WORKING GROUP 2 DEMAND RESPONSE PROGRAM EVALUATION –PROGRAM YEAR 2004

SUB-METERING SUMMARY REPORT

Prepared for

Working Group 2 Measurement and Evaluation Committee

Prepared by

QUANTUM CONSULTING INC.
2001 Addison Street, Suite 300
Berkeley, CA 94704

P1996

June, 2005

TABLE OF CONTENTS

1. INTRODUCTION

1.1 THE SUB-METERING ELEMENT OF 2004 WG2 EVALUATION

This summary report presents key findings from the sub-metering element of the 2004 Working Group 2 (WG2) Demand Response (DR) Program Evaluation. The overall WG2 evaluation study included a wide variety of evaluation research activities focused on California's Demand Bid Program (DBP), Critical Peak Pricing (CPP), Demand Reserves Program (DRP), and Interruptible programs. The overall WG2 evaluation results were published in December 2004 (WG2 2004).[1] This report addresses only the sub-metering element of the evaluation project.

The sub-metering element of the evaluation was established to provide a more in-depth understanding of DR program participant behavior - beyond what is revealed by analysis of revenue meter data, or by what can be learned about participants' DR strategies and behaviors from traditional survey methods. Key aspects of the sub-metering element of the evaluation are summarized below:

- **Twelve sites were included in the sub-metering portion of the 2004 evaluation**. These sites span each of the three primary price-responsive DR programs (i.e., CPP, DBP and DRP), a variety of business types and end uses, and each of the state's major investor-owned utilities (IOUs) (i.e., SCE, PG&E, SDG&E).

- **Individual reports have been prepared for each of the sub-metering sites**. These individual reports detail the characteristics of each site, their DR strategies, the end uses monitored, and provide comparisons of revenue meter load reduction results with estimates developed from the sub-metering data.

- **This summary report provides an integration of findings from across the 12 sites monitored**, as well as lessons learned from the sub-metering recruitment process.

- **Appendix J of the December 2004 final WG2 evaluation report provides a detailed summary of the methodology and procedures** used to design and implement the sub-metering project**.**

Specific elements of the sub-metering task included:

- Developing a detailed screening process that resulted in a sample that includes a variety of customer types, programs, and DR strategies.

---

[1] Quantum Consulting, 2004. *Working Group 2 Demand Response Program Evaluation – Program Year 2004*, prepared by Quantum Consulting, Inc. for Working Group 2 Measurement and Evaluation Committee, December. http://www.fypower.org/pdf/DR_EVAL.pdf

- Development and execution of detailed sub-metering plans for each of the study participants.
- Primarily remote (dial up) collection of sub-hourly equipment and circuit data.
- In-depth interviewing with each of the study participants.
- Analysis of individual equipment and circuit loads, as well as customer strategies and observed behavior.
- Preparation of the individual site reports and this summary report.

1.2 WHY SUB-METERING?

Although many large customer demand response programs have been in existence for some time, the customer market for price-responsive DR is still in a relatively nascent stage. Few customers have a detailed understanding of the composition of their hourly loads or have the ability to easily and precisely control those loads. This is borne out by the results of the 2004 WG2 Evaluation, as well as other recent related research led by the California Energy Commission, the Public Interest Energy Research Program (PIER), Lawrence Berkeley National Laboratory, the Demand Response Research Center, and others.

Although much can and has been learned about how customers do or don't respond to DR program offerings through traditional evaluation approaches that do not include sub-metering customer loads, sub-metering offers a level of information and insight into customer activity that is difficult if not impossible to glean from other evaluation approaches. For example, using revenue meter data and customer self reports, the overall WG2 2004 Evaluation results provide a great deal of useful information on total program impacts as well as distributions of impacts across individual customers. However, information on the underlying sources of customer impacts, the sophistication and robustness of their DR implementation strategies, the degree to which they carried out their strategies, and the underlying reasons why they did or did not carry them out, is more limited. The sub-metering element of the evaluation was envisioned and designed to provide additional insights into these more detailed customer-specific issues.

The ability to analyze participants' loads at an equipment or circuit level provides significantly more information that can be used to enhance understanding of customers' DR strategies and their ability to effectively participate in DR events. The inclusion of sub-metering data in the analyses of participant performance is also useful to understanding how curtailed end uses contribute to load reductions at the revenue meter. Sub-metering data can be used to develop bottom-up estimates of DR impacts for sub-metered participants that can be compared to estimates of impacts measured by revenue-meter interval data. Comparing these results improves understanding of the relative accuracy of different revenue-meter impact estimation methods, which complements the results published in the December WG2 2004 Evaluation report.

Analysis of sub-metering data also significantly improves understanding of the strengths and weaknesses of customers' curtailment strategies and helps to illuminate barriers associated with the execution of these strategies. For example, the sub-metering data allows closer tracking and analysis of participants actions over time. When conducted over multiple events, this analysis can yield a great deal of information about the *evolution* of customer's applied DR strategy.

While each sampled site reveals only one participant's experience, the integration of findings from this research reveals a number of findings that would likely not be obtainable by other means. These enhanced findings, when combined with the overall evaluation results, provide important input for program design and ongoing DR policy development. This research also makes significant contributions to DR research in the commercial and industrial sectors by adding twelve sub-metering sites to the small but growing number of in-depth case studies and monitoring projects carried out in related studies.[2] This combined body of work offers considerable potential for improving program offerings and enhancing the technical and organizational knowledge of active and prospective DR program participants.

1.3 SUB-METERING STUDY OBJECTIVES

Four broad objectives were initially identified for the sub-metering element of the 2004 WG2 Evaluation. These were to:

- Develop findings on what works and what doesn't to help improve program participation and forecasts of DR potential.
- Develop sub-metering-based estimates of DR impacts and compare with whole-meter estimates.
- Develop in-depth understanding of real and perceived end use service/demand response tradeoffs.
- Integrate results into the PIER DR Database.[3]

---

[2] "Development and Evaluation of Fully Automated Demand Response in Large Facilities" Piette, M. A., O. Sezgen, D. Watson, N. Motegi, (Lawrence Berkeley National Laboratory), C. Shockman (Shockman Consulting), L. ten Hope (Program Manager, Energy Systems Integration CEC). CEC-500-2005-013. January 2005

"Measurement and Evaluation Techniques for Automated Demand Response Demonstration" Motegi, N., M.A. Piette, D.S. Watson, and O. Sezgen, Lawrence Berkeley National Laboratory. Proceedings, ACEEE 2004 Summer Study on Energy Efficiency in Buildings: Breaking out of the Box, August 22-27, 2004, Asilomar, Pacific Grove, CA. Washington D.C. American Council for an Energy-Efficient Economy. LBNL-55086. August 2004.

The Demand Response Research Center is currently operating a pilot project to examine Automated Critical Peak Pricing for Large Commercial facilities. For more information on this project and the above two citations, see http://drrc.lbl.gov/drrc.html

[3] PIER has developed a DR database that is intended as a repository for DR-related data collected through a variety of individual DR evaluation projects. This data can then be leveraged for further research by PIER and the Demand Response Research Center.

Building on the original project objectives, the following set of research questions were considered during the analysis process:

- What DR strategies work, which don't, and why? What are the weak points in the customers' participation processes? Are there differences in real and perceived effects of DR strategies? What are specific program, institutional, and technical barriers to event participation?

- What are the true costs and benefits of participation? What can be done to help customers bear the costs that prevent them from participation?

- What are the successful manual and automated DR strategies? What level of automation is appropriate for different customer and end use types? Are customers actively seeking to automate? What are the primary constraints to improving automation?

- Do customers possess all the knowledge they need to carry out effective DR actions? What more do they need to know? Where might they obtain this knowledge? Are they aware that they may need more knowledge or tools to participate more effectively?

- What do customers understand about their baselines? How are they impacted by baseline estimates? What is the variability in their daily load shapes? Can they obtain baseline data when they need it? Do current baseline methods create opportunities for free riders? Are customers aware of that potential?

Section 3 provides a summary of the key findings from this research. This summary report does not repeat the detailed site-specific findings that are included in the individual site reports. In addition, the sub-metering results address many but not all of the research questions listed above. This is primarily due to limitations in the study sample and the limited number of real program events that were called in 2004.

2. ORIGIN AND CHARACTERIZATION OF THE SUB-METERING SAMPLE

The sub-metering recruitment and installation process, which did not begin in earnest until May 2004, was challenged by the need to have sub-metering equipment installed in time to capture DR events for the summer of 2004. The steps for recruiting customers into the sub-metering sample entailed obtaining participant lists from the utilities, conducting detailed telephone screening, carrying out on-site surveys, developing metering plans, and installing the monitoring equipment.

The following screening criteria were developed and applied to the spring 2004 population of CPP, DBP, and later, DRP, participants:

1. Customer had to be highly likely to opt-in for DR events.
2. Customer had to indicate they would shed multiple loads at a site.
3. Customer had to fit within our quota for a diverse mix of business types and customer sizes.
4. Customer had to fit within our quota for a mix of end uses and shed strategies.
5. Customer's characteristics had to enable cost-effective monitoring of loads and energy services of interest.

These criteria were intentionally biased in favor of a sample that included participants that are most likely to actually take DR actions and would utilize more complex DR strategies relative to participants who might only activate back-up generation or shut down one major type of load within their facility. Consequently, the first two criteria were applied as pass-fail decisions, whereas application of the third and fourth required considerably more scrutiny. Inherent in the third and fourth criteria was the intention to seek a reasonably representative distribution of the program population by utility and program.

As shown in Exhibit 1, 19 sites made it through the initial telephone screening. Of these, seven were subsequently rejected for sub-metering. Very often candidates had either not developed a DR strategy at all or had not developed it to a level where it could be efficiently executed in the event of a day-ahead or day-of event notification. Others were uncertain of whether their intended curtailment strategy would meet minimum program requirements (e.g., achieving at the 100 kW minimum reduction for DBP events), whether participation would justify the costs, or whether load reductions could carried be out without significant disruptions to site operations or occupant comfort or productivity.

Exhibit 1
Distribution of Onsite Surveys and Installed Sites by Sector, Utility and Program

	Total Sites	By Sector		By Utility			By Program		
		Com.	Ind.	PG&E	SCE	SDGE	DBP	CPP	DRP
Total Onsite Surveys Completed	19	10	9	7	8	4	13	5	1
Onsite Survey Sites Rejected	7	4	3	2	3	2	5	2	0
Metering Installations Completed	12	6	6	5	5	2	8	3	1
Installed Sites with Summer '04 Events Captured	6	2	4	3	2	1	2	3	1
Installed Sites with Summer '04 Events Not Captured	3	3	0	0	2	1	3	0	0
Installed Sites Not Participating in Summer '04 Events	3	1	2	2	1	0	3	0	0

Throughout the recruitment process, recruitment efforts were continuously redirected to attain a broad sample of sites across the three utilities, DR programs, customer types, and affected end uses. As a number of industrial facilities were among the first sub-metering sites recruited, the focus of recruitment was shifted to commercial and institutional sites. Exhibit 1 includes the final distribution of the sub-metering sample across these categories and indicates the number of sites in each category where sub-metering data is available from DR events in summer 2004.

More details on the recruitment process itself can be found in Appendix J of the December WG2 2004 Evaluation report.

Exhibit 2 summarizes some of the key site characteristics of the sub-metering sample, along with metering installation and event dates. Given the timing and challenges of the recruitment and data collection process, it was fortunate that 2004 DR events were concentrated in the latter half of the summer. These circumstances allowed one or more DR events to be captured for half of the twelve sampled sites. For the three of the sites where DR events were not captured, sub-metering installations were either not in place in time to capture events, or the customer chose not to participate in later events that would have otherwise been monitored. Three sites in the sample did not participate at all in the one or two DR events of summer 2004 that occurred in their respective programs despite their initial assurances that they would participate. Even so, the proportion of sub-metering customers that took action during 2004 DR events is many times higher than the rate of action found for the entire participant population.

Exhibit 2
Summary of Selected Sub-metering Site Characteristics and DR Event Monitoring

Customer	Utility & Program	Business Type & Size (1000 sq.ft.)	Curtailed End Uses (1: Primary EUs 2: Secondary EUs)	Level of Auto-mation	Instal-lation Date (2004)	Event Dates (2004)	Event Partic-ipation	Events Monit-ored
Installed Sites Participating in Summer 2004 - Events Captured (Sites 1 to 6)								
SITE 1: Product Repackaging Facility	PGE /CPP	Industrial / Packaging (64)	1: HVAC, Lighting 2: Process Equip.	Manual	7/30	8/27 9/8 9/9 9/10 10/13	NO NO YES YES YES	NO NO YES YES YES
SITE 2: Agricultural Product Processing, Packing & Cold Storage Facility #1	PGE /CPP	Industrial / Ag Process (250)	1: Cold Storage 2: Lighting, Process Equip.	Manual	6/11	8/27 9/8 9/9 9/10 10/13	YES YES YES YES YES	YES YES YES YES YES
SITE 3: Baking & Frozen Storage Facility	PGE /CPP	Industrial / Food Process (135)	1: Freezers 2: HVAC, Lighting, Process Equip.	Manual	6/24	8/27 9/8 9/9 9/10 10/13	YES YES NO NO YES	YES YES NO NO YES
SITE 4: Agricultural Product Processing, Packing & Cold Storage Facility #2	SCE /DBP	Industrial / Ag Process (174)	1: Cold Storage 2: Process Equip.	Manual	5/28	6/9 9/23	YES YES	YES YES
SITE 5: Multi-Building Office Complex #1	SCE /DBP	Commercial Office (1,000)	1: HVAC (AHUs) 2: Lighting, Fountain Pumps	Fully Autom-ated	8/13	6/9 9/23	YES YES	NO YES
SITE 6: Multi-Building Office Complex #2	SDGE /DRP	Commercial Office (278)	1: HVAC 2: Lighting, Elevators	Partially Autom-ated	8/27	9/28 (facility test)	YES	YES
Installed Sites Participating in Summer 2004 - Events Not Captured (Sites 7, 8 & 9)								
SITE 7: Multi-Building Office Complex #3	SCE /DBP	Commercial Office (192)	1: HVAC 2: Common Lighting	Partially Autom-ated	7/31	6/9 9/23	YES NO	NO YES
SITE 8: Office Building & Call Center	SDGE /DBP	Commercial Office (288)	1: HVAC 2: Lighting	Partially Autom-ated	8/26 installed; 9/23 data	5/03 6/30 9/7	NO YES NO	NO NO NO
SITE 9: University Campus	SCE /DBP	Institutional / Educ. (720)	1: HVAC 2: Lighting, Pumps, Freezers, etc.	Partially Autom-ated	Not Complete	6/9 9/23	YES NO	NO NO
Installed Sites Not Participating in Summer 2004 Events (Sites 10, 11 & 12)								
SITE 10: Glass Processing Facility	SCE /DBP	Industrial / Material Process (128)	1: Process Equip 2: Other Process Equip.	Manual	7/12	6/9 9/23	NO NO	N/A
SITE 11: Corporate Office & Laboratory	PGE /DBP	Commercial Office (242)	1: HVAC (AHUs) 2: Exhaust Fans	Partially Autom-ated	8/28	7/26	NO	N/A
SITE 12: Food Production & Frozen Storage Facility	PGE /DBP	Industrial / Food Process (70)	1: Freezers 2: Other Process	Manual	6/1	7/26	NO	N/A

3. KEY 2004 SUB-METERING FINDINGS

This section summarizes the key findings drawn across the individual sub-metering analyses and reports. Key findings are organized by categories related to the research questions presented previously.

3.1 SUCCESSFUL AND UNSUCCESSFUL DR STRATEGIES, MEASURES, AND PRACTICES

HVAC was the primary curtailed end use for commercial sites. There were only a few categories of curtailed end-uses within the sub-metering sample. In all six commercial sites, curtailment of HVAC systems was the primary, and often exclusive, source of planned load reductions. Each site tended to have a unique process for obtaining load reductions from their HVAC systems and had varying degrees of success. Only two of six commercial sites that actually curtailed HVAC system loads were unable to meet minimum (DBP) bid requirements of 100 kW through HVAC curtailments. Successful HVAC demand reductions from the remaining four commercial sites ranged between 100 kW and 460 kW. These impacts ranged from 6 to 31 percent of event-day peak loads and averaged 19 percent across the four sites. Sub-metering revealed how each type of curtailed HVAC equipment contributed to overall load reductions, and indicated which HVAC strategies worked well and which did not.

Lighting impacts were minimal. Load reductions from secondary measures in the sampled sites of the commercial sector are small as compared with HVAC and are typically obtained from lighting and a host of other ancillary loads (e.g., elevators, fountain pumps, etc.). Lighting curtailments, if used at all, were typically confined to common areas, and often required separate, manual controls to activate curtailments (in 4 of 6 commercial sites). Building managers in this sample appeared reluctant to impact tenants with lighting sheds and were far more focused on managing complex HVAC systems during DR events. The potential for lighting reductions may have been limited in this sample by concerns over tenant and general occupant reaction and limited ability to remotely and precisely control lighting levels. The sub-metering results showed how rarely lighting measures where actually deployed during DR events despite the frequency with which they appeared in planned DR strategies.

Cold storage systems produced significant, successful curtailments. Based on the ease of recruitment and observations from DR events, agricultural and food processing facilities with cold storage systems were generally successful program participants. These facilities tended to use their cold storage systems as the primary source of load reductions by cutting off compressors and letting product temperatures float for a limited period. Product temperatures were monitored in several of these cases and found to remain within the participants' identified tolerance levels. Cold storage system loads are typically manually controlled, and can often provide greater load reductions as an individual measure than those derived from multiple, manually-controlled process loads.

Moderate Batch Process Impacts. It is assumed that batch processing industries have greater control over process equipment curtailments relative to their counterparts in continuous process industries, yet within the sub-metering sample, the use of batch process loads for curtailments was not as extensive as expected. All of the six industrial sites in the sub-metering sample utilized a form of batch processing, yet in all but one site, process load curtailment was not the primary end use by which load reductions were to be obtained. In five industrial sites, cold storage or HVAC system curtailments were the primary curtailed load, and in the only industrial site where all planned curtailments were process loads, no DR actions were taken in summer 2004. For the customers in our sample, daily (and seasonal) variability in process loads were often observed for individual pieces of equipment, thereby introducing considerable variability in the estimated baseline loads as seen at the revenue meter. In some cases, load impacts from the curtailment of a process load that was not operating during some or all baseline days would be measured at a lower value than the actual load impacts on the event day. Participating customers were typically unaware that curtailing process loads may be rendered neutral by baselines circumstances, yet in some instances, had they not acted, baselines methods would have penalized them with negative load impacts.

Scheduling and daily production requirements were the primary obstacles to deploying process load curtailments. Yet, with enough notice and production flexibility, site managers were often effective in planning and deploying process load curtailments during DR events, but these efforts were rarely the primary contributor to load reductions at the revenue meter. For the first of several curtailments at one site, work shifts were actually modified in order to execute a curtailment of all available process loads. In this event, the customer exceeded their first, experimental bid by a factor of three. Yet, in successive events at this facility, many of the process load curtailments were jettisoned as the customer came to increasingly rely on cold storage system curtailments in successive events.

Seasonality and work shifts significantly limited curtailment potential for some customers. In the analysis of successive DR events for commercial and industrial sites, it was observed that times of reduced facility demand were associated with reduced impacts. Seasonal production cycles and daily work shifts cycles tend to impact some process industries' (e.g., food production related) ability to shed load during certain summer months, during certain hours of DR events, or when summer peak loads otherwise occur. Seasonal shifts in production affect the extent to which load reductions can be obtained from these types of process loads. In peak production periods there is tendency to operate processes without interruption. Conversely, during periods of reduced production, process lines have a higher probability of being shut down. One agricultural processing and cold storage site obtained off-season load reductions that were less than a third of the load reduction observed during their peak season. For similar reasons, daily work shifts affect the ability to obtain load reduction potentials at certain times of the day. For example, in several cases, work shifts ended in early or mid-afternoon, and load reductions from process loads were eliminated or substantially diminished.

Limited use of Back-Up Generators. During recruitment, sites that planned to use only back-up generators (BUG) as their *exclusive* means of obtaining load reductions were eliminated as sub-metering candidates. Many of these sites were known to have successfully participated in DR events. Back-up generators were present in half of the commercial and industrial facilities in the sub-metering sample and some of these customers indicated they might be used during DR events in conjunction with other DR actions. In practice, however, there were no instances of their use during DR events. This suggests that BUGs for load reductions are typically deployed as an exclusive measure, and may rarely be combined with other load shedding measures. One reason for this may be the complexity of synchronizing the integration of load shedding measures with onsite generation.

3.2 EFFECTIVENESS OF MANUAL VERSUS AUTOMATED DEMAND RESPONSE

Widely varying levels of automation among commercial and industrial sites. Levels of automation differed significantly between the commercial and industrial sectors. All of the six industrial sites in the sub-metering sample utilized manual controls to activate and control load reductions. The six commercial sites, primarily offices, were found to have varying levels of controls, primarily conventional energy management systems (EMCS), as discuss further below.

Industrial applications used manual controls. Manual implementation of load curtailment in the sampled industrial process facilities was highly effective, particularly when there was one primary load or a small set of loads to control. Note that these facilities were characterized as primarily batch processes not continuous. There were no cases of automated controls being used for load curtailment.

Limited use of automated controls for HVAC. All of the commercial sites had EMCS systems for controlling HVAC systems, though the systems and their operators varied in their level of sophistication. However, many of these systems featured patch-worked integration with remnant legacy control systems, did not store or trend data points, and most did not include control of other building systems (e.g., lighting). The level of commercial building automation did not necessarily improve the probability or effectiveness of DR event participation within our sample. One customer with a moderately sophisticated EMCS needed assistance in determining how they could achieve a minimum bid of 100 kW and did not have the means to control their HVAC system in a comprehensive manner - thereby limiting their load reduction potential and the ability to participate. HVAC systems are complex, with design and control features that can counteract singular measures (e.g., raising chilled water temperature setpoints without controlling supply fans on air-handling units). Consequently, integrated HVAC system strategies that are pre-tested tend to be more effective.

Virtually no use of automated systems for lighting. Among the six commercial sites, lighting DR measures were identified as secondary measures in the DR strategies of five sites. As discussed above, lighting measures were often on a manual or separately controlled, automated system (separate from HVAC EMCS), lacked precision (limited or no ability to *partially* reduce lighting within usage areas), were not often deployed in actual events, and were often confined to common areas. Lighting measures may not generate impacts commensurate with the time and attention required to execute curtailments in commercial buildings.

3.3 CONSTRAINTS AND LIMITATIONS ON PARTICIPATION

Commercial concerns over tenant and occupant impacts. Commercial office participants were particularly averse to impacting their tenants. Several of the commercial sites never actually experimented with HVAC curtailments prior to DR events, presumably out of aversion to possible tenant impacts and complaints. Others terminated HVAC curtailments prior to the end of DR events, thereby falling short minimum DR bid requirements. Similarly, the few DR lighting measures that were seen in sub-metered events were typically undertaken in common areas. Office HVAC DR measures were terminated or avoided altogether if they were likely to increase indoor temperatures to a point where occupants would notice them. Notably, however, there were no instances where site managers reported complaints of occupancy discomfort during DR events in commercial sites. In one instance of a DR event in an industrial process facility, a facility manager reported receiving complaints of high indoor temperatures on the production floor where temperature setpoints had been raised as a DR measure. Unlike his counterparts in commercial office sites, the facility manager felt these complaints were minor and had no bearing on production or worker productivity. This consistent with the finding from the baseline survey in the overall 2004 WG2 Evaluation study that found commercial sector customers were much more concerned about occupant comfort impacts of DR than were industrial customers.

Limitations of Notification Process. For several customers, the notification process was a significant barrier to participation. Site operations personnel are exceedingly busy managing industrial production on the plant floor or managing different aspects of commercial building operations. Many cases of non-participation in the monitored DR events were simply a function of the manager being unavailable to receive or respond to notifications. Especially for industrial participants, longer event notification periods are helpful, if not essential. Although there are cases of industrial process sites that need an hour or less to respond, the necessary changes to production schedules and work shifts more often take between 8 and 24 hours to plan. Curtailment of industrial processes was very much subject to the time between notifications and events, and the customer's flexibility to modify process and production lines during an event. The level of flexibility was determined by a host of factors, most notably production schedules and deadlines, interdependencies between processes within the plant, ease of (manual) control, work shifts and other labor impacts. While there are cases of large and significant process load curtailments in our sample, there are a greater number of instances where process curtailments were not undertaken due to production requirements that could not be quickly rescheduled. Aside from more advance notification, other improvements may be made to the process including more widespread use of notifications that are distributed to several individuals involved with site energy management.

Declining institutional "memory" for executing DR strategies. Infrequent DR events appear to reduce the participants' probability of event participation and their ability to deploy a planned DR strategy during events. There were several instances where site energy managers incorrectly recalled which events they participated in or what measures they took during past events. This reinforces findings in the overall evaluation that the institutional "memory" for executing DR strategies erodes over time if there are no or very few opportunities for participation.

Diminishing trend in participation. The three commercial office sites that had an opportunity to participate in more than one DR event either did not participate in the last event, or obtained load reductions that were lower than the prior event. One fully automated site deviated from its' automated, pre-programmed HVAC load reduction strategy and curtailed small loads that it had not previously indicated in its original DR strategy. The other two sites did not participate in the last event because they did not receive notifications in time to shed, or did not receive them at all.

Change in site ownership and staff. Change in site ownership or in the personnel responsible for operating DR strategies significantly erode the institutional capability to participate in DR events and deploy an effective DR strategy. In one case, change in site ownership terminated customer participation in the DRP program, and the site energy manager responsible for the DR strategy was dismissed. Change in site personnel can also translate into missed notifications because contact information has not been updated or knowledge of the applied DR strategy is lost. At another site, event notifications were missed as they were sent only to the departed personnel.

Limited building operator knowledge of DR strategies. Based on observations of the candidate recruitment process and among sampled sites, building operators level of knowledge of how to operate DR measures and the likely impacts on energy services varies considerably, but it is often limited. Many candidates in the sub-metering recruitment process were screened out because they were found to lack coherent DR strategies.

Limited ability to quantify costs of participation. Customers in the sample did not have any reliable and comprehensive process for the accounting of participation costs and benefits (incentives), and relied on a more intuitive assessment of the potential costs of disrupted production and tenant dissatisfaction. In one instance, a frozen food processing facility (Site 12) that did not participate in summer 2004 DR events, expressed a specific need to for a process or tool to help them assess the costs and benefits of curtailment actions specific to their facility.

Need for assistance but not in form of 2004 Technical Assistance Program. Despite the general need for information on DR measures, costs and benefits, none of the sub-metering candidates or sampled sites were known to have utilized the Technical Assistance Program (TAP). Among the sampled sites, customers indicated they were either unaware of the program, felt they did not need it, believed that it would not address their specific information needs, or thought that there were prohibitive institutional constraints or financial risks associated with it. Requests for advise on appropriate DR measures and strategies were commonly encountered throughout the sample recruitment process, and among five sites that received onsite surveys. In that the sub-metering sample sought to identify the most probable program participants, it is posited that there is a considerable demand for technical assistance with developing and operating DR strategies in the program population.

3.4 ACCURACY, EFFECTIVENESS, AND PROBLEMS WITH BASELINE METHODS

The findings below refer to the two principal baseline load estimation methods described and analyzed in the 2004 WG2 Evaluation report. Readers unfamiliar with these baselines and issues associated with their accuracy should see Sections 6 and 7 of the December WG2 Evaluation report.

Sub-metering results reinforce finding that 10-day adjusted baseline method is more accurate than 3-day method. The sub-metering results provide strong evidence that both the three-day and ten-day adjusted baselines can be inaccurate under different circumstances. However, the three-day baseline appears to be much less accurate on average than the 10-day adjusted method. The ten-day baseline was found to more closely track the trajectory of event-day loads in non-event hours for the sub-metered loads. In many cases, the use of revenue meter data with the 3-day settlement provides a false indication of impacts that is revealed by analyzing the sub-metered data. In all but a few instances, the three-day baseline overestimated load reduction impacts relative to absolute load reductions observed on the event days. The ten-day adjusted baseline commonly measured smaller load impacts relative to the three-day. The ten-day baseline was subject to under-estimation of load impacts in cases where the customer initiated load reductions more than an hour in advance of the start of the event period. An inaccurate baseline creates opportunities for free riders when it consistently over-estimates event day loads. The sub-metering data provides evidence of possible free riding in a few cases.

Customers' need to better understand baseline methods. Baseline methods are critical to assessing benefits of DR strategies, and among participants of the sub-metering sample, two customers expressed the specific need to better understand their baselines, or had a need to know what they looked like as a part of their process for taking action during DR events. Two of the sampled customers complained of delays in receiving bill credits, such that they experienced a very long delay in finding out how much they benefited from participation in a specific event. Customers need to understand how baseline methods are applied, especially for customers who intend to shed a small proportion of their total load; in these cases, unrecognized baseline inaccuracies may severely, and randomly, penalize or reward participating (or non-participating) customers. The need to better understand and guage baselines was evident in the common discrepancies between load intended (bid) and actual load reductions. Improved baseline recognition could help participants consider not taking action when baselines are likely to work against them in terms of reducing incentives to where they do not justify actions taken. Alternately, baseline recognition may motivate participants to take action when the baseline will clearly capture the impacts of curtailments.

The baselines of sub-metered loads showed the variability of curtailed loads in terms of their contribution to revenue meter impacts, both within event periods and between events. The graphic comparison of baselines and event-day loads also revealed instances when end-uses intended for curtailment are not curtailed and actually detract from load reductions at the revenue meter. For HVAC loads, event day and baseline load comparisons are useful in terms of better understanding the effects of weather on HVAC curtailments. For example, there is one clear instance of a curtailed HVAC load that did not contribute significantly to the overall facility load reduction due largely to a mis-estimation associated with the three-day baseline method.

The comparison of event day and baseline loads also revealed instances when customers were unable to maintain curtailments on specific end uses and when specific loads were curtailed in advance of the DR event start times or other curtailed loads. In each case, sub-metered baselines provide key insights into challenges of curtailing specific end uses and the sequence of measures that the customer used in activating their DR strategy.

There was also one case of a fully-automated commercial site[4] that, in the latter of two events, chose to forego their automated HVAC curtailment routine (as shown in the sub-metering data) in favor of manually curtailing ancillary loads that had not previously been identified as DR measures (specifically, lighting and fountain pumps). This customer was also known to have sophisticated real-time metering capabilities, and is presumed to have the ability to observe calculated baselines along with actual event-day loads. The reason for this customer's deviation from a highly automated DR strategy is not known, yet it is plausible that, during a prior event, the customer learned of the opportunity present in the three-day baseline method to earn incentives without taking significant actions. In the latter of two events, this customer bid a curtailment that was much greater than any observable load drop seen on the event day, and did not use the primary HVAC component of their strategy.

[4] Within the sample, this customer had the most sophisticated EMCS, including pre-programmed, three-tiered DR curtailment sequence. The development of DR capabilities within this customer's EMCS were subsidized by the AB 970-funded Demand Reduction Program.