRANCISCO CA 94102-3214
Email: dnl@cpuc.ca.gov
Status: STATE-SERVICE

JOHN LAUN
APOGEE INTERACTIVE, INC.
1220 ROSECRANS ST., STE 308
SAN DIEGO CA 92106
Email: jlaun@apogee.net
Status: INFORMATION

DONALD C. LIDDELL ATTORNEY
DOUGLASS & LIDDELL
2928 2ND AVE
SAN DIEGO CA 92103
Email: liddell@energyattorney.com
Status: APPEARANCE

JODY S. LONDON
JODY LONDON CONSULTING
PO BOX 3629
OAKLAND CA 94609
Email: jody_london_consulting@earthlink.net
Status: INFORMATION

RICHARD MCCANN
M.CUBED
2655 PORTAGE BAY ROAD, STE 3
DAVIS CA 95616
Email: rmccann@umich.edu
Status: INFORMATION

SUSAN MCNEILL
PACIFIC GAS AND ELECTRIC COMPANY
PO BOX 770000, B8M
SAN FRANCISCO CA 94177-0001
Email: sem4@pge.com
Status: INFORMATION

EDWARD V. KURZ ATTORNEY
PACIFIC GAS AND ELECTRIC COMPANY
77 BEALE ST
SAN FRANCISCO CA 94105
FOR: Pacific Gas and Electric
Email: evk1@pge.com
Status: INFORMATION

CLARE LAUFENBERG
CALIFORNIA ENERGY COMMISSION
1516 NINTH ST, MS 46
SACRAMENTO CA 95814
Email: claufenb@energy.state.ca.us
Status: STATE-SERVICE

JOYCE LEUNG
SOUTHERN CALIFORNIA EDISON COMPANY
6060 J IRWINDALE AVE.
IRWINDALE CA 91702
Email: joyce.leung@sce.com
Status: INFORMATION

KAREN LINDH
LINDH & ASSOCIATES
7909 WALERGA ROAD, NO. 112, PMB 119
ANTELOPE CA 95843
Email: karen@klindh.com
Status: INFORMATION

MARK S. MARTINEZ
SOUTHERN CALIFORNIA EDISON
6060 IRWINDALE AVE., STE J
IRWINDALE CA 91702
Email: mark.s.martinez@sce.com
Status: INFORMATION

KEITH R. MCCREA ATTORNEY
SUTHERLAND, ASBILL & BRENNAN, LLP
1275 PENNSYLVANIA AVE., NW
WASHINGTON DC 20004-2415
FOR: CA Manufacturers & Technology Assn.
Email: keith.mccrea@sablaw.com
Status: APPEARANCE

MIKE MESSENGER
CALIFORNIA ENERGY COMMISSION
120 ADELAIDE ST WEST STE 1600
TORONTO ON M5H 1T1 CANADA
Email: mike.messenger@powerauthority.on.ca
Status: INFORMATION

CHARLES MIDDLEKAUFF ATTORNEY
PACIFIC GAS AND ELECTRIC COMPANY
77 BEALE ST
SAN FRANCISCO CA 94105
FOR: Pacific Gas and Electric Company
Email: crmd@pge.com
Status: INFORMATION

KAREN N. MILLS ATTORNEY
CALIFORNIA FARM BUREAU FEDERATION
2300 RIVER PLAZA DRIVE
SACRAMENTO CA 95833
FOR: California Farm Bureau Federation
Email: kmills@cfbf.com
Status: APPEARANCE

WARREN MITCHELL
THE ENERGY COALITION
15615 ALTON PARKWAY, STE 245
IRVINE CA 92618
Email: wmitchellrunner@socal.rr.com
Status: INFORMATION

Joy Morgenstern
CALIF PUBLIC UTILITIES COMMISSION
ENERGY RESOURCES BRANCH
505 VAN NESS AVE AREA 4-A
SAN FRANCISCO CA 94102-3214
Email: jym@cpuc.ca.gov
Status: STATE-SERVICE

DAVID MORSE
1411 W, COVELL BLVD., STE 106-292
DAVIS CA 95616-5934
Email: demorse@omsoft.com
Status: INFORMATION

JEFF NAHIGIAN
JBS ENERGY, INC.
311 D ST
WEST SACRAMENTO CA 95605
Email: jeff@jbsenergy.com
Status: INFORMATION

STACIA OKURA
RLW ANALYTICS, INC.
1055 BROADWAY, STE G
SONOMA CA 95476
Email: stacia.okura@rlw.com
Status: INFORMATION

PETER OUBORG ATTORNEY
PACIFIC GAS AND ELECTRIC COMPANY
77 BEALE ST, B30A
SAN FRANCISCO CA 94105
Email: pxo2@pge.com
Status: INFORMATION

ROGER PELOTE
WILLIAMS POWER COMPANY
12736 CALIFA ST
VALLEY VILLAGE CA 91607
Email: roger.pelote@williams.com
Status: INFORMATION

CARLOS F. PENA
SEMPRA ENERGY LAW DEPARTMENT
101 ASH ST
SAN DIEGO CA 92101
Email: cfpena@sempra.com
Status: INFORMATION

BRUCE PERLSTEIN
PACIFIC GAS AND ELECTRIC COMPANY
245 MARKET ST
SAN FRANCISCO CA 94105
Status: INFORMATION

CLARK E. PIERCE
LANDIS & GYR
246 WINDING WAY
STRATFORD NJ 8084
Email: clark.pierce@us.landisgyr.com
Status: INFORMATION

KA-WING MAGGIE POON
GO1, QUAD 2B
2244 WALNUT GROVE AVE.
ROSEMEAD CA 91770
Email: ka-wing.poon@sce.com
Status: INFORMATION

TED POPE PRESIDENT
ENERGY SOLUTIONS
1738 EXCELSIOR AVE.
OAKLAND CA 94602
Email: ted@energy-solution.com
Status: INFORMATION

SNULLER PRICE
ENERGY AND ENVIRONMENTAL ECONOMICS
101 MONTGOMERY, STE 1600
SAN FRANCISCO CA 94104
Email: snuller@ethree.com
Status: INFORMATION

DAVID REED
SOUTHERN CALIFORNIA EDISON
6060 IRWINDALE AVE., STE. J
IRWINDALE CA 91702
Email: david.reed@sce.com
Status: INFORMATION

MICHAEL ROCHMAN MANAGING DIRECTOR
SPURR
1430 WILLOW PASS ROAD, STE 240
CONCORD CA 94520
Email: Service@spurr.org
Status: INFORMATION

Lisa-Marie Salvacion
CALIF PUBLIC UTILITIES COMMISSION
LEGAL DIVISION
505 VAN NESS AVE RM 4107
SAN FRANCISCO CA 94102-3214
FOR: DRA
Email: lms@cpuc.ca.gov
Status: APPEARANCE

REED V. SCHMIDT VICE PRESIDENT
BARTLE WELLS ASSOCIATES
1889 ALCATRAZ AVE
BERKELEY CA 94703
Email: rschmidt@bartlewells.com
Status: INFORMATION

LINDA Y. SHERIF ATTORNEY
CALPINE CORPORATION
3875 HOPYARD ROAD, STE 345
PLEASANTON CA 94588
Email: sherifl@calpine.com
Status: APPEARANCE

JEFF SHIELDS UTILITY SYSTEMS DIRECTOR
SOUTH SAN JOAQUIN IRRIGATION DISTRICT
11011 E. HWY 120
MANTECA CA 95336
Email: jshields@ssjid.com
Status: INFORMATION

JOE PRIJYANONDA
GLOBAL ENERGY PARTNERS, LLC
3569 MT. DIABLE BLVD., STE 200
LAFAYETTE CA 94549
Email: cpjoe@gepllc.com
Status: INFORMATION

L. JAN REID
COAST ECONOMIC CONSULTING
3185 GROSS ROAD
SANTA CRUZ CA 95062
Email: janreid@coastecon.com
Status: INFORMATION

Jason R. Salmi Klotz
CALIF PUBLIC UTILITIES COMMISSION
ENERGY RESOURCES BRANCH
505 VAN NESS AVE AREA 4-A
SAN FRANCISCO CA 94102-3214
Email: jk1@cpuc.ca.gov
Status: STATE-SERVICE

GAYATRI SCHILBERG
JBS ENERGY
311 D ST, STE A
WEST SACRAMENTO CA 95605
FOR: TURN
Email: gayatri@jbsenergy.com
Status: INFORMATION

MARGARET SHERIDAN
CALIFORNIA ENERGY COMMISSION
DEMAND ANALYSIS OFFICE
1516 NINTH ST, MS-22
SACRAMENTO CA 95814
Email: msherida@energy.state.ca.us
Status: INFORMATION

NORA SHERIFF ATTORNEY
ALCANTAR & KAHL, LLP
120 MONTGOMERY ST, STE 2200
SAN FRANCISCO CA 94104
FOR: Energy Producers & Users Coalition
Email: nes@a-klaw.com
Status: APPEARANCE

CARL SILSBEE
SOUTHERN CALIFORNIA EDISON
GO1, RP&A
2244 WALNUT GROVE AVE
ROSEMEAD CA 91770
Email: carl.silsbee@sce.com
Status: INFORMATION

KEN SKINNER VICE PRESIDENT, COO
INTEGRAL ANALYTICS, INC
312 WALNUT ST, STE 1600
CINCINNATI OH 45202
Status: APPEARANCE

KATHRYN SMITH ANALYST
SAN DIEGO GAS AND ELECTRIC COMPANY
8306 CENTURY PARK COURT
SAN DIEGO CA 92123
Email: ksmith2@semprautilities.com
Status: INFORMATION

ANNIE STANGE
ALCANTAR & KAHL
1300 SW FIFTH AVE., STE 1750
PORTLAND OR 97201
Email: sas@a-klaw.com
Status: INFORMATION

LISA TAKEUCHI
PACIFIC GAS AND ELECTRIC COMPANY
77 BEALE ST
SAN FRANCISCO CA 94105
Email: latd@pge.com
Status: INFORMATION

SHARON TALBOTT
EMETER CORPORATION
ONE TWIN DOLPHIN DRIVE
REDWOOD CITY CA 94065
Email: sharon@emeter.com
Status: INFORMATION

KAREN TERRANOVA
ALCANTAR & KAHL, LLP
120 MONTGOMERY ST, STE 2200
SAN FRANCISCO CA 94104
Email: filings@a-klaw.com
Status: INFORMATION

PATRICIA THOMPSON
SUMMIT BLUE CONSULTING
2920 CAMINO DIABLO, STE 210
WALNUT CREEK CA 94597
Email: pthompson@summitblue.com
Status: INFORMATION

VICKI L. THOMPSON ATTORNEY
SAN DIEGO GAS & ELECTRIC COMPANY
101 ASH ST
SAN DIEGO CA 92101
FOR: San Diego Gas & Electric Company
Email: vthompson@sempra.com
Status: APPEARANCE

LUKE TOUGAS
PACIFIC GAS AND ELECTRIC COMPANY
PO BOX 770000, MC B9A
SAN FRANCISCO CA 94177
Email: LATc@pge.com
Status: INFORMATION

Rebecca Tsai-Wei Lee
CALIF PUBLIC UTILITIES COMMISSION
ELECTRICITY RESOURCES & PRICING BRANCH
505 VAN NESS AVE RM 4209
SAN FRANCISCO CA 94102-3214
Email: wtr@cpuc.ca.gov
Status: STATE-SERVICE

ANDREW ULMER STAFF COUNSEL
CALIFORNIA DEPARTMENT OF WATER RESOURCES
3310 EL CAMINO AVE, STE 120
SACRAMENTO CA 95821
FOR: California Department of Water Resources
Email: aulmer@water.ca.gov
Status: STATE-SERVICE

ROGER VAN HOY
MODESTO IRRIGATION DISTRICT
1231 11TH ST
MODESTO CA 95354
Email: rogerv@mid.org
Status: INFORMATION

Christopher R Villarreal
CALIF PUBLIC UTILITIES COMMISSION
DIVISION OF STRATEGIC PLANNING
505 VAN NESS AVE RM 5119
SAN FRANCISCO CA 94102-3214
Email: crv@cpuc.ca.gov
Status: STATE-SERVICE

EDWARD VINE
LAWRENCE BERKELEY NATIONAL LABORATORY
BUILDING 90-4000
BERKELEY CA 94720
Email: elvine@lbl.gov
Status: INFORMATION

DANIEL M. VIOLETTE
SUMMIT BLUE CONSULTING
1722 14TH ST, STE 230
BOULDER CO 80302
Email: dviolette@summitblue.com
Status: INFORMATION

ROBIN J. WALTHER, PH.D.
1380 OAK CREEK DRIVE., 316
PALO ALTO CA 94305
Email: rwalther@pacbell.net
Status: INFORMATION

JOY A. WARREN ATTORNEY
MODESTO IRRIGATION DISTRICT
1231 11TH ST
MODESTO CA 95354
Email: joyw@mid.org
Status: INFORMATION

JAMES WEIL DIRECTOR
AGLET CONSUMER ALLIANCE
PO BOX 37
COOL CA 95614
Email: jweil@aglet.org
Status: INFORMATION

LESLIE WILLOUGHBY
SAN DIEGO GAS AND ELECTRIC COMPANY
8305 CENTURY PARK CT.
SAN DIEGO CA 92123
Email: lwilloughby@semprautilities.com
Status: INFORMATION

SHIRLEY WOO ATTORNEY
PACIFIC GAS AND ELECTRIC COMPANY
77 BEALE ST, B30A
SAN FRANCISCO CA 94105
FOR: Pacific Gas and Electric
Email: saw0@pge.com
Status: APPEARANCE

DON WOOD
PACIFIC ENERGY POLICY CENTER
4539 LEE AVE
LA MESA CA 91941
Email: dwood8@cox.net
Status: INFORMATION

VIKKI WOOD
SACRAMENTO MUNICIPAL UTILITY DISTRICT
6301 S ST, MS A204
SACRAMENTO CA 95817-1899
Email: vwood@smud.org
Status: INFORMATION

ERIC WOYCHIK
COMVERGE, INC.
9901 CALODEN LANE, STE 1
OAKLAND CA 94605
FOR: Comverge, Inc.
Email: ewoychik@strategyi.com
Status: APPEARANCE

ERIC C. WOYCHIK
STRATEGY INTEGRATION LLC
9901 CALODEN LANE
OAKLAND CA 94605
FOR: Comverge, Inc.
Email: eric@strategyi.com
Status: APPEARANCE

JOSEPHINE WU
PACIFIC GAS AND ELECTRIC COMPANY
PO BOX 770000, MAIL CODE B9A
SAN FRANCISCO CA 94177
Email: jwwd@pge.com
Status: INFORMATION

DAVID M. WYLIE, PE
ASW ENGINEERING
2512 CHAMBERS ROAD, STE 103
TUSTIN CA 92780
Email: dwylie@aswengineering.com
Status: INFORMATION

JOY C. YAMAGATA
SAN DIEGO GAS & ELECTRIC/SOCALGAS
8330 CENTURY PARK COURT
SAN DIEGO CA 91910
Email: jyamagata@semprautilities.com
Status: INFORMATION