California ISO
Your Link to Power

California Independent
System Operator Corporation

California ISO

Quarterly Report on Market Issues and Performance

REVISED December 23, 2009

Prepared by: Department of Market Monitoring

TABLE OF CONTENTS

Executive Summary

This quarterly report covers the second three months of the California ISO's new nodal market (July – September, 2009), which correspond to the third quarter of 2009 (Q3). The report provides an overview of general market performance, as well as more detailed analysis of a variety of special market issues or areas for market improvements. In terms of general market performance, the new ISO markets are continuing to perform well and have improved in Q3. Most notably:

- The day-ahead Integrated Forward Market (IFM) has been very stable and competitive.
- Market activity in the Residual Unit Commitment (RUC) market has been minimal due to high levels of load scheduling in the IFM and sufficient Resource Adequacy (RA) capacity in RUC.
- The five-minute Real Time Dispatch (RTD) market has improved as a result of several software and operational changes implemented by the ISO to reduce the frequency and magnitude of extremely high or low prices that are not reflective of actual real-time supply and demand conditions While extreme RTD prices continue to occur in some intervals, these prices tend to reflect short-term supply and demand conditions, such as ramping constraints and sudden unit outages.
- Local Market Power Mitigation (LMPM) procedures have remained effective in both the day-ahead IFM and real-time markets.

Provided below is a summary of the more detailed analysis of special market issues and areas for market improvements identified in this report.

Real Time Market Performance

During Q3, the performance of the ISO's energy markets improved in terms of several key measures of market performance: (1) the competitiveness of overall prices in the 5-minute RTD market, (2) the lower frequency and magnitude of price spikes in the RTD not reflective of fundamental market conditions, (3) improved price convergence between the sequential energy markets (IFM, Hour Ahead Scheduling Process (HASP) and RTD), and (4) reduced price volatility.

Despite these improvements, significant systematic price divergence has continued to occur at times, particularly between the HASP and RTD. This price divergence has been coupled with a trend for the ISO to export relatively large quantities of additional energy in the HASP (at low prices), and then dispatch additional energy within the ISO in RTD (at significantly higher prices). This pattern of "selling low" in HASP and "buying high" in RTD has continued to create substantial revenue imbalances that are recovered based on each participant's metered loads through Real Time Energy Imbalance Energy Offset charges. Chapter 1 of this report includes a discussion of some of the potential root causes of these trends, and some of the potential solutions being implemented or explored by the ISO to reduce these price divergences. The Department of Market Monitoring (DMM) believes that the price divergence between HASP and RTD represents one of the most critical areas for further improvement in the ISO's new market software and processes.

Local Market Power Mitigation

The new ISO markets are functioning competitively and the local market power mitigation (LMPM) procedures are working effectively to mitigate any uncompetitively high market bids when they are needed to relieve congestion on uncompetitive constraints. Two specific aspects of LMPM examined in Chapter 2 of this report include the following:

- ***Mitigation of Exceptional Dispatch***. For the first four months of the ISO's new market, all Exceptional Dispatches (EDs) for energy (above a unit's minimum operating level) were subject to price mitigation. When mitigated, ED energy is paid the higher of the resource's nodal Locational Marginal Price (LMP) or its Default Energy Bid (DEB). Starting in August, however, EDs for energy are only subject to mitigation if made to relieve congestion for *non-competitive* constraints or for seasonal environmental constraints known as "Delta Dispatch". All other EDs are paid their unmitigated bid price. DMM has found that this more limited mitigation of EDs has had a relatively low impact on costs due to a combination of two factors: (1) the volume of exceptional dispatches for energy since this change took effect has been relatively small, and (2) unmitigated bid prices paid for most ED energy have not been significantly higher than the market LMPs and/or the DEBs that would be paid if the EDs were subject to mitigation. DMM will continue to monitor the potential for local market power by units receiving EDs so that appropriate changes in operating practices or market rules might be implemented if costs of such EDs became excessive.

- ***Failures of LMPM in HASP*.** Prior to the start-up of the ISO's new LMP market, one major issue identified by DMM was the relatively high frequency with which the LMPM process was not applied to bids used in the in RTD due to various problems or failures occurring during the HASP process. The HASP process is where the LMPM procedures are applied to mitigated bids used in the real time energy market. Thus, if the HASP LMPM process is not run, bids used in the 5-minute RTD market are unmitigated. During Q3, the frequency of failures in the pre-RTM LMPM process has been relatively low and has trended downward. There have been only limited price impacts resulting from failures in the pre-RTM LMPM procedures. DMM has reviewed instances where LMPM failed, and has determined that there have been numerous hours of LMPM procedure failure that were not reviewed for price impacts by the ISO's price correction team. Review by DMM indicates that it is unlikely that LMPM failures in Q3 had a significant impact on market outcomes. However, DMM is recommending that the ISO improve the price correction process to ensure that all hours in which LMPM procedures fail in HASP are thoroughly reviewed for price impacts.

Ancillary Services Markets

The ancillary services markets have generally performed well since the start of the ISO's new market design. Prices in the day-ahead and real-time ancillary services markets have been reasonable and highly competitive, with day-ahead A/S prices somewhat higher than in real-time. In Chapter 3 of this report, we examine two issues involving how the ancillary service markets interact with the energy markets:

- ***Contingency-Only Reserves*.** The first issue is the procurement of spinning and non-spinning reserve that is designated as *contingency-only* – i.e. capacity that can only be dispatched for energy in the case of a contingency event or an imminent or actual system emergency. Since the start of the ISO's new market, a very high portion of spinning and non-spinning reserve procured in the IFM has been designated by participants as contingency-only (e.g. about 70 percent during many hours). In addition, all incremental reserve procured in the Real Time Pre-Dispatch Process (RTPD) performed every 15 minutes prior to the RTD is automatically designated as contingency-only.[1] Even in cases when the ISO has enough reserve to meet its system requirements, this can create price spikes during periods where supply is tight, particularly in transmission constrained load pockets. This can occur when a relatively large amount of contingency-only reserve is located in a load pocket or anywhere on the grid where this capacity would be particularly effective when dispatched as energy to meet a local constraint. While this may not significantly impact prices with a high degree of frequency, this can increase prices dramatically when supply is tight and penalty prices on constraints are setting prices that could be relieved with a relatively small amount of additional supply that is being held as contingency-only reserve. In Chapter 3, we suggest that the ISO consider several ways in which this issue might be addressed.

- ***Reserve Scarcity Pricing*.** This report also examines the relationship between real-time energy and ancillary services as it relates to the ISO's scarcity pricing proposal.[2] Specifically, we highlight a disconnect between the real-time ancillary services and RTD energy prices that may dampen price signals in the 5-minute RTD during instances where scarcity pricing would be triggered if these two markets were directly linked. Although energy and ancillary services are co-optimized in the RTPD run performed every 15 minutes, these energy prices are not financially binding for energy – with RTD energy being settled based on prices resulting from the subsequent 5-minute RTD process. Thus, as part of the longer-term market design process, we recommend further consideration of the potential for a scarcity pricing mechanism that would more directly affect real-time energy prices, such as co-optimization of ancillary services and energy in the 5-minute RTD market.

Exceptional Dispatch

Exceptional Dispatch (ED) is a term used to describe manual dispatches performed by an ISO operator in cases where unit commitments and energy dispatches made by the market software did not fully address a particular reliability need. Since the start of the ISO's new market design, the use of ED has raised concerns, particularly among generation owners, about the efficacy of the new market software and impact these manual dispatches may have on market prices. The reasons that EDs are necessary were explained at length in DMM's previous *Quarterly Report* (Q2 Report*).*[3] In addition, in its September 2, 2009 Order, the Federal Energy Regulatory Commission (FERC) established more detailed ED reporting requirements for the ISO.[4]

---

[1] Moreover, as discussed in Chapter 3, since the ISO software requires that all spinning or non-spinning reserve being supplied by a single generating unit be either contingency-only or non-contingent, in cases when any incremental spinning or non-spinning reserve is procured from a unit in the real-time ancillary services market, any of that same reserve product that was procured from that unit in the day-ahead is also automatically designated as contingency-only.

[2] *Final Draft Proposal: Reserve Scarcity Pricing Design*, October 5, 2009, http://www.caiso.com/243e/243ecc4d2d490.pdf

[3] *Quarterly Report on Market Issues and Performance*, July 30, 2009; covering April through June, 2009. http://www.caiso.com/2425/2425f4d463570.html

[4] *Order Accepting Tariff Revisions, Subject to Modification,* 128 FERC P 61,218 (2009**).**

In Chapter 4 of this report, we provide some updated analysis of ED trends and follow-up on three specific recommendations made by DMM in our Q2 report. A summary of actions that have been taken of these three prior recommendations is provided below:

- ***Perform a comprehensive review of operational procedures and other criteria for determining exceptional dispatch***. In the last week of July, the ISO formed an ED "strike team" to focus on potential improvement to practices and software to reduce EDs, particularly with respect to unit commitments made in the day-ahead timeframe.[5] This strike team also focused on improving the consistency and logging of ED data, and providing more accurate and timely feedback on ED trends to operations staff. The team also monitored the impacts of new RUC capacity nomograms designed to meet reliability requirements previously met by committing additional units via ED either before or after the IFM. The result of these efforts – combined with the new RUC capacity nomograms discussed below – appears to have reduced EDs in late July and August. However, as discussed in Chapter 4, the amount of capacity committed via EDs increased again in late August and September due to other factors, such as the need to protect against contingencies related to fires in Southern California and a significant prolonged outage on the Southwest Power Link (SWPL), which affected numerous reliability requirements within the ISO and required a significant de-rating of import capacity on the Palo Verde branch group.

- ***Explore and implement options for incorporating into the market model the reliability constraints driving exceptional dispatch.*** In Q3 (July 27), the ISO implemented capacity nomograms in the RUC process that reflect capacity needs incorporated in the G-217 (South-of-Lugo) and G-219 (Orange Country) operating procedures, which were found to be driving a large portion of unit commitments in the Southern California area. The ISO is also finalizing a RUC nomogram to reflect a third major operating procedure that covers the San Diego area (G-206). Since minimum load energy and other capacity from units committed in RUC is not available in the IFM market, DMM has recommended that these constraints be incorporated in the IFM market model if possible. This will reduce excess generation in the real-time markets (HASP and RTD) resulting from minimum load committed after the IFM, and will also provide resources needed for these constraints with additional opportunity for market revenues in the IFM. The ISO is currently developing procedures to incorporate these capacity constraints in the IFM, and has indicated these may be completed by the end of 2009.

- ***Consider new market products that might mitigate the need for exceptional dispatch***. As described in the ISO's most recent 120-day report to FERC, the ISO has committed to a process over the next nine months to consider potential new products. However, the ISO believes that it would be more appropriate to have a full year of operational experience and information before determining what, if any, specific new products or market design enhancements can most effectively mitigate the volume of future EDs. DMM considers this approach prudent ─ particularly in light of the operational and software improvements that have been implemented or will be implemented in the near future to reduce EDs. DMM also notes that by continuing to identify ways to incorporate constraints requiring EDs into the market model, the ISO is continuing to develop information that will be valuable and necessary as part of the process of considering new potential products.

---

[5] See Memorandum to ISO Board of Governors, Jim Detmers, September 2, 2009, *Re: Briefing on Exceptional Dispatch*, http://www.caiso.com/241e/241eb60ca5f0.pdf

Biasing of Transmission Constraints

Since the start of the ISO's new market design, one of the key "levers" that may be used by ISO operations to help manage reliability and congestion within the ISO grid has been to bias (or adjust) the limit used by the market software to limit modeled flows allowed on each constraint. For example, two common reasons for biasing of a constraint are to (1) adjust for discrepancies between calculated market flows and measured or predicted actual flows; or (2) allow a reliability margin for certain flowgates in case of a sudden change in system conditions.

Chapter 5 of this report provides (1) a detailed description of the various reasons that constraints may be biased, (2) statistical analysis of the frequency and degree to which constraints have actually been biased in Q3, and (3) several case study examples of the use and impacts of biasing on specific constraints. Key findings of this analysis include the following:

- During the first few months of the ISO's new market design, one of the root causes of numerous RTD price spikes was that constraints were sometimes biased down to a significantly lower level in 5-minute RTD process than in the real-time pre-dispatch (RTPD) and real-time unit commitment (RTUC) process. Such discrepancies can increase RTD price spikes by preventing additional short-start resources from being available to relieve constraints in RTD. However, DMM found that in Q3, biasing in the RTPD/RTUC and RTD processes has generally been highly consistent, indicating that efforts to improve the consistency of biasing in these sequential real-time market processes have been successful.

- In real-time, constraint biasing tended to be used to increase – rather than decrease – the market limit on constraints in order to avoid "phantom" congestion (i.e., congestion that would occur in the RTM software when observed flows in real-time were below the constraint's actual operating limit).

- The practice of biasing constraint limits is used very infrequently in the day ahead market, since review by the ISO's Operations Engineers has typically concluded that biasing in the IFM or RUC would not tend to avoid "phantom congestion" in the day-ahead market or mitigate the potential for congestion in the real-time market. DMM's review of data on the biasing of constraints in real-time and congestion that occurred in the IFM confirms that flowgates that were biased up in real-time were very rarely congested in the IFM.

Based on analysis in this report and DMM's ongoing monitoring of this issue, we provide the following recommendations:

- Given the dynamic nature of discrepancies between modeled and actual flows – and the significant impact that biasing can have on market outcomes – the ISO should continue to place a high priority on continuing to refine the use of constraint biasing in the day-ahead and real-time processes as it gains more experience and data in this area. For example, more automated statistical metrics that correlate the degree of biasing and congestion in the various sequential markets may be helpful in tracking trends and identifying potential areas for improvement as conditions change.

- While we have observed consistent biasing across the real-time markets in Q3, applying a bias is a manual process that takes some time and must be repeated for the different real-time markets. Thus, DMM suggests that use of the bias might be made more effective by developing a tool for operators that better facilitates applying bias across the ISO's two real-time market models (RTPD and RTD) on a consistent basis.

- Overall market transparency and the ability for participants to "self-manage" congestion can be improved by providing timely data to market participants on the application of bias and un-enforcing constraints in market operations. DMM understands this issue will be addressed as part of a more comprehensive stakeholder process on public data release to be initiated in Q4.

Resource Adequacy

The Resource Adequacy (RA) program is a key component of the ISO market that is designed to ensure there will be sufficient generation capacity to meet demand, particularly under peak load conditions. Under the RA program, all load-serving entities (LSEs) must arrange enough RA generation and demand response capacity to meet 115 percent of their forecast peak demand in each month (based on a 1-in-2 year load conditions). The 115 percent requirement is designed to include the additional operating reserve needed above peak load (about 7 percent), plus an allowance for outages and other resource limitations (about 8 percent).

Most resources counted toward this RA requirement are required to be made available to the ISO markets for each hour of the month that the resource is physically available. However, since the ISO has limited information upon which to verify the physical availability of many RA resources, the actual overall availability of RA resources during peak periods when this capacity is needed most for reliability and market performance ultimately depends on the amount of RA capacity that is scheduled or bid by participants into the ISO markets.

In Chapter 6, we examine the actual availability of the RA resources to the various ISO markets (IFM, RUC and RTM) during the highest 140 load hours in Q3, which includes all hours with loads over 40,000 MW. The overall average availability of RA resources was relatively high during these hours (about 91 percent in the IFM and 88 percent in RUC). This represents an overall availability just slightly below the 92 percent level that is implicitly incorporated in RA program requirements.[6] DMM notes that under higher loads that equal or exceed the 1-in-2 year load conditions used in setting RA requirements, this difference could have a significant impact on ISO market performance and system reliability. DMM also believes these findings reinforce the need to maintain or even improve overall availability of RA resources, and for the ISO to continue to consider future refinements to the RA process and the ISO's recent RA Standard Capacity Product (SCP) tariff provisions applicable to some RA resources. For example, refinements to the SCP to measure the amounts of all RA capacity actually made available to the ISO markets through bids or self-schedules may help ensure that the required overall level of availability of RA resources can be maintained.

[6] 115 percent RA requirements less 7 percent operating reserve = 108 percent. Thus, after accounting for operating reserve, just over 92 percent of remaining RA resources would be necessary to meet the 1-in-2 year peak load used in setting the RA requirement.

1 Energy Market Performance

This chapter focuses on price convergence across the ISO's three energy markets: the Day-Ahead Integrated Forward Market (IFM), the Hour-Ahead Scheduling Process (HASP), and the five-minute Real-Time Dispatch (RTD). In addition, we provide a review of trends in RTD price volatility. As discussed in this section, while price convergence has improved and price volatility decreased substantially in Q3 relative to the first three months of the ISO's new market, significant systematic price divergences have continued to occur at times, particularly between the HASP and RTD markets. This price divergence has been coupled with a trend where the ISO decrements or exports relatively large quantities of energy in the HASP (at low prices), and then dispatches additional energy within the ISO in RTD (at significantly higher prices). This pattern of "selling low" in HASP and "buying high" in RTD, has continued to create substantial revenue imbalances that are recovered based on each participant's metered loads through Real Time Energy Imbalance Energy Offset charges. This section includes a discussion of some of the potential root causes of these trends, and some of the potential solutions being implemented or explored by the ISO to reduce these high uplift charges.

1.1 Overview

The performance of the ISO's energy markets improved during the second three months (Q3) of the ISO's new nodal market design in terms of several key measures of market performance: (1) the competitiveness of overall market prices and outcomes, (2) the frequency and magnitude of price spikes not reflective of fundamental market conditions, (3) price convergence between the sequential energy markets, and (4) price volatility.

In Q3, the performance of the ISO's real-time market (RTM) for energy improved significantly as a result of a variety of steps taken toward the end of Q2 and beginning of Q3 that decreased the frequency and magnitude of price spikes not reflective of fundamental market conditions. Three of these changes that appear to have had very significant impacts include the following:

- In early June, the pricing run of the RTD software was modified to allow transmission constraints to be exceeded by 5 MW instead of the previous threshold of .1 MW during the first 5-minute interval of the RTD optimization. This modification allows extra slack on a constraint that may not be fully resolved in a single 5-minute interval, but would otherwise have a significant impact on prices due to ramping constraints enforced in the RTD software.

- Starting August 1, the RTD software was modified to represent how regulating reserve is used to balance short-term high-frequency load fluctuations. This modification allows limited relaxation of the power balance constraint through a lower scheduling run penalty price.[7] These modifications would account for the effect of regulation ramping capability that will naturally be provided by resources providing regulation via Automated Generation Control (AGC).

[7] Prior to relaxing the power-balance constraint in the scheduling run at a penalty price of $6500, the power-balance is allowed to relax at a price slightly above the bid cap in cases of acute under-generation conditions and slightly lower than the bid floor for acute over-generation conditions. This relaxation is only for a limited quantity of megawatts reflective of a portion of awarded regulation capacity to account for the effect of regulation ramping capability that will naturally respond to meet load in real-time.

- In Q3, the ISO also implemented a tool that allowed phasing in bias (of load and transmission limits) across several market intervals rather than in one interval. This allows the market to adjust to new targets and limits more gradually (generally over a 15 minute period) and reduces the frequency of extreme prices and their impact on price convergence that otherwise would occur due to sudden "shocks".

However, as discussed in this chapter, the degree of price convergence between the ISO's sequential energy markets (IFM, HASP and RTD) in Q3 represents a major source of potential improvement to the ISO's overall energy market performance.

1.2 Price Convergence

One of the key measures of overall performance of the ISO's energy markets (IFM, HASP, and RTD) is the degree to which prices across these markets converge. A high degree of price convergence is an indicator of market efficiency, as it suggests that resource commitment and dispatch decisions are being optimized across the markets within the ISO, as well as between the ISO and neighboring control areas. In addition, as noted above, price divergence in the HASP and RTD can create substantial "uplifts" that must be recovered from LSEs through Real Time Energy Imbalance Energy Offset.

Price convergence can be measured and analyzed in a variety of ways. One approach is to examine the extent to which average prices converge over a period of time. In the first few months of the ISO's new market, average IFM prices tended to be consistently lower than RTD prices, and average HASP prices tended to be consistently lower than both IFM and RTD prices. However, over the first six months of this new market, price convergence in these three markets has improved substantially, even in the presence of extraordinary grid conditions, such as the significant transmission outages that occurred in September.

As shown in Figure 1.1, convergence of IFM and RTD prices in the Southern California Edison (SCE) load aggregation point (LAP) improved significantly in Q3 during peak hours (HE 7-22 Monday through Saturday), with average monthly SCE LAP prices approximately equal to IFM prices in July, and about 10 percent above IFM prices in August and September. However, during off-peak hours (HE1-6 and 23-24, and all day Sundays), RTM prices were systematically lower than IFM prices, with off-peak RTD prices averaging about 25 percent less than IFM prices during Q3.

While convergence of HASP prices with IFM and RTD prices also improved in Q3, average HASP prices continued to be systematically lower than the IFM and RTD prices during peak hours. During peak hours, HASP prices were about 20 percent lower than RTD prices in July and August, and about 4 percent lower in September. In the off-peak hours, HASP prices were about 32 percent and 13 percent higher than RTD prices in July and August, respectively, before falling dramatically lower than RTD prices in September. Low off-peak HASP prices in September in the SCE and San Diego Gas and Electric (SDG&E) regions are skewed by HASP interval prices below -$1,100/MWh in three hours on September 29 and 30, when the hour-ahead dispatch optimization predicted excess generation trapped within Southern California. Because HASP LAP prices have no settlement impact, they were not corrected by Market Services. Other sources of more systematic divergences between HASP and RTD prices are discussed in detail later in Section 1.3.1 of this chapter.

Figure 1.1 Comparison of SCE LAP Prices

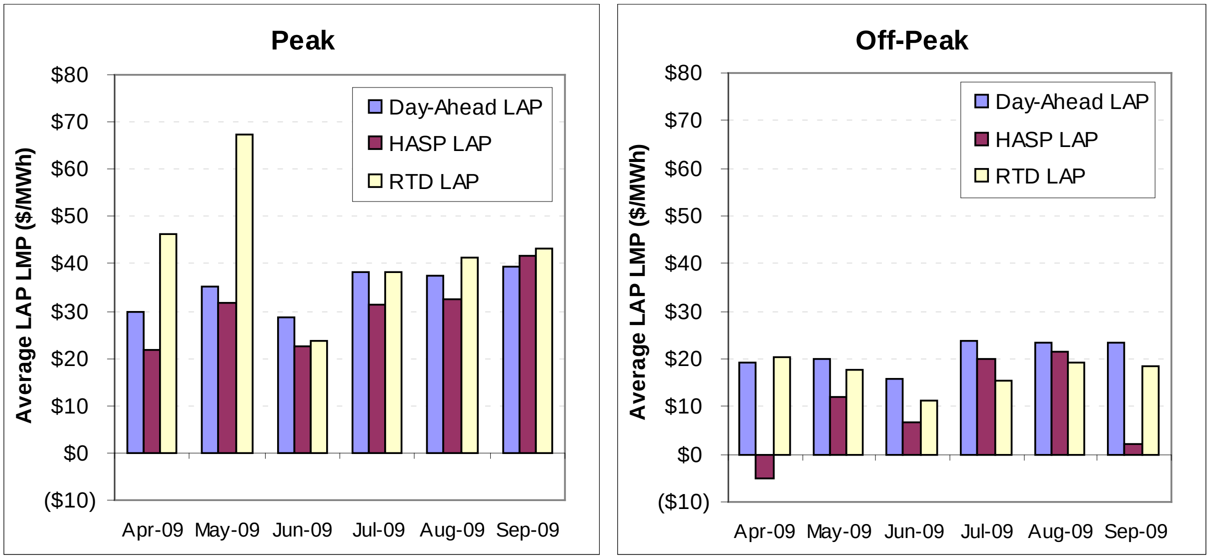
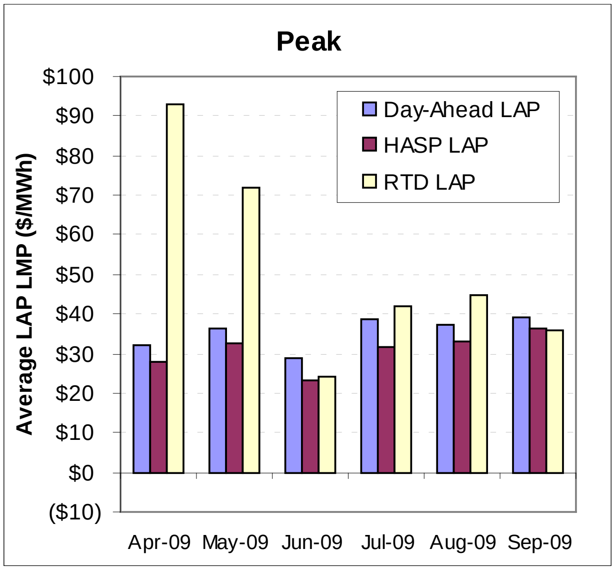


Figure 1.2 and Figure 1.2 below show similar charts for the SDG&E and Pacific Gas and Electric (PG&E) LAPs, respectively. SDG&E and PG&E average prices show similar patterns to that of SCE in July and August. In September, SDG&E prices tracked closely with SCE prices, again with low off-peak HASP prices due to the anomaly on September 29 and 30, while the PG&E LAP exhibited a pattern of higher prices due to a Path 15 outage during the same period that resulted in congestion. During peak hours in the PG&E LAP, HASP prices were approximately 11 percent higher than IFM prices, while RTD prices exceeded IFM prices by approximately 22 percent. During off-peak hours in September, HASP prices in the PG&E LAP were approximately 17 percent lower than both IFM and RTM prices, but were not as dramatically low as the HASP prices in the SCE and SDG&E LAPs.

Figure 1.2 Comparison of SDG&E LAP Prices



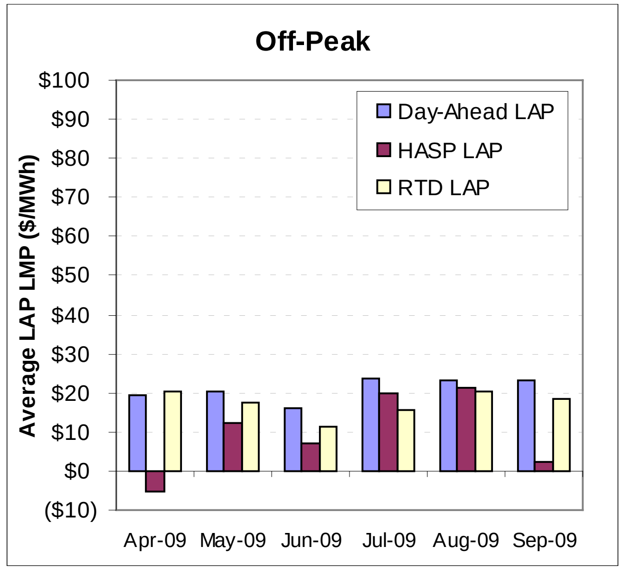
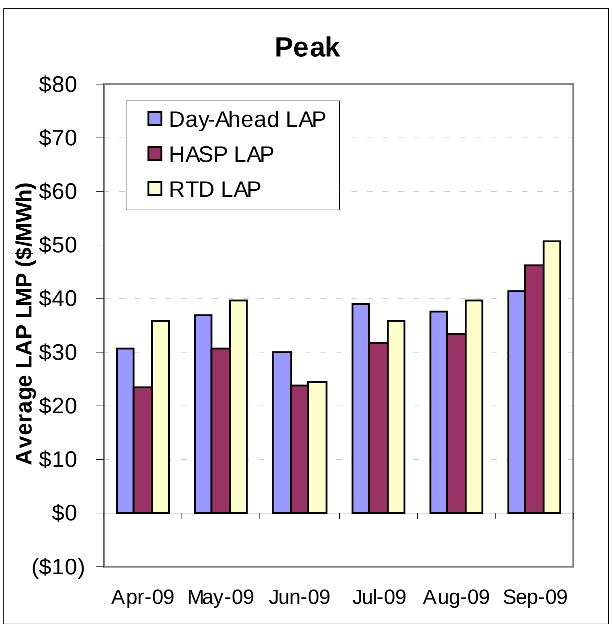
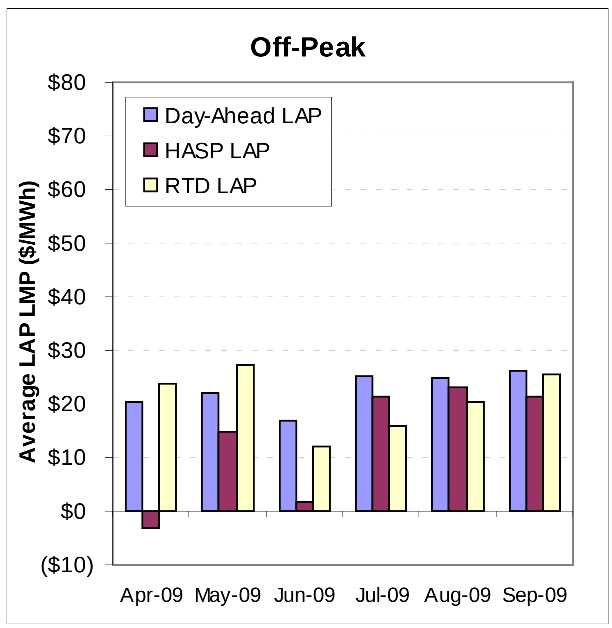


Figure 1.3 Comparison of PG&E LAP Prices





1.3 Day-Ahead and Real-Time Market Prices

A second measure of price convergence is the distribution of the price differences between markets, such as the day-ahead IFM and RTD prices. Figure 1.4 through Figure 1.6 below show the distribution of price differences between real-time (RTD) and day-ahead IFM LAP prices (LAP LMP_{RT} – LAP LMP_{DA}) by month and period of day for SCE, SDG&E, and PG&E, respectively.

A consistent trend throughout all LAPs and periods is a decrease in the frequency of real-time prices that were significantly lower than day-ahead prices. In the initial months of market operation, problems with the load forecasting tool resulted in frequent over-scheduling in the day-ahead. The ISO's day-ahead forecasts have since improved, resulting in fewer over-generation conditions that required large volumes of power to be sold in real-time at low prices.

The distribution of price differences in SCE, SDG&E, and PG&E are shown in the figures below. Differences have generally decreased in magnitude, except for the peak period in July. The anomalous July peak-hour difference's upper tails in SCE and SDG&E are due to price spikes affecting Southern California caused by congestion on Path 26, which affects prices in both SCE and SDG&E LAPs, and by congestion on the SDG&E-CFE branch group, which affects only the SDG&E LAP. Overall, the distribution of price differences narrowed in Q3, indicating better convergence of prices between the two temporal markets.

Figure 1.5 and Figure 1.6 show similar summaries for SDG&E and PG&E, respectively. Prices in these LAPs have much the same distribution pattern as SCE, except for July where the PG&E LAP LMP was less affected by congestion on Path 15 and SDG&E-CFE Import Branch Group.

Figure 1.4 Distribution of SCE LAP Price Differences Between IFM and RTD

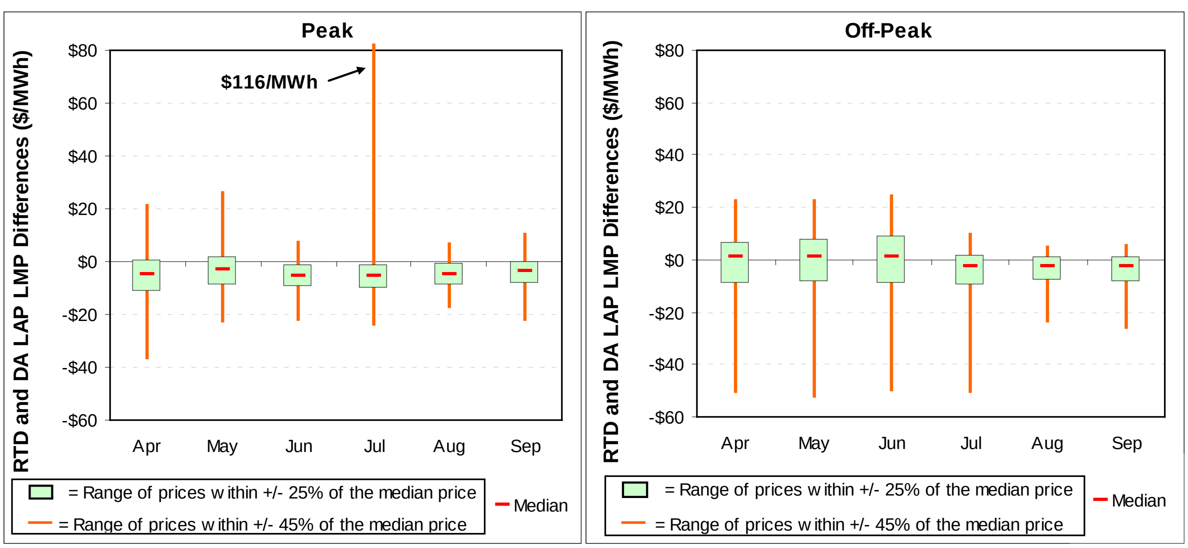


Figure 1.5 Distribution of SDG&E LAP Price Differences Between IFM and RTD

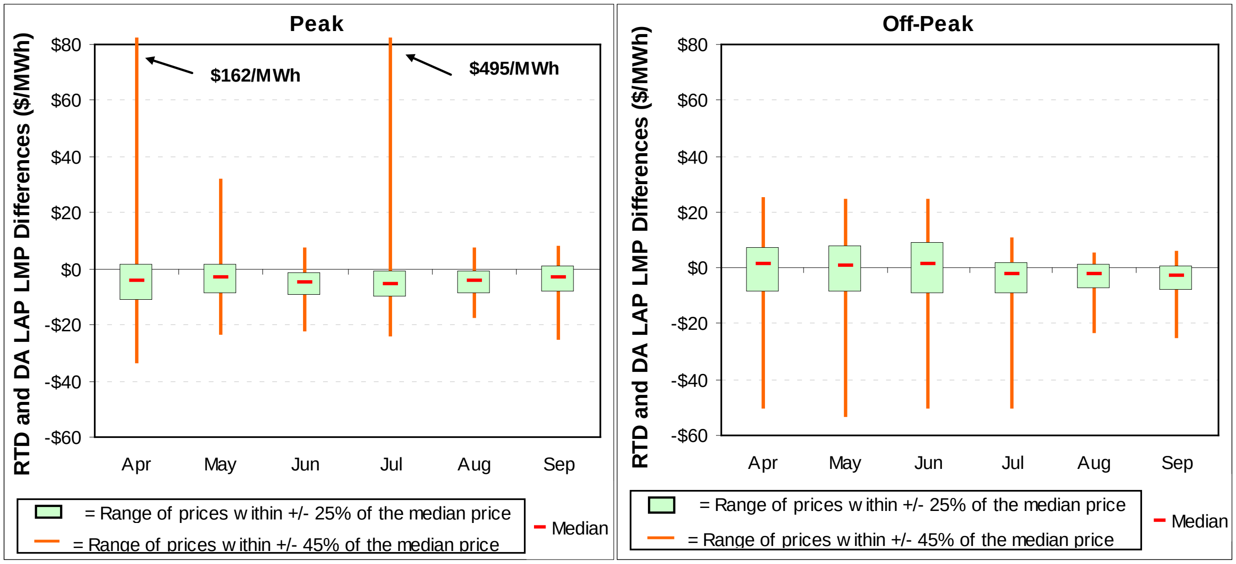
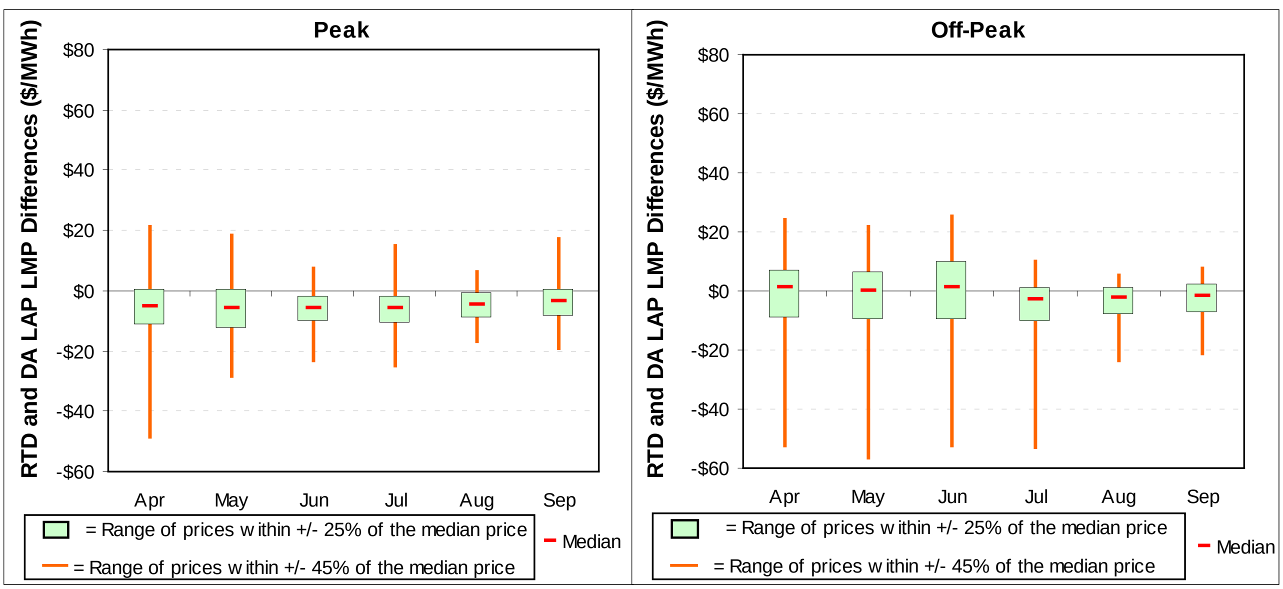


Figure 1.6 Distribution of PG&E LAP Price Differences Between IFM and RTD



1.3.1 HASP and Real Time Market Prices

While price convergence between the ISO's different energy markets improved in Q3, prices in the HASP continued to be systematically lower than prices in both the day-ahead and real-time markets. HASP prices are used financially only to settle prices for incremental changes in hourly imports and exports on the inter-ties (i.e., relative to final day-ahead schedules). However, the lack of convergence between hourly HASP prices and RTD prices used to settle generation dispatched within the ISO in the 5-minute RTD market can have a significant impact on real-time energy uplift charged to load-serving entities (LSEs) within the ISO. The magnitude and root causes of uplift charges to LSEs created by this trend towards "selling low" in the HASP and then "buying high" in the RTM are explained and analyzed in several recent ISO whitepapers.[8]

A significant portion of the divergence between HASP and RTM prices can clearly be attributed to a combination of (1) price spikes in the RTM markets during a relatively small number of peak hours, and (2) very low prices in the HASP during an even smaller number of off-peak hours. The degree of this impact is further illustrated in Figure 1.7 and Figure 1.8, which compare average hourly prices for the PG&E and SCE LAPs in the RTM with HASP prices for the major inter-ties into the PG&E and SCE LAPs (Malin and Palo Verde, respectively) *with* and *without* very high or low RTD and HASP prices. The dotted lines in these figures show average prices based on all hours, while the solid lines represent average prices after screening out prices for about 3 percent of hours in Q3 when hourly LMPs in these markets exceeded $100/MW or fell below -$30/MW.

However, as shown in Figure 1.7 and Figure 1.8, during peak hours, even after excluding the 3 percent of hours with relatively high RTD prices or low HASP prices, HASP prices used to settle imports/exports tend to be systematically lower than RTD prices used to settle generation dispatched within the ISO in the 5-minute market. During off-peak hours, HASP prices for imports on Malin are systematically higher than RTD prices for the PG&E LAP, while HASP prices on Palo Verde are roughly equal to RTD prices for the SCE LAP.

This more systematic trend towards "selling low" in the HASP and then "buying high" in the RTM to replace the energy decremented in HASP (as well as meet additional demand to meet ISO load in the RTD) is illustrated in Figure 1.9. Again, the hourly average price and dispatch data in Figure 1.9 exclude the approximately 3 percent of hours when hourly prices in the HASP or RTD exceed $100/MW or drop below -$30/MW in order to illustrate how these dispatches and prices appear to be driven by more systematic factors beyond extreme prices in a relatively few hours. The net dispatch in the HASP averaged about -700 to -1,000 MW during most hours in Q3, while the net dispatch in the RTD averaged about +300 to +900 during most hours. On average, the price at which this energy was sold in HASP was significantly lower than the price at which additional energy was purchased in RTD (e.g., averaging $5 to $10 during many hours). While congestion and changes in system conditions between HASP and RTD could contribute to this trend, the consistency and magnitude of the trend of "selling low" in the HASP and then "buying high" in the RTM to replace the energy decremented in HASP provides a

---

[8] *Issue Paper: Analysis of Real-time Imbalance Energy Offset (CC 6477)*, revised August 26, 2009, http://www.caiso.com/2416/2416e7a84a9b0.pdf.

Straw Proposal: Mitigation and Allocation of Real Time Imbalance Energy Offset Costs (CC 6477), prepared September 23, 2009, http://www.caiso.com/2432/2432e7916dfa0.pdf strong indication that other more systematic factors are contributing to this trend. Several key factors that appear to be contributing to this trend are discussed later in this section.

Figure 1.7 Comparison of Malin Price in HASP and PG&E LAP Price

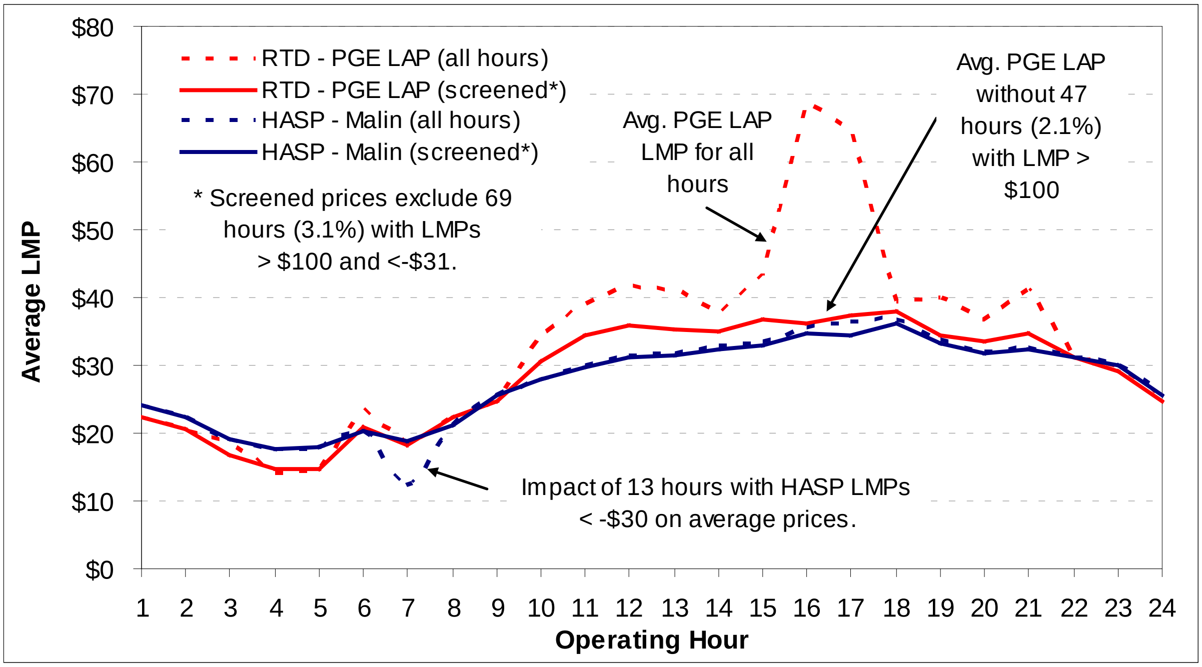


Figure 1.8 Comparison of Palo Verde Price in HASP and SCE LAP Price in RTD

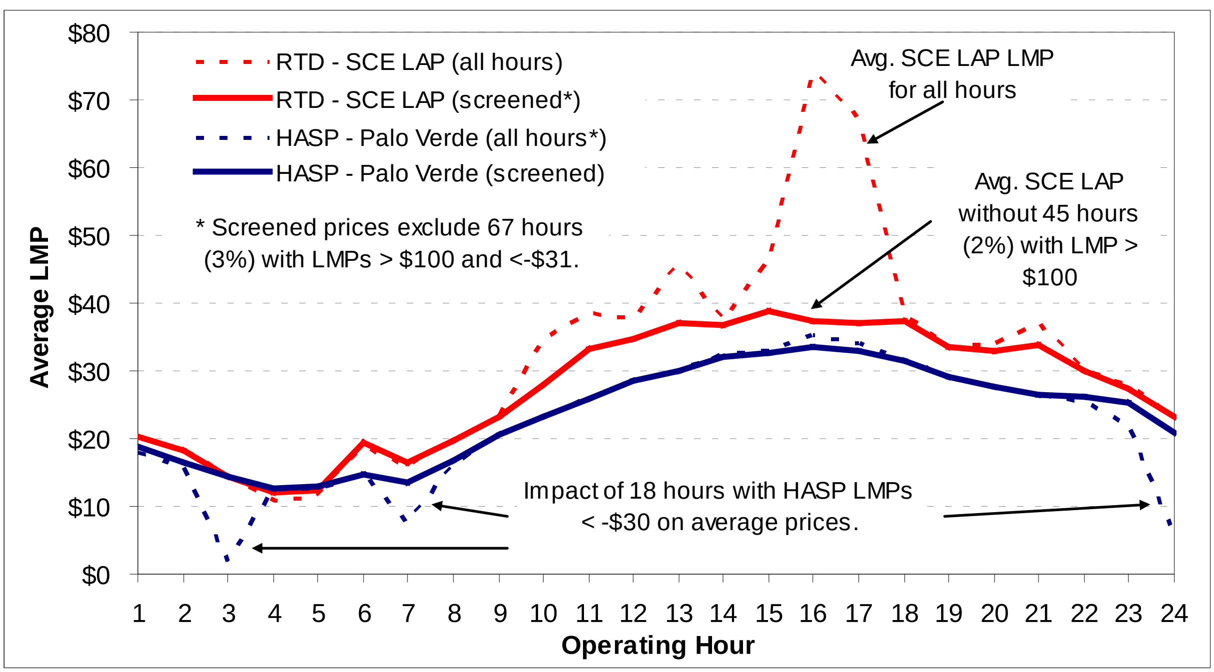
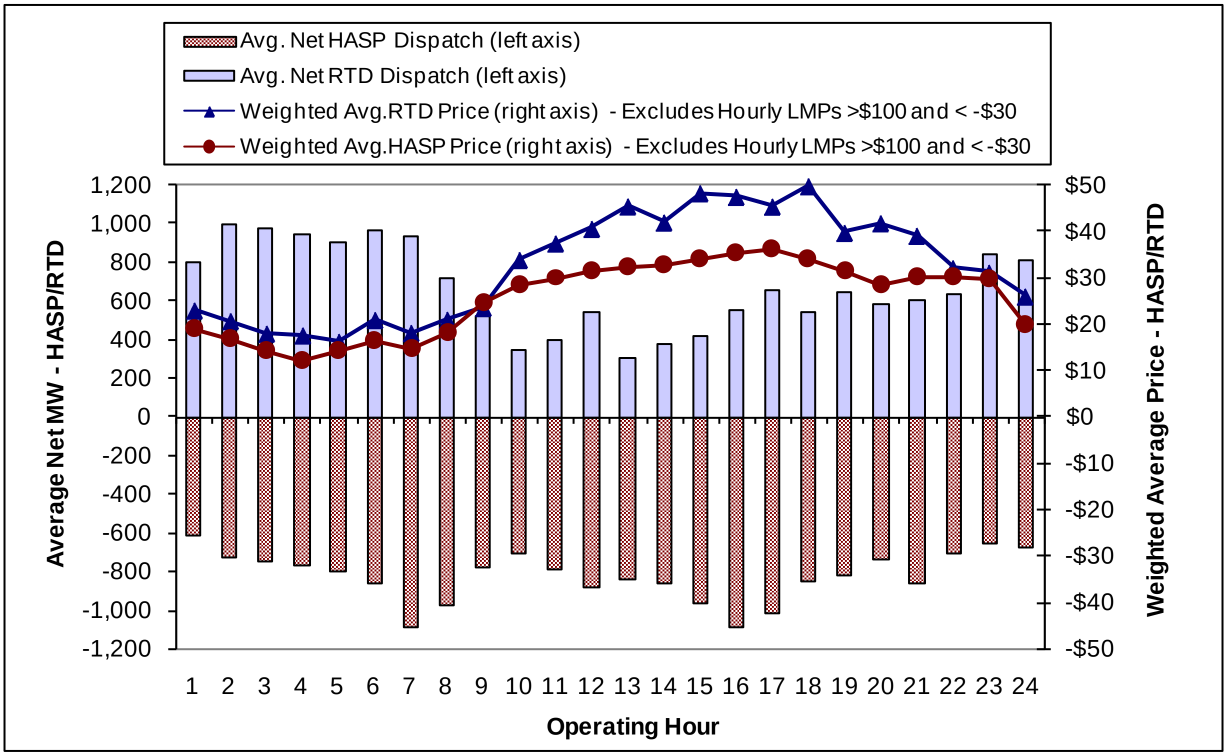


Figure 1.9 Average Net Energy Dispatches and Prices in HASP and RTD



The pattern of "selling low" in HASP and "buying high" in RTD has continued to create substantial revenue imbalances that are recovered from market participants with metered load through Real Time Energy Imbalance Energy Offset charges (which are included in CC 6477). Figure 1.10 shows total CC 6477 charges by month over the first six months of the ISO's new market, broken out by hours in which relatively high price RTD spikes occurred (i.e., hourly prices in the SCE or PG&E LAP >$100) or negative HASP prices (<$0 on Malin or Palo Verde). As shown in Figure 1.10, these charges have averaged over $10 million per month over this six month period, and continued to exceed $10 million in August and September. These charges are presented in percent of hours in Figure 1.11 to give some perspective as to how frequently hours associated with high charges are occurring. In this figure we see that:

- About 40 percent of CC 6477 charges have been incurred in only 7 percent of hours in which relatively high price RTD spikes occurred (i.e., hourly prices in SCE or PG&E LAP >$100).

- About 28 percent of CC 6477 charges have been incurred in only 4 percent of hours in which negative prices have occurred in HASP (on Malin or Palo Verde).

- The remaining 32 percent of CC 6477 charges have been incurred in the other 89 percent of hours when such extremely high RTD or low HASP prices did not occur.

The ISO has estimated that during hours in April when the bulk of CC 6477 charges have been incurred, about 70 percent of these charges are attributable to extremely high RTD prices or low HASP prices, while about 20 to 30 percent may be attributable to the fact that uninstructed deviations by load are charged based on average hourly RTD prices, while generation is settled on 5-minute prices.[9] Analysis by DMM indicates that during all hours of Q3, at least half of CC 6477 may be attributed to the pattern of "selling low" in HASP and "buying high" in RTD.

Figure 1.10 Total Imbalance Energy Offset Charges (CC 6477)

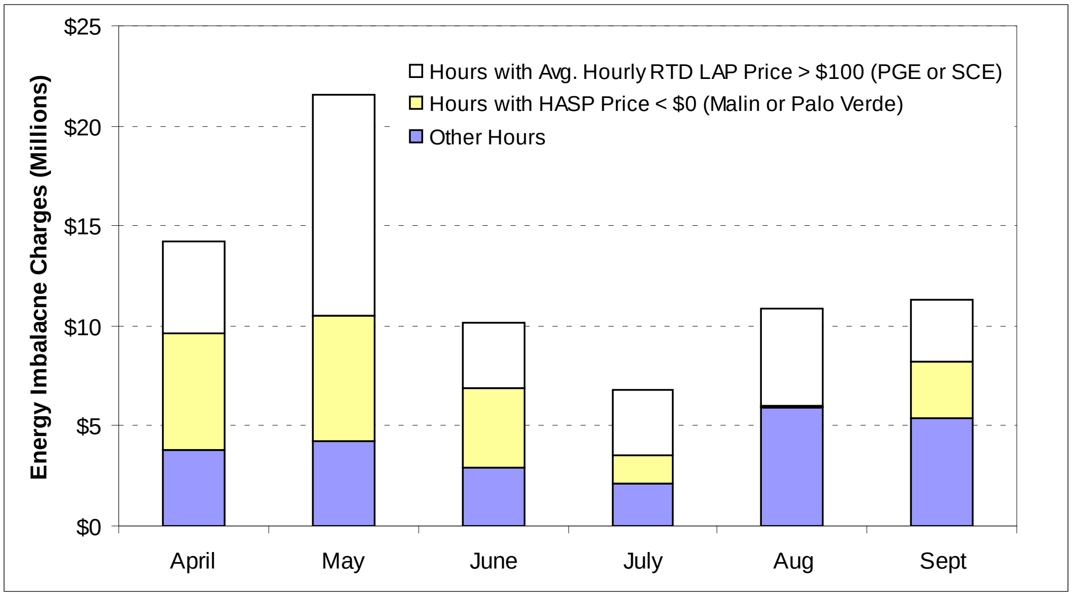
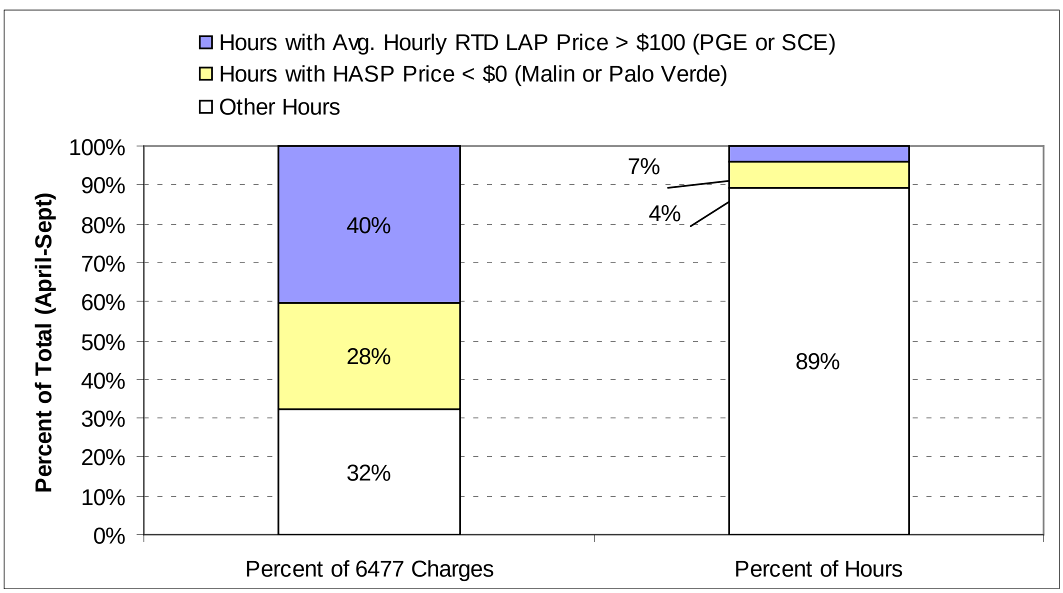


Figure 1.11 Imbalance Energy Offset Charges (CC 6477) – Percentage of Charges April – September, 2009



[9] See page 6 of *Issue Paper: Analysis of Real-time Imbalance Energy Offset (CC 6477)*, revised August 26, 2009, http://www.caiso.com/2416/2416e7a84a9b0.pdf.

1.3.2 RTD Price Spikes

Factors that have caused many of the extremely high RTD prices contributing to CC 6477 charges include the following:

- Limitations of upward ramping capacity needed to meet increases in demand and changes in inter-tie schedules between hours.

- The sudden loss of major generating units or transmission lines after execution of HASP.

- Biasing down of the flow limits on major internal paths in RTD by operators after execution of HASP (or to a lower level than the bias used in the HASP market model). A more detailed discussion of how this issue has been addressed in Q3, by making any bias placed on transmission constraints more consistent between the HASP and RTD, is provided in Chapter 5 of this report.

- Relatively large, sudden increases in the RTD forecast (including manual biasing of the RTD forecast up) in response to observed or anticipated system conditions in real-time. For example, this can occur when the ISO switches the basis for its RTD load forecast from one “similar day” to a different typical day during the course of a day. Such adjustments may provide a better overall load forecast, but can cause sudden price spikes as the system responds to the new forecast. A discussion of steps the ISO is taking to address this issue is provided in Section 1.3.5of this chapter.

Although these factors may have a significant impact on prices during a relatively small number of hours, RTD prices during these hours can be extremely high due to the impact of constraint violations or relatively extreme re-dispatch solutions made in RTD to avoid such violations. While the ISO has implemented a number of measures that have reduced the frequency and magnitude of extreme price spikes in RTD since the start of the ISO’s new market design, significant price spikes continue to occur in RTD during a relatively small percentage of hours.

1.3.3 Extremely Low HASP Prices

Factors that have caused many of the extremely low HASP prices contributing to CC 6477 charges include the following:

- During some hours, excessive self-scheduling and congestion on inter-ties have driven HASP prices to very low (negative) levels.

- In some cases, if operators anticipate over-generation conditions in RTD, they may bias the HASP load down by several hundred megawatts to increase exports (or reduce imports) in HASP.

- During some off-peak hours, over-generation has occurred due to a relatively high level of generation being scheduled in the IFM, combined with uninstructed generation and additional energy from units committed after the IFM through RUC or exceptional dispatch. In over-generation scenarios, HASP prices can go extremely low (much lower than the RTD price) due to mathematical modeling issues. Specifically, the use of lossless shift factors, combined with additional constraints modeled in the HASP optimization such as ancillary service requirements and hourly “block” inter-tie schedules, can cause HASP prices well below the -$30 bid floor. While RTD prices in these periods may be low, they are generally not far below the -$30 bid floor due to differences in actual system conditions and the RTD optimization model.

1.3.4 Other Factors Contributing to Systematic Price Differences

As shown in Figure 1.7 and Figure 1.8, while extremely high RTD and extremely low HASP prices account for a significant portion of the difference in RTD and HASP prices, the HASP prices still tend to be systematically lower during other hours as well. Other reasons that may contribute to this more systematic price difference include the following:

- Systematic under-forecasting of load in HASP during hours when loads are increasing, and over-forecasting of load in hours when loads are decreasing.
- Negative uninstructed generation in RTD.
- Modeling of all inter-tie schedules and bids as fixed hourly blocks, without the 20-minute ramping period actually applied in RTD.
- Use of a 15-minute optimization in HASP versus a 5-minute optimization in RTD.

The first two factors noted above can be quantitatively assessed based on historical data, as discussed below:

- ***Systematic Forecasting Errors in HASP*.** The HASP forecast for each operating hour must actually be developed 75 minutes prior to the start of each operating hour, such that the HASP forecast actually represents a forecast of load over one to two hours in the future. Shortly after the start of the ISO's new market, inaccuracies in the load forecasting tool (STLP) embedded in the Siemens HASP and RTD software led the ISO to revert to use of the ISO's previous hour-ahead load forecasting tool (ALFS). While this modification significantly improved the HASP and RTD load forecast, Figure 1.12 illustrates the systematic difference that continues to result from the current method used to forecast loads in the HASP and RTM. As shown in Figure 1.12, the HASP load forecast tends to *underestimate* the RTM forecast during hours that load is *increasing* (HE 6-16), and tends to *overestimate* the RTM forecast during hours that load is *decreasing* (HE 1-5, and 17-24). The systematic *underestimation* of loads during HE 6-16 would tend to *decrease* HASP prices relative to RTM prices during these hours. During other – primarily off-peak – hours, the systematic *overestimation* of loads in HASP may tend to increase HASP prices relative to RTM prices. However, during these off-peak hours, there is often a significant supply of low cost energy that reduces the impact of any *over-forecasting* in HASP in terms of *depressing* the HASP price relative to the RTM price. Thus, the net effect of the systematic forecasting differences in HASP and RTM is likely to be an overall increase in RTM prices relative to HASP prices.

- ***Negative Uninstructed Generation in RTD*.** Another factor that may contribute to HASP prices that are systematically lower than RTD prices is under-generation (i.e., negative uninstructed deviations). Analysis by DMM indicates there is a slight trend toward negative uninstructed energy on a system-wide level. While periodic major unit outages contribute to this trend and can cause significant price spikes in the RTM, the cumulative impact of smaller divergences from schedules and dispatch instructions by multiple generating units may also contribute to divergences in HASP and RTD prices. These deviations likely increase RTD prices relative to HASP prices, since the HASP market software "assumes" that all units respond to the 15-minute advisory dispatch instructions developed in the HASP optimization, while in RTD generation may not respond to dispatch instructions as perfectly as is assumed in the HASP optimization As shown in Figure 1.13, under-generation from schedules and dispatch instructions in RTD average about 100 to 200 MW during peak hours, especially during hours when the average divergence between HASP and RTD has been the greatest in Q3.

Figure 1.13 shows the correlation that appears to exist between (1) underestimation of loads in HASP plus under-generation by resources in the RTM, and (2) the difference in HASP and RTM prices during Q3. The prices in Figure 1.13 exclude the relatively high and low hourly prices excluded in Figure 1.7 and Figure 1.8 (> $100 and < -$30). Thus, data in Figure 1.13 suggest that the combination of underestimation of loads in HASP plus under-generation by resources in the RTM seems to be a major driver of the systematic difference in HASP and RTM prices during Q3, particularly during peak hours (HE 6-16). As discussed below, DMM believes that other factors – including some significant differences between the HASP and RTM market software – may be the root cause of much of the difference between HASP and RTM prices.

Two other factors that have been identified as significant potential root causes of the trend of "selling low" in HASP and "buying high" in RTD involve modeling simplifications or differences made in the current HASP optimization process relative to the more detailed RTD optimization model:

- ***Hourly Inter-tie Schedules in HASP*****.** In the HASP optimization process, all inter-tie schedules and bids are represented as hourly "block" schedules (i.e., without any ramping between operating hours). In the actual RTD, however, the net change in inter-tie schedules resulting from the HASP must be ramped in during the period 10 minutes before and 10 minutes after the start of each operating hour. This modeling simplification in HASP creates a systematic discrepancy between the HASP optimization used to determine which HASP import/export bids are accepted, and the actual RTD market. Specifically, this simplification causes HASP to "underestimate" the actual ramping that will be needed in the RTD during this 20-minute ramping period.

- ***15-minute HASP Optimization versus 5-minute Optimization in RTD*****.** The HASP optimization process models the real-time market based on a 15-minute optimization period (i.e., the HASP optimizes over a two hour forward looking period in eight 15-minute intervals). However, the actual RTD is optimized every five minutes. This modeling simplification in HASP creates another systematic discrepancy between HASP and RTD. Specifically, the use of a 15-minute optimization in HASP causes HASP to "overestimate" the actual ramping capability that will be available on a 5-minute basis in RTD. For example, the use of 15-minute optimization intervals in HASP enables the HASP optimization to dispatch generation within the ISO over a 30-minute period to meet changes in the hourly inter-ties (starting 15 minutes prior to each hour and ending 15 minutes after each hour), while in practice this change must be met over a 20-minute period. Thus, HASP effectively underestimates the need for ramping in RTD (or overestimates ramping capability in RTD) to meet changes in hourly block schedules by a factor of 50 percent.

The combined effect of these two modeling discrepancies is likely to cause the HASP optimization to tend to systematically under-estimate actual RTD prices that will result from

HASP schedules, and therefore tends to export additional energy on the inter-ties (or reduce imports) in HASP at a price that tends to be systematically lower than RTD.

Figure 1.12 Comparison of HASP and RTD Forecast

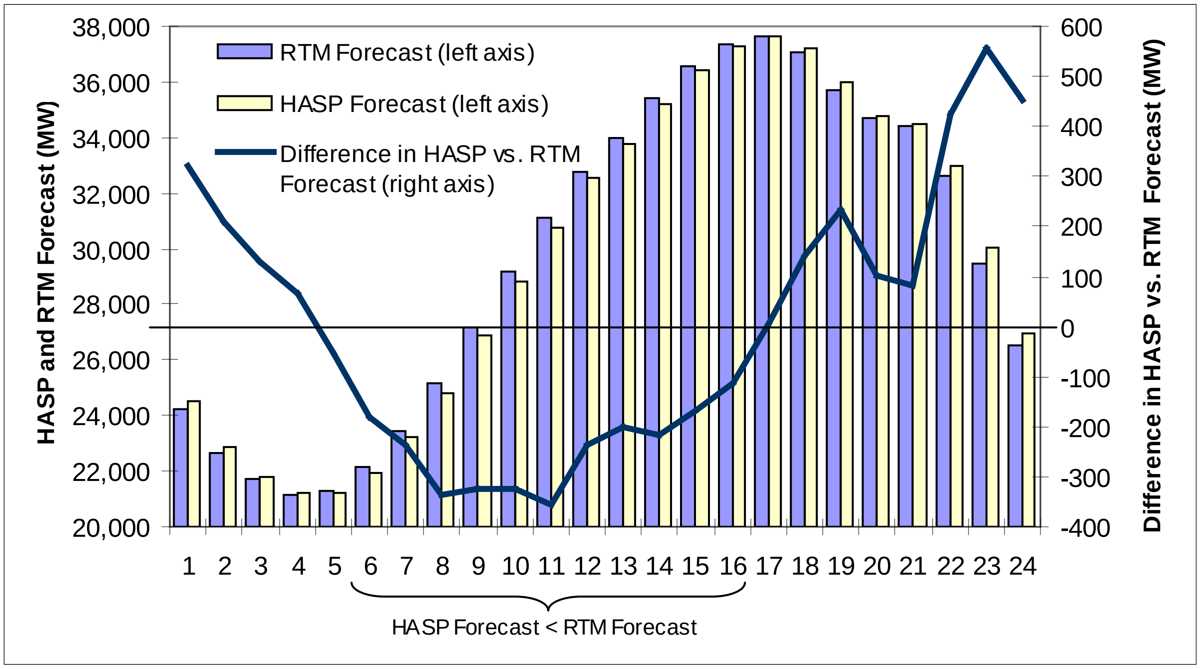
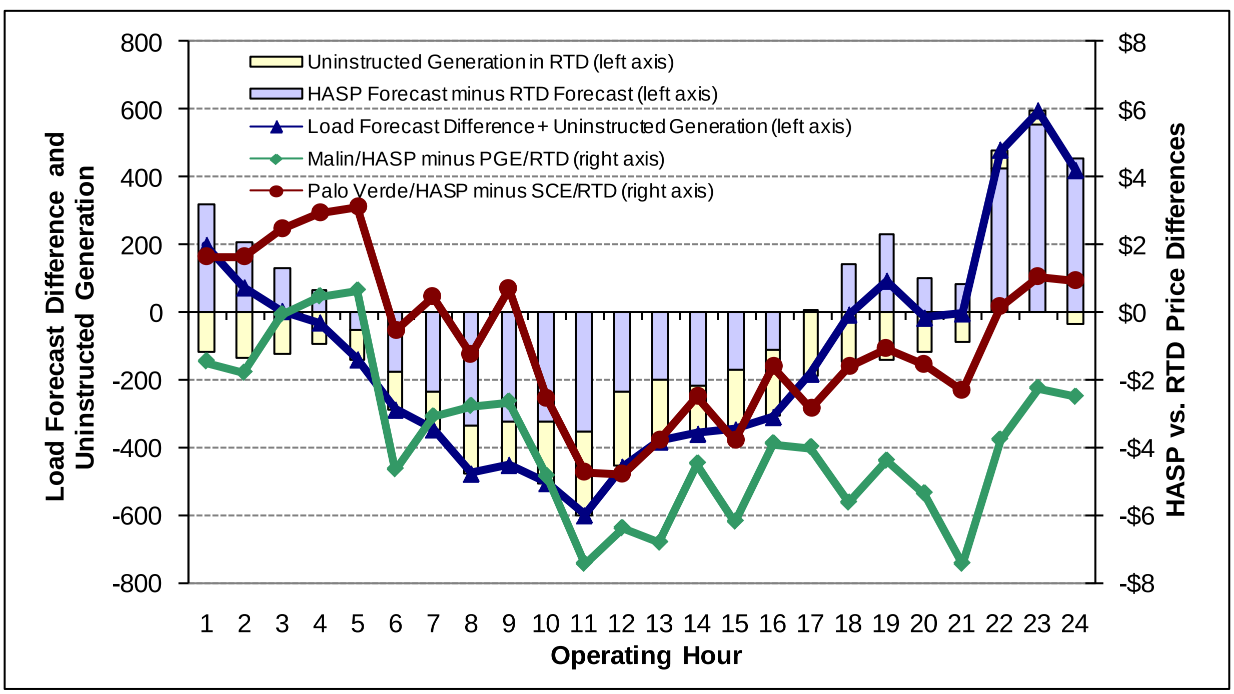


Figure 1.13 Comparison of HASP and RTD Forecast



1.3.5 Actions to Mitigate Root Causes of Systematic Price Divergence

The ISO is currently taking several steps to mitigate some of the key root causes of the more systematic divergence in HASP and RTM prices:

- The ISO currently has a new short-term forecasting tool under development that is designed to provide a more accurate and consistent forecast for both HASP and RTM. In addition, this new forecast will specifically be designed to provide forecasts at the 15-minute and 5-minute level of granularity over the approximately two hour forecasting timeline needed for the HASP and RTM.[10] Implementation of this new forecasting tool is anticipated in early 2010.

- In Q3, the ISO assessed a variety of options that might mitigate the impacts of the differences in ways that inter-tie schedules and ramping of resources are modeled in HASP compared to RTD. As an initial step, the ISO is developing enhancements that would modify HASP to account for the imbalance energy difference that arises due to the fact that HASP does not model how changes in net hourly inter-tie schedules are ramped in over a 20-minute period each operating hour.

- The ISO has also developed a more systematic procedure to perform biasing of load and transmission constraints on branch groups. The issue of biasing transmission constraints is discussed in Chapter 5 of this report.

DMM believes these steps may significantly increase convergence of HASP and RTM prices. However, a continued effort should be made to identify other ways in which the pattern of "selling low" in HASP and "buying high" in RTD can be addressed. For example, DMM believes that convergence might be improved by other adjustments to account for the 15-minute optimization and assumptions of "perfect response" to dispatch instructions used in HASP, compared to actual resource availability and performance under the 5-minute optimization used in RTD.

[10] The ALFS forecasting tool currently being used actually produces a 30-minute forecast, so that the more granular 15- and 5-minute forecasts needed for the HASP and RTM software are developed by interpolating from this 30-minute forecast.

1.4 Price Volatility

This section provides a brief summary of some observed trends in price volatility in the HASP and real-time markets. One notable trend is a tightening of price dispersion, particularly in the off-peak periods. In Q2, heavy decremental dispatch in the off-peak, typically at low or even large negative prices, was a persistent feature of the HASP and real-time markets; this was a relatively rare phenomenon in Q3.Figure 1.14 reflects the distribution of HASP LAP prices for the SCE area for April through September. In peak hours in Q3, HASP LAP prices have shown a fairly tight distribution, with the middle 90 percent of prices (orange vertical lines) ranging from approximately $20/MWh to approximately $45/MWh in July and August, and between $20/MWh and just below $60/MWh in September. July and August 2009 were unseasonably mild, with only 974 intervals in the two months with load in excess of 40,000 MW. September 2009 was more typical of a summer month in terms of temperature and load, and there were 576 intervals with load above 40,000 MW, compared to 306 such intervals in September 2008. Prices in the off-peak hours have had a distribution of similar shape, but with a range of approximately $0 to $40/MWh.

Figure 1.15 and Figure 1.16 show similar graphs of HASP price distributions in the SDG&E and PG&E LAPs, respectively.

Figure 1.14 SCE HASP LAP Price Distributions (April – September)

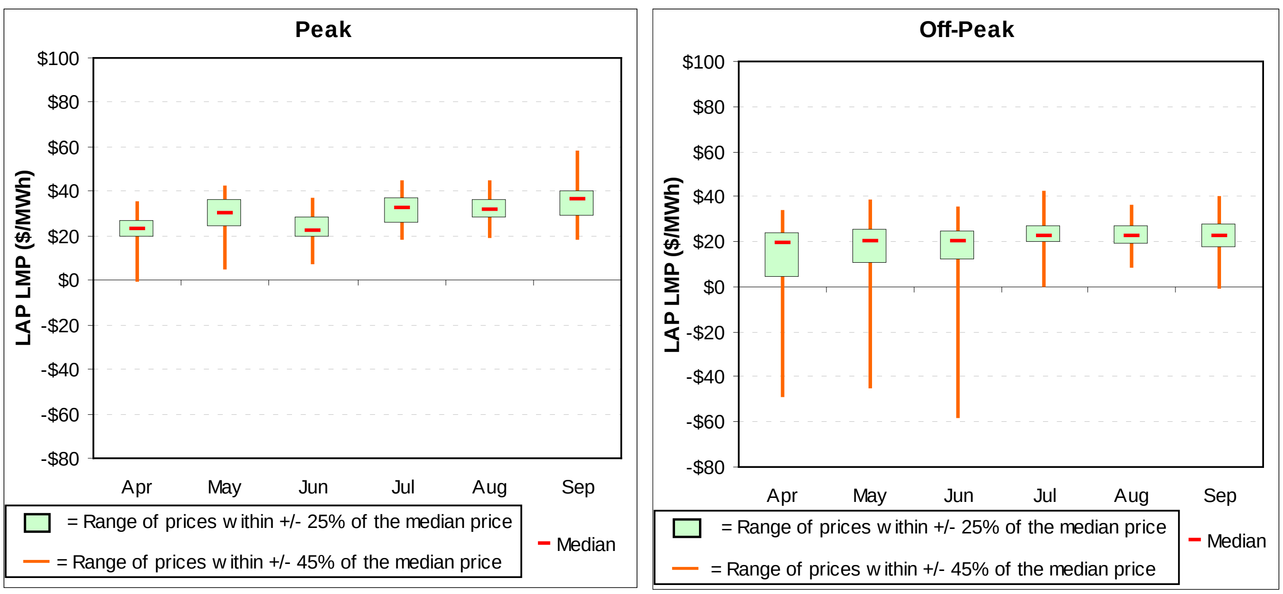


Figure 1.15 PG&E HASP LAP Price Distributions (April – September)

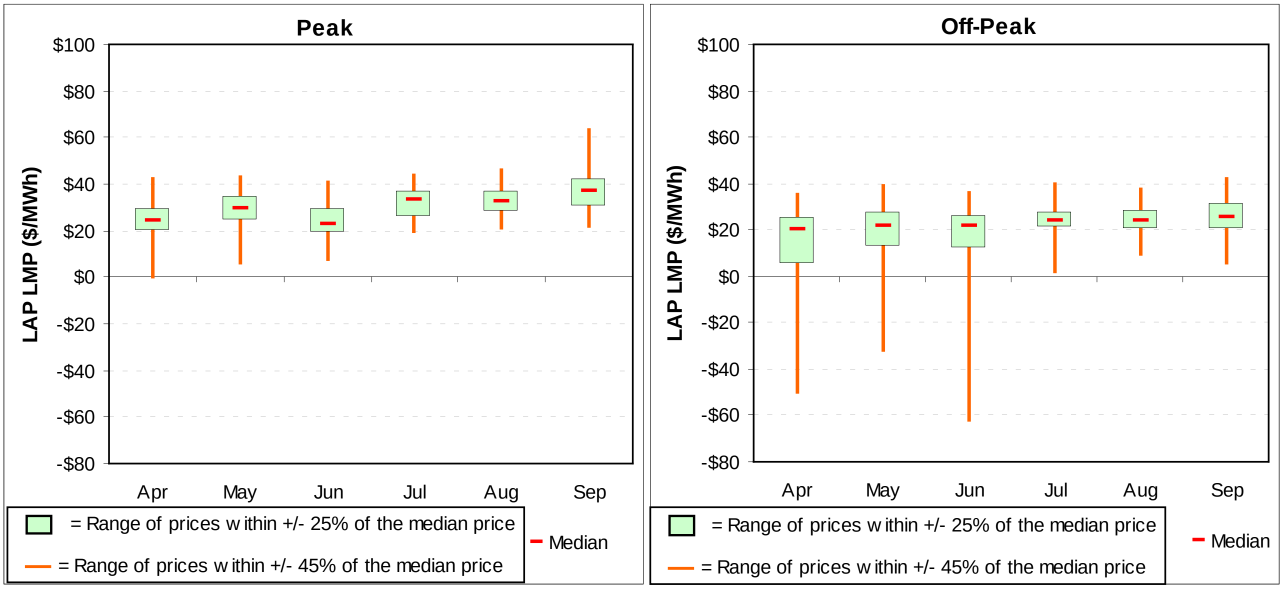


Figure 1.16 SDG&E HASP LAP Price Distributions (April – September)

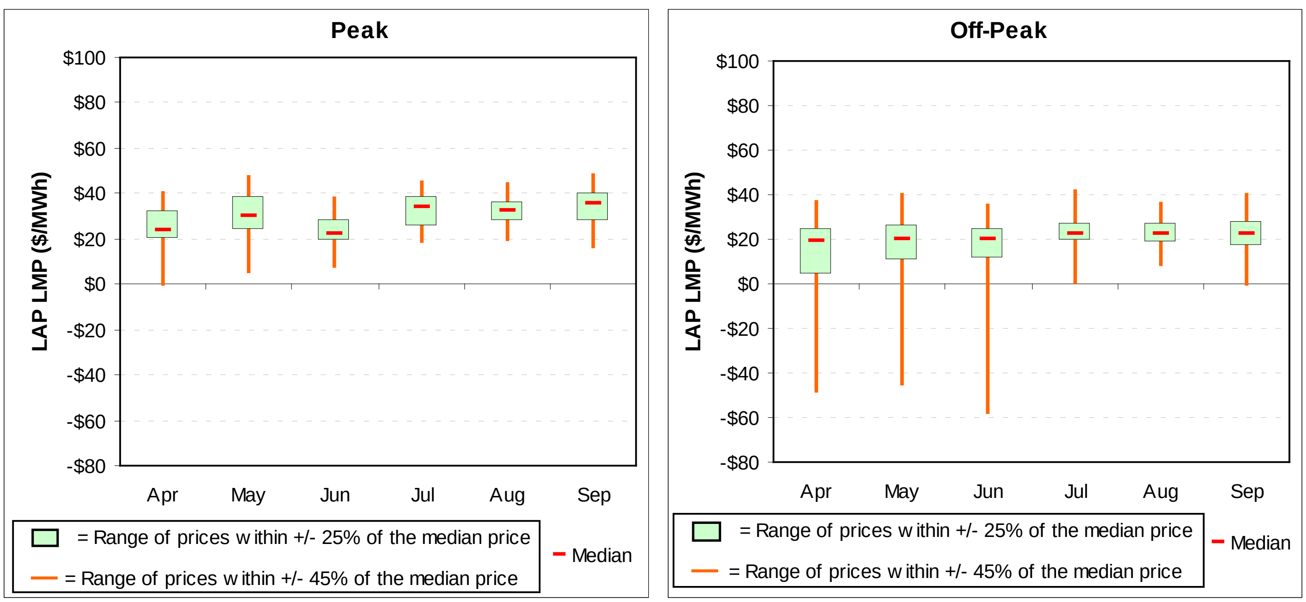


Figure 1.17 below provides a box-whisker plot representation of the distribution of SCE RTD LAP prices for April through September. Again, the trend of the lowest off-peak LAP prices' tail decreasing persists in the real-time market, although the long tail persisted through July in the real-time market. All three regions show a cluster of low prices at or around the bid floor of -$30/MWh, plus or minus loss factors, as more severe over-scheduling persisted in early morning hours through July. This has been an ongoing issue for several years (pre-dating the new market) due to units committed for capacity requirements and operating at minimum load. Reducing the amount of unit commitment that is performed outside of and after the IFM may help reduce the occurrence of this. Peak prices were considerably less volatile, with 90 percent of intervals within a $25 range in July and August, and within a $35 range in September.

Figure 1.18 and Figure 1.19 show similar graphs of RTD price distributions in the SDG&E and PG&E LAPs, respectively. The RTD LAP price distributions for PG&E are similar to those for the southern LAPs, as seen in Figure 1.19 below. One notable difference can be seen when comparing the peak-hour price distributions in September. Path 15 was derated by 1500 to 2500 MW for much of September 8 to 14 for a planned outage of the Los Baños-Midway #2 500kV line, resulting in higher prices in the PG&E area for that week.

Figure 1.17 SCE RTD LAP Price Distributions (April – September)

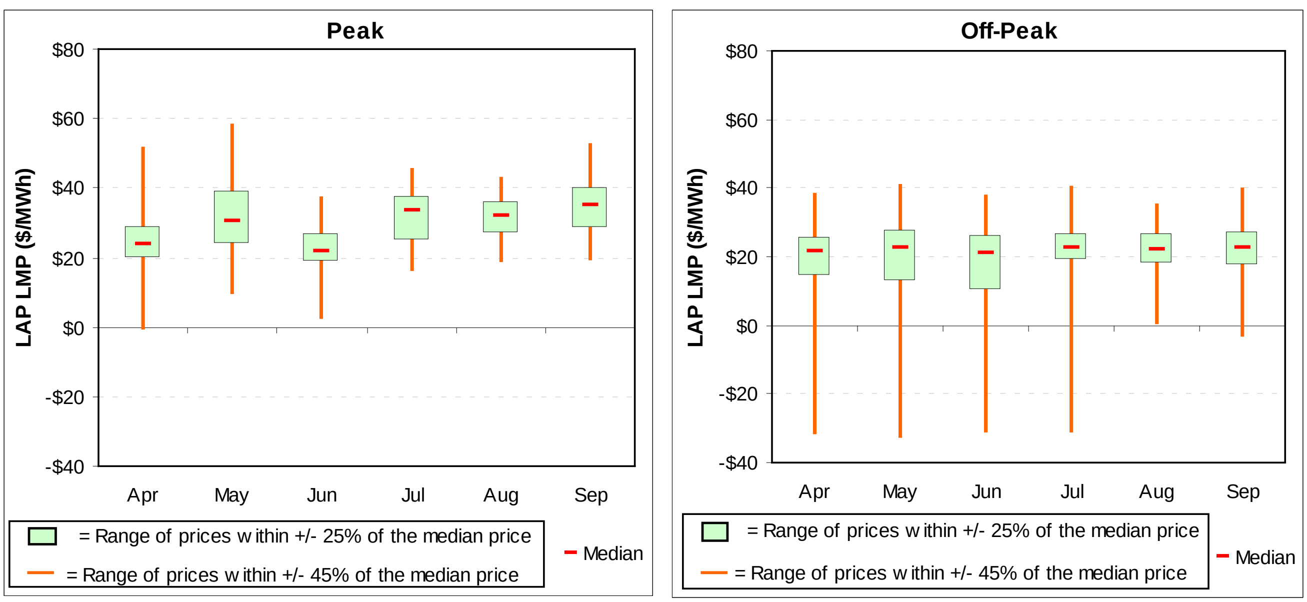


Figure 1.18 SDG&E RTD LAP Price Distributions (April – September)

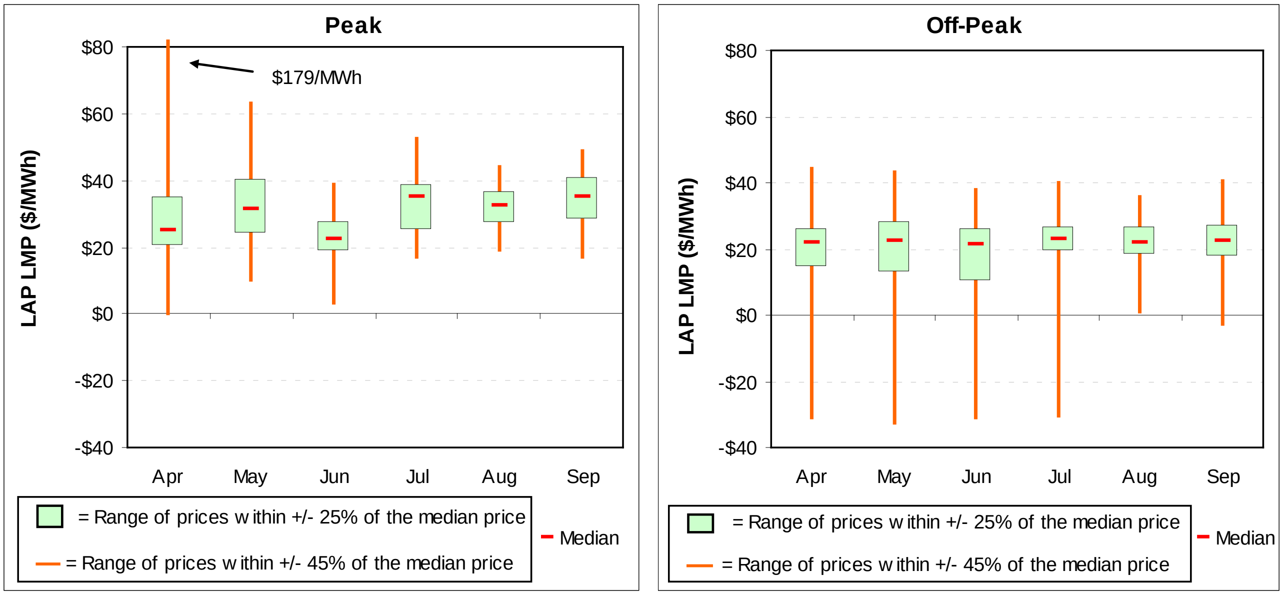


Figure 1.19 PG&E RTD LAP Price Distributions (April – September)

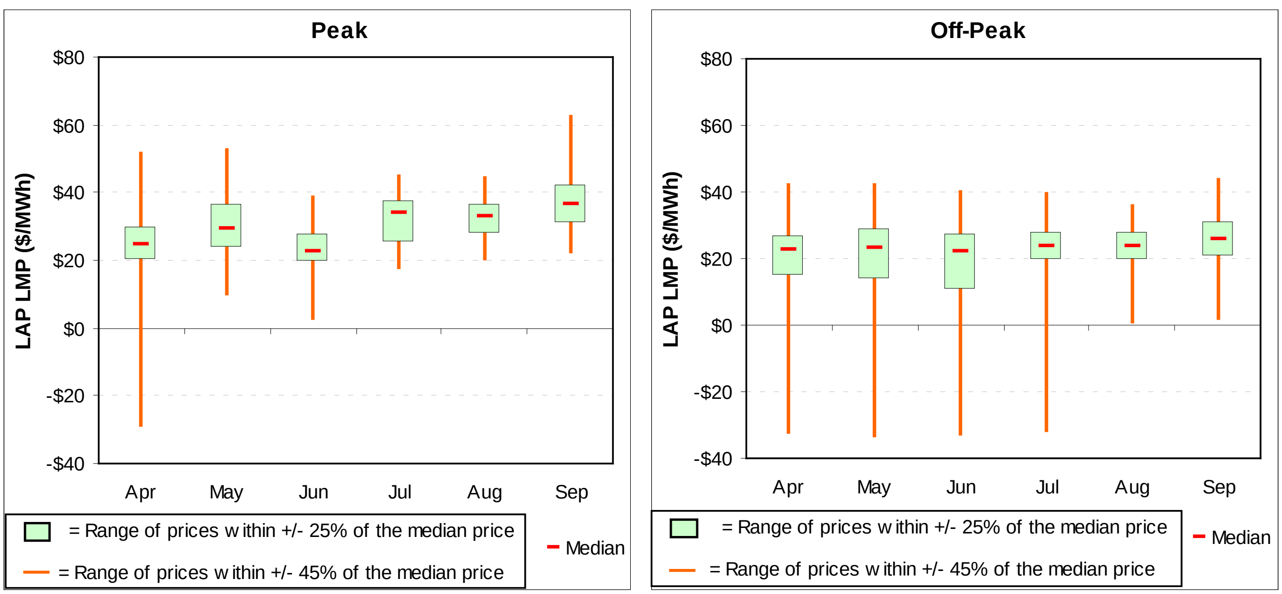
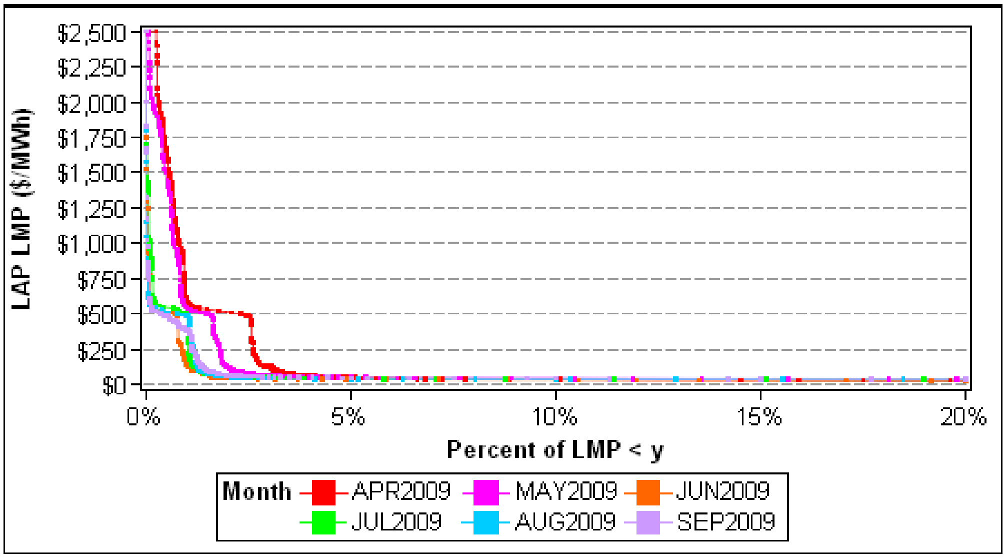


Figure 1.20 shows the top 20 percentiles of LAPs by month as duration curves. Price spikes in excess of $250/MWh in April and May 2009 represented 2 to 3 percent of real-time intervals overall, with most such spikes near or exceeding $500/MWh. Positive spike frequency then retreated in June to approximately 1.5 percent of intervals. In the third quarter, spike frequency was similar to that in June, with all months clustered in the neighborhood of 1.5 percent of all intervals.

Figure 1.20 LAP Duration Curves by Month: Top 20 Percentiles



Price spikes in the negative direction decreased in frequency during Q3. Real-time LMPs below the bid floor of -$30/MWh exceeded five percent of intervals in each month in Q2, and in no month in Q3. July saw approximately 4.5 percent of intervals at or below the -$30 floor, whereas the figure was below 2 percent for both August and September. The challenges of September, including congestion resulting from the Station Fire and the SWPL outage, both contributed to an increase in negative spike frequency over August.

Figure 1.21 LAP Duration Curves by Month: Bottom 20 Percentiles

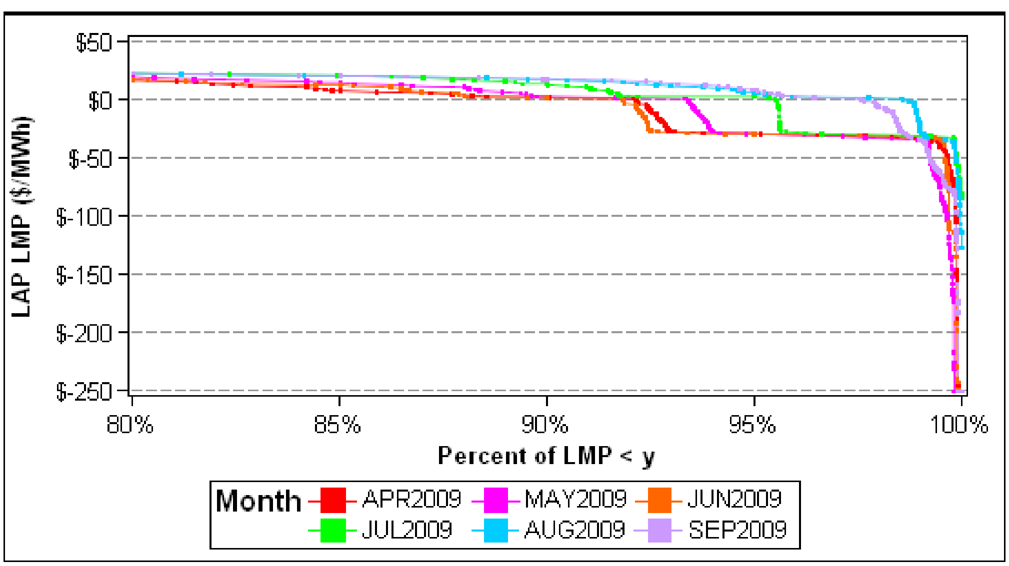


Figure 1.22 shows the daily frequency of high prices, by price level, for RTD LAP LMPs. Fewer extreme prices (>= $1,000/MWh in dark blue) occurred in late Q2, with a series in mid-July, followed by a period of fewer spikes until they increased again in late August through September, a challenging period on the grid. On July 14, several unit trips and an EMS failure resulted in a prolonged price spike near $500/MWh and a few intervals with higher prices. On July 19, several unit trips and congested import transmission combined to cause a series of price spikes. The Station Fire and consequent transmission damage resulted in many intervals near $500/MWh during the weeks beginning August 23 through September 6. The Southwest Power Link (SWPL) and Path 15 outages both contributed to price spikes for much of mid-September, often near or in excess of the $500/MWh price cap. On September 18, an inaccurate load forecast resulted in day-ahead scheduled energy 3000 MW short of load at the peak. Meanwhile, the SWPL outage and transmission damage from the Station Fire all combined to cause intermittent price spikes over a five-hour period.

Figure 1.22 RTD Positive LAP Price Spike Frequency

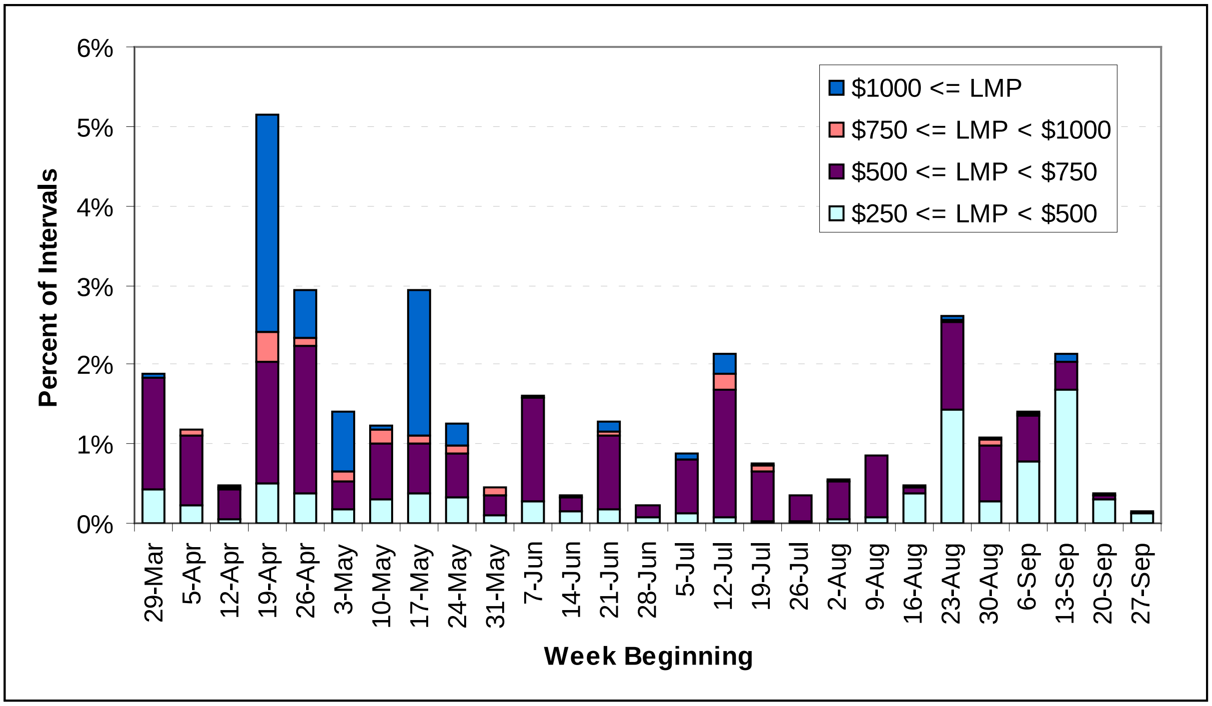
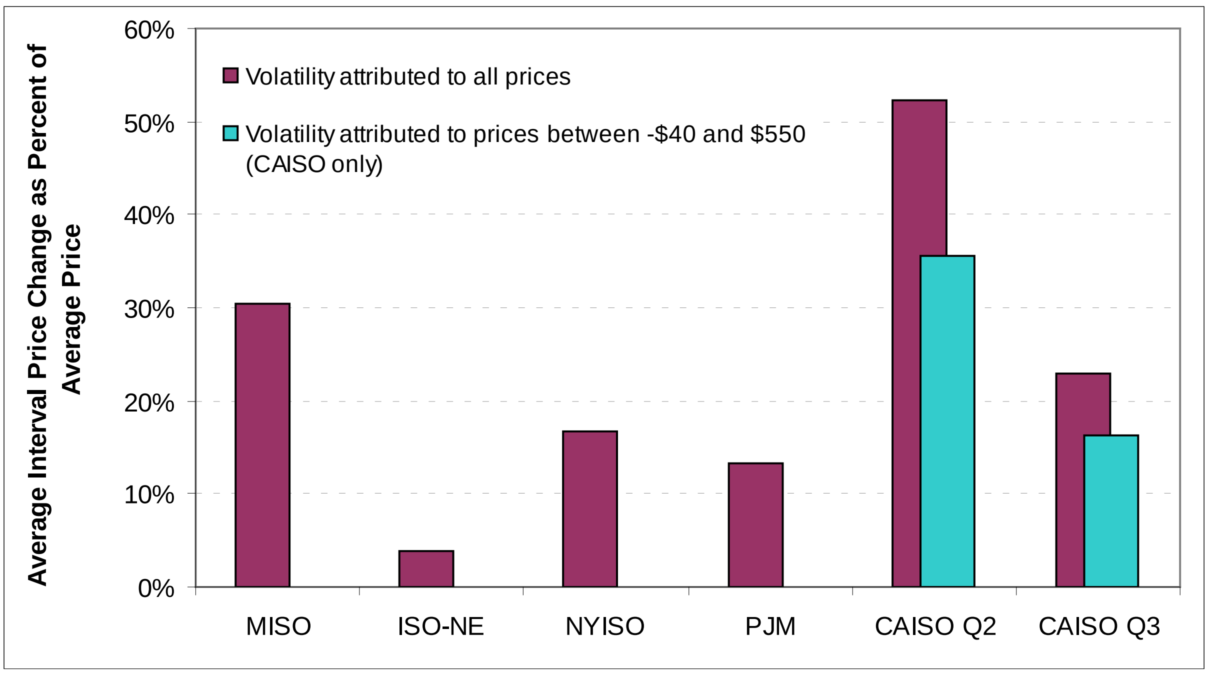


Figure 1.23 provides a different perspective on price volatility by showing the extent to which prices change from one 5-minute interval to the next. It provides an indication that after an initial period of wild price swings, price volatility is now in a range similar to that of other ISOs. This metric is a calculation of the average interval price change (in absolute value) expressed as a percentage of the average price. We calculate this metric by taking the arithmetic average of the three Default LAP prices (SCE, SDG&E, and PG&E) across all intervals in each quarter, and comparing it to the same metric for other ISOs with nodal pricing for all of 2007.[11] The volatility metric for other ISOs ranges from roughly 5 percent (ISO New England) to 30 percent (Midwest ISO). The volatility metric for the California ISO is divided into two contributing factors. The blue portion of the California ISO bars denotes the contribution to volatility excluding extreme or outlier prices; that is, it includes only prices within the range of -$40/MWh to $550/MWh.[12] The maroon portion includes the entire set of prices, and thus is more comparable to the metric used for the other ISOs. The non-outlier contribution is roughly 36 of the total 52 percent for the volatility metric in Q2, and 16 of the 23 percent in Q3.

Figure 1.23 Real Time LAP Price Volatility across ISOs



[11] The data shown for other ISOs are from the 2007 State of the Market Report for the Midwest ISO, prepared by Potomac Economics. The metrics for the other ISOs are calculated using several hub prices for each ISO – see Figure 35, page 46 of the report.

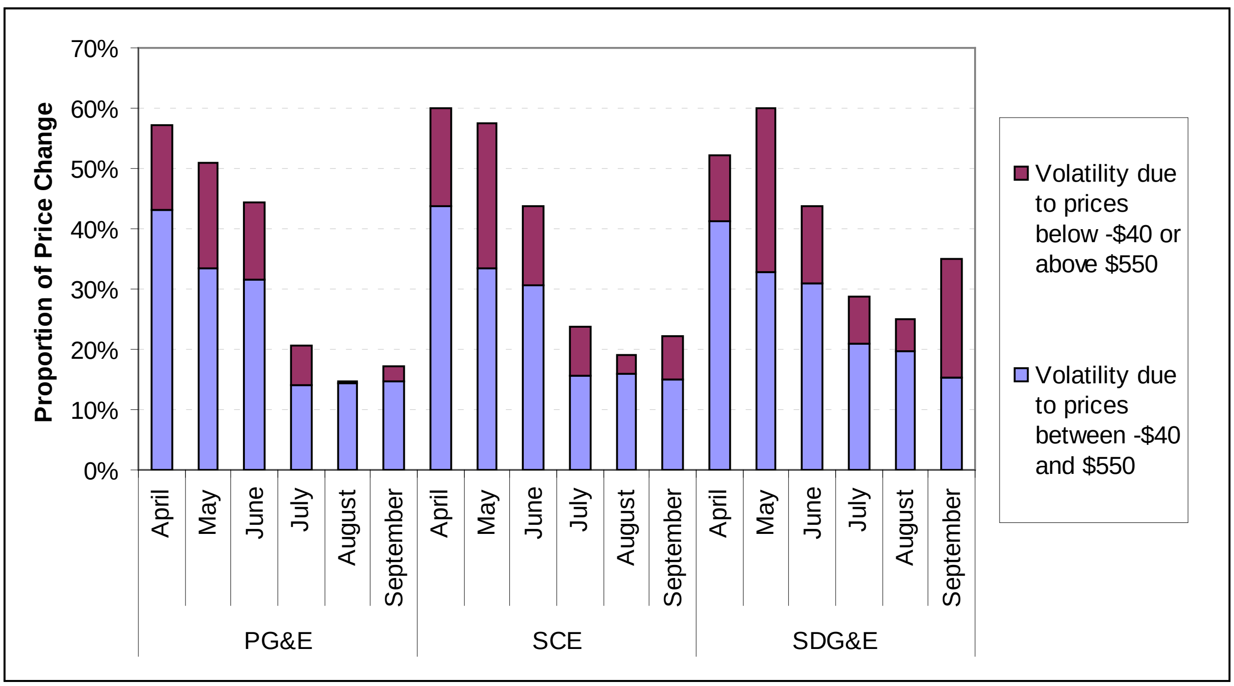
(http://www.potomaceconomics.com/uploads/midwest_presentations/2007_State_of_the_Market_Report-Full_Text_07-08.pdf).

[12] These values were chosen to reflect the current minimum and maximum bid limits of -$30 and $500 plus some additional margin to account for losses on the LMP.

As evident in Figure 1.23, the interval-to-interval volatility of the California ISO 5-minute prices for the first three months of the new market is substantially greater than what has been observed in ISOs with mature LMP markets; however, volatility in Q3 was comparable to that seen in other ISOs. Differences in observed interval-to-interval price volatility across various ISOs are likely due to important differences in particular aspects of each ISO's real-time market design and optimization features as well as differences in market fundamentals and characteristics such as dependency on inter-tie schedules, daily load profiles, and internal resource mix. In light of these factors, we do not necessarily view the comparison of price volatility shown in Figure 1.23 as a simple "less is better" exercise – note that an extended period of $500/MWh prices itself is not volatile by this measure. Rather, it should be used as a basis to determine which aspects of the California ISO real-time market design contribute to price volatility, and to assess whether these features are desirable or require modification.[13]

The same volatility metric is presented in Figure 1.24 below for the three LAP areas separately, by month. The overall trend from April through September is a decrease in volatility between Q2 and Q3 across all LAPs. The exception was an increase in extreme price volatility in September in the SDG&E LAP, due largely to approximately four extraordinarily volatile days during periods of fire-related and other outages.

Figure 1.24 Monthly Average ISO Real Time Price Volatility



[13] For example, different ISOs have different means for utilizing energy from regulation reserves to manage periodic shortages of ramping energy. Many of the price spikes occurring in the ISO's 5-minute dispatch market (RTD) are due to shortages of ramping energy; therefore, comparing how RTD utilizes energy from regulation reserves to practices in other ISOs might reveal opportunities for market enhancements that could appropriately reduce price volatility. In other cases, there may be differences in the California ISO real-time market design that produce greater price volatility but are desirable.

In summary, this analysis shows that price volatility in the California ISO's real-time market was considerably lower in Q3 relative to Q2, and is trending (overall) further in the direction of being comparable to volatility in other ISOs. Specifically, HASP and RTD LAP prices showed tighter price distribution with fewer extreme prices, particularly in the off-peak and in regard to negative prices. Price volatility in the RTD market, from one 5-minute interval to the next, remains higher than what is observed in most other ISOs. Comparing volatility across all prices to volatility omitting extreme prices, it is clear the significant contribution to this measure made by a small proportion of intervals with high prices. Otherwise, prices in the RTD market during Q3 exhibited a volatility that is highly comparable to that in the other ISOs.

2 Market Competitiveness and Mitigation

This chapter provides an assessment of the overall competitiveness of the ISO's Integrated Forward Market (IFM) and real-time market (RTM), and provides analysis of several key provisions of the local market power mitigation (LMPM) provisions included in the ISO's new market design. Key finding of this chapter include the following:

- Prices in the ISO's IFM during each month of Q3 continued to be approximately equal to prices we estimate would result under perfectly competitive conditions, based on competitive benchmark prices DMM develops by re-simulating the IFM with bids reflecting each unit's actual marginal cost.

- Price spikes in the RTM dropped significantly in Q3 relative to the first three months of the ISO's new market, so that average RTM prices in Q3 have also converged to be approximately equal to the competitive benchmark prices developed by DMM.

- Starting in July, resources have had the option to have the Default Energy Bids (DEBs) used in LMPM based on an LMP-based option. While a significant number of resources initially selected this LMP-based option, no resources remained under this DEB option by the end of Q3. This trend can be attributed to the fact that DEBs that resulted under this option were relatively low due to the low LMPs during many hours these resources were in operation.

- Starting in August, all Exceptional Dispatches (EDs) for energy were only subject to mitigation if made to relieve congestion for *non-competitive* constraints or for a limited category of seasonal environmental constraints known as "Delta Dispatch". All other EDs were paid their unmitigated bid price. However, DMM has found that the incremental cost impact of this more limited mitigation of EDs has been relatively low due to (a) the relatively small volume of ED for energy made since this change took effect, and (b) the fact that bid prices paid for most ED energy have not been significantly higher than the market LMPs or DEBs for these resources.

- During Q3, the frequency of failures in the pre-RTM LMPM process has been relatively low, and has trended downward. Review by DMM and the ISO's price correction team indicates that the price impact of failures in the pre-RTM LMPM procedures has been very limited. However, DMM determined that during several of the 19 hours when HASP LMPM procedures were not run in Q3, no review of the price impacts was performed by the ISO's normal price correction process. Although the overall impact of potential price increases due to the lack of mitigation during these hours appears to be minimal, DMM is recommending that the ISO improve the price correction process to ensure that all hours in which LMPM procedures in HASP fail are thoroughly reviewed for price impacts.

2.1 Competitive Benchmark

To assess the competitiveness of the day-ahead market, DMM runs two simulations using its stand-alone copy of the IFM software. The first run is a re-run of the IFM using data for the applicable IFM Saved Case (the ISO's archive of market and system inputs and settings saved after completion of the final IFM market run). Results of this initial re-run are benchmarked against actual market results to validate that the DMM stand-alone system is accurately reproducing results of the actual market software.[14] In cases where the stand-alone system does not produce comparable results, results for these days are excluded from the analysis.[15]

The second run of the stand-alone IFM software is designed to represent a perfectly competitive scenario which provides a *competitive benchmark* against which the re-run of actual IFM prices can be compared. In this second run, bids for gas-fired generating resources are replaced with their respective Default Energy Bids (DEBs), which are designed to represent each unit's actual variable or opportunity costs.[16] This run reflects the assumption that under perfectly competitive conditions, each resource would bid at their marginal operating or opportunity costs. The percentage difference between actual market prices and prices resulting under this competitive benchmark scenario represents the *price-cost markup* or *competitive baseline index* for the IFM. Generally, DMM considers a market to be competitive if the index indicates no more than a 10 percent mark-up over the competitive baseline.

Figure 2.1 through Figure 2.3 show monthly summary results of this competitive baseline analysis for each of the three LAPs in the system. The light blue bar (left axis) represents the weighted average price for each LAP for the days that were re-run using actual IFM market inputs (IFM Actual). The darker blue line (left axis) shows the weighted average price for each LAP for these same days based on the re-run performed using DEBs for gas-fired generation (Competitive Baseline). The red line in each figure (right axis) represents price-cost mark-up, or the percentage difference between actual prices and the prices under the competitive baseline. As illustrated in these figures:

- In July, the monthly price-cost mark-up ranged from -.2 percent to -.8 percent across the three LAPs.
- In August, the price-cost mark-up averaged about -1 percent across all three LAPs.
- During September, the average mark-up ranged from -.3 to -1.5 percent across the three LAPs.

Overall, the competitive index indicates that monthly LAP prices are within competitive ranges through the first six months of the ISO's new market. The competitive index for the third quarter

---

[14] Results of the market software and DMM's stand-alone version can vary for several reasons. First, since these two systems are managed and updated independently, the DMM system may sometimes be running with a somewhat previous version of the actual IFM software. In addition, differences may occur due to changes in one or more settings that may have been made between the pre-IFM MPM, IFM and RUC runs. Data archived in Saved Cases represent settings used in the final RUC run. Thus, if any changes in settings (such as the MIP gap, for example) are made between the pre-IFM MPM, IFM and RUC runs during actual market operations, a re-run based on the settings used in the final RUC run that are archived in the Saved Case data may not duplicate the actual IFM results.

[15] For this 3rd Quarter 2009 report, results were excluded for 8 out of 31 days in July; 9 out of 31 days in August; and 13 out of 30 days in September. DMM's goal is for the portion of re-runs that do not accurately replicate market outcomes (and are therefore excluded from such analyses) to decrease as updates to the IFM software decline, and DMM is able to successfully perform a greater portion of re-runs with a smaller lag time from the date of actual market operations.

[16] Under the market power mitigation provisions of the ISO's tariff, cost-based DEBs are increased by 10 percent to reflect potential costs that may not be entirely captured in the standard fuel and variable cost calculations upon which cost-based DEBs are based (Section 39.7.1.1). Units such as use-limited resources may also have a DEB that reflects their opportunity costs under the negotiated cost option of the ISO tariff (Tariff Section 39.7.1.3, and *Business Practice Manual for Market Instruments*, Version 1, Revised: Mar 26, 2009, D-3 to D-4).

of 2009 shows slightly negative price-cost mark-ups, which are attributable to the fact that a significant amount of generators bid slightly below their DEBs. Since cost-based DEBs include a 10 percent adder above fuel and variables costs, these relatively small negative mark-ups are not indicative of uncompetitively low prices, and simply reflect the fact that actual bids for many units cover fuel and variables costs, but do not include the additional 10 percent multiplier included in DEBs.

Meanwhile, the increase in average cost from June to July in both the actual IFM and the competitive baseline scenario results can be explained by an increase in spot market prices for natural gas and the increase in demand during these periods. For the June to July period, spot market prices for natural gas averaged about 7 percent more in July (about $3.62/mmBtu in July compared to $3.37/mmBtu in June), while actual IFM prices during the days included in the competitive baseline analysis increased by about 35 percent (about $26/MW in June compared to about $35/MW in July). Higher average costs can also be attributed to the increase in peak and average demand, which were about 11 percent and 15 percent higher, respectively, in July compared to June. The slight increase from August to September in both the actual IFM and the competitive baseline scenario results can be explained by a small increase in spot market prices for natural gas of about 4 percent (from about $3.72/mmBtu in September compared to $3.57/mmBtu in August). Additionally, peak demand in September (the 2009 summer peak demand month) was 3 percent higher compared to August which again would generally result in higher average monthly costs, as higher cost supply was needed to meet demand.

Figure 2.1 PG&E LAP Competitive Baseline Index (April – September, 2009)

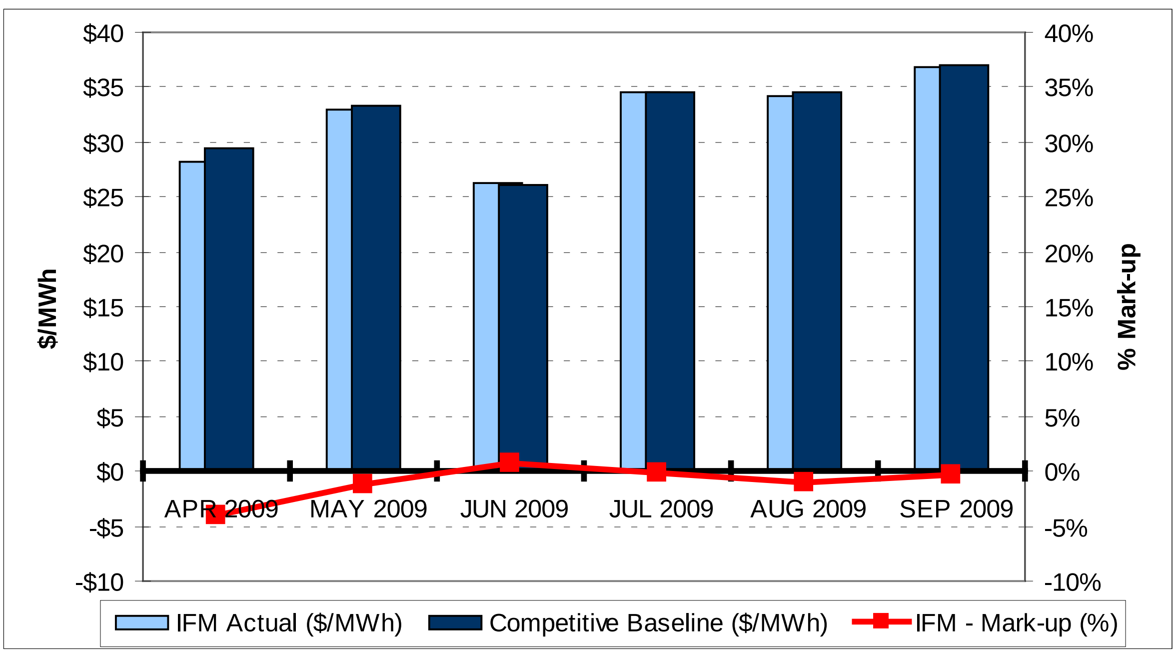


Figure 2.2 SCE LAP Competitive Baseline Index (April – September, 2009)

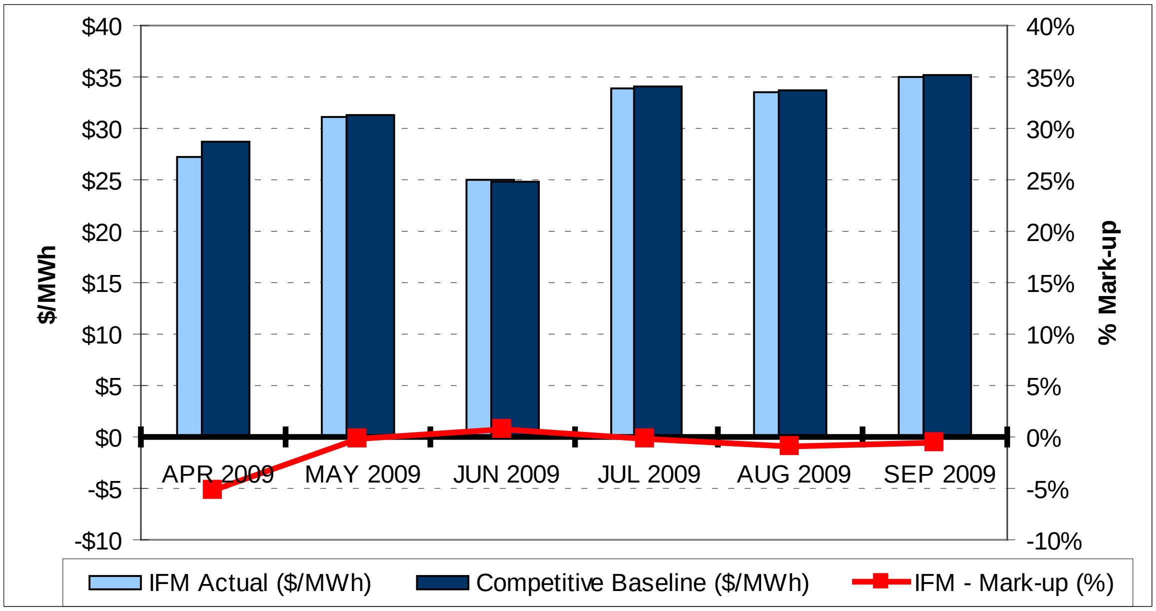


Figure 2.3 SDGE LAP Competitive Baseline Index (April – September, 2009)

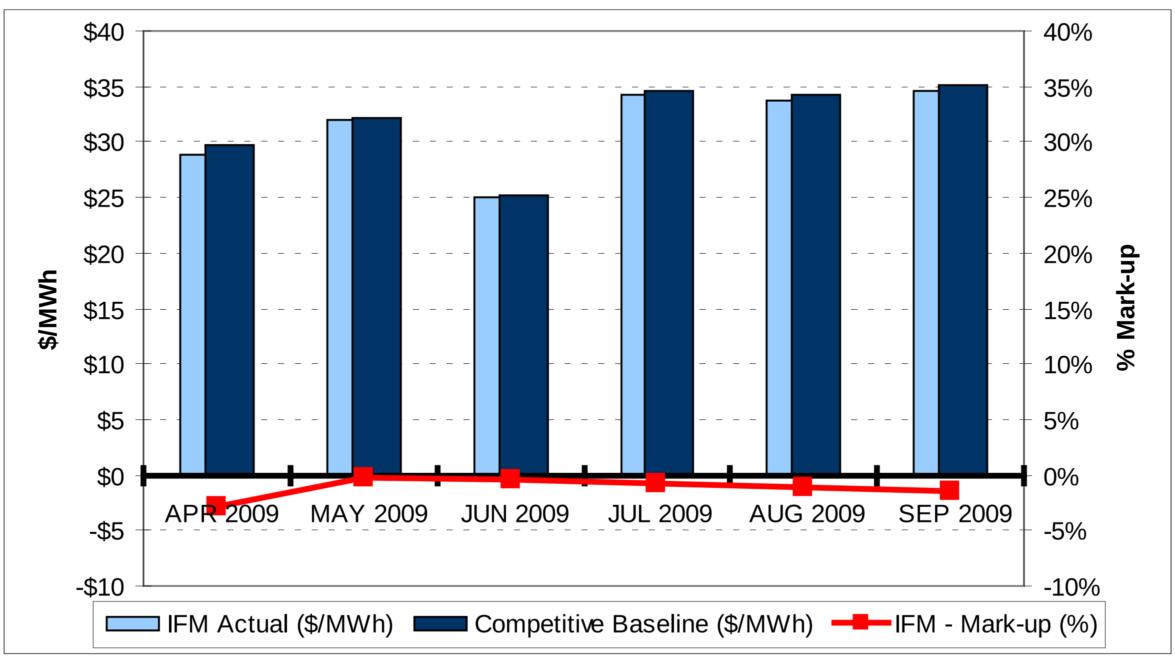
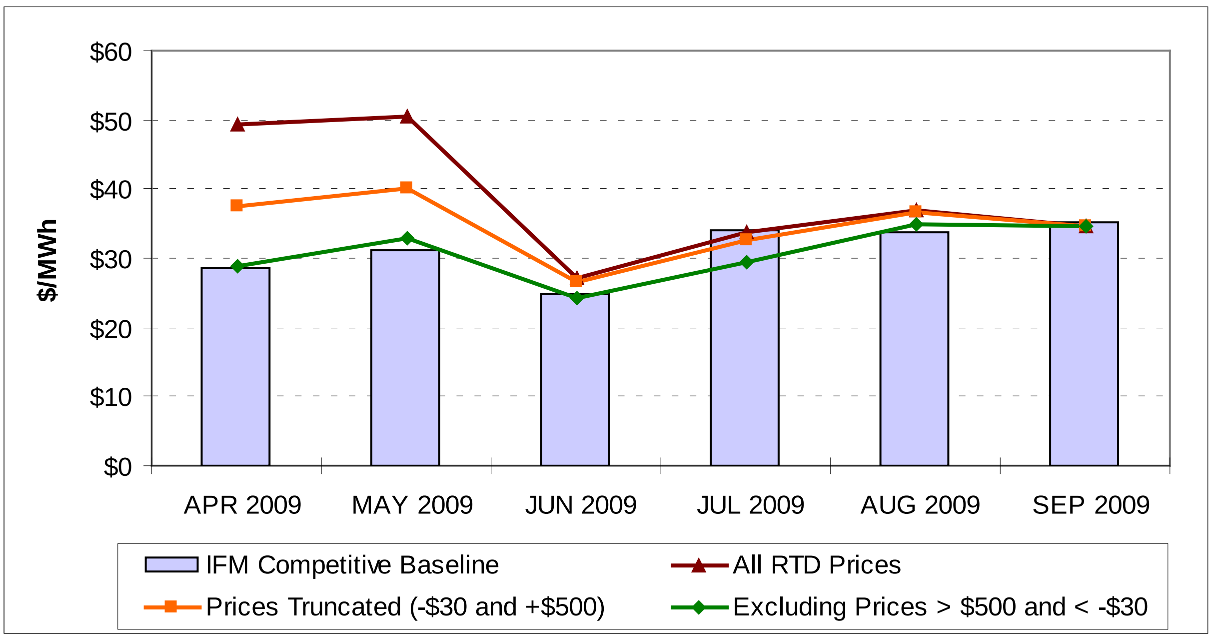


Figure 2.4 compares the competitive baseline price for the SCE LAP to three different averages of 5-minute real-time SCE LAP prices. As shown in Figure 2.4, when extremely high or low 5-minute prices (greater than $500 or less than -$30) are excluded, average real-time prices for each of the three months are essentially equal to the competitive baseline estimate. For purposes of this comparison, DMM believes it is appropriate to exclude such extreme prices on the grounds that they reflect 5-minute operating constraints that cannot be captured in the competitive baseline estimate, which is produced using the day-ahead market software.

Figure 2.4 also provides two additional comparisons based on real-time prices with less screening of extreme prices, including one that includes all 5-minute prices but truncates extreme prices at the bid caps (purple line), and a second comparison that includes all 5-minute prices with no prices excluded or truncated (black line). As shown in Figure 2.4, these other two comparisons were significantly higher than the competitive baseline in April and May, then converged to the competitive baseline from June to September. This convergence of IFM and RTD prices reflects the fact that there were much fewer extreme real-time prices in the June to September months.

Figure 2.4 Comparison of SCE LAP Competitive Baseline to Real Time Prices



2.2 LMP-Based Default Energy Bids

Starting in July 2009, resources had the option of having their DEB calculated using the LMP-based approach. The DEBs for units under this option are set by averaging the lowest quartile of LMPs for the time periods in which the unit was dispatched over the previous 90 days.[17] In calculating the LMP-based DEB, dispatches (and the corresponding LMPs) during all peak hours (7-22) are used to calculate a DEB for peak hours, while dispatches and LMPs during other hours are used to calculate a DEB for off-peak hours. The LMP-based DEB is calculated separately for the IFM and RTM. Thus, each unit under this option has a total of four DEB bid curves: IFM peak and off-peak, and RTM peak and off-peak.[18]

Figure 2.5 and Figure 2.6 summarize the number of units and the amount of capacity by fuel type assigned to the LMP-based DEB for July through September. When sufficient data first became available for the LMP-based option to become effective in July 2009, a total of 83 resources, representing nearly 10,000 MW generating capacity, had DEBs set using the LMP-based option. However, by September 2009, there were no resources with DEBs being set under the LMP-based option.[19] This trend can be attributed to the relatively low DEBs that resulted under the LMP-based option for most units, given the relatively low LMPs that have been observed in the ISO markets during many hours.

For gas-fired units, the relatively low LMP-based DEBs can be illustrated by comparing these LMP-based DEBs to each unit's actual variable operating costs as calculated under the cost-based DEB option based on each unit's heat rate, variable operating and maintenance cost, and spot market gas prices (plus 10%). Figure 2.7 and Figure 2.8 show this comparison for the eight gas-fired units that selected the LMP-based option in July 2009. As shown by this analysis, the LMP-based DEBs for gas-fired units are generally lower than the DEBs that would be used under the cost-based DEB option. Although the LMP-based DEB for peak periods in the IFM would be higher for some units under this option, the LMP-based DEB for these units would generally be lower during off-peak hours in the IFM, and lower for both peak and off-peak hours in the RTM. Since units selecting the LMP-based option are required to have DEBs calculated using this option for both peak and off-peak hours in the IFM and RTM, this likely explains the shift away from using the LMP-based DEB as the primary DEB option after August 2009.

For non-gas units, DMM does not have cost data that can be directly compared to each unit's LMP-based DEB. However, as shown in Figure 2.9 and Figure 2.10, a review of LMP-based DEBs for these non-gas units suggests that the relatively low LMP-based DEBs – particularly for off-peak hours and in the RTM – also explains the shift away from using the LMP-based DEB as the primary DEB option after August 2009.

---

[17] Pursuant to the ISO Tariff (39.7.1.6), the LMP-based DEB became available only after the first 100 days of the new market (MRTU) since a history of LMP observations from the new market is required to calculate this DEB option.

[18] The ISO's BPM on Market Instruments requires that in order to calculate an LMP-based DEB bid segment, the unit must have been dispatched at least x times during the previous 90 days. In the event that a unit selects the LMP-based option, but there is not sufficient dispatch data to calculate an LMP-based bid segment, the DEB for that segment is based on the other two DEB options (cost-based or negotiated) in the order of preference that was selected by the unit owners.

[19] In September, one resource still had the LMP-based option designated as its first choice for setting its DEBs, but this resource had not been dispatched during enough intervals to qualify for having its DEBs calculated using the LMP-based option.

Figure 2.5 Number of Units under LMP-Based DEB Option by Fuel Type

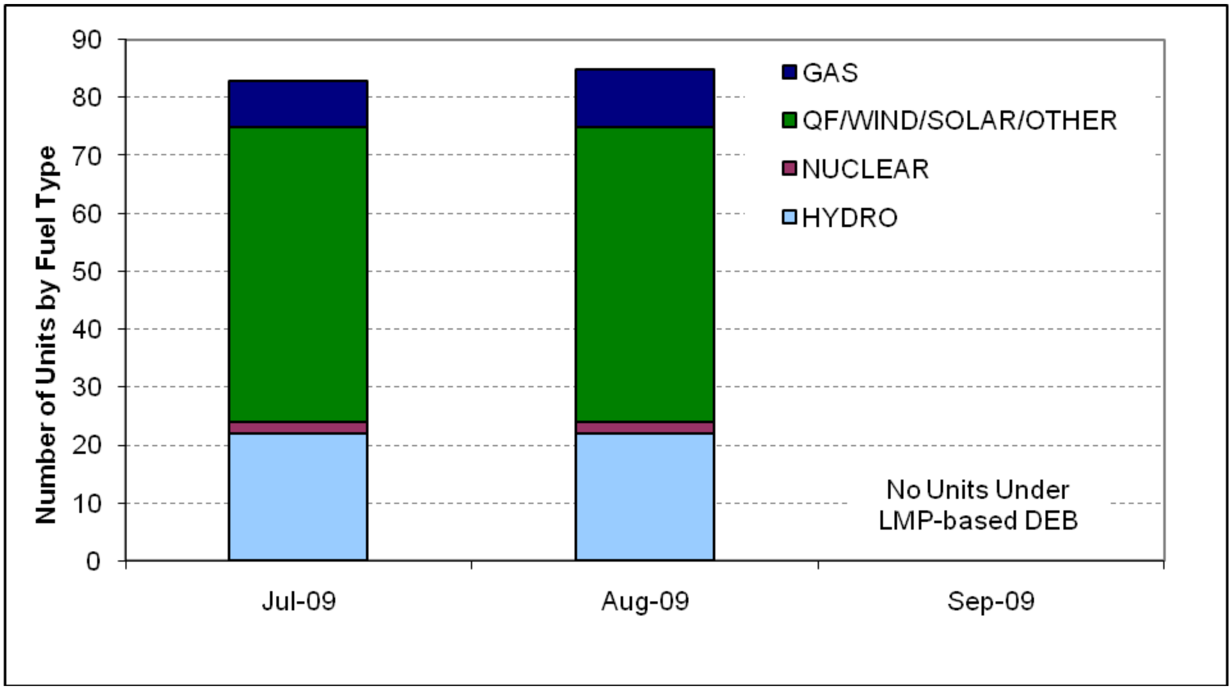


Figure 2.6 Capacity Under the LMP-Based DEB Option by Fuel Type

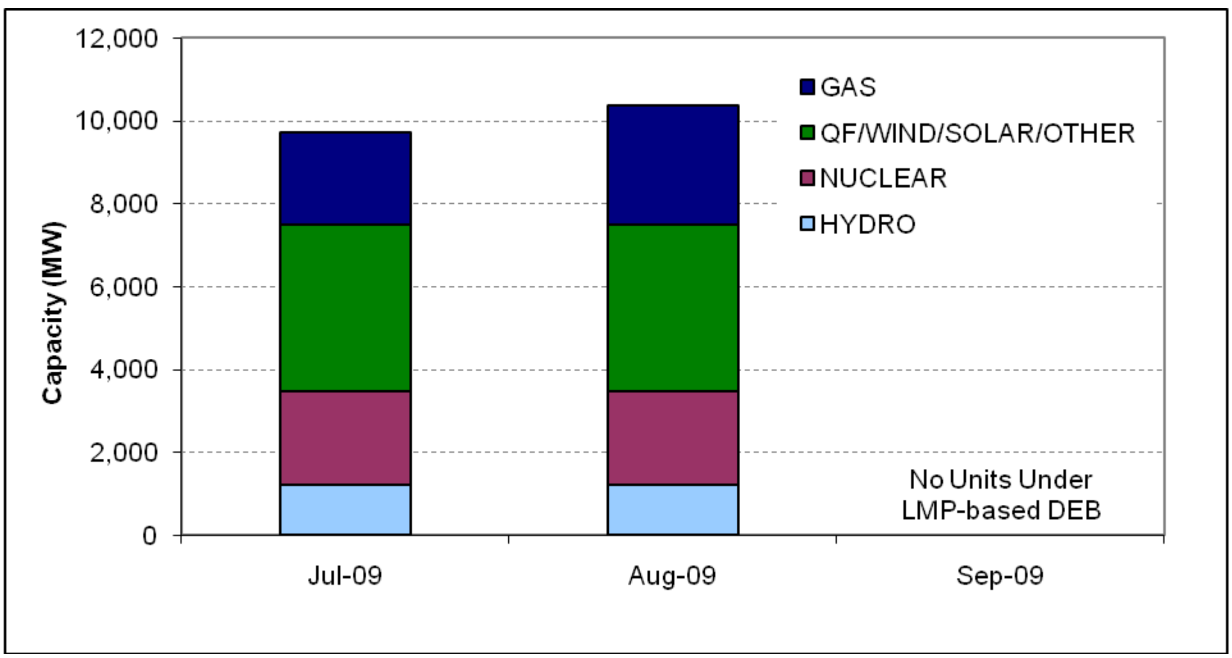


Figure 2.7 Comparison of Variable Cost to Day-Ahead LMP-Based DEB (Gas-fired Units Only)

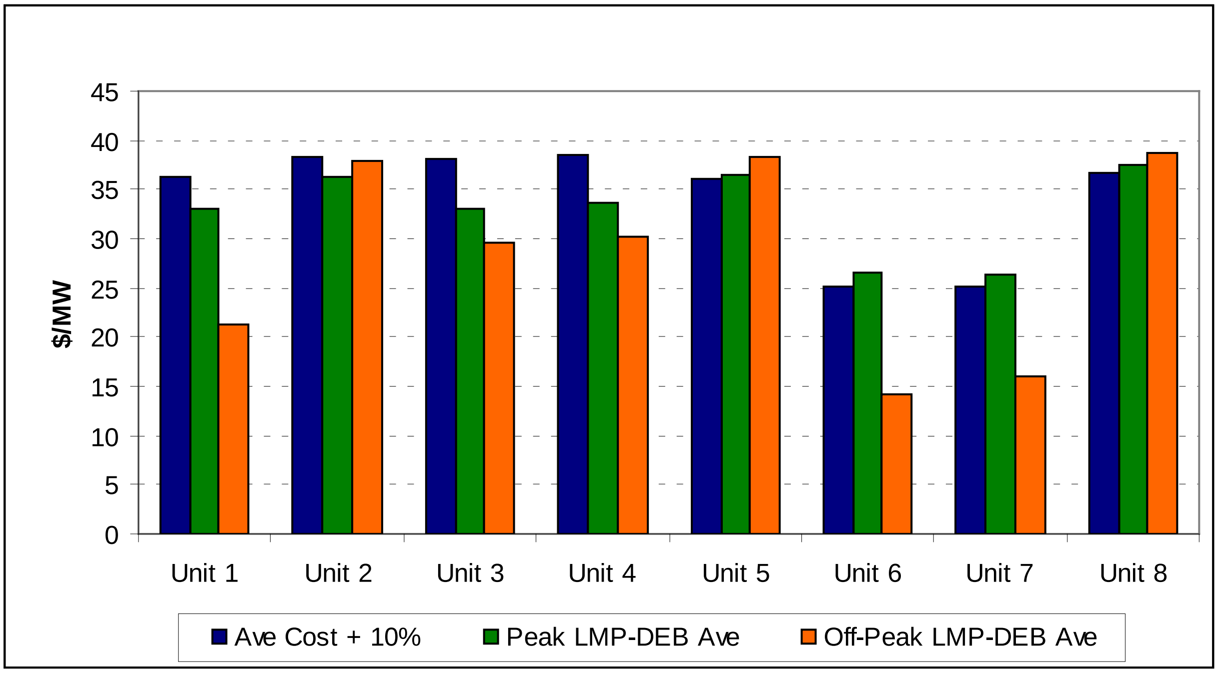


Figure 2.8 Comparison of Variable Cost Estimate to Real-Time LMP-Based DEB (Gas-fired Units Only)

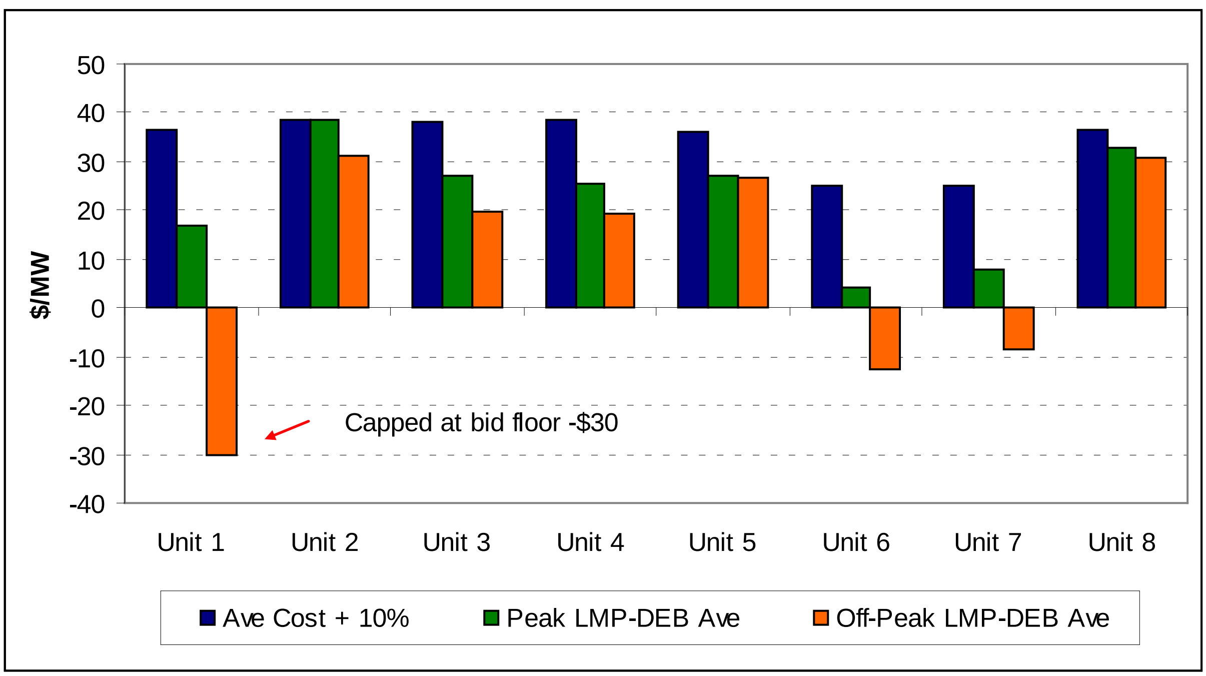


Figure 2.9 Average Day-ahead LMP-Based DEB by Fuel Type (Non-Gas Units) August 2009

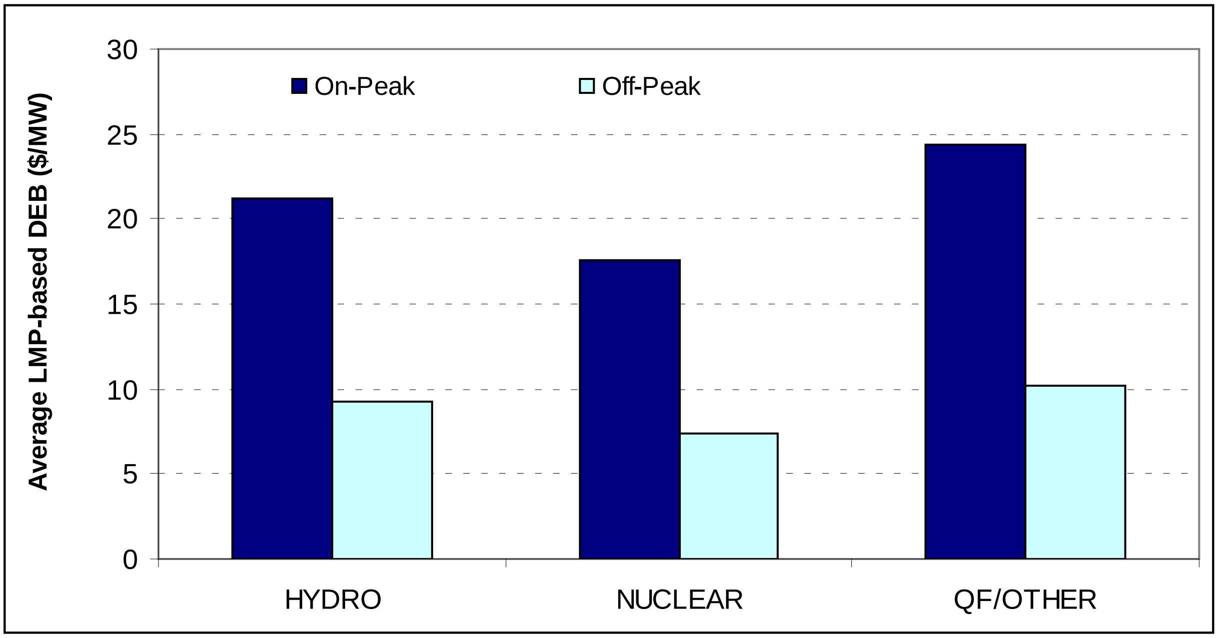
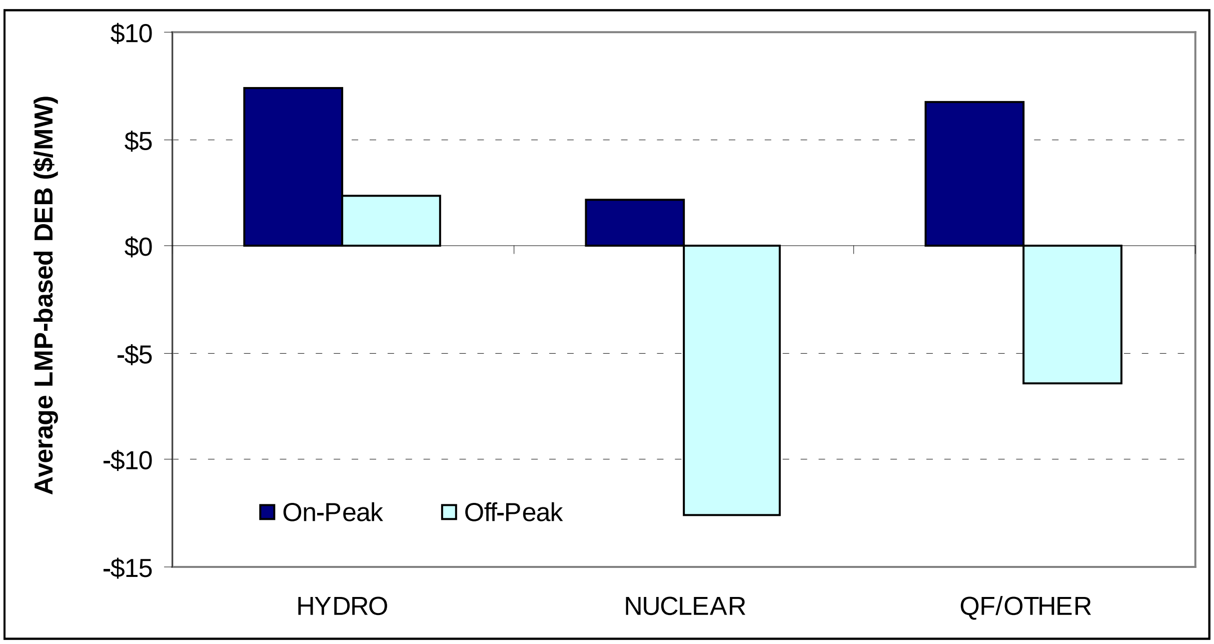


Figure 2.10 Average Real-time LMP-Based DEB by Fuel Type (Non-Gas Units) August 2009



2.3 Mitigation of Exceptional Dispatches

Under FERC's February 20, 2009 Order,[20] all exceptional dispatch (ED) instructions for energy (above minimum operating levels) were subject to mitigation for the first four months of the ISO's new nodal markets (April through July, 2009). However, FERC's February 20 Order directed that after the initial four month period, mitigation would only be applied to two categories of ED for incremental energy:[21]

- Exceptional dispatches to mitigate congestion on constraints deemed to be "non-competitive" under the Competitive Path Analysis (CPA) performed by the ISO as part of its local market power mitigation procedures; and

- Dispatches for Delta Dispatch.

Thus, starting on August 1, all other categories of ED for incremental energy will not be subject to bid mitigation, and will be eligible to be paid the higher of (a) the market LMP or (b) their market bid price. If an ED is mitigated, the generator will be paid the maximum of (a) the LMP or (b) the unit's Default Energy Bid (DEB).

As noted in its Q2 Report, DMM was concerned that if the ISO continued to issue exceptional dispatches for substantial volumes of energy for reasons that were not logged as being for a specific non-competitive constraint, there could be the potential for significant volumes of high cost EDs. For example, ED for additional energy logged for general reasons such as "Ramp Rate" or "Transmission Outage" are no longer subject to mitigation. As this more limited mitigation took effect in Q3, DMM worked closely with Operations staff to clarify mitigation rules, to identify the potential cost implications of unmitigated EDs, and to establish adequate logging practices for distinguishing between EDs for *competitive* and *non-competitive constraints.*

Based on DMM's analysis of ED dispatches in Q3, a relatively small portion of ED energy dispatched after the more limited mitigation that took effect in August has been logged for non-competitive paths and will therefore not be subject to mitigation.[22] However, this more limited mitigation has not had a significant impact on overall costs for a combination of several reasons:

- First, the total volume of ED for energy has been relatively low since August, as discussed in Chapter 4 of this report, which provides a detailed analysis of the volume and causes for ED for energy in Q3.

- Second, analysis by DMM indicates that over 50 percent of ED energy has cleared "in sequence" (i.e., had a bid price less than the market LMP).

---

[20] February 2009 Order, 126 FERC ¶ 61, 150.

[21] February 2009 Order at P 74.

[22] Analysis in this section is based on DMM's estimate of the payments for ED energy that should ultimately be made *with* and *without* mitigation. DMM's review indicates that mitigation of ED has not yet been applied as part of the ISO's settlement process. ISO Market Services is currently seeking to apply the post-August 1 bid mitigation rules to final August invoices. Mitigation rules in effect for the April through July period will be applied when Market Services has the ability to rerun settlements. However, as discussed later in this section, DMM's analysis suggests that the actual impact of these adjustments is likely to be relatively small both before and after changes in mitigation rules took effect in August 2009.

- Finally, bid prices for ED energy that is out-of-sequence (OOS) and not subject to mitigation have not been extremely high relative to each unit's DEB.

This third trend is illustrated in Figure 2.11, which shows the average cost of ED energy each month in Q3 *with* and *without* mitigation. In August, the average bid price for ED energy that was OOS energy and not subject to mitigation was about $50/MW, compared to an average DEB of about $39/MW, representing an incremental cost of about $11/MW or 28 percent of the average DEB for this ED energy. In September, the average bid price for ED energy that was OOS energy and not subject to mitigation was about $65/MW, representing an incremental cost over the average DEB for this energy of about $15/MW or 30 percent.

Figure 2.12 shows the total estimated cost of EDs for energy for each month in Q3, *with* and *without* mitigation. As shown in Figure 2.12, even if all ED energy was mitigated, total costs for ED energy would have been about $246,000 lower in August and only $230,000 lower in September.

Figure 2.11 Estimated Average Price of Out-of-Sequence (OOS) Exceptional Dispatch Energy Before and After Mitigation

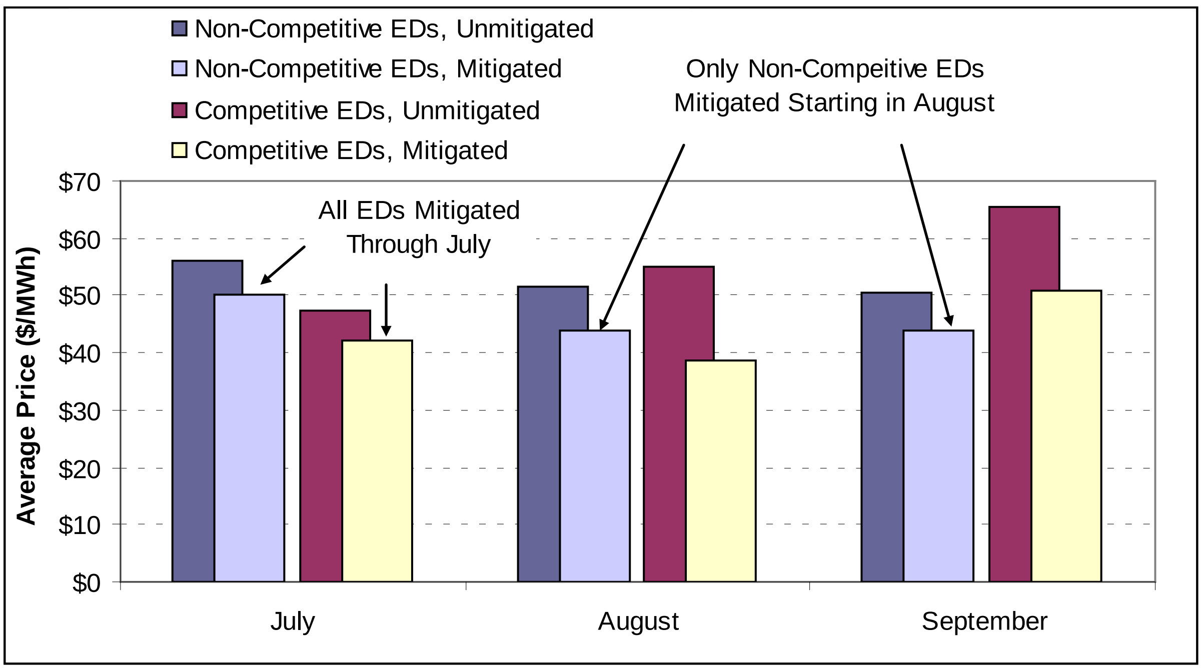
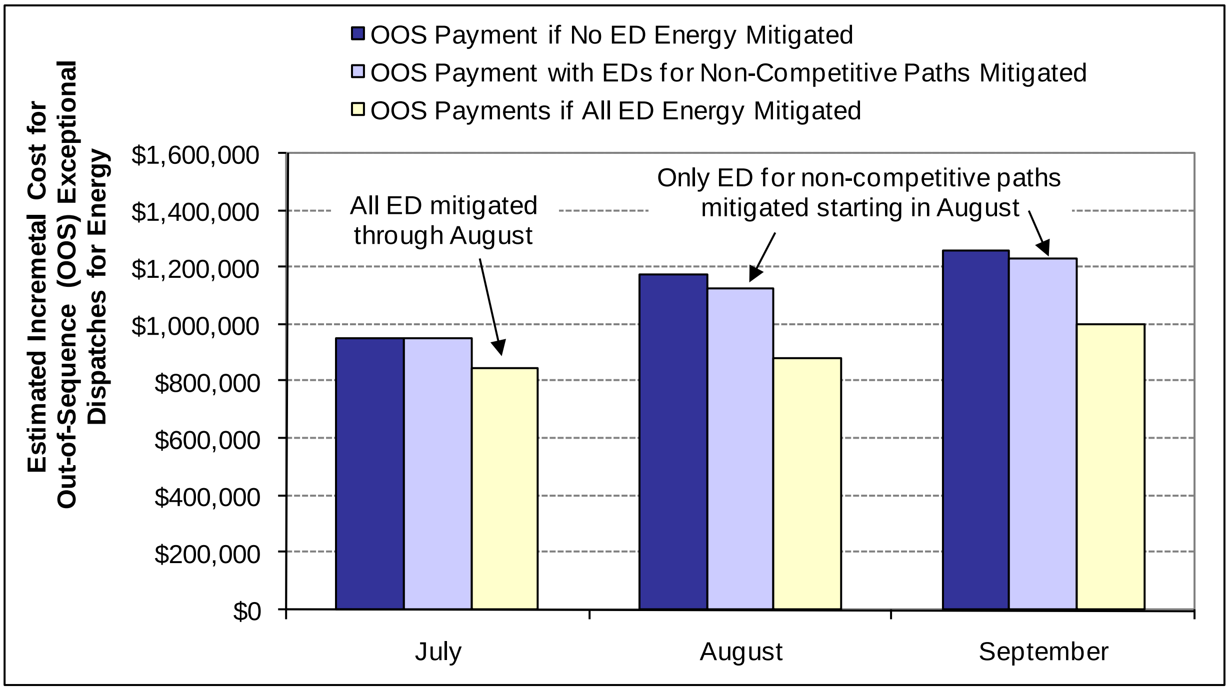


Figure 2.12 Estimated Total Incremental Cost of Out-of-Sequence (OOS) Exceptional Dispatch Energy Before and After Mitigation



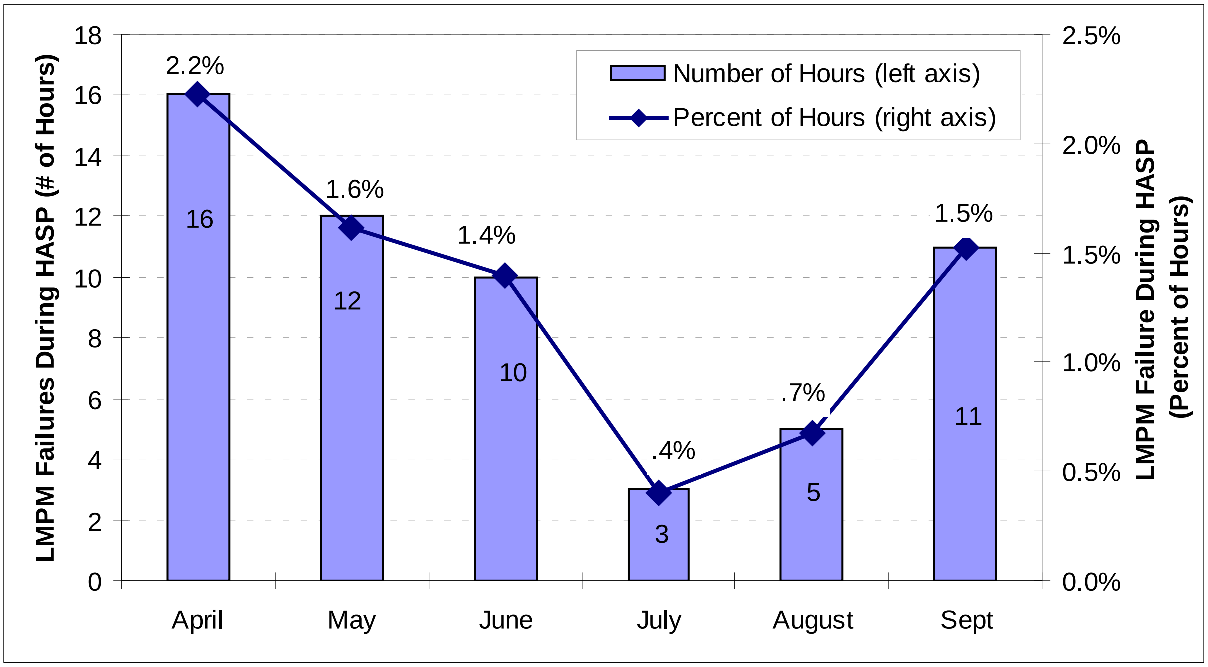
2.4 LMPM Failures During HASP

Prior to the start-up of the ISO's new LMP market, one of the major issues identified by DMM was the relatively high frequency with which the pre-RTM LMPM process was not run due to various problems or failures occurring during the HASP process – during which the pre-RTM LMPM process is performed. We recommended that the ISO closely monitor this issue and seek to reduce the frequency of pre-RTM LMPM failures due to problems in the HASP. In addition, we recommended that the ISO develop a process for assessing the market impact of any failures of the LMPM procedures on prices in the RTM and performing price correction, as appropriate.

During the first six months of the ISO's new market, the frequency of failures in the pre-RTM LMPM process has been relatively low, and has trended downward. As shown in Figure 2.13, the portion of hours that the LMPM process has failed to run in the HASP has continued to drop in July and August, before rising slightly in September, when the HASP LMPM procedures were not run 11 times or about 1.5 percent of hours.

In addition, review by DMM and the ISO's price correction team indicates that the price impacts of failures in the pre-RTM LMPM procedures has been very limited. As discussed in its Q2 report, DMM has reviewed instances of HASP failures and worked with the ISO's price correction team to establish more automated and standard criteria for determining if price correction may be needed in cases when the pre-RTM LMPM procedures are not run. Based on review by the price correction team, no price corrections were made for any of the hours when HASP LMPM procedures were not run that were due to the lack of mitigation during these hours. However, review by DMM determined that for some of the 19 hours when HASP LMPM procedures were not run in Q3, no review of the price impacts was performed by the ISO's price correction team.[23] Although the overall impact of potential price increases due to the lack of mitigation during these hours appears to be minimal, DMM is recommending that the ISO improve the price correction process to ensure that all hours in which LMPM procedures in HASP fail are thoroughly reviewed for price impacts.

Figure 2.13 Frequency of LMPM Failures During HASP



[23] DMM's review indicates there was only one hour in which the pre-IFM LMPM procedures were not run that does not appear to have been reviewed by the price correction team where this may have had a significant impact on price (9/29 HE 1). During this hour the hourly average PG&E LAP price in RTD was about $42, compared to $24 the hour before and $33 the hour after.

3 Ancillary Services

This chapter provides a high-level description of ancillary services procurement, profiles of ancillary service market outcomes for the third quarter, and two specific issues related to the real-time market for ancillary services and energy. The ancillary service markets have generally performed well since the start of the ISO's new market design. Prices in the day-ahead and real-time ancillary service markets have been reasonable and highly competitive, with day-ahead ancillary service prices somewhat higher than in real-time. However, in this chapter, we examine two issues involving how the ancillary service markets interact with the energy markets:

- ***Contingency-Only Reserves*.** The first issue is the procurement of spinning and non-spinning reserve that is designated as *contingency-only* – i.e. capacity that can only be dispatched for energy in the case of a contingency (i.e., major transmission or generation outage) or an imminent or actual system emergency. Since the start of the ISO's new market, a high portion of spinning and non-spinning reserve procured in the IFM has been designated by participants as contingency-only. In addition, all incremental reserve procured after the IFM in the real-time pre-dispatch process (RTPD) run every 15-minutes is automatically designated as contingency-only.[24] Even in cases when the ISO has enough reserve to meet its system requirements, this can create price spikes during periods where supply is tight, particularly in transmission constrained load pockets. This can occur when a relatively large amount of contingency-only reserve is located in a load pocket or anywhere on the grid where this capacity would be particularly effective if dispatched as energy to meet a local constraint. While this may not significantly impact prices with a high degree of frequency, this can increase prices dramatically when supply is tight and penalty prices on constraints are setting prices that could be relieved with a relatively small amount of additional supply that is being held as contingency-only reserve. At the end of this chapter, we suggest that the ISO consider several ways in which this issue might be addressed.

- ***Reserve Scarcity Pricing*.** This chapter also examines the relationship between real-time energy and ancillary services as it relates to the ISO's scarcity pricing proposal.[25] Specifically, we highlight a disconnect between the real-time ancillary service and RTD energy prices that may dampen price signals in the 5-minute RTD during instances where scarcity pricing would be triggered if these two markets were directly linked. Although energy and ancillary services are co-optimized in the RTPD run performed every 15 minutes, these energy prices are not financially binding – with the subsequent 5-minute RTD runs being financially binding for energy. Thus, we recommend further analysis of the potential for the scarcity pricing mechanism to affect real-time energy prices, including the longer-term possibility of deploying co-optimization in the 5-minute RTD market.

[24] In addition, since the ISO software requires that all spinning or non-spinning reserve being supplied by a single generating unit be either contingency-only or non-contingent, in cases when any incremental spinning or non-spinning reserve is procured from a unit in the real-time ancillary services market, any of that same reserve product that was procured from that unit in the day-ahead is also automatically designated as contingency-only.

[25] *Final Draft Proposal: Reserve Scarcity Pricing Design*, October 5, 2009, http://www.caiso.com/243e/243ecc4d2d490.pdf

3.1 Background and Overview

Ancillary services are procured regionally from pre-defined regions, shown in Figure 3.1, in the day-ahead and real-time pre-dispatch (RTPD) market runs to minimize concentrated procurement. Currently, there are ten pre-defined regions, but to date, only four have been enforced: System (CAISO), System Expanded (CAISO Expanded), South of Path 26 (SP26), and South of Path 26 Expanded (SP26 Expanded). The expanded regions are identical to the granular regions but also include any inter-ties with one end in the granular region.

Each of the four types of ancillary services (regulation up, regulation down, spin, and non-spin) has a different minimum requirement constraint that must be met in the day-ahead market. Any capacity procured in the RTPD market is *incremental* to the day-ahead awards and is procured to either (1) replace capacity that is no longer available due to outages and de-rates, or (2) to meet an increase in market requirement (e.g., due to an increase in the demand forecast). Furthermore, the system minimum requirement for each service is distributed as minimum sub-regional requirements for each enforced sub-region. Due to the nesting of regions, capacity procured in a more granular region also aids in meeting the minimum requirement of the outer region.

Figure 3.1 California ISO Ancillary Service Procurement Regions

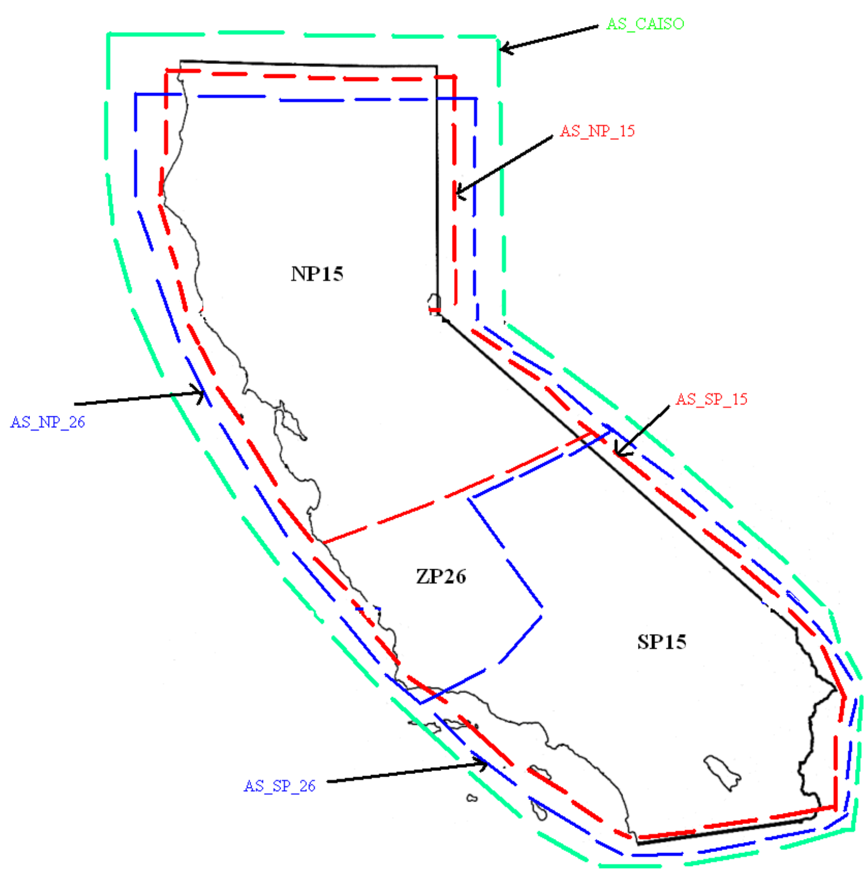


Figure 3.2 shows the average hourly RTPD procurement by region for each service type. The capacity represented in the charts includes the day-ahead awards that cleared the real-time market in addition to the real-time incremental procurement. Within each service, the regional procurement is incremental of any capacity procured from a more granular region that helped meet both regional requirements. For example, procurement in the SP26 Expanded region is net of all capacity procured in the SP26 region, as the capacity in SP26 helped meet both of these requirements.

Overall, there have been no intervals since April 1, 2009, when the ancillary service market was deficient in meeting any of these minimum requirements. Most of the capacity was procured in the SP26 and System regions from generation internal to the ISO control area, with minimal procurement from inter-tie resources (represented by the MW from the SP26 Expanded and System Expanded regions).

Regulation Down

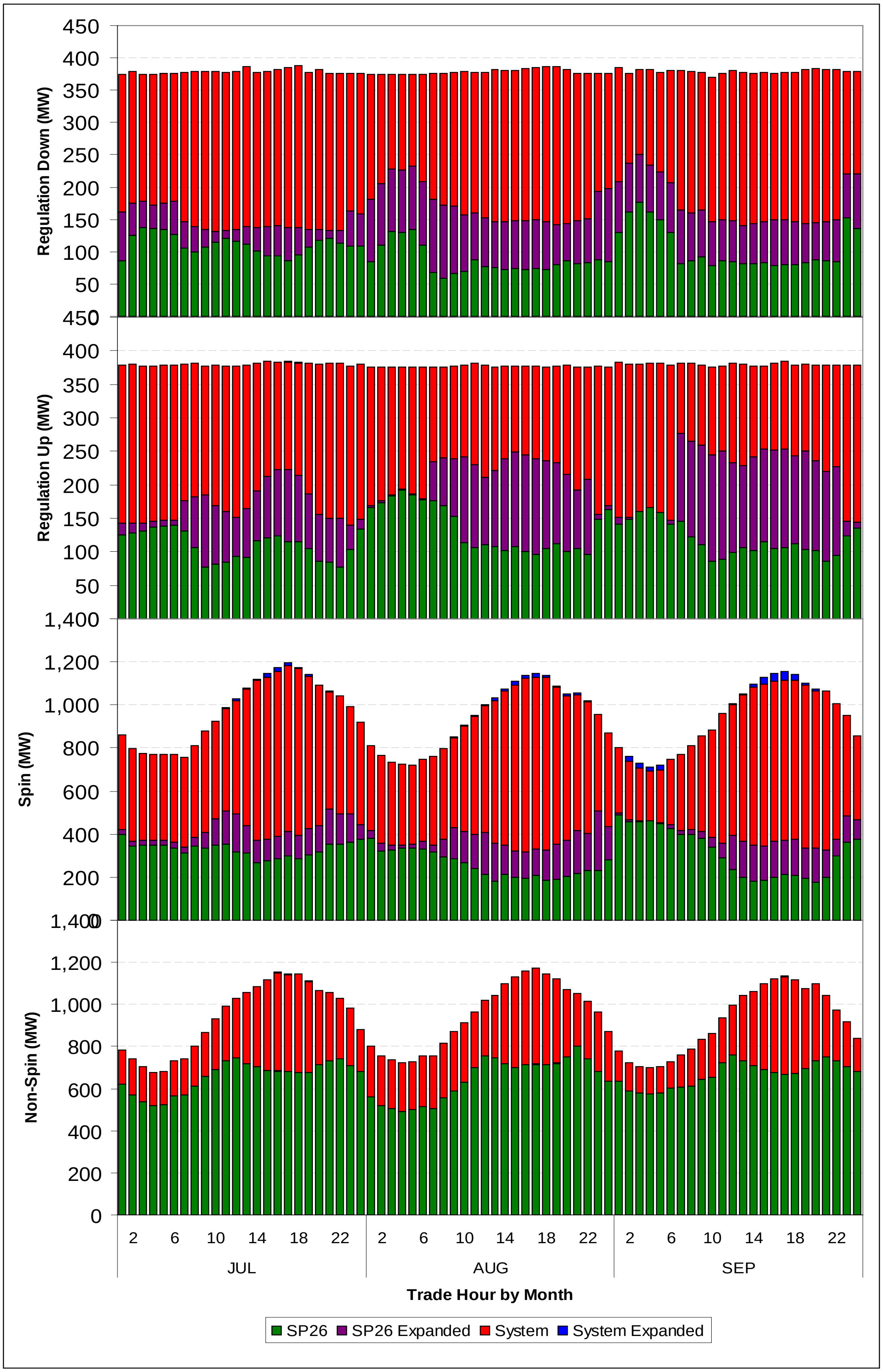
The requirement for regulation down remained at 375 MW during the third quarter,[26] distributed among the regions as follows: 2.67% each for the SP26 and System regions, 35% in the SP26 Expanded region, and 100% in the System Expanded region. While the minimum requirement for the SP26 and System regions is only 2.67%, most of the capacity has been procured from within those two regions (relying lightly on expanded regions). This indicates that downward regulation capacity from internal generation has been relatively abundant and more economical than capacity from imports. While regulation down capacity from inter-ties was minimal, there was an increasing trend of capacity procurement from inter-ties in the SP26 Expanded region across the months, as shown in Figure 3.2. While the minimum requirement for downward regulation has been consistently 375 MW, there were a few intervals where the average hourly procurement was greater than 375 MW. This is due to over-procuring capacity, which impacted all four ancillary service types, and will be discussed more thoroughly later in this section.

Regulation Up

During the three months represented in Figure 3.2, the regulation up market requirement and regional distribution is identical to that of regulation down. Here, again, most of the capacity is procured from the SP26 and System regions with more capacity from the SP26 region during the off-peak hours. The upward trend of procurement of upward regulation on inter-ties in the SP26 Expanded region across the months is more pronounced for regulation up when compared to regulation down. There were only two hours during Q3 where 25 MW of capacity was procured from the System Expanded region that did not also belong to an inner region. Regulation up also had a few intervals where the average hourly procurement was greater than 375 MW. Higher procurement of regulation up is due in part to the over-procurement issue previously noted, as well as the cascading of higher quality services. Regulation up is the highest quality upward ancillary service; therefore, if it is economical, the market will procure more than the minimum amount to help meet the requirement of lower quality reserves, namely spin and non-spin.

[26] Since April 1, 2009, the regulation requirement fluctuated from 350 MW to 500 MW, but remained at 375 MW during the third quarter.

Figure 3.2 Ancillary Service Procurement by Region



Spinning and Non-Spinning Reserve

The market requirement for spin and non-spin is a percentage of the system load, and thus mirrors the daily load pattern. The distribution of spin requirement among the regions remained a constant percentage of the total requirement, as follows: 17.5% in SP26 region, 35% in SP26 Expanded region, 50% in System region, and 100% in System Expanded region. Here, again, most of the capacity came from internal generation with a slight increasing reliability on capacity from the inter-ties. Spin capacity was procured from inter-ties in the System Expanded region typically at the top of the morning load pull and across the peak hours, especially in September.

The regional distribution of the requirement for non-spin remained identical to that of spin. Non-spin procurement was strictly from internal generation with the exception of a few megawatts on one day in July across the peak hours that came from inter-tie resources. Non-spin is the only ancillary service with most of the capacity being procured in the SP26 region; the other services have a more even distribution of capacity among SP26 and System regions.

A second facet of the spin and non-spin reserves is the ability for a scheduling coordinator to designate the capacity as non-contingent or contingency-only through the day-ahead bid. All capacity procured in the RTPD is contingency-only.

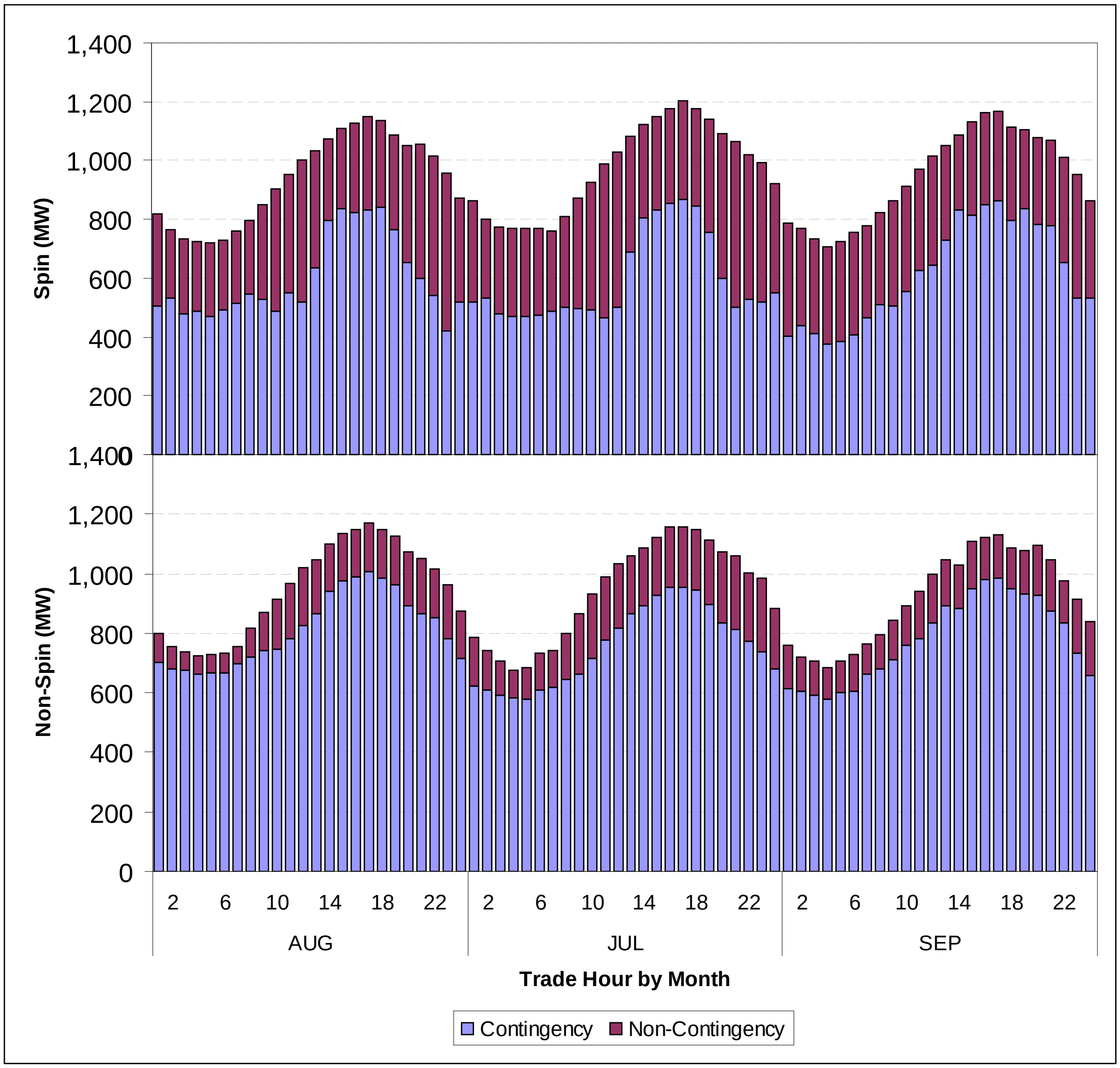
Figure 3.3 shows the average hourly megawatts designated as contingency-only for spin and non-spin across the three months.

The contingency-only capacity for spin fluctuated from 50% to 75% of total procured capacity. During the peak hours, the percentage of contingency-only capacity increased relative to off-peak hours, most notably in July and August. Non-spin contingency-only capacity remained relatively consistent at approximately 80% of total capacity. The percentage of contingency-only capacity did slightly increase during the peak hours within each month, and also showed a minimal increasing trend across the months, reaching almost 85% in September for most hours.

Regional Ancillary Service Shadow Prices

Regional ancillary service shadow prices (RASSPs) are non-zero when the regional minimum requirement constraint is binding, and are always zero when the regional minimum requirement constraint is not binding. The RASSPs reflect the marginal bid price, and, as a result of the energy and ancillary service co-optimization, the Lost Opportunity Cost (LOC) of selling ancillary services rather than energy. The market clearing price received by each unit that sells ancillary services (ASMP) is the summation of the RASSPs for the services sold across the ancillary service regions in which the unit resides. For example, the ASMP of a unit in the SP26 region is equal to the summation of RASSPs for the SP26, SP26 Expanded, System, and System Expanded regions. It then follows that units within the most granular region will always receive the highest ASMP and all units belonging to the same set of regions, for the same service, will receive the same ASMP. Due to cascading of higher quality reserves for lower quality reserves when economical, the highest quality reserve, regulation up, will always have the highest RASSP. Figure 3.4 shows the average hourly regional ancillary service shadow price by commodity across the reporting months. Overall, the average hourly RASSPs were below $5/MW with a couple exceptions in the System Expanded region for regulation down and regulation up.

Figure 3.3 Contingent versus Non-Contingent Procurement



The System Expanded and SP26 Expanded regions were the most frequently binding regions across all commodities. The SP26 region was binding less than 1% of the intervals across all commodities and there was not one interval during which the System region was binding. The minimum requirements for the System and SP26 regions were a small percentage of the total requirement and therefore were easily met, resulting in non-binding constraints and consequently no shadow value or price applied. Most or all of the requirement for the SP26 Expanded and System Expanded regions were often met by capacity procured from the nested regions, as previously noted, and therefore resulted in binding constraints and non-zero shadow values. When capacity from nested regions exceeded their respective minimum requirements, the remaining capacity would be procured from the expanded regions up to their respective minimum requirement, resulting in binding constraints. Furthermore, when the expanded region requirements were met by capacity in nested regions, no additional capacity would be procured from the expanded regions but their respective minimum requirements would become binding. In summation, due to the ample supply of low priced ancillary service capacity from internal generation, the expanded region constraints are more often binding than the nested regions.

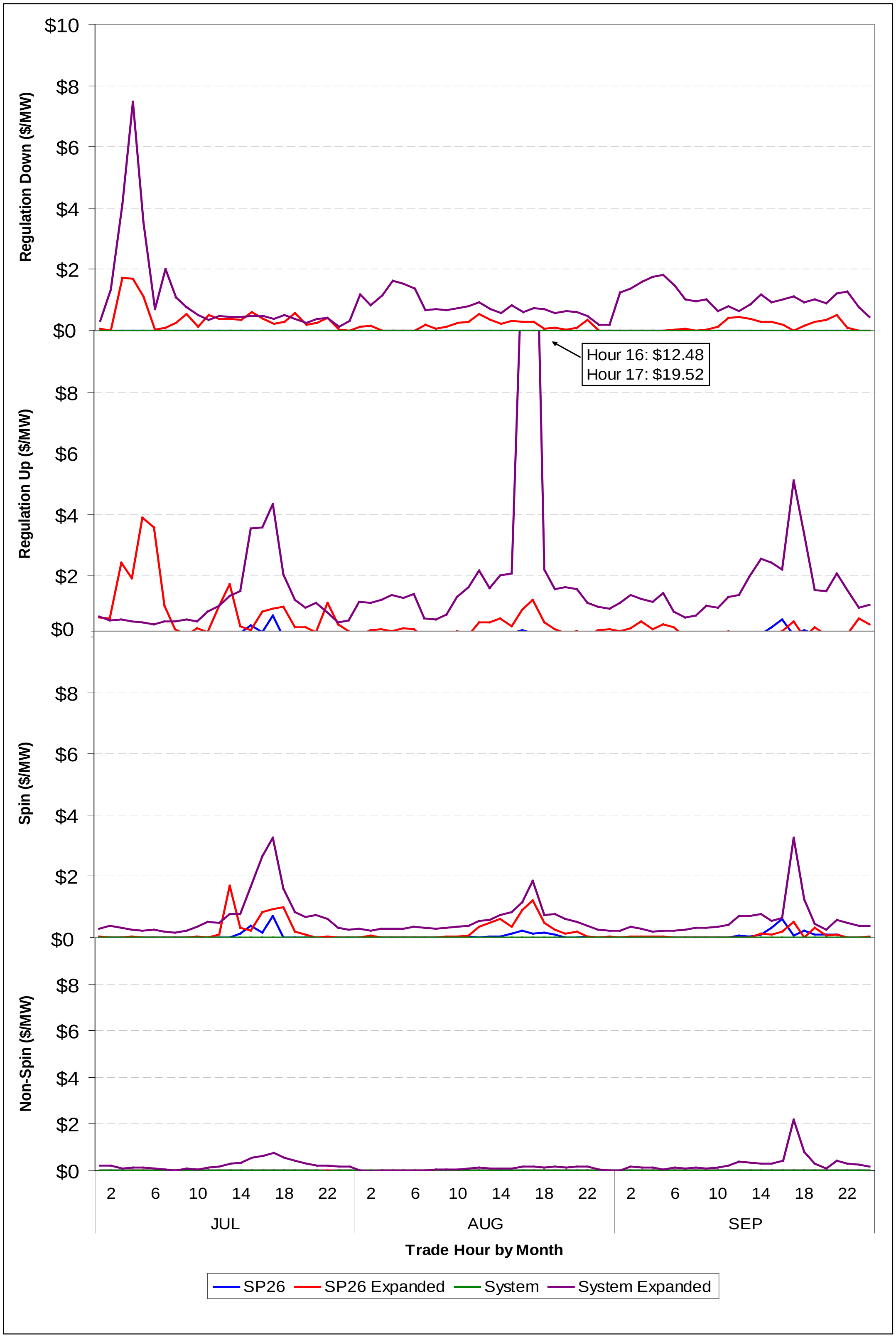
Regulation down tended to have higher RASSPs during the morning hours when it was in higher demand due to the low load and all online units sitting at minimum load, most notably in July. Overall there was a slight upward trend of RASSPs across the months, but they remained under $3/MW, reflecting ample capacity at low prices and low opportunity cost.

Overall, regulation up RASSPs were less than $5/MW with a few exceptions. The RASSPs were higher during the peak hours, reflecting the higher energy prices and opportunity cost during that time of the day. The SP26 region was binding a few times during the peak hours. This is most likely due to higher energy prices in the south making it more economical to procure capacity from the System region, causing the minimum requirement for the SP26 region to bind.

The RASSPs for spin were higher during the peak hours of the day, reflecting higher lost opportunity cost, and the SP26 region tended to bind during those hours for the same reasons it binds in regulation up. Overall, the RASSPs for spin were lower than that of regulation up due to the cascading of higher quality reserves for lower quality.

The only binding region for non-spin from April through September was the System Expanded region. This is due to the fact that 100 percent of the requirements were always met with capacity in the SP26 and System regions. The SP26 Expanded region was never binding because enough capacity was procured in the SP26 region to not only surpass its minimum requirement but also that of the SP26 Expanded region. As with spin, prices were higher during peak hours, and non-spin RASSPs were lower than regulation up and spin as a result of the cascading products.

Figure 3.4 Real-Time Pre-Dispatch Regional Ancillary Services Shadow Prices (RASSP) by Commodity



3.2 Procurement of Contingency-Only Reserves

In the day-ahead market, in which 100 percent of the market ancillary service requirement is met, scheduling coordinators can bid in capacity as *non-contingency* or *contingency-only*. Any incremental capacity procured in the RTPD market is automatically flagged as contingency-only capacity. Furthermore, any incremental capacity procured in real-time from a unit that also sold non-contingent ancillary services in the day-ahead will trigger conversion of all of that unit's capacity within that service to contingency-only.

During real-time, non-contingent spin and non-spin capacity is included in the energy bid stack and can be dispatched economically as long as the total operating reserve remains above the 7 percent North American Electric Reliability Corporation (NERC) standard. If the percentage starts to drop, ISO grid operators may instruct the market software to skip the spin and non-spin bids to maintain operating reserves. Contingency-only capacity cannot be dispatched for energy in real-time except for a contingency (i.e., major transmission or generation outage) or due to an imminent or actual system emergency. Therefore, the more contingency-only capacity that the ISO procures, the less dispatchable capacity is available for real-time energy (again, provided the NERC standard for operating reserve is met). Under tight supply conditions, this may have an impact on real-time energy prices.

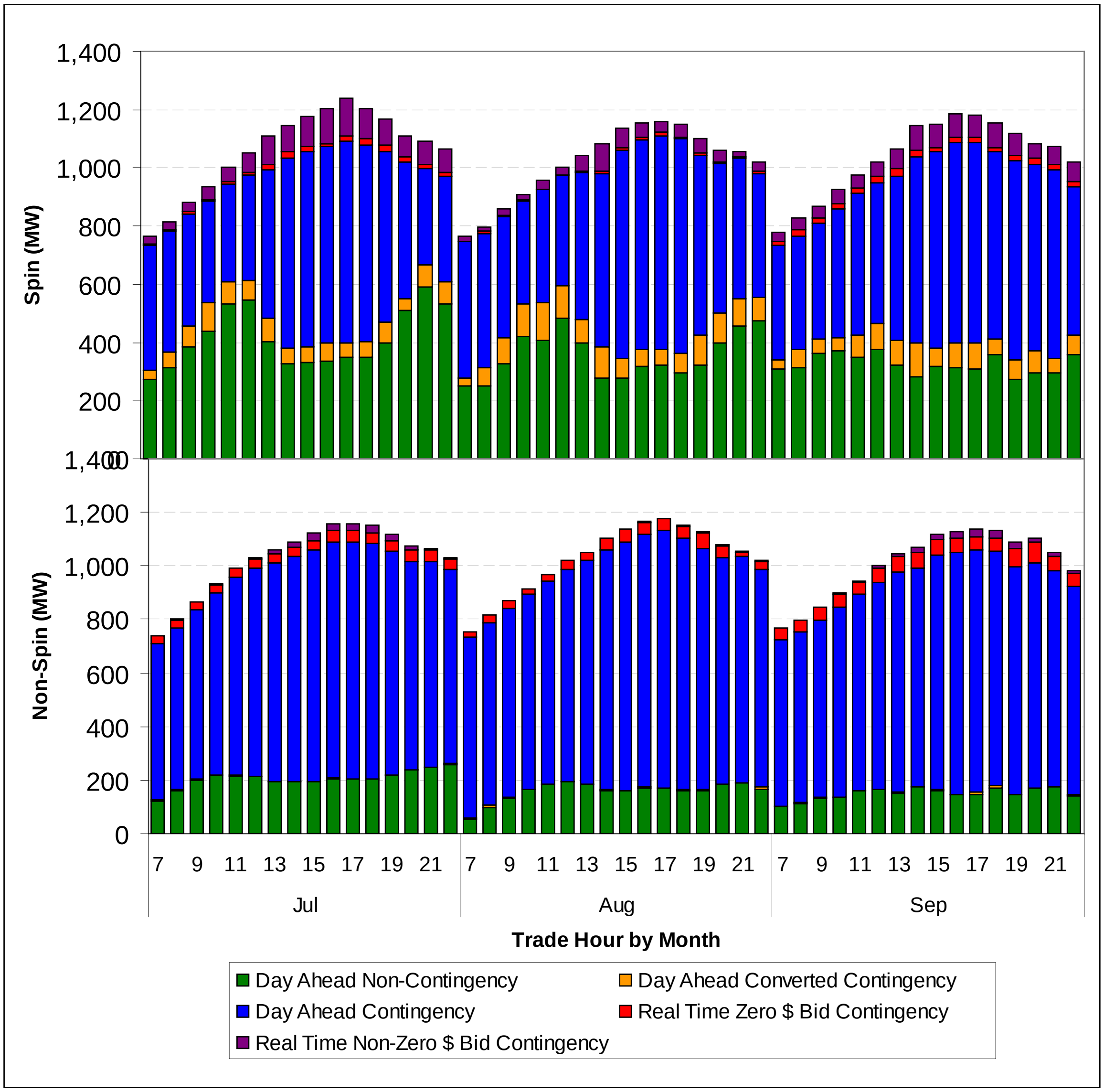
Figure 3.5 shows the total spin and non-spin capacity procured during the peak hours each month, broken down into the following categories:

- Day-ahead non-contingent (as designated by the scheduling coordinator),
- Day-ahead contingent (as designated by the scheduling coordinator),
- Day-ahead non-contingent converted to contingent (due to incremental procurement of additional ancillary services in the RTPD by the ISO),
- Real-time contingent procured due to zero dollar bids, and
- Real-time contingent procured from resources with non-zero dollar bids.

As previously noted, approximately 50 to 75 percent of spin and 80 to 85 percent of non-spin capacity was designated as contingency-only. Most of the spin and non-spin capacity was bid in the day-ahead as contingency-only, as shown in Figure 3.5. However, on average, an additional 150 MW of spin and 60 MW of non-spin capacity was flagged contingency-only for one of three reasons:

- First, even though 100 percent of the requirement was met in the day-ahead, additional capacity was procured in real-time due to either an increase in market requirement, units with day-ahead awards unable to provide that capacity in real-time, or it was more economical to procure additional spin to substitute for non-spin. In either case, any incremental capacity procured in real-time was automatically designated as contingency-only, represented by the purple and red bars in Figure 3.5.

Figure 3.5 Contingency Status of Spin and Non-Spin Procurement for Peak Hours 7 – 22



- Second, any incremental capacity procured in real-time from a unit with day-ahead non-contingent capacity in that service was automatically converted to contingent, shown by the orange bars in Figure 3.5.
- Lastly, the real-time ancillary service market has been over-procuring additional self-scheduled or zero dollar bid-in capacity above the minimum market requirement at no additional cost, represented by the red bars[27] in Figure 3.5.

[27] Not all capacity represented by the red bars is due to over-procurement. When additional capacity is needed in RTPD, some of that capacity may also be met by zero dollar bids.

With respect to spinning reserve, key trends illustrated in Figure 3.5 include the following:

- Most of the incremental contingency-only spin capacity was due to procurement of non-zero dollar bids. This was most often the result of three factors: (1) increased market requirement from day-ahead to real-time (e.g., due to a higher load forecast), (2) units with day-ahead awards not able to provide the capacity in real-time (e.g., due to a forced outage), and (3) procuring more spin to substitute for non-spin when economical.

- Less than 10 percent of the total spin capacity was day-ahead non-contingency capacity that was converted to contingency-only from the incremental procurement.

- Minimal amounts of spinning reserve were procured in real-time from additional self-schedules or zero dollar bids, indicating that over-procurement at no additional cost to the market was minimal for spin. As of September 23, additional logic was implemented in the RTPD market run to resolve the over-procurement issue and has since then been effective in this regard.

With respect to non-spinning reserve, key trends illustrated in Figure 3.5 include:

- Almost 78 percent of non-spin capacity was bid in the day-ahead market as contingency-only.

- Minimal amounts of non-spin were procured in real-time from non-zero dollar bids, indicating that typically either (1) most of the units with day-ahead awards were available in real-time, or (2) even with an increase in market requirement, most of the incremental megawatts were met with zero dollar bids or higher quality reserves.

- The incremental amount of non-spin procured in real-time did not tend to come from units with day-ahead non-contingent capacity, as indicated by the lack of orange bars in Figure 3.5.

- The over-procurement of capacity in real-time from additional self-schedules and zero dollar bid in capacity had the largest impact on non-spin, as shown by the red bars. While the red bars may also include incremental megawatts procured for legitimate reasons, on a day-to-day basis, the over-procurement of ancillary services was concentrated in non-spin as well as regulation down, which is not shown here.

There are two major concerns with real-time procurement of spin and non-spin: (1) the market rule which converts a unit's day-ahead non-contingent capacity to contingency-only capacity, and (2) the potential for over-procuring capacity above the minimum requirement. Both of these issues increase the amount of contingency-only reserves, restricting the amount of capacity available for imbalance energy in the event supply is tight (especially in constrained areas). Non-contingency capacity may be dispatched economically as energy in real-time, given the total operating reserves remain above 7 percent, while contingency-only capacity is reserved for system contingencies.[28] There can be situations in which the system or a constrained area is stressed due to high loads and tight supply, dispatching most or all available energy and non-contingency reserves. While some contingent reserves should be kept aside for contingencies and system emergencies, converting over 200 MW (150 MW spin and 60 MW non-spin) to contingency-only reduces the RTD bid stack, increasing the likelihood that higher priced energy will be dispatched sooner. Having that extra capacity not held up as reserves may, in some situations, mitigate the impact an unforeseen event has on the energy market in terms of prices. The latter issue was of major concern to the ISO, and, as of September 23, 2009, a new maximum requirement is being enforced in real-time to minimize over-procuring ancillary service capacity.

3.3 Relationship Between Ancillary Service and Energy Prices in Real Time

This section focuses on the relationship between ancillary service prices and energy prices in the real-time market, in the context of scarcity price signals communicated between the two products. One of the major enhancements of the ISO's new nodal market design is that it co-optimizes the procurement of energy and ancillary services. Co-optimization considers the lost opportunity cost of providing one product (energy or ancillary service) over the other when determining prices.[29] The benefit of co-optimization is that the market outcomes more closely reflect both the cost of production and the cost of providing one product in lieu of the other, and in doing so results in a more efficient least cost procurement of both products.

The scarcity pricing mechanism is designed to trigger higher prices in ancillary services when supply is insufficient to meet a minimum level of procurement, signaling the scarcity of the product through the (administratively set) higher market price. In addition to higher ancillary service prices, the scarcity pricing mechanism may also convey higher prices to the energy market through the co-optimization. In cases where both energy and ancillary services are scarce, the high price set in ancillary services through scarcity pricing can also influence the energy price, as the opportunity cost of not providing a high-priced scarce ancillary service is captured in the energy price at a pricing node where either could be procured.[30] Scarcity pricing will be executed in both the day-ahead market and the real-time market.

---

[28] It is DMM's understanding that operators do have the ability to dispatch contingency-only capacity for specific individual resources in cases when the ISO has an excess of reserves system-wide, but may need to mitigate very localized constraints that could be most effectively relieved by dispatching a relatively small quantity of contingency-only resources. However, DMM recognizes that it may be difficult in practice to implement or effectively utilize this ability due to the very dynamic nature of real-time conditions and the manual process of "switching" an individual unit's bids from contingency-only to non-contingent.

[29] For example, take a 100 MW unit that bid in 90 MW of energy at $20/MW and 20 MW of spin at $5/MW. If the unit is awarded 90 MW of energy and 10 MW of spin, there is no opportunity cost because the unit did not forego any bid in energy for spin capacity. However, if the unit was awarded 80 MW of energy with an energy LMP of $50 and 20 MW of spin, there would be a $30 opportunity cost equal to the difference of the energy LMP at that unit's PNode and its energy bid price ($30=$50-$20). Furthermore, assuming this unit was marginal for spin, the spin ASMP would be $35 = $5(marginal spin bid price)+$30(opportunity cost). For more discussion on and examples of the mechanics of co-optimization, see "Reserve Scarcity Pricing Design Revised Numerical Examples" at http://www.caiso.com/1f65/1f65dabe49d90.pdf.

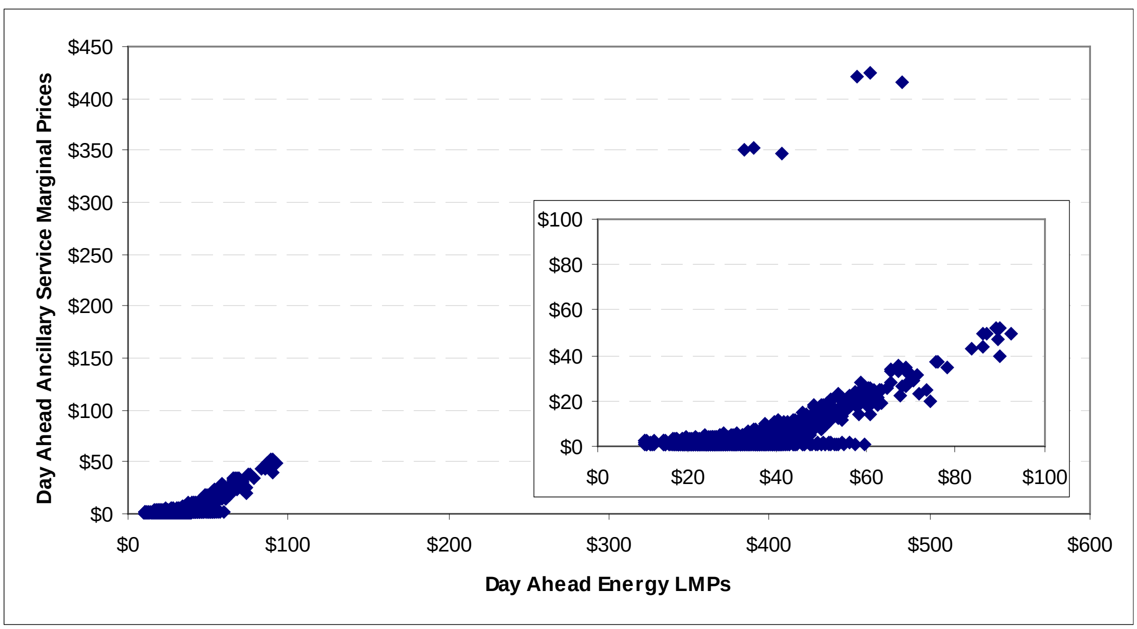
[30] For a more detailed discussion of the ISO's scarcity pricing proposal, see "Final Draft Proposal Reserve Scarcity Pricing Design" at http://www.caiso.com/243e/243ecc4d2d490.pdf.

The remainder of this section examines the relationship between the ancillary service price and the energy price in the real-time market prior to implementing scarcity pricing, for the purpose of better understanding how higher prices in ancillary services may be transmitted to the energy market through co-optimization in real-time. Before proceeding, some additional discussion of the scarcity pricing mechanism is necessary.

Ancillary services are procured, and co-optimized with energy, in both the day-ahead and real-time markets. The co-optimization in the day-ahead market is somewhat simpler, since the day-ahead market has only one financially binding market run where both energy and ancillary services are procured together. The real-time market for these products is slightly more complicated in this regard since ancillary services and energy are procured in two separate market runs.

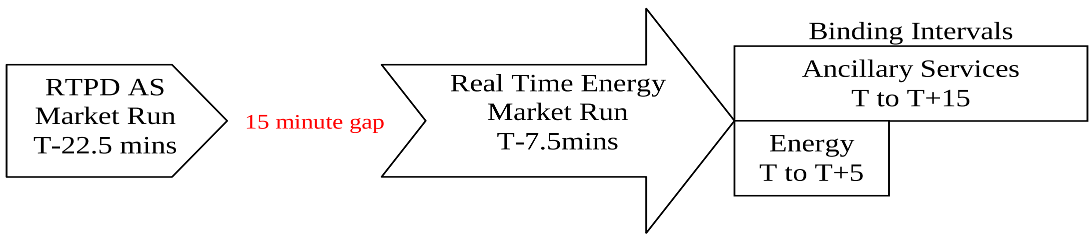
A consequence of the co-optimization is that we would expect to see a positive correlation between ancillary service prices and energy prices at nodes where ancillary services are procured. First we review the day-ahead market prices, where co-optimization of energy and ancillary services is done within a single market run where both products are financially binding. Figure 3.6 below shows the correlation of energy and ancillary services from the day-ahead market. The prices represented in Figure 3.6 are taken at PNodes where resources sold ancillary services, and reflect the corresponding PNode energy LMP used in calculating the lost opportunity cost of providing ancillary service. For the day-ahead market run, there is a strong positive correlation between the two sets of prices; when energy prices increase, the prices of ancillary services also increase, reflecting the lost opportunity cost of providing one product over the other. The opposite is also true for an increase in ancillary service prices.

Figure 3.6 Day-Ahead Ancillary Service Marginal Prices versus Energy LMPs



While 100 percent of the expected ancillary service requirements are procured in day-ahead, a small amount of incremental ancillary services are procured in real-time, if needed, for 15-minute intervals in the real-time pre-dispatch run (RTPD) which occurs 22.5 minutes prior to the binding interval (see Figure 3.7 below for timeline). The RTPD market run has all the final real-time bids and performs a full energy and ancillary service optimization, including co-optimization between the two products; however, only the ancillary service procurement is financially binding from this market run. The energy procurement from RTPD is not financially binding, except for hourly HASP intertie schedules. The real-time dispatch (RTD) market is deployed 7.5 minutes prior to the binding interval, resulting in binding energy schedules for a 5-minute period. When running the RTD, the awarded ancillary services are already determined. Non-contingency reserves may be dispatched to energy while reserves flagged as contingency will not be converted to energy until an event or operator has activated contingency reserves. The disconnect in real-time between procurement of ancillary services in RTPD and procurement of imbalance energy in RTD weakens the link between these two products and impedes price signals reflecting higher cost or scarcity in one product from being reflected in the other product.

Figure 3.7 Real-Time Markets Timing Diagram



For example, the binding results for the third 5-minute interval in hour 2 are comprised of ancillary service awards that were generated in hour 1, while the energy awards were generated at the beginning of hour 2, with a 15 minute gap between the market runs. Furthermore, the ancillary service awards are also binding for the two subsequent 5-minute intervals while new energy schedules are generated from two additional real-time market runs.

Through co-optimization, we expect to see a high correlation between ancillary service and energy prices within the RTPD market run. Figure 3.8 shows the correlation between real-time ancillary service prices and the non-binding RTPD energy prices resulting from co-optimization for July, August, and September 2009. There is a strong positive correlation between RTPD energy LMPs and upward ancillary service[31] ASMPs, especially during times when both products are competing for the capacity. The data enclosed by the larger red oval represent instances in which the two products are competing for capacity, and the Lost Opportunity Cost (LOC) is reflected in both prices; it would be during these situations we would expect scarcity pricing to be triggered.

The smaller red oval encircling the data along the horizontal axis shows instances where the high energy price did not strongly impact the ancillary service prices. Based on a review of

[31] Only upward ancillary service prices are shown because the lost opportunity cost for regulation down has a slightly different calculation due to the nature of regulation down.

these potentially anomalous cases, DMM has identified several reasons why this pricing pattern can occur:

- ***Ramp Constraints****.* Due to ramping constraint for a unit in real-time, a unit's full energy bid may not be not available in RTPD. However, a unit may still be able to provide some ancillary service capacity over the additional ramping horizon upon which the unit's available ancillary service capacity is calculated. In these instances, the additional capacity was awarded to ancillary services, but did not receive LOC because that capacity could not be used as energy.

- ***Ancillary Service Capacity Not Bid in As Energy****.* Some Resource Adequacy (RA) unit units that were not under an RA obligation for their full capacity (referred to as "partial" RA units) may bid energy and ancillary services such that the sum of the maximum megawatts for both products would not be more than the unit's entire capacity. In this case, the products would not have to compete for capacity and no LOC would be calculated (i.e., since capacity bid in as ancillary services was not simultaneously bid in as energy).

- ***Maximum Limits Established by Exceptional Dispatch***. In some cases, DMM found that units with a maximum generation limit established through an Exceptional Dispatch (ED) were awarded energy up to the maximum ED limit, with the remaining capacity being awarded as ancillary services with no LOC. There is no LOC because the capacity beyond the maximum ED value is protected with a penalty price greater than the energy LMP. However, the capacity awarded as ancillary services would not be available in a real-time contingency run as dispatchable energy because the ancillary service capacity is beyond the maximum ED value.

- ***Congestion****.* In some case, a unit's day-ahead energy schedule may be backed down to help relieve congestion. The amount of capacity backed down was then awarded as ancillary services and no LOC was calculated.

DMM is conducting further review of potentially anomalous results such as those described above to more fully understand the cause for such results and determine if any refinements to the software may be appropriate to address any scenario that might cause inappropriate ancillary service prices.

The smaller circles will be discussed in the context of Figure 3.9. The correlation in RTPD is not as linear as in the day-ahead market for a few reasons. First, procurement in real-time pre-dispatch is generally low since 100 percent of the requirement is procured in day-ahead; therefore, the market clears at lower bid-in prices on average. Secondly, due to a lower demand for upward ancillary service capacity in real-time, there tends to be ample supply from units with low bid prices and no opportunity cost; therefore, there is no opportunity cost reflected in the ancillary service price. Overall, when energy prices increase past a minimal level in the RTPD market run, more opportunity cost for energy is reflected in the ancillary service market clearing price.

Figure 3.8 Comparison of Real-Time Pre-dispatch Ancillary Service Marginal Prices with Energy LMPs

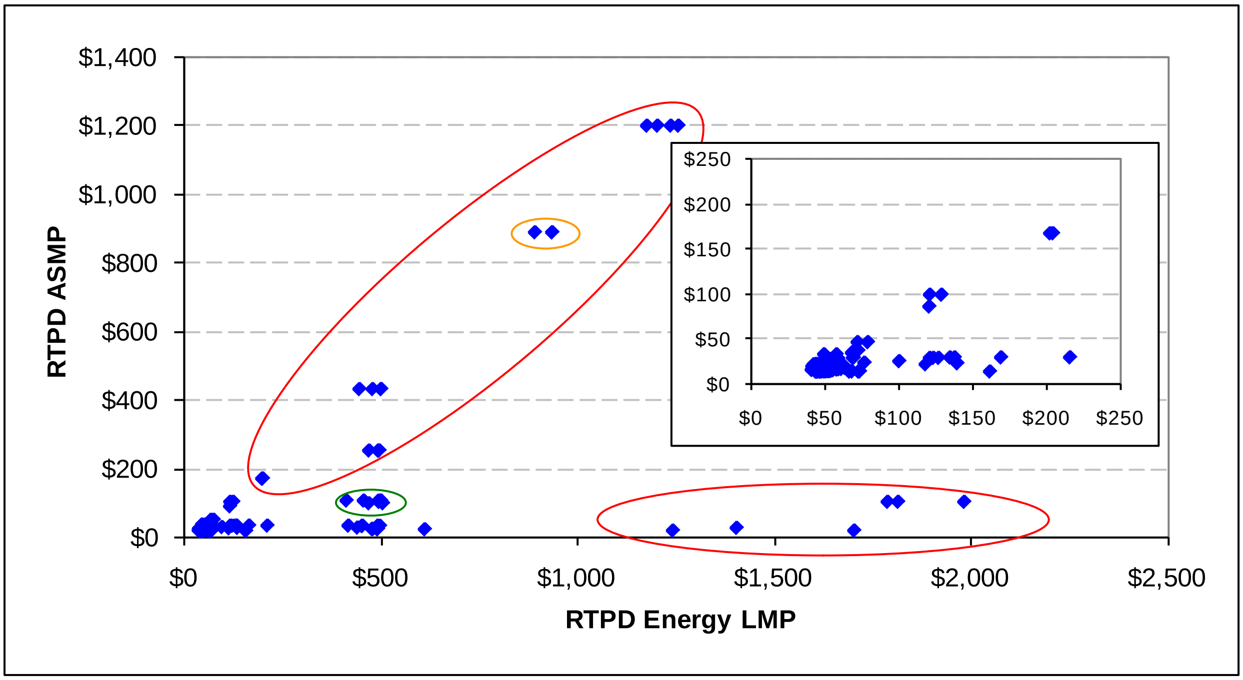
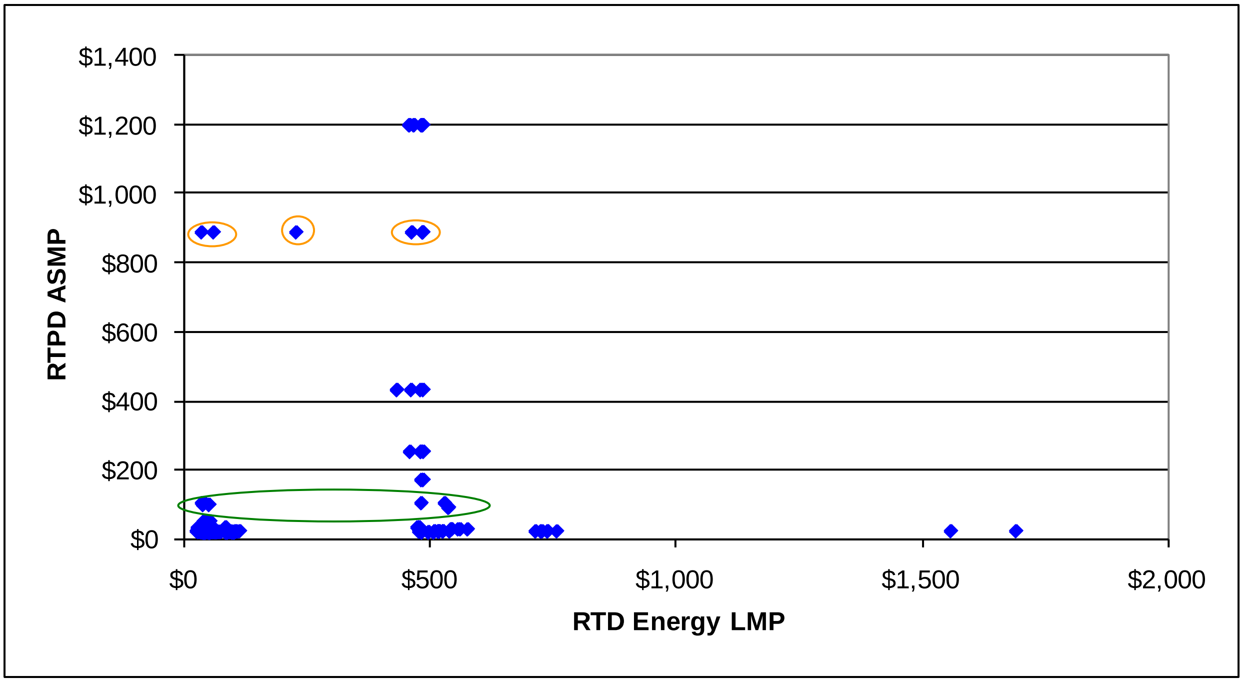


Figure 3.9 below shows the RTPD ancillary service prices plotted against the RTD energy prices at PNodes where ancillary services were procured in RTPD. This combination reflects the financial settlement prices for real-time ancillary services and energy. There is a lack of correlation between the two sets of prices, indicating the lost opportunity cost of providing ancillary services, in terms of foregone net revenues from RTD energy sales not being reflected in the price of ancillary services. Conversely, higher prices in RTPD ancillary service procurement in RTPD are not being reflected in the RTD energy prices. For example, the data points enclosed by the orange circle in Figure 3.8 show RTPD ancillary service and energy prices of $900/MW and $850/MWh respectively. In RTD, the ancillary service prices remained the same, but the RTD energy prices ranged from $50/MWh to $500/MWh, which are shown in Figure 3.9. Similar scenarios hold true for the other circled data points, further emphasizing the disconnect between RTPD and RTD energy prices.

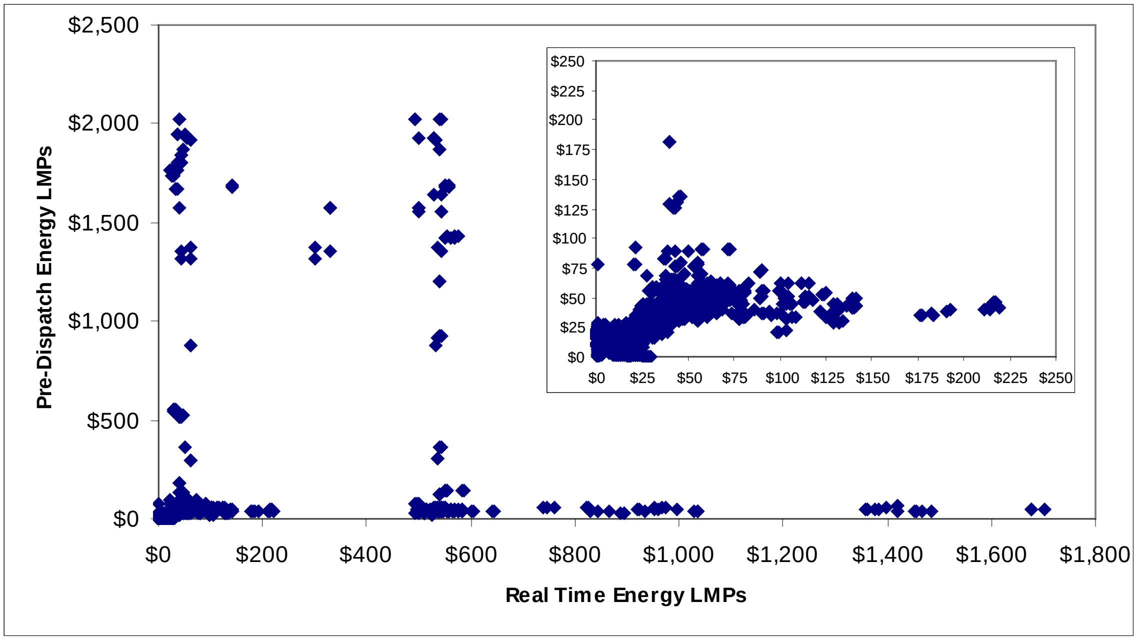
Figure 3.9 Comparison of Real-Time Pre-dispatch Ancillary Service Marginal Prices with Real-Time Energy LMPs



Since go-live there have been few high prices for ancillary services in RTPD relative to the frequency of higher real-time energy prices. This is in part due to the fact that 100 percent of the ancillary service requirement is met in day-ahead, reducing demand for ancillary services in real-time. In addition, when there has been moderate demand for ancillary services in real-time and high energy prices, enough capacity has been bid in at relatively low prices at PNodes with lower energy LMPs.

Despite the lack of direct linkage between ancillary service prices in RTPD and energy prices in RTD, we would expect to see some correlation in prices between the two market runs given the similarity in system conditions, bids considered, and market structure as well as the close proximity in time. Figure 3.10 below shows the energy LMPs of the three LAPs from RTPD and RTD market runs for July 2009. There is a lack of correlation between the two sets of energy prices when prices are greater than $50/MWh in either market, further emphasizing the disconnect between the two markets especially during times of system stress and/or higher demand.

Figure 3.10 Correlation of Energy LMPs – Real-Time Pre-dispatch (15-min) and Real-Time (5-min) LMPs



There are a few explanations for why these two markets diverge:

1. The priority of constraints between RTPD and RTD differ, which is reflected by different protective penalty prices. Therefore the same constraints may not be violated, and, furthermore, the same violated constraint will produce a different price. Also, RTPD is able to commit units while RTD cannot.

2. There is a difference in the time horizons and what the market is able to see 22.5 minutes out versus 7.5 minutes out. Unforeseen events may occur in real-time and are able to be modeled in the software for the real-time run but not the real-time pre-dispatch run, resulting in divergent prices.

3. Different load forecasts are used.

4. Transmission biases may vary (see section on transmission biasing and un-enforcing in this report for more discussion on this topic).

5. The real-time market often sees intermittent price spikes due to a lack of ramping capability. The real-time pre-dispatch market uses a 20-minute ramp while the real-time only a 10-minute ramp; therefore, the ramping constraint is less likely to bind in the real-time pre-dispatch market. Please see the section on price convergence in a prior section of this report for a more detailed discussion.

Due to the divergence of real-time pre-dispatch and real-time energy prices, the binding ancillary service prices may not be reflecting the true opportunity cost of providing energy in RTD. Another way of stating this is that energy and ancillary services are not co-optimized in real-time across their respective financially binding markets. This has implications when considering implementation of a scarcity pricing mechanism that is intended to measure scarcity from ancillary service procurement and communicate that scarcity through price linkage into the imbalance energy market. The current proposed design for scarcity pricing is to trigger high ancillary service prices when there is a regional shortage of supply to meet market requirements in RTPD. It further states that both the ancillary service and energy prices will rise during true scarcity as a result of the market co-optimization. However, as a result of the temporal disconnect between these two markets, the financially binding energy prices in RTD will not directly reflect the scarcity pricing of ancillary services in RTPD when triggered.

As previously mentioned, the current scarcity pricing proposal is designed to trigger scarcity when the regional procurement in real-time pre-dispatch is less than the minimum requirement. In practice, there are two possible ways to trigger scarcity: off the real-time pre-dispatch procurement, or when the actual operating reserves in real-time dip below the 7% or 5% NERC standard. Each method has one major drawback; neither of them allows the market to increase energy prices when scarcity is triggered. Triggering scarcity in real-time pre-dispatch will not simultaneously increase the energy prices due to the disconnect between the two markets and prices previously discussed. While real-time measurements of actual operating reserves may more accurately identify periods of scarcity, by the time this measurement occurs the energy and ancillary service prices will have already been calculated and thus will not reflect the cost of the shortage.

Finally, it should be noted that in RTD, there are numerous other penalty price mechanisms which – in effect – trigger scarcity pricing for energy when shortages of energy occur in the real-time markets. Thus, while under an ideal market design scarcity pricing for ancillary services and energy may be directly linked, the current market design does incorporate elements of scarcity pricing for both energy and ancillary services.

3.4 Recommendations

To mitigate the potential impact of contingency-only reserves on energy prices, DMM recommends that the ISO:

- Consider accommodating both contingent and non-contingent reserve from an individual resource in the market software so that day-ahead awards are not automatically converted to contingent reserve when additional capacity is purchased in real-time.

- Continue the practice of allowing operators to dispatch limited quantities contingency-only capacity for specific individual resources in cases when the ISO has an excess of reserves system-wide, but may need to mitigate localized constraints that could be effectively relieved by dispatching a relatively small quantity of contingency-only resources. We recognize that opportunities for this may be limited due to the highly dynamic nature of this type of situation and the manual nature of the process of removing the contingency-only designation for a unit.

- Consider imposing a limit on the amount of contingency-only ancillary service capacity that can be procured to ensure reserves are not held unless a system contingency is declared. However, we recognize that this approach could raise ancillary costs and could not be highly effective at targeting cases where a relatively small reduction contingency-only reserves within a load pockets could provide effective relief of a very localized constraint.

Finally, to improve price signals between real-time ancillary services and energy under scarcity pricing, DMM suggests that the ISO:

- Consider – as a longer term design change – implementation of procurement of ancillary services in the 5-minute RTD market co-optimized with energy. We recognize that this may represent a significant software design change, and that the priority placed on making such a change would need to be based on consideration of the benefits and costs of such a change relative to other potential market and software enhancements.

4 Exceptional Dispatch

Exceptional Dispatch (ED) is a term used to describe manual dispatches performed by an ISO operator in cases where unit commitments and energy dispatches made by the market software did not fully address a particular reliability need. Exceptional dispatches are generally considered to be an undesirable but necessary feature of ISO operations, as they reflect operating constraints that cannot be fully enforced within the automated economically dispatched market, and thus require manual intervention. The reasons that EDs are necessary were explained at length in DMM's previous *Quarterly Report* (Q2 Report*).*[32]

In this chapter, we provide an updated analysis of ED trends in Q3,[33] and follow-up on three specific recommendations made by DMM in our Q2 report. In Q3, the ISO took a number of steps to reduce the use of ED through changes in software and operational practices. These actions reduced EDs starting in July through August. However, the amount of capacity committed via EDs increased again in late August and September due to other factors, including several major fires and a major transmission outage in Southern California.

4.1 Summary of Exceptional Dispatch

On a day-ahead basis, the ISO issues EDs to commit non-short start units at their minimum operating level for the next operating day. These day-ahead commitments may be made either before or after the day-ahead IFM and RUC processes are completed. In the real-time market, the ISO also issues EDs for additional energy (above minimum load). These real-time EDs for energy can be issued to units committed on a day-ahead basis through ED, as well as units that are committed through self-schedules or the day-ahead market process. The following charts show day-ahead and real-time exceptional dispatch frequencies in Q3.

[32] *Quarterly Report on Market Issues and Performance*, July 30, 2009; covering April through June, 2009. http://www.caiso.com/2425/2425f4d463570.html

[33] This Q3 report builds on the analysis presented in DMM's Q2 Report, and employs a similar, and occasionally refined, methodology. Thus, we note where the methodology has substantively changed from that described in the Q2 report. Charts and underlying data are based on the best-available information and methodology and supersede any previously-released materials. Underlying data and methodologies are continuously subject to review and improvement to ensure the most accurate information is provided.

Figure 4.1 Frequency of ED for Day-Ahead Commitment by Reason (Q3 2009)

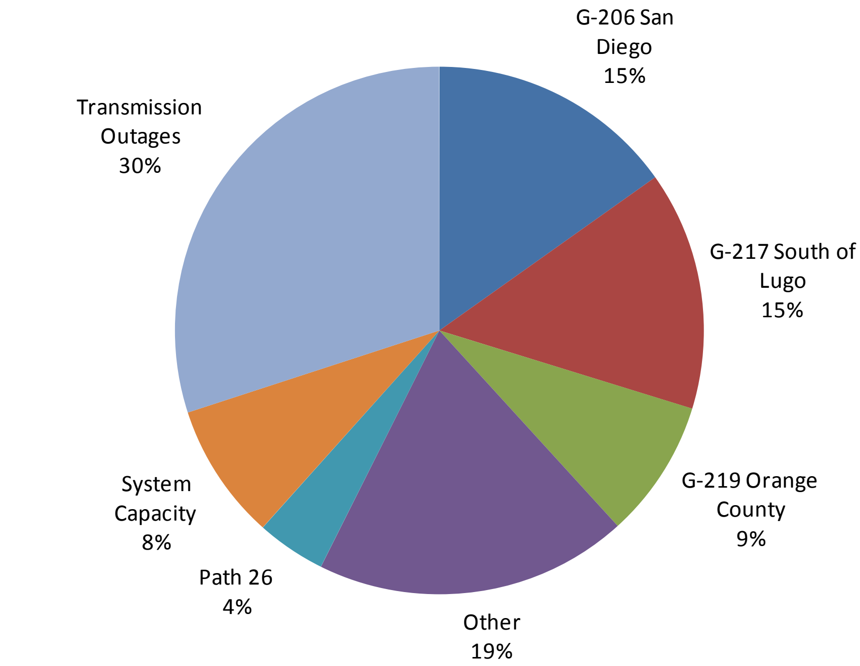
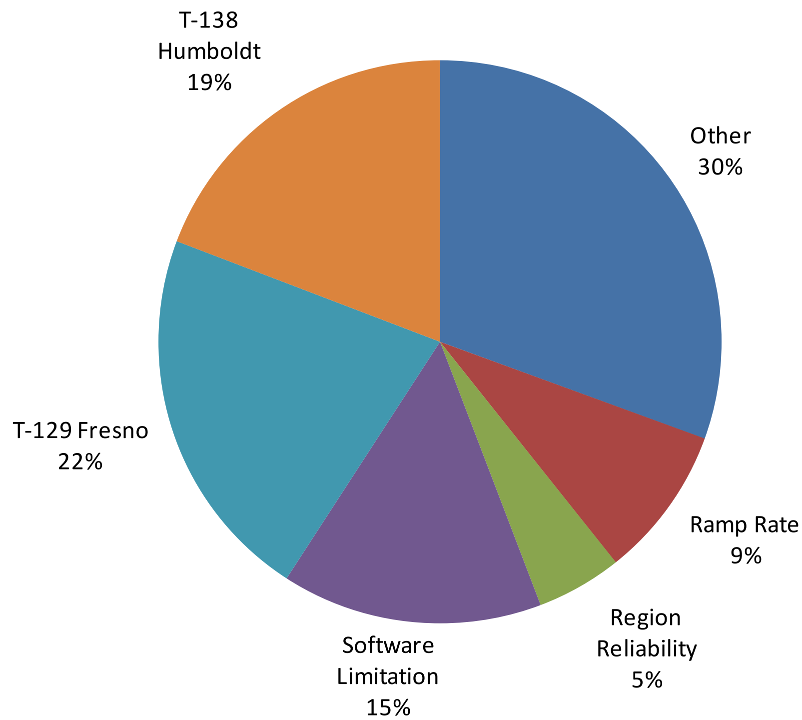


Figure 4.2 Frequency of ED for Real-Time Energy by Reason (Q3 2009)



During Q3, the ISO issued approximately 779 distinct day-ahead ED commitments, compared to approximately 523 day-ahead ED commitments issued in May and June. About 298 ED commitments made in Q3 were issued between July 1-26. ED commitments decreased after July 27, as two of the generation procedures that have required a substantial portion of ED unit commitment were incorporated into the RUC market through capacity constraints (G-217 (South-of-Lugo) and G-219 (Orange County)), but increased in frequency during the forced SWPL outage in September. Figure 4.1 summarizes the portion of day-ahead unit commitments made in Q3 via ED for different specific reasons. As shown in Figure 4.1:

- The G-217 (South-of-Lugo) and G-219 (Orange County) operating procedures account for 15 percent and 9 percent of ED commitments for the entire quarter, respectively, but accounted for approximately 178 (59%) of the 298 commitments prior to July 27. As discussed in this chapter, the ISO implemented capacity nomograms reflecting these constraints in the RUC process on July 26, 2009.

- Approximately 15 percent of ED commitments were issued for the G-206 (San Diego) operating procedure. The ISO is currently developing a RUC nomogram to reflect this third major source of ED commitments, and expects to implement this in the near future.

- Transmission outages accounted for approximately 30 percent of day-ahead unit commitments via ED. A major cause for these commitments was an outage on the Southwest Power Link (SWPL) from September 11 until September 24. SWPL is a 500kV transmission line that connects the Hassayampa and North Gila substations in Arizona, and serves as a key conduit of power from generation in the Southwest to the San Diego area. Because the Palo Verde line is used as a contingency for SWPL, the total transfer capability into California from the Southwest was derated by approximately 2000 MW. The outage also limited other transfer capabilities in the region and had unique voltage requirements that were not included in historically established nomograms. This required daily commitment for the duration of the outage of approximately 11 individual resources on average, depending on daily peak load.

- Commitments to protect against contingencies related to Path 26 and system level conditions respectively accounted for approximately 4 and 8 percent of ED commitment events.

During Q3, the ISO issued approximately 448 distinct real-time ED energy dispatches, compared to approximately 679 distinct ED energy dispatches in May and June. Figure 4.2 summarizes the portion of real-time ED for energy made in Q3 for various specific reasons. As shown in Figure 4.2:

- The leading reason for ED energy dispatches (by frequency of units per day) was the T-129 procedure for the Fresno area. This nomogram is restricted, so a detailed discussion of the cause for these EDs is not provided. However, most of the dispatches are relatively small – typically in the neighborhood of 5 MW or less.

- The second leading reason for ED energy dispatches (by frequency of units per day) was the T-138 procedure for the Humboldt area. This small load pocket in the northwest corner of California is connected to the rest of the grid by a 60kV transmission line, which is frequently congested and must be relieved by relatively small real-time ED energy dispatches. Modeling this region also poses unique challenges, as real-time conditions are not always visible to the ISO. The dispatch of resources within Humboldt County continues to challenge grid operations, as has been the case since Reliability Must-Run (RMR) contracts supporting the area expired at the end of 2007.

- Another leading reason for real-time ED energy dispatches was to move generators committed either by ED or RUC to dispatchable levels above minimum load (Ramp Rate). One of the major events contributing to this category of ED for real-time energy in Q3 was the Station Fire. This fire also contributed to ED for real-time energy logged as being attributable to Path 26 and Regional Reliability.

- A significant portion of real-time EDs are logged as "Software Limitations". These typically refer to instructions that correct unusual resource-specific issues. Such limitations might cause the market software to view a resource as being in a state that differs from its actual condition, such as a unit that cannot be instructed by the market because it is not equipped with the proper automation technology, or a pump storage unit whose state (generation or pump) cannot be determined by the market accurately. These are discussed in greater detail in the *Q2 Report.* Another category of software limitation that was prevalent in late July to early August includes EDs used to "bridge" across daily RUC commitments. After G-217 and G-219 were implemented in RUC, operators observed that the RUC algorithm frequently turned resources on at approximately HE 6:00, and/or shut resources down late in the evening, as they were needed primarily during peak hours. Because an operation day's RUC market run is independent of the following day's RUC market run, the following day's RUC optimization recognized a shut-down unit's minimum down time, and thus often started a different resource. This frequent startup and shutdown pattern is costly and imposes wear and tear on resources. Operators thus began the practice of keeping units on overnight to "bridge" across the two different market runs, in order to avoid costly shutdowns and startups. These bridging EDs typically are included in the reason category of "software limitation" because they reflect the inability of the market software to optimize across inter-temporal market runs.

- The large real-time "Other" category includes any reasons totaling less than 5 percent of the quarterly total real-time ED volume. The largest single reason in the "Other" category was the T-170 Mirage-Tamarisk nomogram, a restricted procedure covering facilities in the Coachella Valley area.

4.2 Exceptional Dispatch Trends

Figure 4.3 and Figure 4.4 summarize weekly trends in ED unit commitments in Q2 and Q3 in terms of the average number of day-ahead unit commitments and total average hourly minimum load energy of these units, respectively. A discussion of these weekly trends is provided below:

- Day-ahead unit commitments via ED and the minimum load energy of these units decreased substantially the week starting July 26, which corresponds to the point at which the ISO implemented capacity nomograms in the RUC process that reflect capacity needs incorporated in the G-217 and G-219 operating procedures.

- In the last week of July, the ISO also formed an ED "strike team" to focus on potential improvement to practices and software to reduce EDs, particularly with respect to unit commitments made in the day-ahead timeframe.[34] This strike team also focused on improving the consistency and logging of ED data, and providing more accurate and timely feedback on

[34] See Memorandum to ISO Board of Governors, Detmers, September 2, 2009, *Re: Briefing on Exceptional Dispatch*, http://www.caiso.com/241e/241eb60ca5f0.pdf

ED trends to operations staff. The team also monitored the impacts of new RUC capacity nomograms designed to meet reliability requirements previously met by committing additional units via ED either before or after the IFM. These efforts – combined with the new RUC capacity nomograms – appear to have reduced EDs in late July and August.

- As shown in Figure 4.4, minimum-load energy of units committed for G-206 (San Diego) also decreased at approximately the same time, because units committed in July and August had lower minimum loads, and a greater proportion of units committed were covered by RMR contracts and thus were not subject to ED tariff provisions.

- ED unit commitment and minimum load energy increased at the end of August due to several fires that affected transmission primarily in Southern California, and again in mid-September following a forced outage of the Southwest Power Link (SWPL) between Hassayampa and North Gila in Arizona. SWPL is a primary conduit of power into San Diego, and its outage also caused a de-rate of the Palo Verde branch group, affecting Southern California generally, as reflected in the increase in commitments for transmission outages in mid-September.

Figure 4.3 Frequency of Day-Ahead Exceptional Dispatch by Reason

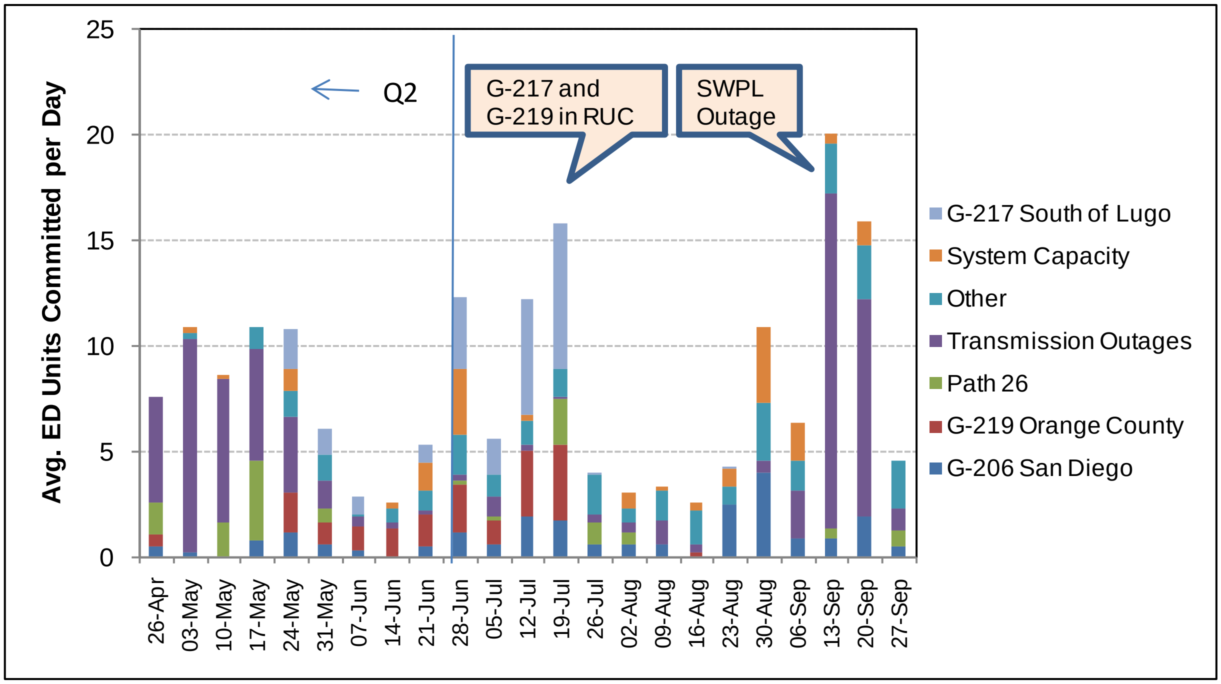


Figure 4.4 Average Energy Volume of Day-Ahead ED by Reason

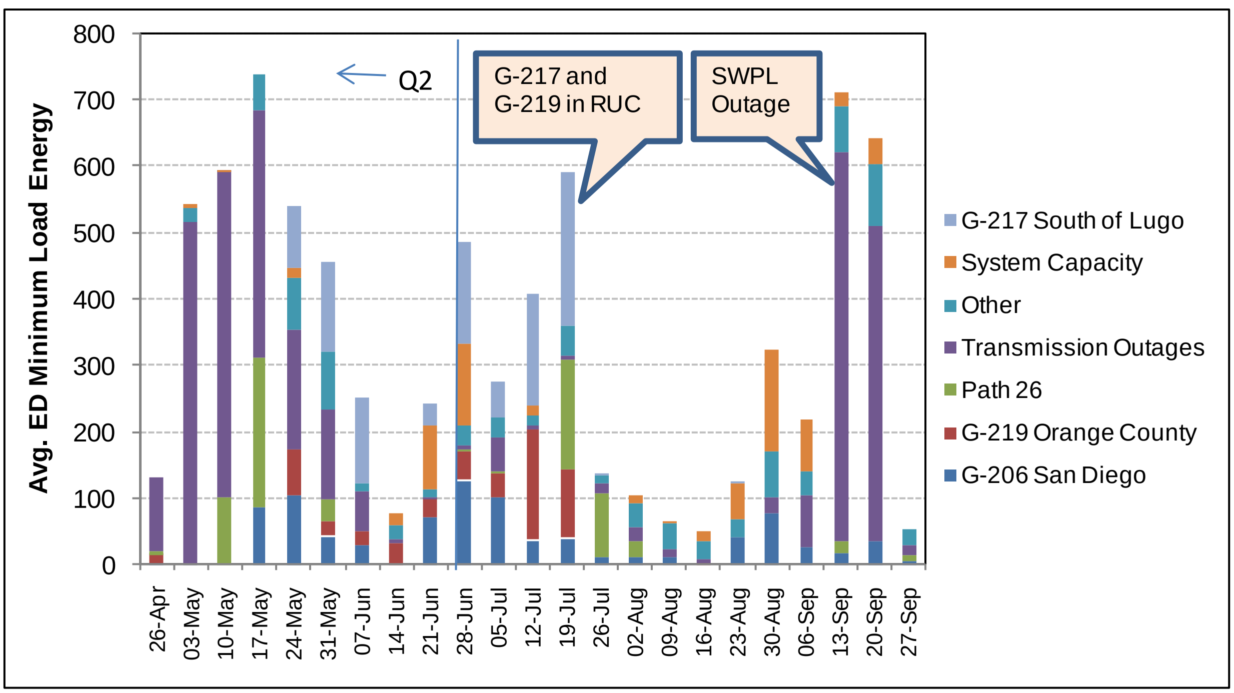


Figure 4.5 and Figure 4.6 summarize weekly trends in ED for real-time energy (above minimum load) in Q2 and Q3 in terms of the average number of units dispatched for real-time energy via ED and the total average volume of real-time ED energy, respectively. As shown in Figure 4.5 and Figure 4.6, there was a notable increase in ED for real-time energy during the Station fire in Southern California. Other overall trends in Q3 include the following:

- The single most frequent cause of ED for real-time energy during the quarter was the T-129 Fresno area nomogram, representing approximately 22 percent of the number of dispatches, but only 6 percent of energy.

- The T-138 procedures (Humboldt County area) represented approximately 19 percent of individual dispatches, and approximately 20 percent of energy volume.

- ED instructions to units committed and operating at minimum load to move them to dispatchable ramp rates represented 9 percent of individual dispatches, but accounted for approximately 15 percent of energy volume.

- Region reliability is a code primarily used to manage around fires affecting transmission and local reliability issues. The category accounted for about 5 percent of dispatches but 35 percent of ED energy volume in Q3, approximately half of which occurred on July 19. The small number of dispatches on this day were in response to a local inaccurate weather forecast that coincided with a transmission outage that limited resources in a load pocket within Southern California.

Figure 4.5 Frequency of Real-Time Exceptional Dispatch Energy by Reason

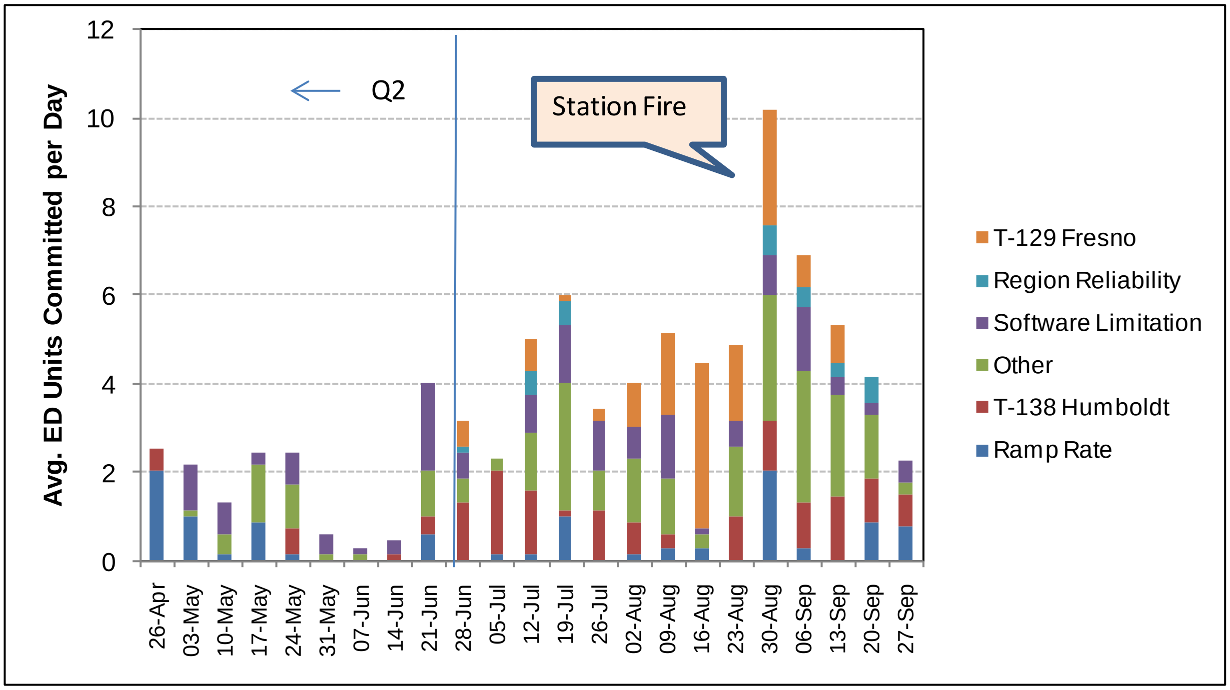
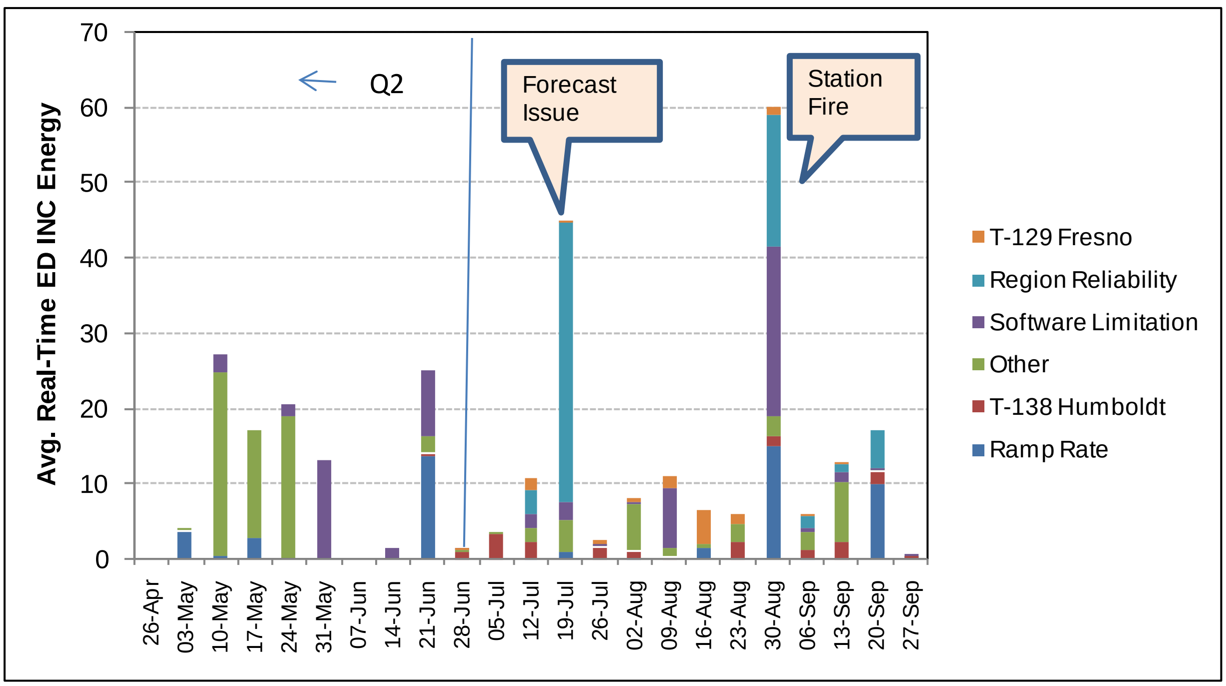


Figure 4.6 Average Energy Volume of Real-Time ED Energy by Reason



4.3 Market Impact of Exceptional Dispatch

It is not possible to estimate directly the market impact of exceptional dispatches, because such an analysis would require the knowledge of the counter-factual of what prices should have been had the reliability constraints necessitating ED been incorporated into the market model. If the constraint information necessary to perform such analysis were available, these constraints would already be modeled by the software, obviating the need for the ED in the first place.

We can, however, discuss the potential impacts conceptually. Since most day-ahead EDs are limited to committing units to their minimum operating levels in the day-ahead market, and such minimum-load energy is not eligible to set price under any case, the market impacts of such dispatches may not be significant. As long as there is a well-founded reliability need for having a unit on-line, the market outcome from having the operator manually force it on in the day-ahead market may not be appreciably different from what would occur if the constraint leading to this action was in the market causing the market to dispatch the unit automatically. However, to the extent exceptional dispatches are overly conservative or the reliability criteria driving the ED can be met by different combinations of unit commitments, having the constraint in the market model and managed by the market optimization would likely produce a more efficient and different market outcome. In this case, prices would be different but not necessarily higher relative to the ED case. Exceptional dispatches for energy above the minimum operating level of the units, which are limited to the real-time market, can distort and suppress market prices if the original market bids for this energy are at or above prevailing market prices. In this case, having the reliability constraint driving these EDs incorporated into the market model would likely lead to higher LMPs at the location of the ED and the surrounding area that defines the reliability constraint.

Figure 4.7 through Figure 4.10 depict hourly profiles of monthly average ED energy by market – day-ahead pre-IFM, day-ahead post-IFM, and real-time – versus actual load, for June through September 2009. Day-ahead ED represents minimum-load energy from ED unit commitment. In addition, the real-time ED energy is separated by in-sequence and out-of-sequence energy. (Other figures in this chapter depict only out-of-sequence energy volume, as this is the quantity that represents energy that is not also dispatched by the market.)

Figure 4.7 Hourly Average Volume of ED Energy vs. Load (June)

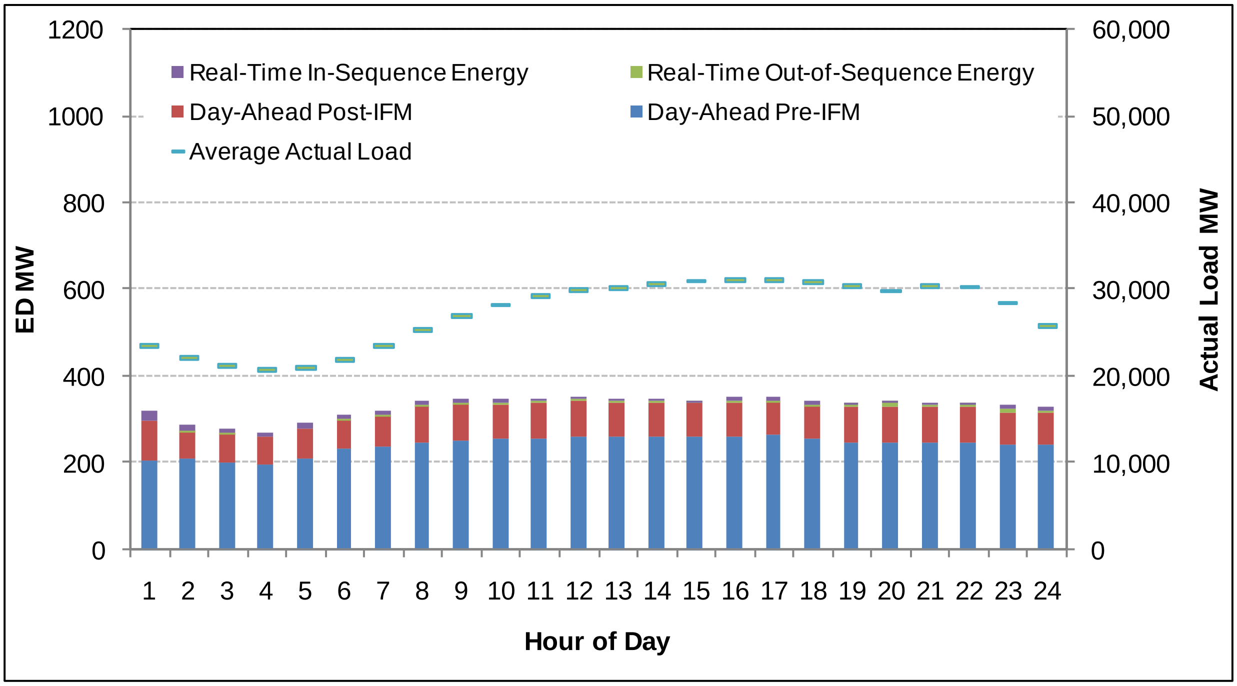


Figure 4.8 Hourly Average Volume of ED Energy vs. Load (July)

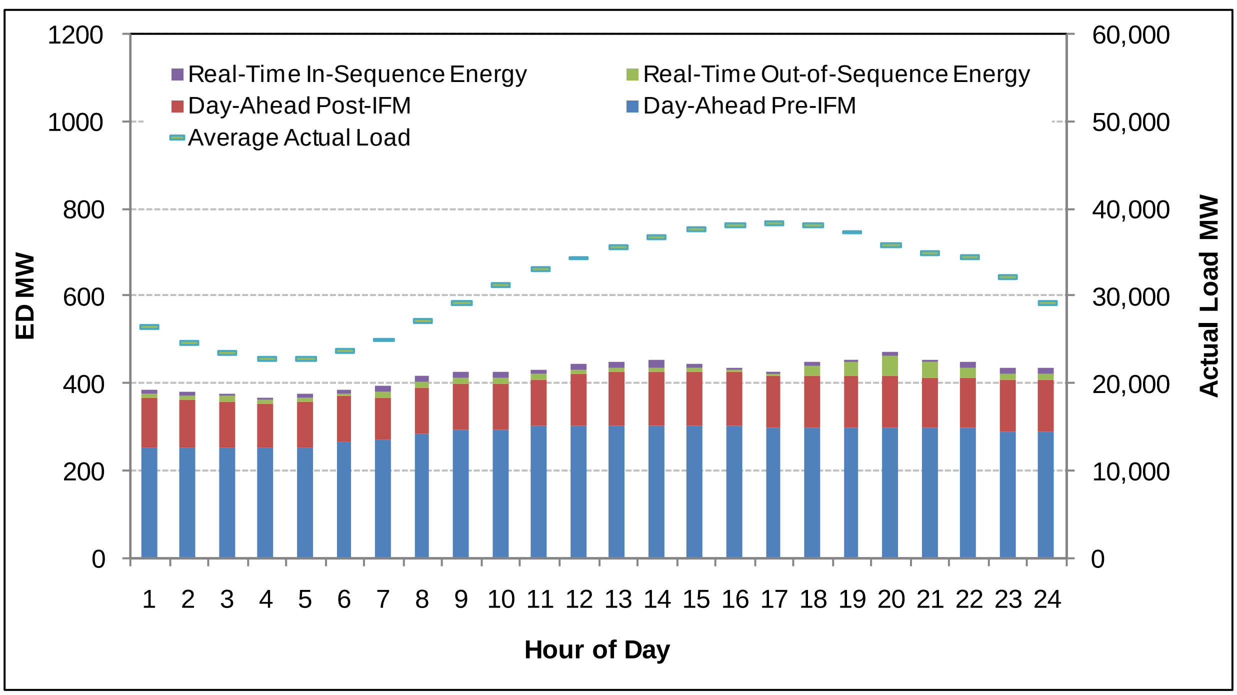


Figure 4.9 Hourly Average Volume of ED Energy vs. Load (August)

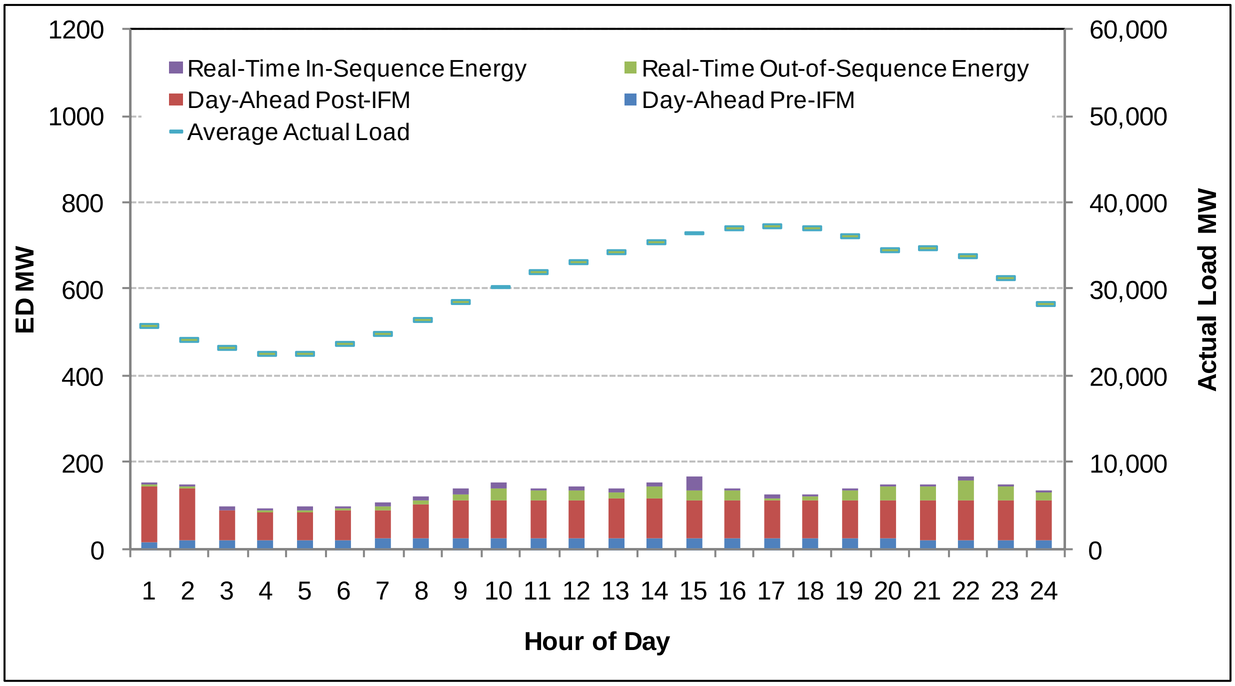


Figure 4.10 Hourly Average Volume of ED Energy vs. Load (September)

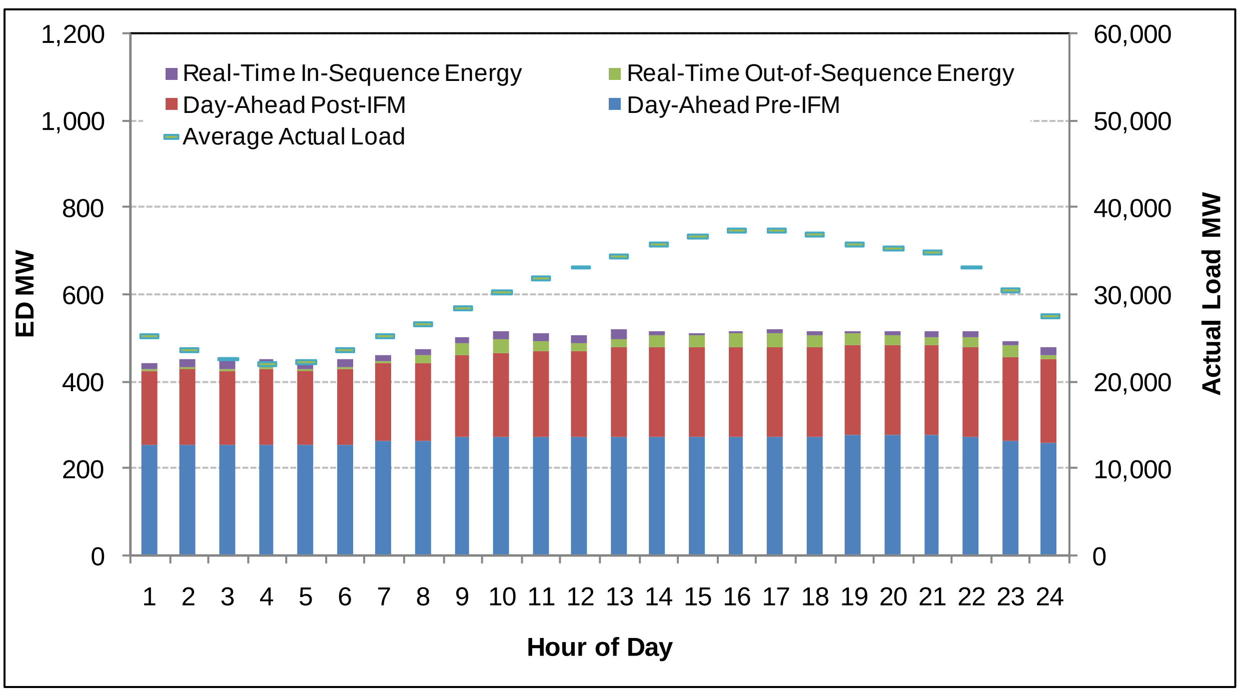
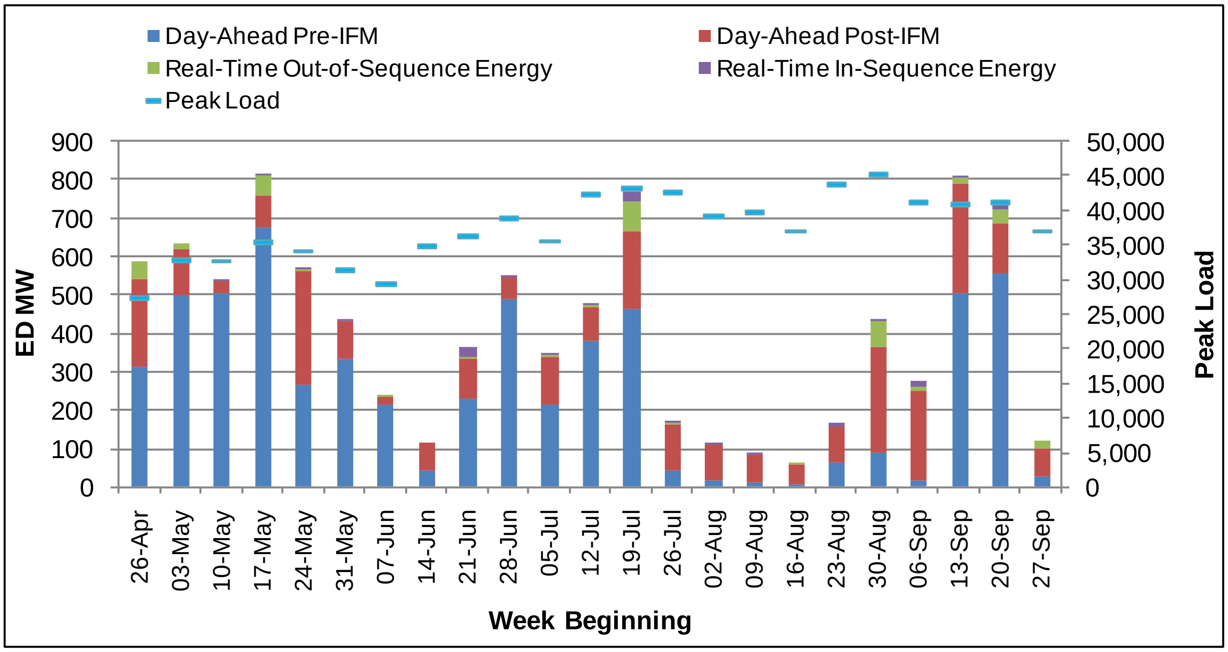


Figure 4.11 provides an average profile of ED volume dispatched during Hour Ending 18, for each week between June and September. This shows that total ED continues to represent a relatively small proportion of total generation. In 97 percent of hours with ED, the total ED volume (day-ahead commitment plus real-time OOS energy) serves less than four percent of load, and in all hours with ED the total ED volume serves less than seven percent of load. The hours in which ED proportion was high typically were at night, when loads are low and units at minimum load represent a larger share of total generation; and in emergency situations. The highest proportion occurred after a Path 26 conductor sagged, forcing a 50 percent derate, on August 7. Several large combined-cycle resources were dispatched up to their maximum generating levels for several hours until the following early morning hours. On a weekly basis, ED proportion was highest during the SWPL outage in mid-September.

Figure 4.11 Weekly Average Volume of ED Energy and Peak Load (HE 18)[35]



4.4 Market Participation by Units committed via Exceptional Dispatch

4.4.1 Unit Loading

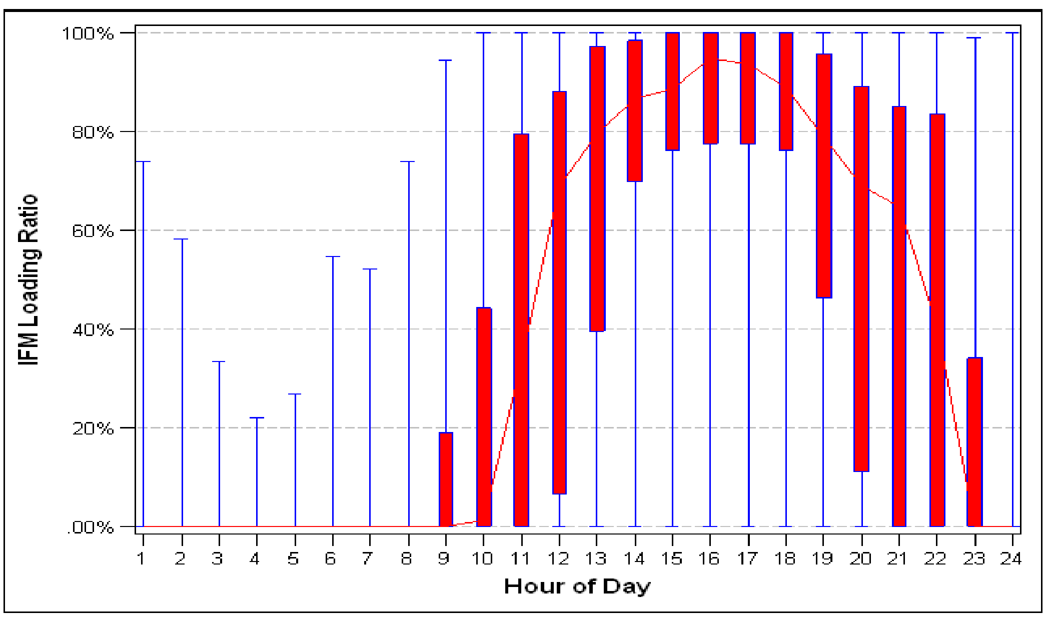
The “Loading Ratio” metric presented in DMM’s Q2 report provides an indicator of market scheduling of resources committed prior to the market in the day-ahead by ED. It is useful in determining the extent to which resources that are committed through ED also benefit from revenue opportunities in the IFM by selling energy above their minimum load. Specifically, it compares hourly IFM schedules of pre-IFM committed units (net of minimum load) to their total resource capacity in each hour of the day. A loading ratio of 100 percent indicates a resource is

[35] Peak load is the highest weekly 5-minute actual load during HE 18.

scheduled in the IFM to its maximum generation capacity; a loading ratio of 0 percent indicates a resource is scheduled at minimum load. Figure 4.12 through Figure 4.14 show the IFM loading ratio for units committed via ED across hours for the periods July 1 through 26 (prior to implementation of G-217 and G-219 in RUC), July 26 through August 31, and September 1 through 30, respectively.

After G-217 and G-219 were implemented in RUC[36] on July 27, average loading ratio decreased. By using the RUC market to commit resources for these two generation capacity requirements, fewer ED were required for this purpose. The RUC market runs after the IFM, so resources that may have been committed through ED before the IFM but are now committed in RUC will not have IFM energy awards. Between July 1 and 26, the middle 50 percent of resources cleared additional energy in the IFM above minimum load beginning Hour Ending (HE) 9, and stayed above minimum load through HE 23, with the median loading ratio of 95 percent across the peak. Beginning July 27 and continuing through August, pre-IFM EDs were primarily for G-206 (San Diego area reliability). During this period, the middle 50 percent remained at minimum load until HE 11 and stayed above that level through HE 22, with a peak median of 78 percent. In September, the SWPL outage required additional online resources that were committed through ED, with the middle 50 percent of resources moving above minimum beginning in HE 10, staying above minimum through HE 23, with the median loading ratio peaking at 93 percent.

Figure 4.12 Distribution of Loading Ratio of Pre-IFM Committed Units by Hour July 1-26, 2009



[36] Implementation of G-217 and G-219 in RUC is discussed in more detail in the next section.

Figure 4.13 Distribution of Loading Ratio of Pre-IFM Committed Units by Hour July 27 – August 31, 2009

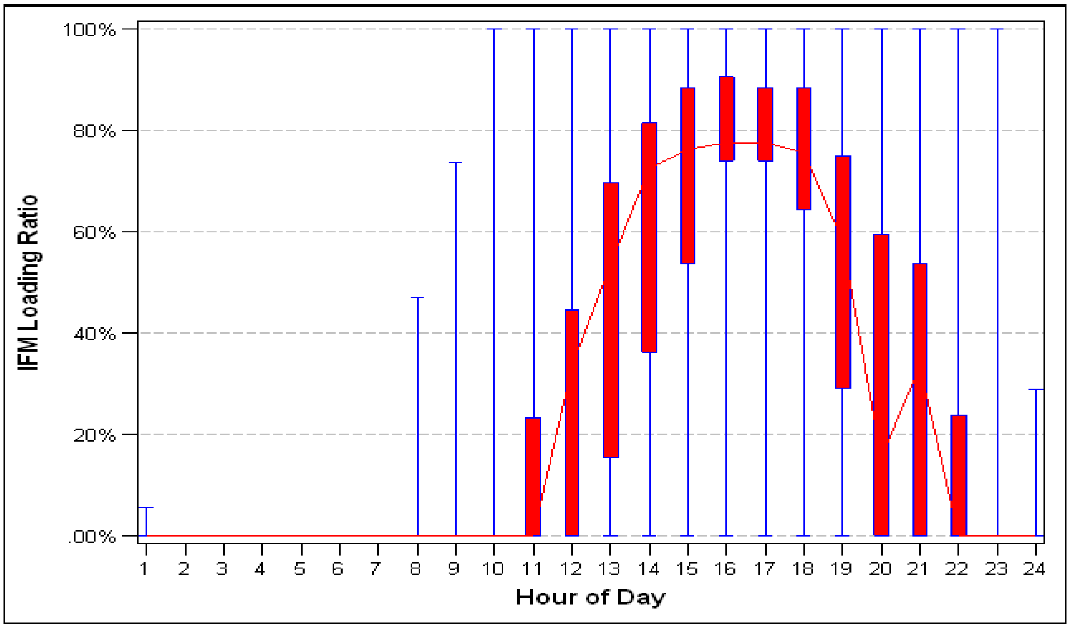
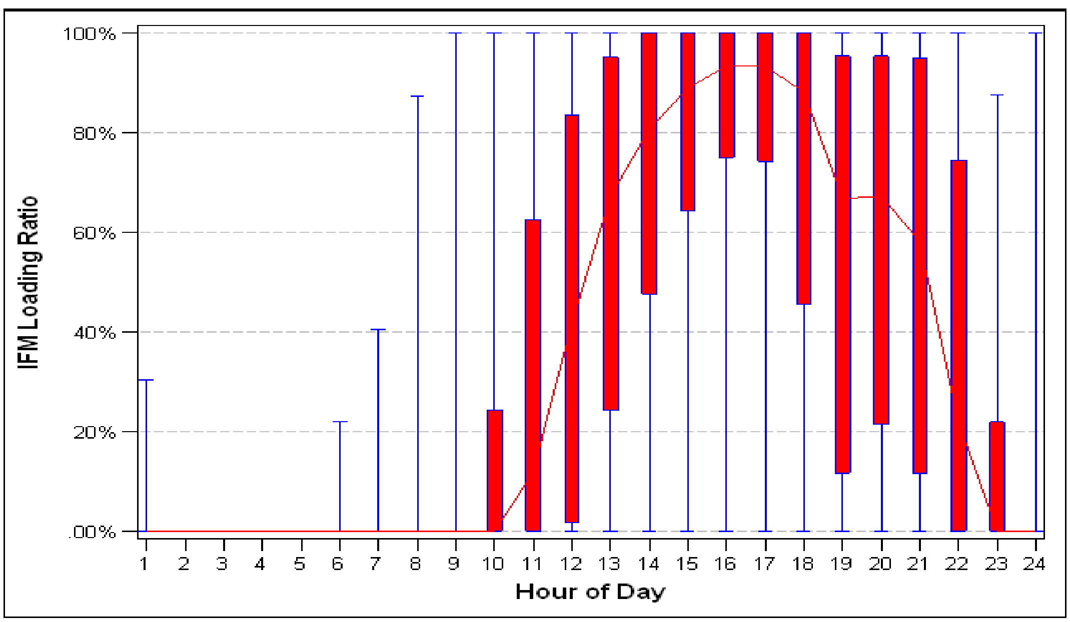


Figure 4.14 Distribution of Loading Ratio of Pre-IFM Committed Units by Hour September 2009



4.4.2 Unit Utilization

The Utilization Proportion of the average resource, as shown in Figure 4.15, provides another view of the ways that ED-committed resources are used by the market. Figure 4.15 shows the average resource's scheduling, incremental dispatch, and ancillary service upward capacity proportions reserved by the market. The day-ahead schedule portion is similar to the Loading Ratio metric described above, but they differ in two ways: (1) The Utilization Proportion is depicted jointly with day-ahead ED volume, and thus begins above 0 percent, whereas the Loading Ratio does not include ED volume; and (2) the Utilization Proportion includes both pre-IFM and post-IFM ED commitment, whereas post-IFM-committed resources have no day-ahead schedules, and thus necessarily have an IFM loading ratio of zero, so they are not included in the aggregated loading ratio metric. The post-IFM resources weigh the Utilization Metric's day-ahead schedule portion downward.

For the quarter, approximately 78 percent of ED-committed resources' capacity is utilized across the peak, of which 16 percent is procured by ED and 62 percent is procured in the markets. The average utilization across the day is approximately 50 percent of capacity, of which 18 percent is procured by ED. Figure 4.15 through Figure 4.17 show the utilization profiles of resources committed by ED for the periods July 1-26, July 27-August 31, and September.

The notable difference among these profiles is in the late July - August profile, in which real-time market volume reaches approximately 21 percent of capacity across the peak, compared to approximately 7 to 8 percent for the other periods. The reason for this difference is that overall volume of committed ED was low during this period, as can be inferred from Figure 4.4 and Figure 4.9 above. Because the July 1-26 period included commitments for G-217 and G-219, and September saw the Station Fire and the SWPL outage, these periods had much higher use of ED unit commitment, whereas the period in between saw unseasonably mild weather and no such events that required heavy unit commitment. By itself, real-time incremental market energy from units committed by ED follows a nearly random pattern across the quarter.

Another exceptional feature in the late July-August chart is the unusual pattern of day-ahead commitment. This also is an artifact of the low volumes during the off-peak periods in late July and August. In these hours, fewer units were committed by ED, as some other units were committed during peak hours only.

Figure 4.15 Average Utilization Proportion for ED-Committed Units July 1-26, 2009[37]

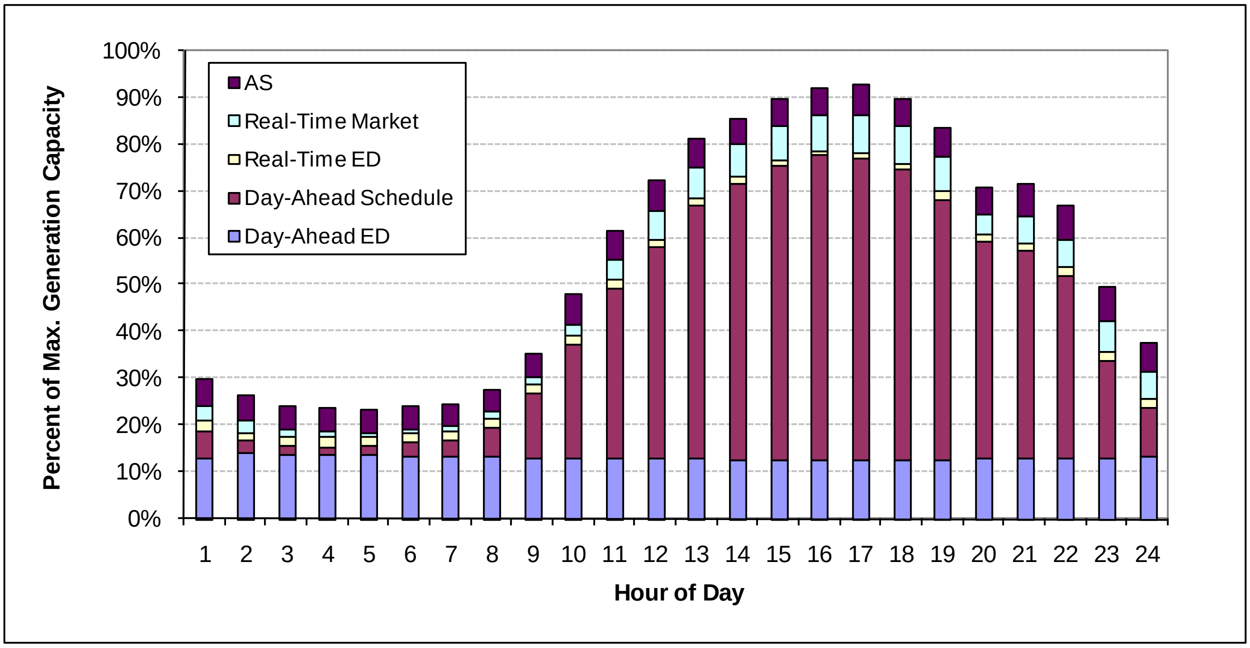
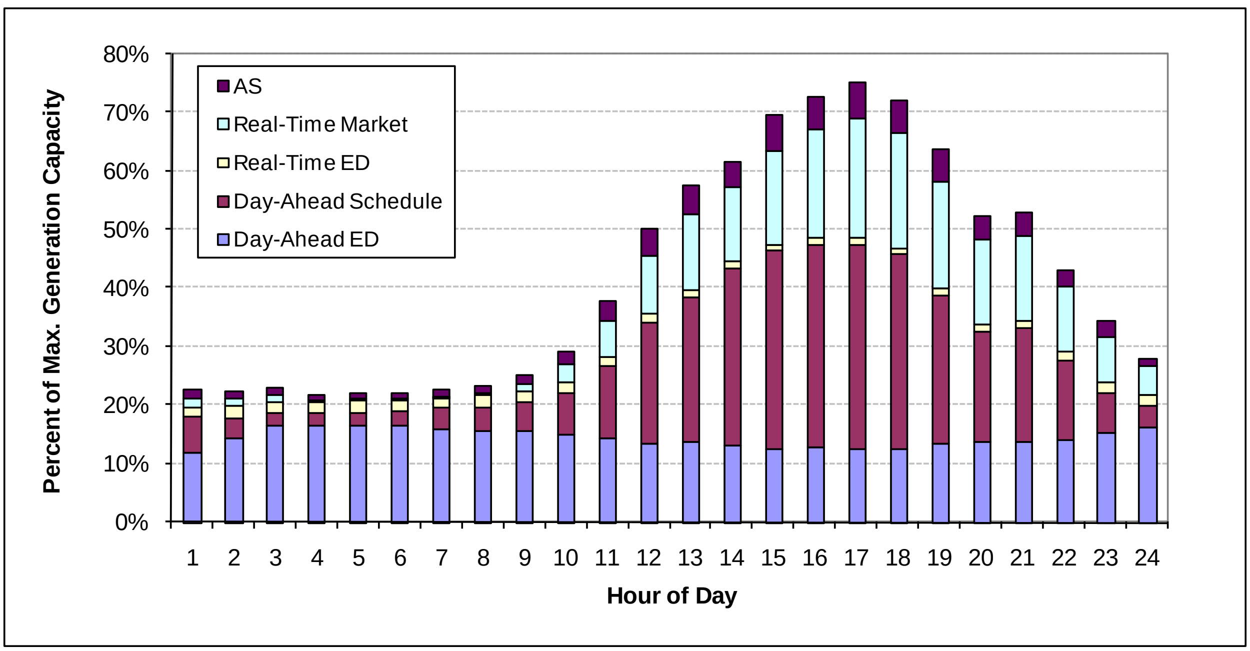


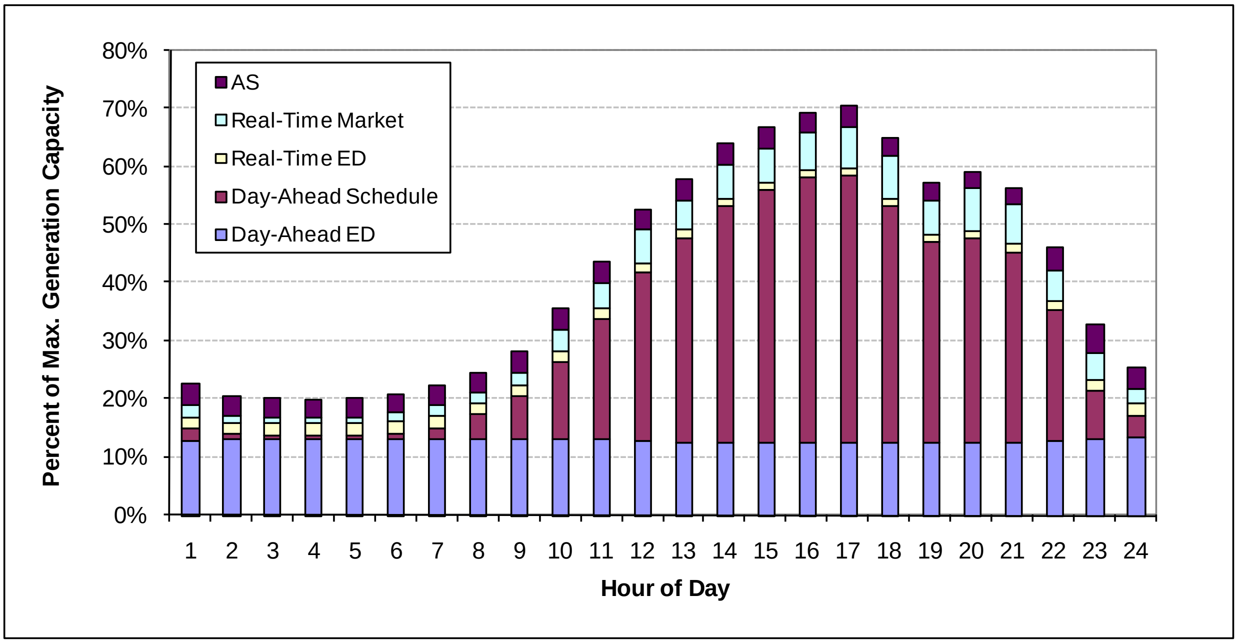
Figure 4.16 Average Utilization Proportion for ED-Committed Units July 27 – August 31, 2009[38]



[37] Excludes decremental real-time market dispatch, exceptional dispatch, and downward regulation.

[38] Excludes decremental real-time market dispatch, exceptional dispatch, and downward regulation.

Figure 4.17 Average Utilization Proportion for ED-Committed Units September 2009[39]



4.5 Capacity Requirements Applied in RUC

The operating procedures for capacity requirements, G-206 (San Diego), G-217 (South-of-Lugo), and G-219 (Orange County), are described at some length in the *Q2 Report.* In general, each itemizes specific combinations of resources that must be retained on-line and operating at or above minimum load to meet voltage support and other capacity requirements for the region it addresses. Resources that satisfy G-219 requirements typically will also contribute to meeting G-217 requirements. G-206 requirements are largely independent of the others, but a greater proportion of capacity that meets G-206 is subject to RMR contract, and thus need not be addressed under the ED tariff provisions. At the time of writing, the G-206 capacity nomogram is in development and will be implemented in RUC in the near future.[40]

The Residual Unit Commitment (RUC) capacity market runs immediately after the integrated forward market (IFM) in the day ahead of operation, and serves only to procure sufficient capacity such that total online capacity meets the day-ahead load forecast for each hour across the actual day of operation (the "power balance constraint"). RUC is an optimization procedure similar to IFM that seeks to minimize total startup and RUC bid costs such that the power balance and other constraints are upheld. Units procured in the RUC market that are not quick-start units will be committed to operate at minimum load for the hours they are needed, up to the full 24 hours of the day, but often for fewer hours.

[39] Excludes decremental real-time market dispatch, exceptional dispatch, and downward regulation.

[40] See presentation at an exceptional dispatch stakeholder meeting, Van Blaricom, September 29, 2009, *Exceptional Dispatch Improvements*, http://www.caiso.com/2434/2434d63470e10.pdf

Beginning trade date July 27, 2009, G-217 and G-219 were successfully implemented into the RUC market, eliminating the need to issue exceptional dispatches to meet these constraints. That is, G-217 and G-219 are additional constraints that the RUC optimization now accounts for in seeking its cost-minimizing solution. G-217 is applied on most, but not all, days; G-219 is applied less frequently, as shown in Table 4.1. DMM views the use of RUC to satisfy G-217 and G-219 as an interim solution, with an ultimate goal of implementing capacity requirements into the IFM. Doing so would further enable generation owners committed for these purposes to also sell energy and ancillary services in the IFM market. The ISO is working internally as well as with the market software vendor to develop and implement an IFM solution.[41]

Between July 1 and 26, the frequency of ED unit commitments for G-217 and G-219 ranged between zero and 13 units per day, and averaged approximately 6 units per day. Beginning July 27, the volume of ED for G-217 and G-219 declined to zero, as they are now all committed in RUC, as shown in Figure 4.18.

Figure 4.18 is actually a “generous” estimate of resources committed by RUC to satisfy G-217 and G-219. Specifically, it is a count of resources that were committed under RUC and that also qualified to satisfy G-217 and G-219, in the hours in which such constraint(s) were applied. However, it is possible that these units were committed in RUC for other reasons, such as to satisfy the power balance constraint, or to satisfy multiple constraints simultaneously. It is not possible to conclusively determine whether or not a resource was committed for a specific nomogram generally; however, it *is* possible to determine that a resource was committed to meet a particular constraint in certain particular situations.[42] We are using the count of resources that qualified to satisfy a constraint in the hours in which the constraint was applied as a proxy for the count of resources that were necessarily committed for that constraint. In other words, if three resources that have no IFM schedule are committed in RUC on a particular day, and two such resources are among the resources that satisfy G-217, and G-217 was enforced on that day, we would assume that these two resources were committed on that day to satisfy G-217.

[41] Ibid.

[42] In 17 hours between July 27 and September 30, the market algorithm found a cost-minimizing solution with capacity commitment short of at least one of the requirements – that is, the constraint is violated in the market. These instances of shortage are identifiable, and such a shortage would indicate that all committed resources that could help to satisfy the capacity constraint are necessary. The list of units committed for G-217 and G-219 only in hours in which the constraints are binding would provide a “conservative” (lower end) estimate of units committed under RUC. Because the number of hours in which the constraints are binding is low compared to the number of hours in which the constraints are applied, commitments limited to those hours would suggest that unit commitment for G-217 and G-219 by RUC is a rare event. We also used a stand-alone version of the market software to compare 13 days in which the constraints were enforced with market runs in which they were unenforced. We found that the application of the constraint resulted in additional unit commitment on 12 of 13 test days, whereas the constraints were binding in single hours on two such days. Given the similarity of magnitude using the “generous” (higher) estimate to the application of the constraints, as well as the frequency of ED unit commitment for the constraints observed prior to July 27, DMM has concluded that the “generous” (higher) estimate provides a more accurate picture of the true count of resources that are committed by RUC to satisfy G-217 and G-219.

Table 4.1 Count of Hours in which G-217 and G-219 Are Applied vs. Are Observed to Be Binding in RUC

Month	G-217		G-219	
	Applied	*Observed Binding*	*Applied*	*Observed Binding*
July	96	5	72	0
August	276	9	264	1
September	156	2	24	0

An artifact of the commitments of G-217 and G-219 in RUC has been occasional anomalous RUC prices. These prices affect relatively minimal volumes; nearly all RUC procurement is from resources under RA or RMR contracts, which are paid $0/MW for RUC capacity, pursuant to CAISO Tariff Section 11.2.2.1.[43] The RUC LMP reached $250/MW in three intervals, -$33.74/MW in one interval, and $115.75/MW in one interval.

In all of these instances, we have found that the price was set not by bids, but by constraint penalty prices used in the software to force the optimization to select priorities among competing constraints. The $250 penalty represents the cost to load per megawatt of violating the G-217 capacity requirement, as this requirement was not met in these intervals. In the instance in which the LMP was -$33.54, the price was set by a violation of the power balance constraint, which at the time had a penalty price of $35,[44] plus a loss component of $1.46/MW. Had the market uncommitted one unit, the G-219 requirement would have been violated at a cost of $250/MW for each megawatt of that resource's minimum operating level. These competing constraints highlight the discrete nature of unit commitment, and the tradeoff between committing an additional unit at the potential cost of excess supply.

[43] These resources receive other capacity payments as part of their RA or RMR contracts, the terms of which include the requirement that the resources bid their available capacity under contract at $0/MW, pursuant to Tariff Section 31.5.1.2.

[44] The software generates a RUC LMP of -$35 as follows: The penalty price of $35 represents the cost to serve an additional MWh of load when additional supply is needed. However, in the case that scheduled generation exceeds the forecast, $35 represents the cost of reducing load by 1 MWh; or, alternatively, the price load must pay to generation to reduce output by 1 MWh. The difference between -$35 and the RUC LMP of -$33.54 represented a loss factor of $1.46.

Figure 4.18 Average Daily Count of Units Committed by ED for G-217 or G-219 vs. Units Potentially Committed in RUC for G-217 or G-219

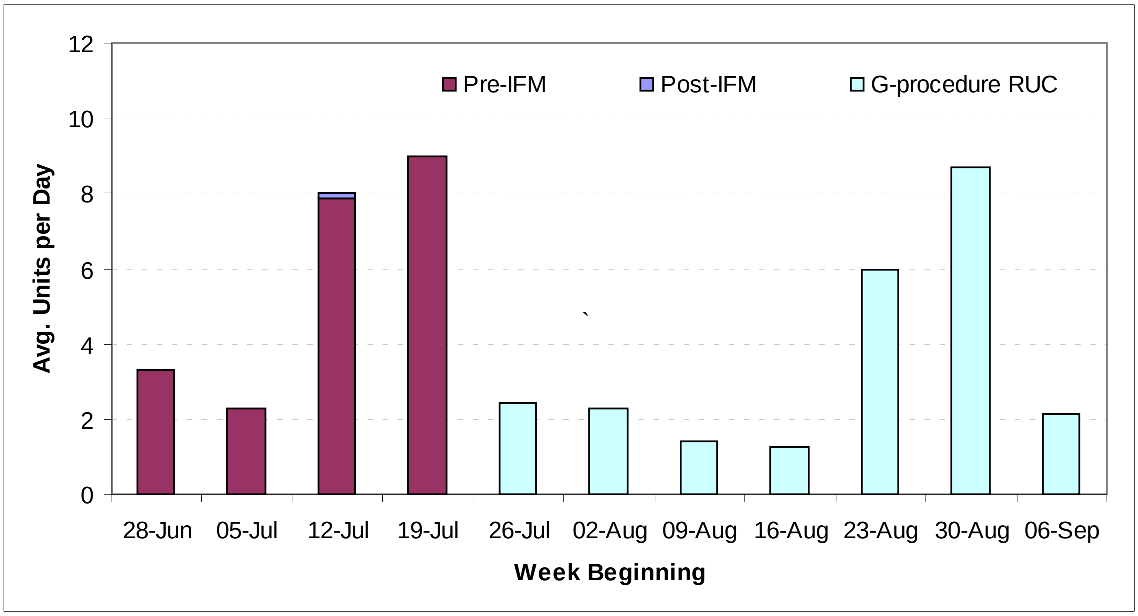


Table 4.2 Hours in Q3 during which RUC LMP below \$0 or above \$50 Due to G-217 or G-219 Constraints

Date	Hour	Non-RA MW	RA MW	Percent of RUC from Non-RA	Non-RA cost	Non-RA Weighted Average LMP	Overall Weighted Average LMP	Constraint Causing Price Excursion[45]
09-Aug	14	0.01	110	0%	$0	($33.54)	$ 0.00	G-219 (not binding)
11-Aug	10	14	2,288	1%	$3,500	$250.00	$ 1.52	G-217 (binding)
12-Aug	24	40.5	2,401	2%	$10,125	$250.00	$ 4.15	G-217 (binding)
26-Aug	1	38.5	2,016	2%	$9,625	$250.00	$ 4.69	G-217 (binding)

[45] Other constraints typically are also binding during these intervals.

4.6 Follow-up On Q2 Recommendation

In our Q2 report, we provided three specific recommendations regarding the use of ED by the ISO. A summary of actions that have been taken of these three prior recommendations is provided below:

- ***Perform a comprehensive review of operational procedures and other criteria for determining exceptional dispatch***. In the last week of July, the ISO formed an ED “strike team” to focus on potential improvement to practices and software to reduce EDs, particularly with respect to unit commitments made in the day-ahead timeframe. This strike team also focused on improving the consistency and logging of ED data, and providing more accurate and timely feedback on ED trends to operations staff. The team also monitored the impacts of new RUC capacity nomograms designed to meet reliability requirements previously met by committing additional units via ED either before or after the IFM. The result of these efforts – combined with the new RUC capacity nomograms discussed below – appears to have reduced EDs in late July and August. However, as discussed above, the amount of capacity committed via EDs increased again in late August and September due to other factors, such as the need to protect against potential transmission outages due to fires in Southern California and a significant prolonged forced outage on the SWPL transmission line.

- ***Explore and implement options for incorporating into the market model the reliability constraints driving exceptional dispatch.*** As noted above, on July 27, the ISO implemented capacity nomograms in the RUC process that reflect capacity needs incorporated in the G-217 and G-219 operating procedures, which were found to be driving a large portion of unit commitments in the Southern California Edison (SCE) area (South-of-Lugo and Orange Country, respectively) The ISO is also developing a RUC nomogram to reflect a third major operating procedure that covers the San Diego area (G-206). However, since minimum load energy and other capacity from units committed in RUC is not available in the IFM market, DMM has recommended that these constraints be incorporated in the IFM market model if possible. This will reduce excess generation in the real-time markets (HASP and RTD) resulting from minimum load committed after the IFM and will also provide resources needed for these constraints with additional opportunity for market revenues in the IFM. The ISO is currently developing procedures to incorporate these capacity constraints in the IFM and expects to have these implemented in late Q4 2009 or early 2010.

- ***Consider new market products that might mitigate the need for exceptional dispatch***. As described in the ISO’s most recent 120-day report to FERC, the ISO has committed to a process over the next nine months to consider potential new products. However, the ISO believes that it would be more appropriate to have a full year of operational experience and information before determining what, if any, specific new products or market design enhancements can most effectively mitigate the volume of future exceptional dispatches. DMM considers this approach prudent – particularly in light of the operational and software improvements that have been implemented to reduce EDs. DMM also notes that by continuing to identify ways to incorporate into the market model constraints that require EDs, the ISO can continue to develop information that will be valuable as part of the process of considering new potential products.

5 Transmission Constraint Enforcement and Biasing

The purpose of this chapter is to provide some additional transparency into the ISO's practice of biasing (or adjusting) transmission limits in the market model. The chapter begins with an introduction to the process for biasing operating limits for flowgates and nomograms in the day-ahead and real-time markets (RTM) and provides a variety of statistics for the constraints that were biased in Q3. The chapter also includes a series of examples of biasing on specific constraints and days that illustrate how biasing of flowgates has been used and the impacts on market outcomes.

This analysis indicates that a total of about 70 flowgates were biased in Q3, with only 22 of those biased in RTD more than 30 percent of the time. There was strong consistency in biasing between the HASP and RTD markets in both frequency and degree of bias. In both these real-time markets, constraint biasing tended to be used to increase – rather than decrease – the market limit on constraints in order to avoid "phantom" congestion (i.e., congestion in the RTM software when observed flows in real-time were below the constraint's actual operating limit). The bias is used very infrequently in the day-ahead market, since review by the ISO's Operations Engineers indicated that biasing was not necessary or could be effectively used to either (1) avoid "phantom congestion" in the day-ahead market or (2) mitigate the potential for congestion that was occurring in the real-time market.

5.1 Background

In general, there are two conditions under which the ISO will bias transmission limits: when flows calculated by the market are not in line with actual flows in real-time, and when grid conditions exist such that a reserve margin must be maintained for reliability.[46]

Biasing to maintain adequate operating margins is a prudent operating practice that was also used by the ISO prior to the launch of the new markets. Under the ISO's prior zonal market structure, in the day-ahead, actual flows on flow-based transmission constraints were not addressed for intra-zonal constraints – only scheduling limits were addressed for inter-zonal transmission constraints. No flowgate biasing was done in the day-ahead. In real-time, for inter-zonal constraints, limits were biased by the operators to compensate for differences between actual flows and scheduled flows, and for intra-zonal constraints adequate margins were maintained through the intra-zonal congestion management process using out-of-sequence (OOS) real-time dispatches. Under the zonal market structure, the costs of OOS dispatches were recovered through uplift charges and did not affect market cleared energy prices.

With the implementation of the new markets based on Locational Marginal Pricing (LMP), the market optimization tools used in conjunction with the Full Network Model (FNM) in the IFM and RTM now perform congestion management through automated processes that calculate locational energy prices that reflect the costs of congestion at such locations. However, for reasons discussed below, the new markets have not eliminated the occurrence of measurable and often predictable differences between actual and market-calculated flows. The process of biasing is, therefore, a necessary operational tool for ensuring that the markets result in schedules and real-time dispatches that more accurately reflect expected real-time flows, respect actual flow

[46] Material in this section is based on discussion with ISO operations staff and Technical Bulletin 2009-07-02, *Process for Biasing Flowgate/Nomogram Operating Limits for Day Ahead & Real Time Markets, July 13, 2009,* http://www.caiso.com/23ea/23eae8aef980.pdf

limits and fully support reliable grid operation. Biasing is not applied to scheduling limits such as inter-ties (ITCs) and market scheduling limits (MSLs); it is applied only to market operating limits for certain branch groups (flowgates/transmission interfaces), as necessary.

One primary driver of biasing is the difference between the actual flows and market flows. This discrepancy can be more severe at the inter-ties. This stems from the fact that the ISO market model does not fully model the network outside of our control area. There are also seasonal unscheduled flow issues that can cause market and actual flows to diverge. A discrepancy in flow at the inter-ties can also create divergence in actual and market flow on internal flowgates as well. Biasing internal flowgates is one way to overcome the discrepancy between the modeled and actual flows at the inter-ties. The ISO has been testing a software feature designed to automatically mitigate these discrepancies at the inter-ties (Compensating Injections). This functionality was not implemented in Q3, but was implemented in early October.

5.1.1 Day Ahead & Real Time Limit Adjustment Level

Flowgates that consistently become binding in real-time and are biased in real-time may need to be biased in the day-ahead runs. Such biasing is needed to provide for better consistency of margin management for these flowgates in the day-ahead so that the congestion/reliability issues are manageable in real-time. Biasing may be necessary to account for the difference in flows between the day-ahead and real-time that are caused by changes in load forecast, generation and transmission, etc. While this is true in principle, in practice the ISO has biased transmission constraints in the day-ahead infrequently compared with the frequency of biasing exercised on the same constraints in real-time. Reasons for the much lower level of biasing in the day-ahead market runs compared with real-time are explained in Section 5.2.1.

Each of the constraints is unique and may require different levels of biasing in the day-ahead, based on the congestion experienced in real-time. The adequate level of adjustment in the day-ahead is based on measurable or predictable difference between actual flows (from telemetry) in the real-time and day-ahead estimated flows. The degree to which these differ may require further review of the historical and day-ahead flow differences. In determining the biasing need, the ISO also considers the conditions leading to flow differences and their interplay with reserve/regulation management and the level of scheduled intermittent resources.

5.1.2 Reasons for Biasing in the Day Ahead & Real Time Markets

The key reasons for biasing operating limits in the day-ahead and real-time markets are:

1) *To align calculated market flows with measurable or predictable actual flows.*

   In the real-time market, flows for certain flowgates may not align closely with real-time telemetry flows. In such cases, the flowgates are biased when the market flows approach binding limits in the market or if the telemetry flows get close to reliability limits in the Energy Management System (EMS). Reasons for flow differences may include: (i) unscheduled flow, (ii) differences in load distribution, (iii) deviations on resources internal to the ISO, and (iv) external network model limitations. Pursuant to good utility practice, efforts are made to minimize the flow differences. However, to the extent a flowgate is susceptible to significant differences between actual and market flow, it is necessary to have a process for monitoring and adjustment of limits in the real-time market on a more frequent basis.

2) *To accommodate mismatch due to inherent design differences of the day-ahead market, real-time unit commitment (RTUC) and the real-time dispatch (RTD) runs.*

A different level of biasing may be required for enforcement of the same constraint in the day-ahead, RTUC and RTD runs due to the difference in dispatch intervals (one hour for day-ahead, 15 minutes for RTUC, and 5 minutes for RTD) and the difference in ramping capabilities of resources in these different dispatch intervals. For example, while RTUC can provide a solution for a 15 minute interval that is 30 minutes into the future using a 15 minute ramp of resources, RTD is run 20 to 30 minutes later and when it gets to dispatch that 15 minute interval it can only use a 5 minute ramp of resources. At that time, it is possible that the exact initial condition predicted by RTUC 20 to 30 minutes prior does not occur. If a constraint is binding in RTUC, then RTD has a high chance of not having the means (resource ramp rate) to respect that limit if no additional margin is available.

3) *To allow reliability margins for certain flowgates.*

Flowgates may need to be biased in real-time to maintain a reliable operating margin for flowgates that are approaching their actual operating limits. There are numerous reasons why operating margins are required in real-time. The operating margin required in real-time is determined using EMS Data and Market Contingency Reserve Awards available in real-time. The following are some of the reasons for biasing based on reliability margins for flowgates:

- *Historical Contingency Reserve Procurement:* To ensure that operating reserves can be delivered in light of (persistent or anticipated) congestion that may otherwise prevent that from happening.
- *Historical AGC (Regulating Reserve) awards:* To account for energy that is likely to be delivered across a flowgate given historical patterns. The market software does not explicitly consider the delivery of regulation reserve in managing congestion. This dispatch may cause market flows to diverge from actual flows.
- *Intermittent resource deviations:* The day-ahead schedules and the real-time actual generation for various intermittent resources can deviate significantly. This can potentially cause congestion and/or reliability issues in real-time.
- *Adverse operating conditions:* Adverse operating conditions, such as fire, may also necessitate the need to temporarily bias flowgates in the real-time market runs. This is usually needed to maintain appropriate operating margin for flowgates impacted directly or indirectly by these adverse operating conditions.

4) *To adjust margins for flowgates impacted by telemetry issues.*

The ISO also biases select flowgates that are impacted by lack of telemetry in the area. This is typically an issue for the 115 kV and below part of the transmission system. Certain pockets for this kV level have little or no telemetry. Therefore the state estimator (SE) solution is impacted by the lack of visibility. Most of these flowgates are typically un-enforced in the market model. However, if a flowgate comes close to its limit in real-time based on the SE solution, the ISO then enforces this flowgate into the market with a margin, as needed.

5.2 Trends

5.2.1 Constraint Biasing in Real Time Market

Flowgates

This section provides a review of the frequency of biased flowgates and nomograms in the Hour Ahead Scheduling Process (HASP) and the 5-minute Real Time Dispatch (RTD) markets. Our analysis shows that the flowgates and nomograms are biased consistently in both these markets.

There are about 5,500 flowgates which are modeled in the ISO's new market. As Figure 5.1 shows, only a very small portion (about 1.2 percent) were biased in Q3. Our analysis indicates about one third of the flowgates shown in Figure 5.1 were biased more than 60 percent of the time in Q3. For these flowgates, the primary reason for the frequent biasing is that there is a significant (and, in these cases, frequent) discrepancy between the market flow and actual or telemetered flows.

For some other major transmission lines, such as Path 26 and Path 15, biasing was necessary in real-time to maintain a reliable operating margin for these flowgates that were approaching their actual operating limits. In most cases, maintaining the reliable operating margin of these major flowgates causes significant congestion which leads to higher energy prices. In the next sections we review the impact of biasing major flowgates on LAP energy LMPs in more detail.

Table 5.1 lists all flowgates that were biased in the RTM in Q3, along with the percentage of hours that each flowgate was biased, and other related statistics (i.e., average, minimum, and maximum percent of actual limit biased during Q3). The statistics presented in Table 5.1 are calculated only on intervals where the bias moved the effective limit off of the actual limit. For the majority of these transmission lines the level of biasing was fixed during the time period in which they were biased. For example, in the RTD market, the Contra Costa-Ross Tap 230 kV line (item 5) was biased to 113 percent of its actual limit in all the intervals in which biasing was applied. On the other hand, for those major paths and branch groups for which the operators used biasing to maintain a reliable operating margin, the level of biasing varied significantly. For example, in the RTD market, for the San Diego CFE import branch group (SDGE_CFEIMP_BG), the level of bias ranged from as low as 45 percent to as high as 110 percent of the branch group's actual operating limit during Q3.

As described in Section 5.1.2, a different level of biasing may be necessary for the same constraint in RTUC compared to RTD. The second numerical column of Table 5.1 also shows the percentage of the hours for which any given flowgate had a different level of biasing applied in RTD compared to RTUC. For example, the Path 15 branch group (item 33 in the table), was biased 19 percent of the time in Q3. During these hours the bias was different between the two market runs almost 8 percent of the time. Several factors that could contribute to different levels of biasing were presented in the previous section.

Our analysis also indicates that biasing is very consistent in both the RTUC and RTD runs. Specifically, the third numerical column of Table 5.1 shows the percentage of hours in which biasing was only applied in the RTD run and not in the RTUC run. During Q3, there were only three flowgates which were biased only in the RTD run (HUMBOLDT_BG, VICTVL_BG and SDGEIMP_BG), and the number of hours where this occurred was extremely low.

Figure 5.1 Percent of Hours Biased in RTD Market – 2009 Q3

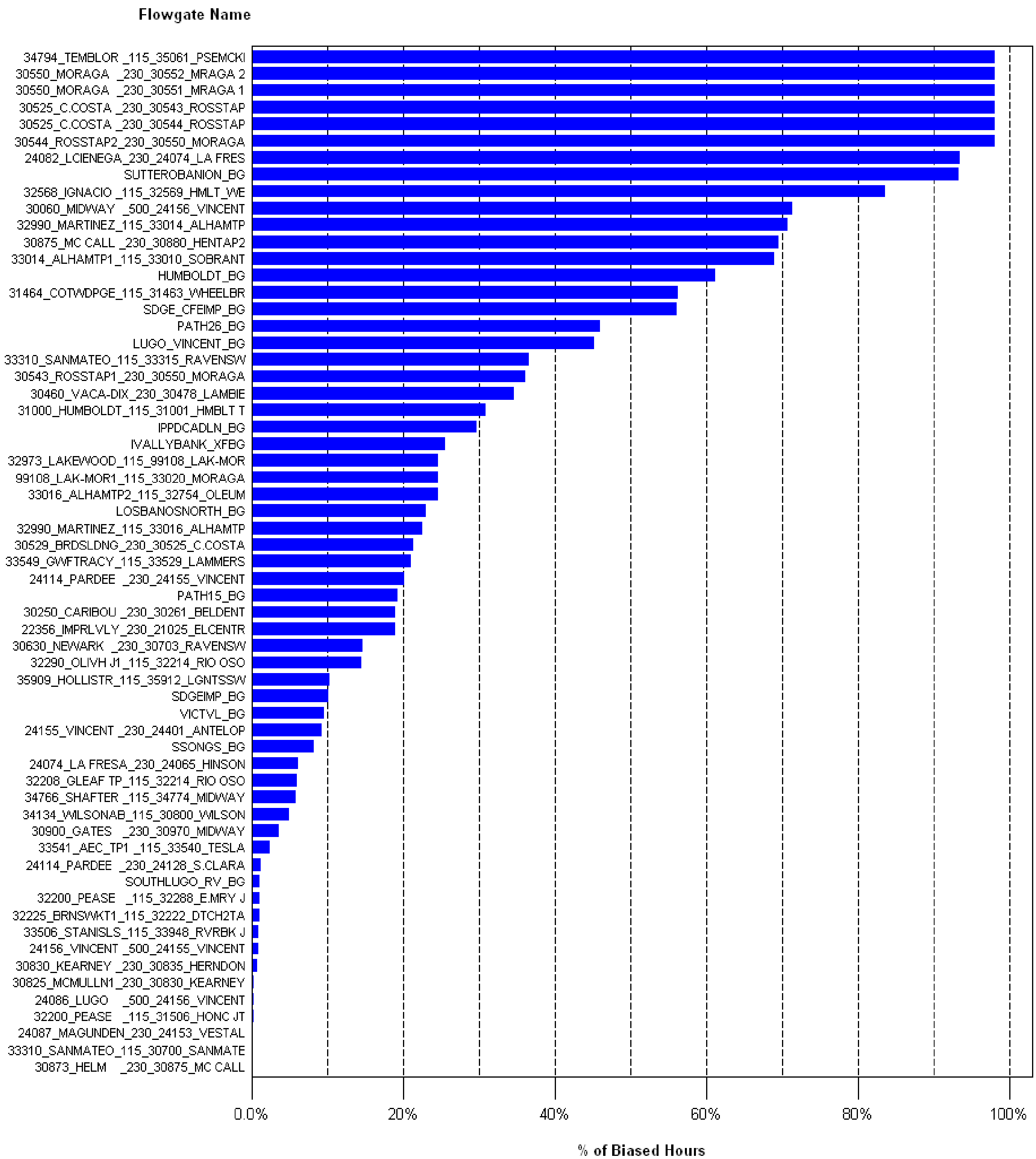
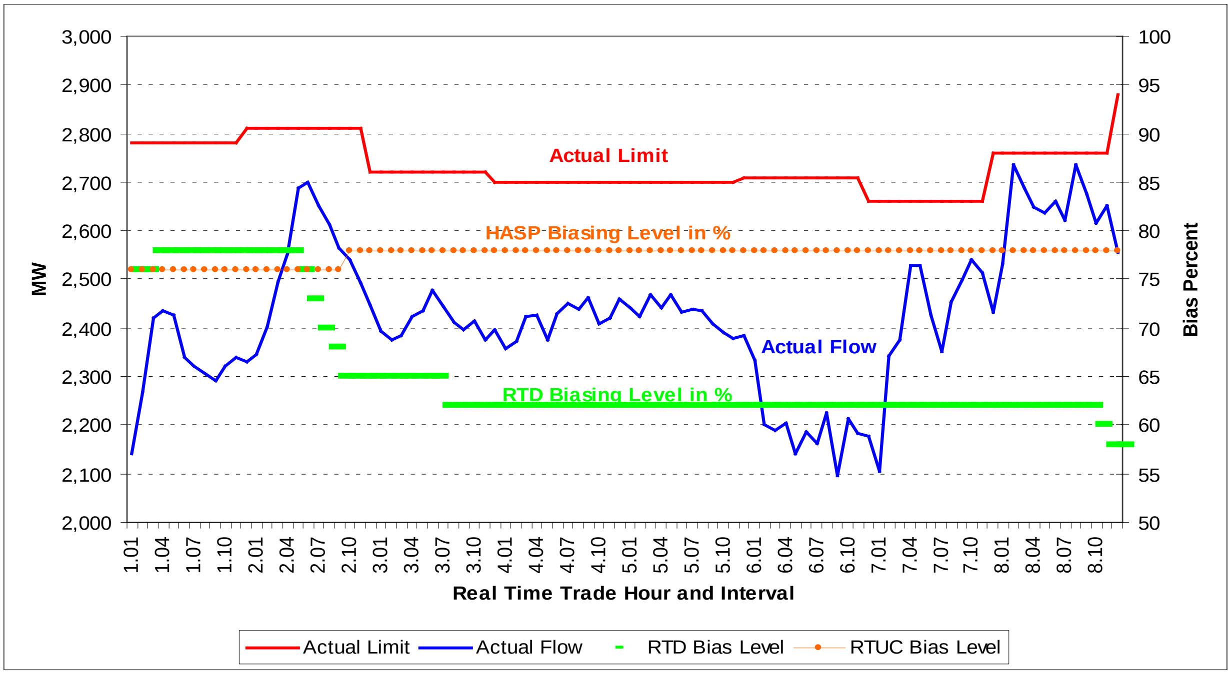


Figure 5.2 Path15 Flow and Biasing on September 30, 2009



Path 15 Flow and Biasing on September 30, 2009

Figure 5.2 provides an example that illustrates how in some cases it may be appropriate and necessary to bias a flowgate at a different level in RTUC and RTD where actual flow and market flow diverge. On the morning of September 30, Path 15 was de-rated due to a scheduled outage of the Diablo-Gates #1 500kV line. The HASP failed for HE 1, as congestion on Path 15 was more severe with a lower bias and the market software was not able to find a solution. In the first half of HE 2, actual (real-time) flow was increasing rapidly and approaching the Path 15 available transfer capability (ATC). Market flow was increasing as well, and both the RTD and RTUC market flows were lower than actual flow. In order to maintain a reserve margin on Path 15 that was consistent with operating conditions given the Diablo-Gates outage and Path 15 de-rate, operators began increasing the bias (decreasing the effective limit) on Path 15 to force the market optimization to dispatch in a way that reduced the actual flow on Path 15. The path was biased to as low as 58 percent in RTD. However, the bias in RTUC was left at 78 percent to prevent additional difficulties with obtaining a solution in that market run. By keeping the biasing level at 78 percent in the RTUC run, the market flows were clearing at about 2,100 MW and lower during the early hours of the morning. Later on in the RTD run, the operators adjusted the biasing level to sustain the safe reserve margin on the path. By increasing the bias in RTD (decreasing effective path limit), the actual flow on Path 15 responded and dropped significantly, to about 70 percent of the path limit, and remained below or near 75 percent of the path limit until HE 8, providing a sufficient reserve margin during those hours. Note also that beginning in HE 8, the actual flow increased to near the path limit despite a significant bias to between 58 percent and 63 percent of the path limit.

Nomograms

Adverse operating conditions, such as fire and planned and forced outages of a significant or critical amount of transmission or generation capacity, may require the ISO to create temporary nomograms to enforce lower limits to operate the market reliably within the guidelines of established operating procedures. Operators also need to bias nomograms briefly in the real-time market runs to maintain appropriate operating margin for areas impacted by the adverse operating conditions. The most significant nomograms which had an impact on the real-time market were those related to the San Onofre-Santiago 220 kV line outage which occurred in early September. More details are provided in Section 5.3.

Un-enforcing Constraints

Another tool available to grid operators for managing flow where the market optimization does not do so accurately is to un-enforce specific transmission constraints where the market may calculate flow that is binding but actual flows are never, or very rarely, close to the constraint limit. This is equivalent to biasing the transmission constraint to a sufficiently high effective limit that it will not be binding, and will not cause re-dispatch to relieve congestion, in the market optimization. Given the similarity between extreme positive bias and un-enforcing of a constraint, similar analysis of un-biasing of constraints would complement the analysis in this section. Unfortunately, it was not possible at the time this report was produced to reduce the set of constraints considered to those that are most relevant. Many of the constraints that are frequently un-enforced are either outside the ISO control area market model and are not relevant to meaningful statistics for analysis of this practice or are smaller components of a larger transmission facility and would overstate calculated statistics. Given this, we do observe a significant number of flowgates that are not enforced in the market model, many of which are at or below the 115 kV transmission system. Additional transparency of un-enforced constraints, including frequency, voltage class, market, and reason for un-enforcement, would provide valuable information to market participants regarding the network model on which the optimization was calculating dispatch and prices.

Table 5.1 RTD Biased Flowgates and Frequency of Biasing with Additional Statistics Q3 2009

Number	Flowgate Name	Hours Which Biasing Was Applied in RTD Run During Q3	Hours Which the Applied Level of Biasing was Different in RTD and RTUC Runs	Hours Which Biasing Was Only Applied in RTD Market and not RTPD Runs	RTD Avg Biasing Limit in Q3	RTD Min Biasing Limit in Q3	RTD Max Biasing Limit in Q3
1	34794_TEMBLOR _115_35061_PSEMCKIT_115_BR_1 _1	98%	1%	Less than 1%	120%	120%	120%
2	30550_MORAGA _230_30552_MRAGA 2M_ 1.0_XF_2	98%	1%		114%	103%	117%
3	30550_MORAGA _230_30551_MRAGA 1M_ 1.0_XF_1	98%	1%		114%	103%	117%
4	30525_C.COSTA _230_30543_ROSSTAP1_230_BR_1 _1	98%	15%		115%	113%	125%
5	30525_C.COSTA _230_30544_ROSSTAP2_230_BR_2 _1	98%	1%		113%	113%	113%
6	30544_ROSSTAP2_230_30550_MORAGA _230_BR_2 _1	98%	1%		113%	113%	113%
7	24082_LCIENEGA_230_24074_LA FRESA_230_BR_1 _1	93%	1%		108%	96%	120%
8	SUTTEROBANION_BG	93%			100%	100%	100%
9	32568_IGNACIO _115_32569_HMLT_WET_115_BR_1 _1	84%	1%		110%	110%	115%
10	30060_MIDWAY _500_24156_VINCENT _500_BR_3 _2	71%	1%		111%	110%	120%
11	32990_MARTINEZ_115_33014_ALHAMTP1_115_BR_1 _1	71%	1%		123%	110%	160%
12	30875_MC CALL _230_30880_HENTAP2 _230_BR_1 _1	70%	1%		111%	100%	115%
13	33014_ALHAMTP1_115_33010_SOBRANTE_115_BR_1 _1	69%	0%		116%	110%	150%
14	HUMBOLDT_BG	61%	4%		145%	10%	175%
15	31464_COTWDPGE_115_31463_WHEELBR _115_BR_1 _1	56%	1%		110%	96%	115%
16	SDGE_CFEIMP_BG	56%	0%		93%	45%	110%
17	PATH26_BG	46%	4%		89%	46%	100%
18	LUGO_VINCENT_BG	45%	1%		105%	95%	112%
19	33310_SANMATEO_115_33315_RAVENSWD_115_BR_1 _1	36%	0%		110%	101%	120%
20	30543_ROSSTAP1_230_30550_MORAGA _230_BR_1 _1	36%	1%		119%	113%	130%
21	30460_VACA-DIX_230_30478_LAMBIE _230_BR_1 _1	35%	1%		120%	105%	120%
22	31000_HUMBOLDT_115_31001_HMBLT TM_ 1.0_XF_1	31%			110%	110%	110%
23	IPPDCADLN_BG	30%			100%	100%	105%
24	IVALLYBANK_XFBG	25%	26%		91%	65%	105%
25	32973_LAKEWOOD_115_99108_LAK-MOR1_115_BR_1 _1	25%	0%		111%	111%	120%
26	99108_LAK-MOR1_115_33020_MORAGA _115_BR_1 _4	25%	0%		111%	111%	120%
27	33016_ALHAMTP2_115_32754_OLEUM _115_BR_1 _1	25%			111%	110%	111%

Number	Flowgate Name	Hours Which Biasing Was Applied in RTD Run During Q3	Hours Which the Applied Level of Biasing was Different in RTD and RTUC Runs	Hours Which Biasing Was Only Applied in RTD Market and not RTPD Runs	RTD Avg Biasing Limit in Q3	RTD Min Biasing Limit in Q3	RTD Max Biasing Limit in Q3
28	LOSBANOSNORTH_BG	23%	3%		80%	10%	100%
29	32990_MARTINEZ_115_33016_ALHAMTP2_115_BR_1 _1	22%	0%		128%	95%	130%
30	30529_BRDSLDNG_230_30525_C.COSTA _230_BR_1 _1	21%	0%		105%	93%	120%
31	33549_GWFTRACY_115_33529_LAMMERS _115_BR_1 _1	21%	0%		112%	110%	120%
32	24114_PARDEE _230_24155_VINCENT _230_BR_1 _1	20%	0%		150%	150%	150%
33	PATH15_BG	19%	8%		82%	5%	99%
34	30250_CARIBOU _230_30261_BELDENTP_230_BR_1 _1	19%			113%	98%	115%
35	22356_IMPRLVLY_230_21025_ELCENTRO_230_BR_1 _1	19%	0%		114%	98%	130%
36	30630_NEWARK _230_30703_RAVENSWD_230_BR_1 _1	15%			112%	110%	112%
37	32290_OLIVH J1_115_32214_RIO OSO _115_BR_1 _1	14%	1%		123%	100%	125%
38	35909_HOLLISTR_115_35912_LGNTSSW2_115_BR_2 _1	10%			105%	105%	105%
39	SDGEIMP_BG	10%	2%	2%	95%	70%	105%
40	VICTVL_BG	10%	1%	5%	104%	80%	130%
41	24155_VINCENT _230_24401_ANTELOPE_230_BR_1 _1	9%	0%		101%	100%	120%
42	SSONGS_BG	8%			90%	85%	95%
43	24074_LA FRESA_230_24065_HINSON _230_BR_1 _1	6%			105%	105%	107%
44	32208_GLEAF TP_115_32214_RIO OSO _115_BR_1 _1	6%			110%	110%	110%
45	34766_SHAFTER _115_34774_MIDWAY _115_BR_1 _1	6%	1%		120%	105%	120%
46	34134_WILSONAB_115_30800_WILSON _230_XF_1	5%	2%		106%	105%	125%
47	30900_GATES _230_30970_MIDWAY _230_BR_1 _1	4%			94%	84%	105%
48	33541_AEC_TP1 _115_33540_TESLA _115_BR_1 _1	2%			110%	110%	110%
49	24114_PARDEE _230_24128_S.CLARA _230_BR_1 _1	1%	17%		89%	65%	200%
50	SOUTHLUGO_RV_BG	1%			103%	103%	108%
51	32200_PEASE _115_32288_E.MRY J1_115_BR_1 _1	1%	4%		115%	115%	115%
52	32225_BRNSWKT1_115_32222_DTCH2TAP_115_BR_1 _1	1%			112%	112%	112%
53	33506_STANISLS_115_33948_RVRBK J2_115_BR_1 _1	1%			110%	102%	110%
54	24156_VINCENT _500_24155_VINCENT _230_XF_1 _P	1%	5%		117%	105%	120%
55	30830_KEARNEY _230_30835_HERNDON _230_BR_1 _1	1%	6%		120%	120%	120%
56	30825_MCMULLN1_230_30830_KEARNEY _230_BR_1 _1	0%			111%	110%	115%
57	24086_LUGO _500_24156_VINCENT _500_BR_1 _1	0%			105%	105%	105%

5.2.2 Consistency of Biasing Between Day Ahead and Real Time Market

Although DMM found that biasing of constraints in the day-ahead process was extremely limited and much lower than the frequency of biasing in the real-time market, a more detailed review of biasing trends in Table 5.1 indicates that these findings are not indicative of a lack of feedback between real-time and day-ahead market operation. Specifically:

- As shown in Table 5.1, almost all the constraints that were biased in the real-time market tended to be "biased up", which means that biasing was needed to avoid “phantom” congestion in real-time (i.e., congestion in the market model when observed flows were below actual limits). Since most of this congestion doesn't consistently appear in day-ahead runs, there may be no need for biasing for these constraints in the IFM or RUC. If congestion appears in day-ahead runs, the ISO’s Operations Engineers evaluate the validity of this congestion and recommend biasing or un-enforcement, as appropriate. DMM’s review of data on the biasing of constraints in real-time and congestion that occurred in the IFM confirms that flowgates that were biased up in real-time were very rarely congested in the IFM.

- Meanwhile, Table 5.1 also shows that very few constraints were actually "biased down" in the real-time market. Biasing down is usually done to maintain adequate reliability margin to ensure line/path loadings stay within their operating limits. These are the constraints that are typically biased in the day-ahead market to ensure congestion and reliability issues are manageable in real-time. Examples of these are the PATH26_BG and the SDGE_CFEIMP_BG.

These results highlight that while it is desirable to ensure that real-time conditions are considered when determining the appropriate level of biasing in the day-ahead market, each constraint is unique and may require different levels of biasing in the day-ahead market.

5.3 Impact of Biasing on LAP Prices

In this section we provide illustrative examples of how congestion and operator biasing of flowgates impacted the energy LAP prices. We focus on days where high positive LAP prices occurred in several consecutive intervals in the real-time market. These extreme high LAP prices are all driven by congestion and corresponding high shadow prices on major flowgates.

Figure 5.3 summarizes congestion on a sample of days that highlight the extent to which prices in Southern California can be impacted by various conditions, such as congestion in RTD on the SDG&E CFE Import Limit (July 11), Path 26 (September 22) and South of Lugo – Rancho Vista (August 24), and limits imposed to manage flow in light of a Songs – Santiago line outage (September 3). As shown in Figure 5.3, congestion on these transmission limits is a primary contributor to higher LAP LMPs in the south.

For the Pacific Gas and Electric (PG&E) LAP, our analysis shows the congestion on Los Banos North branch group can significantly impact PG&E LAP LMPs, resulting in higher prices. More details on this are provided later in this section, as shown in Figure 5.4. Although Path 15 congestion also contributes to high LAP prices in the PG&E area, these examples focus on the use of transmission limit bias on the Los Banos North branch group.

Figure 5.5 through Figure 5.8 provide more detailed data and descriptions of events in these hours, and show how biasing of transmission limits was used to manage various issues for certain instances shown in Figure 5.3 and Figure 5.4.

Figure 5.3 Impact of Congestion on RTD Prices - Southern LAPs (Select Days)

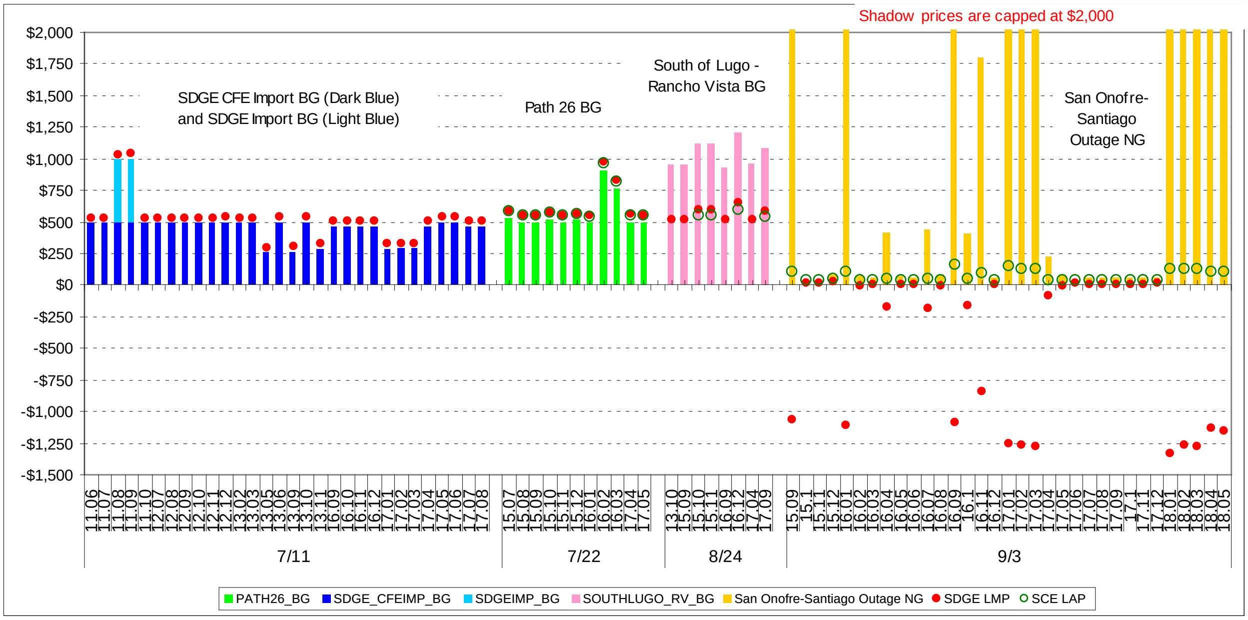


Figure 5.4 Impact of Congestion on RTD Prices – PG&E LAP (Select Days)

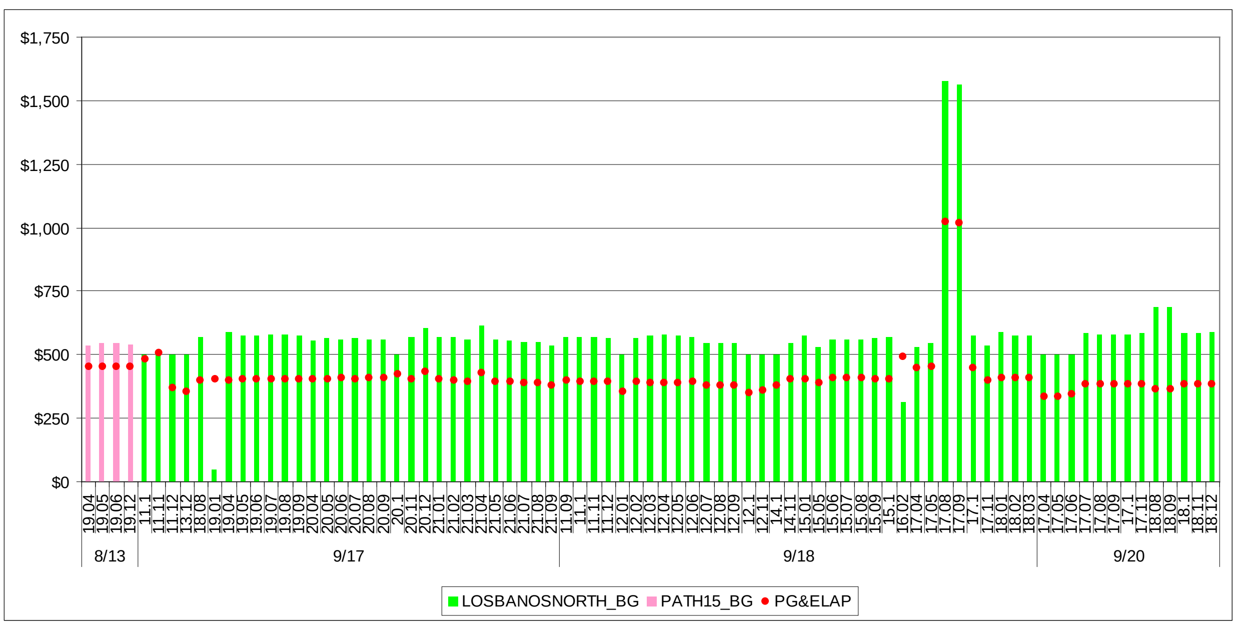
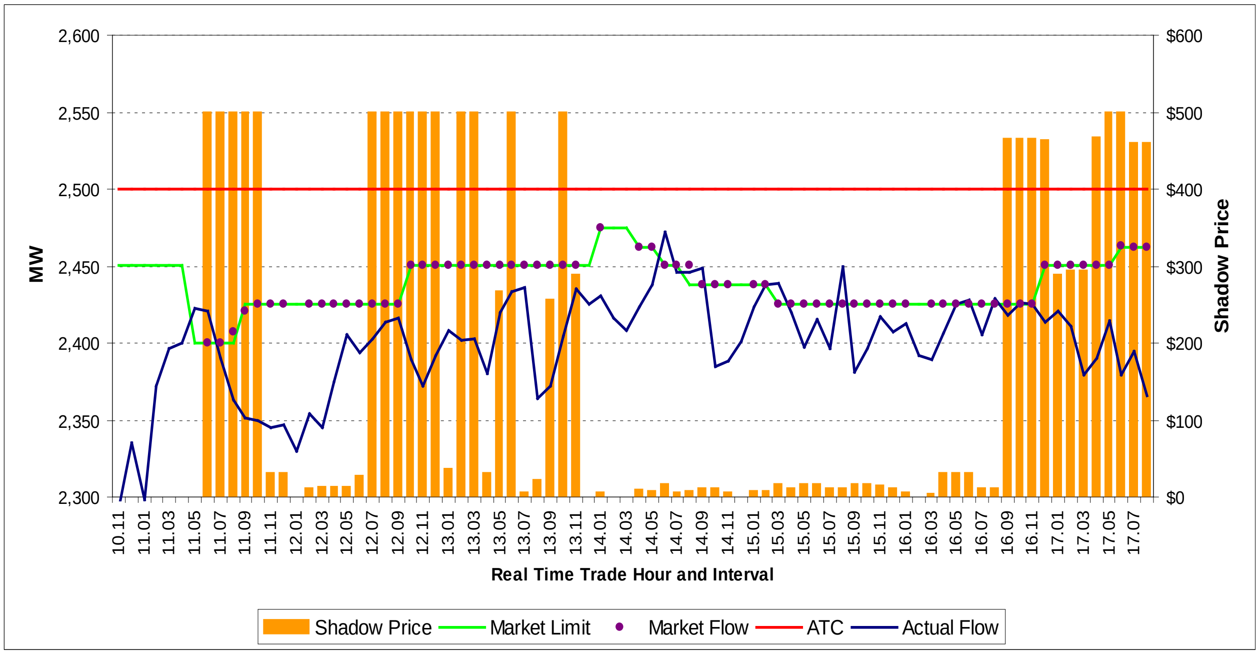


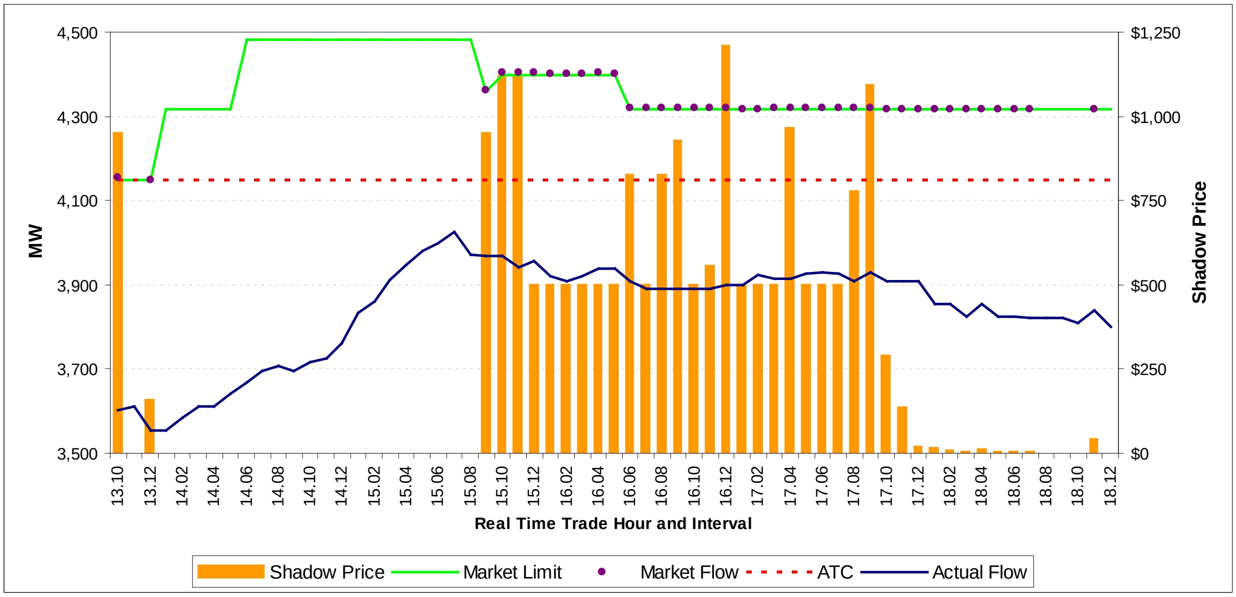
Figure 5.5 Flows and Shadow Prices of SDGE CFE Import BG on July 11



Example 1: High SDG&E LAP Prices due to Congestion on the San Diego CFE Import Branch Group (July 11)

On this day, Otay Mesa - Tijuana 230kV was cleared for scheduled work. The outage started from Hour Ending (HE) 7 and lasted till HE 19. Due to this outage, SDGE/CFE Total Import was de-rated to 2,500 MW from its normal rating of 2,800 MW. Also, San Diego load was increasing over the projected forecast for that day. As shown in Figure 5.3, SDG&E LAP prices spiked to $500/MWh and higher for several intervals in the real-time market on this day. Due to the outage and de-rate of the CFE import, ISO operators applied a downward bias of two percent to maintain a reliability margin. Beginning in HE 11, the actual flow was rapidly increasing and reached 95 percent of the actual operating limit. At this time, operators biased the branch group further down to 96 percent of the de-rated limit. This lower bias level led to the first round of high shadow prices resulting from the bias (Figure 5.5). Operators adjusted the level of bias several times from HE 11 to HE 17 to keep the market optimization dispatching in a way that kept the actual flow below 2,500 MW. Toward the end of this period, in HE 17, the actual flow dropped below 2,400 MW and operators raised the biased market limit for the CFE import back to 98 percent.

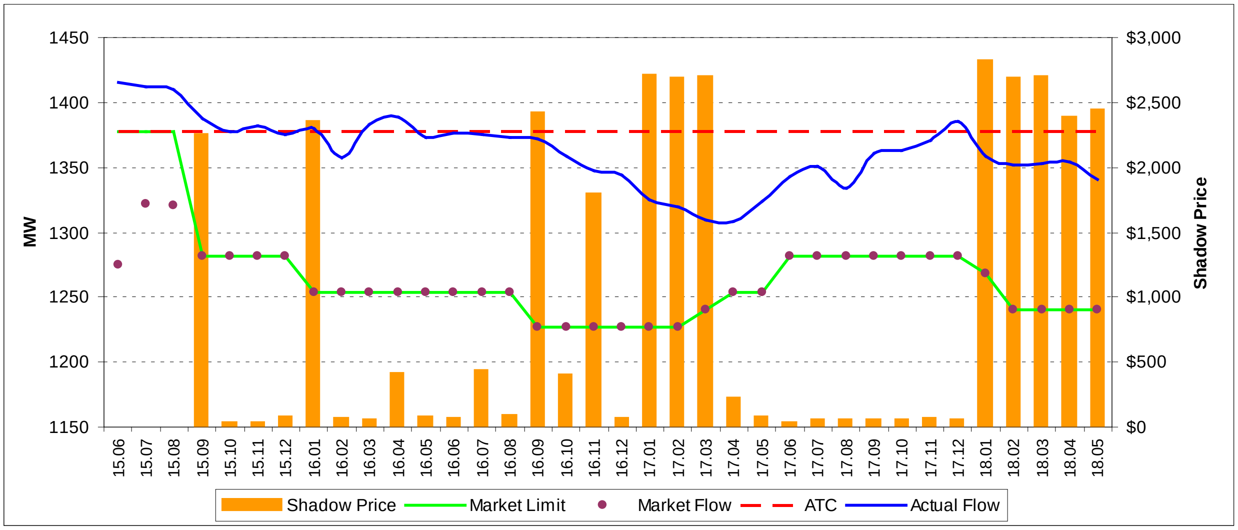
Figure 5.6 Flows and Shadow Prices for the South-of-Lugo to Rancho Vista BG on August 24



Example 2: High SDGE and SCE LAP Prices Due to Congestion on the South of Lugo to Rancho Vista Branch Group (August 24)

This example shows the use of bias to increase the effective limit of a constraint to mitigate a circumstance where the market optimization determined there was congestion but actual flows were well below the actual limit. During the peak hours of August 24 the South of Lugo to Rancho Vista branch group was limited to 4,150 MW, below its normal rating of 5,900 MW, due to scheduled work on this station. This limitation — in conjunction with higher than forecasted load for SCE — triggered high energy prices in SP26 for several intervals in the real-time market, as previously shown in Figure 5.3. Figure 5.6 shows that almost all of the high shadow prices occurred between HE 15 and HE 17, which correspond to the high energy LAP LMPs for SCE shown in Figure 5.3. Near the end of HE 13, operators increased the effective limit by introducing a positive bias since the actual flow was well below the actual operating limit. This alleviated the “phantom congestion” being observed by operators (as shown by the lack of shadow values on the constraint, represented by bars). From this time forward, actual flow increased but remained well below the actual limit. For non-binding intervals (not shown), the market flow was increasing as well and approaching the actual limit. As actual flow became closer to the actual limit (HE 15 interval 7), grid operators lowered the bias slightly and the flowgate became congested two intervals later in HE 15 interval 9. As shown by the dots representing market flow, the market flow was at the biased limit (green line) and was well above the actual flow (blue line). At this point the flowgate became congested and the market produced shadow values for the constraint (bars). The effective limit was kept around 4,300 MW to 4,400 MW (compared to the 4,150 MW limited rating) by using bias to maintain a reserve margin given the difference in market and actual flows. In this example, operators were able to use the bias to manage the difference between the market calculated flow and actual flow, alleviate false congestion prices that would otherwise have been produced in this circumstance, and maintain a reserve for margin reliability purposes.

Figure 5.7 Flows and Shadow Prices of SONGS-Santiago Outage Nomogram on September 9

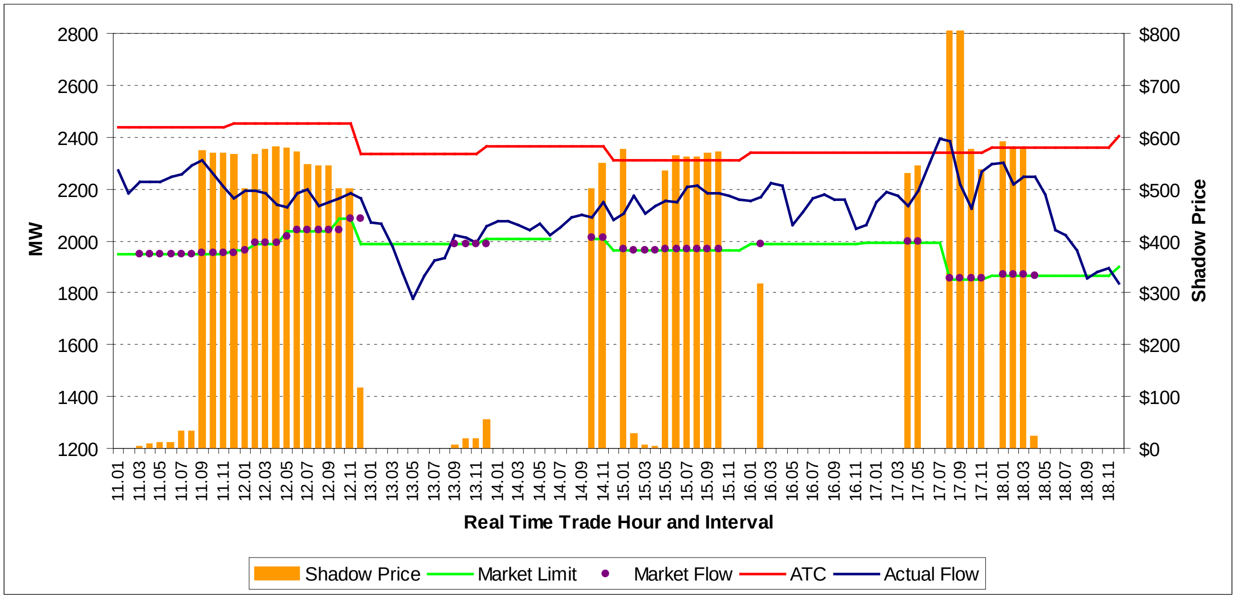


Example 3: Low SDGE LAP Prices Due to San Onofre-Santiago Outage Nomogram (September 9)

This example shows the use of bias to decrease the effective limit of a constraint to mitigate a circumstance where the market-calculated flow was below the actual flow and the market was not detecting that the limit was actually binding or exceeded. On that day, San Onofre-Santiago No. 2 220 kV line cleared for scheduled work during the peak hours. ISO Operating Engineers developed a temporary nomogram to include in the network model with the operating limit of 1,378 MW to reflect the decreased flow limits resulting from this outage.

As Figure 5.7 shows, in HE 15 the actual flow had exceeded this limit. In order to reduce the actual flow, operators biased the nomogram limit down to force the market optimization to dispatch to a lower limit and cause the actual flow to reduce to a level that respects the actual limit. In this circumstance, the market would not have determined the constraint was binding and would not have calculated a shadow price that would have impacted energy prices. This would have been a false "non-signal" under these circumstances. The constraint biasing was effective through the remainder of this period and, as seen in Figure 5.7, in HE 17 the actual flow began to decrease below the actual limit. Operators decreased the amount of downward bias on the limit to account for this and, subsequently, in HE 18 again increased the amount of bias in response to the relationship between actual flow and the actual limit.

Figure 5.8 Flows and Shadow Prices for the Los Banos North BG on September 18



Example 4: High PG&E LAP prices Due to Congestion on the Los Banos North Branch Group (September 18)

In this example, actual flow on the Los Banos North branch group was higher than the market flow, and actual flow was approaching the nomogram limit. Grid conditions that contributed to the increased challenge in managing congestion on this branch group include the Moss Landing-Metcalf 500kV line outage that resulted in de-rates on both Path 15 and the Los Banos North branch group. Multiple lines were congested, with some actually overloaded on this day, creating difficulty managing the congestion as one line would overload to resolve congestion on another. In addition to transmission de-rates, the system was approximately 3,000 MW under-scheduled at the peak. In the nomogram, limit was biased down by roughly 450 MW in HE 11 and became binding in HE 11 interval 3. Actual flow trended down through the middle of HE 13 and the biased limit was increased somewhat. During these earlier hours, the biasing did force congestion with shadow prices up to nearly $600/MW. This had a significant impact on the PG&E LAP LMP, with periods at $500/MW. Maintaining this bias level was effective in keeping the actual flow from exceeding the actual nomogram limit through HE 16. In HE 17 the actual flow increased and exceeded the actual branch group limit. As seen in Figure 5.8, operators further biased the market limit down to get the market software to dispatch to a lower limit and keep the actual flow below the actual limit.

5.4 Recommendations

Based on DMM's review of the issue of constraint biasing, we are providing the following recommendations:

- Given the dynamic nature of discrepancies between modeled and actual flows – and the significant impact that biasing can have on market outcomes – the ISO should continue to place a high priority on refining the use of constraint biasing in the day-ahead and real-time processes. As the ISO gains additional experience and data on discrepancies between modeled and actual flows, this may be utilized to improve how the potential reliability and market impacts of biasing are balanced. For example, more automated statistical metrics that correlate the degree of biasing and congestion in the various markets may be helpful in tracking trends and identifying potential areas for improvement as conditions change.

- While we have observed consistent biasing across the real-time markets in Q3, applying a bias is a manual process that takes some time and must be repeated for the different real-time markets. Thus, DMM suggests that use of the bias might be made more effective by developing a tool for operators that better facilitates applying bias across markets.

- Overall market transparency can be improved by providing timely data to market participants on the application of bias and un-enforcing of limits in market operation. DMM understands this issue will be addressed as part of a stakeholder process starting in Q4 2009.

6 Resource Adequacy

This section provides an analysis of the availability of Resource Adequacy (RA) supply to the ISO markets during the 140 highest peak load hours of July through September 2009 (Q3), corresponding to all hours with loads over 40,000 MW. The overall average availability of RA resources was relatively high during these hours: about 91 percent in the IFM and 88 percent in RUC. This represents an overall availability just slightly below the 92 percent level that is implicitly incorporated in RA program requirements.[47] However, DMM notes that under higher loads that equal or exceed the 1-in-2 year peak forecast used in setting RA requirements, this difference could have a significant impact on ISO market performance and system reliability. DMM believes these findings reinforce the need for the ISO to continue to consider future refinements to the ISO's RA Standard Capacity Product (SCP) tariff provisions. Refinements to the SCP provision aimed at measuring the amounts of all RA capacity actually made available to the ISO markets through bids or self-schedules could ensure the sufficient level of overall availability of RA resources can be maintained.

6.1 Background

The Resource Adequacy program is a key component of the ISO market that is designed to ensure there will be sufficient generation capacity to meet demand, particularly under high peak load conditions. Under the RA program, load-serving entities (LSEs) generally must arrange enough RA generation and demand response capacity to meet 115 percent of their forecast peak demand in each month (based on a 1-in-2 year peak forecast). The 115 percent requirement is designed to include the additional operating reserve needed above peak load (about 7 percent), plus an allowance for outages and other resource limitations (about 8 percent).

About half of the generation resources counted toward this RA requirement are required to be made available to the ISO markets for each hour of the month that the resource is physically available. Exceptions to this "all hours" must-offer requirement include hydro resources, non-dispatchable resources, and "use-limited" thermal resources, which are to be made available to the ISO markets consistent with their operating limitations. Use-limited thermal resources generally have environmental or regulatory restrictions on the hours they can operate, such as a maximum number of operating hours in a month or year (e.g., the 360 hour per year operating limit placed on many peaking units within transmission constrained areas imposed under air permitting regulations). Market participants submit "use plans" for use-limited RA resources to the ISO that describe these restrictions and outline the planned operation of these units.

Market participants make RA resources available to the ISO markets by submitting economic bids or self-schedules to the Integrated Forward Market (IFM) and, depending on the type of resource, to the Residual Unit Commitment (RUC) process and to the Real Time Market (RTM).

- For just under half of RA capacity (including over 23,000 MW of non-use-limited gas-fired generation), the ISO automatically creates the required IFM energy or RUC bids if a bid or

[47] 115 percent RA requirements less 7 percent operating reserve = 108 percent. Thus, after accounting for operating reserve, just over 92 percent of remaining RA resources would be necessary to meet the 1-in-2 year peak load used in setting the RA requirement.

self-schedule is not submitted by the market participant. Bids are not submitted for any capacity that is unavailable due to a scheduled outage, forced outage or de-rate, as reported through the ISO's outage reporting system (SLIC).

- If these resources are committed in the IFM or RUC, then they continue to have a must-offer obligation in the RTM, with the ISO automatically creating the required RTM energy bid if the capacity is not scheduled or bid by the participant.

- In addition, all non-use-limited short-start units that are RA resources are also required to be bid into the RTM, so that the ISO automatically creates the required RTM energy bid if the capacity is not scheduled or bid by the participant.

However, for the other half of the RA resource fleet, the ISO does not create a bid if one is not submitted by a market participant.

- The ISO does not create bids for about 6,400 MW of hydro resources and over 900 MW of use-limited thermal units, since the RA program assumes that market participants will manage availability of these resources and submit bids and self-schedules to make them available to the ISO consistent with their operating restrictions.

- The ISO also does not create bids for about 10,000 MW of non-dispatchable generators, which include nuclear, qualifying facilities (QFs), wind, solar and other miscellaneous resources.

- Currently, the ISO does not create bids for import resources, which accounted for over 4,000 MW of RA capacity in Q3 2009.[48]

Under California Public Utilities Commission's (CPUC) rules, a resource must be available at least 140 hours over the summer months of July through September to be counted as RA.[49] The RA program presumes that market participants will manage the use of resources that cannot be available in all hours of a month to make them available to the ISO during the peak load hours. Since a generator must be able to operate in at least 140 hours over July through September to be counted as an RA resource, we have evaluated the availability of RA generation during the 140 hours during these months with the highest peak loads (i.e., all hours with peak load over about 40,000 MW). While CPUC requirements do not require that RA resources be available during these specific 140 peak hours and participants do not have perfect foresight about which hours will have the highest loads over the summer, we have chosen to assess RA availability during these peak 140 hours in order to provide results that are – in aggregate – roughly comparable to basic market design assumptions that appear to underlie the 140 hour requirement incorporated in the CPUC's RA requirements (i.e., that this 140 hour requirement will provide a high level of availability during peak hours when most RA capacity is needed to ensure reliability).

[48] Although the ISO does not currently create bids for RA import resources if not submitted by the market participant, DMM understands that the ISO is currently implementing the process for doing so. DMM believes that it is important that this is in place as soon as possible, and preferably by next summer.

[49] The CPUC requires that RA resources be available at least 210 hours during the months of May through September based on the resources being available 30, 40, 40, 60, and 40 hours during each of these months, respectively.

6.2 *Analysis of Resource Adequacy Availability*

Figure 6.1 provides an overview of monthly RA requirements, monthly peak load and the frequency of the 140 highest load hours (with load over 38,000 MW) that occurred during July through September 2009.

- The red and yellow lines (plotted against the left axis) compare the monthly RA capacity with the peak load that actually occurred during each of these months. As shown in Figure 6.1, the ISO's total RA capacity was approximately 54,000 to 58,000 MW during these months, which exceeded the monthly peak load in July and August by about 30 percent, and the monthly peak load in September by about 17 percent. The relatively high margins in July and August reflect that RA requirements are designed to meet 115 percent of a 1-in-2 year load forecast, and that peak loads in these months were not unusually high. The lower margin in September reflects the fact that highest peak load occurred in this month, when RA requirements were actually lower.

- The bars in Figure 6.1 show the number of the top 140 load hours during Q3 that occurred during each of these months. These represent the specific 140 hours upon which the analysis in this chapter presented below are based.

The fact that the actual summer peak and a high portion of the highest load hours each summer may not come in the month with the most RA capacity underscores the need for all RA capacity to be made available to the ISO markets, particularly in the peak load hours.

Figure 6.1 Monthly Total RA Capacity, Peak Load, and Peak Load Hours July-Sep 2009

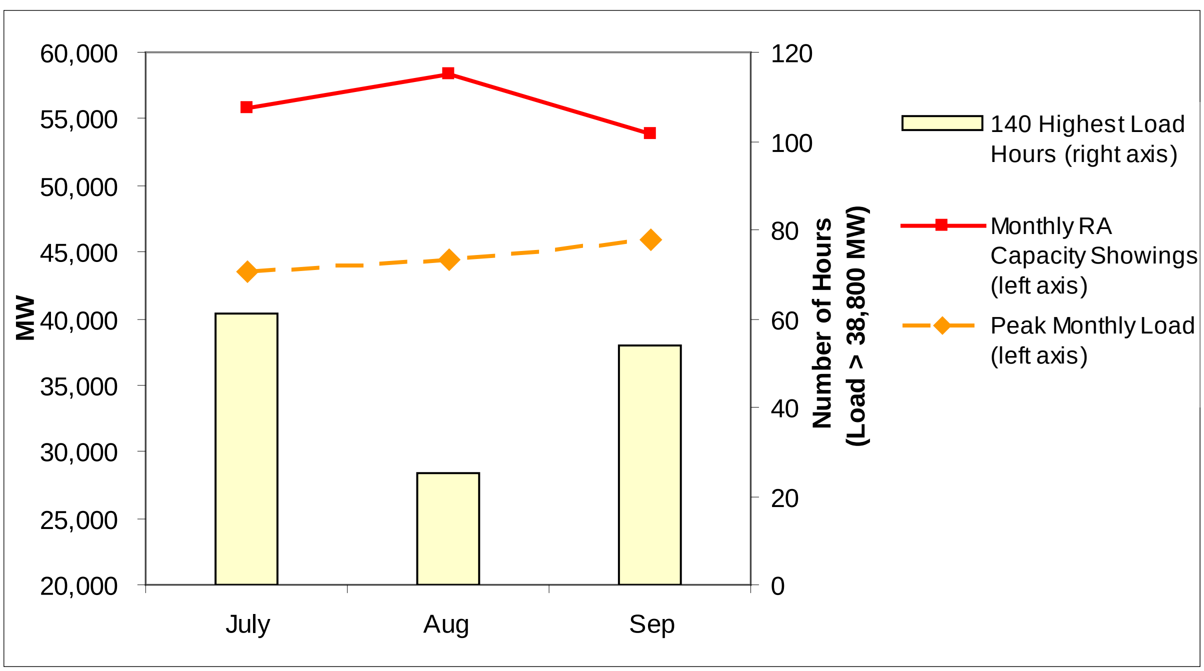


Figure 6.2 summarizes the amount of RA capacity for which bids and self-schedules were available to the ISO markets (IFM, RUC and RTM) during these 140 peak hours in term of an "availability duration curve", as explained below:

- The left vertical axis of Figure 6.2 shows the total RA capacity available to the ISO markets over the 140 highest load hours between July and September (ranked in descending order of total RA megawatt bid or scheduled in each of these three markets).[50]
- On the right vertical axis, Figure 6.2 shows the RA capacity available to the ISO markets as a percentage of the average overall RA capacity over these peak hours.
- The horizontal axis of Figure 6.2 shows the number of hours that the RA capacity listed on the left and right vertical axes was available to the IFM, RUC, and RTM.
- The IFM bids and self-schedule amounts shown include bids and self-schedules for energy and ancillary services for RA capacity.
- The RUC bid amounts shown include RUC bids for RA capacity, as well as the amounts of energy or ancillary services from RA capacity that cleared in the IFM.
- The RTD bid amounts shown include energy bids and self-schedules for energy from RA capacity submitted to the RTD, as well as RA capacity included in an IFM energy schedule.

The approximately 46,300 MW of RA capacity included in this analysis excludes about 9,700 MW of RA capacity for which this analysis cannot be performed or is not highly meaningful (such as RA resources representing "liquidated damages contracts", RA capacity from Reliability Must Run resources, RA requirements met by demand response programs, and load-following metered subsystem resources).[51]

Figure 6.2 shows that a relatively high proportion of RA capacity was available to the IFM, RUC, and RTD during the 140 summer peak load hours.

- In the IFM, bids and self-schedules for RA resources averaged about 91 percent, with a range of about 97 to 85 percent during these highest 140 load hours.
- In RUC, the amount of RA capacity available averaged 88 percent of the average overall RA capacity, with a range of about 94 to 81 percent across the 140 hours. The lower amount of RA capacity available to RUC than the IFM reflects the fact that market participants did not submit RUC bids for some resources that they bid or scheduled in the IFM.
- In the RTM, the amount of RA capacity averaged only about 80 percent of RA capacity, and varied from 85 to 73 percent. This lower amount and variability of RA capacity made available to the RTM is due to not all RA resources being committed in the IFM or RUC, as

[50] Figure 6.2 does not include approximately 9,700 MW of the overall ISO RA capacity. Figure 6.2 does not include RA capacity from some import and liquidated damages contracts that do not have specific ISO "resource IDs," which make it possible to track submitted bids and self-schedules. Figure 6.2 also does not include RA capacity from Reliability Must Run resources, demand response resources, and load-following metered subsystem resources, for which the lack of a submitted bid or schedule does not necessarily make the resource unavailable.

[51] See Footnote 50.

well as different amounts of RA resources being committed in the IFM in different hours. As discussed below, bids and self-schedules were submitted for a relatively high proportion of the RA capacity that was available in the RTD.

Figure 6.2 RA Bids and Self-Schedules Available to the IFM, RUC, and RTM 140 Highest Peak Load Hours (July-Sep 2009)

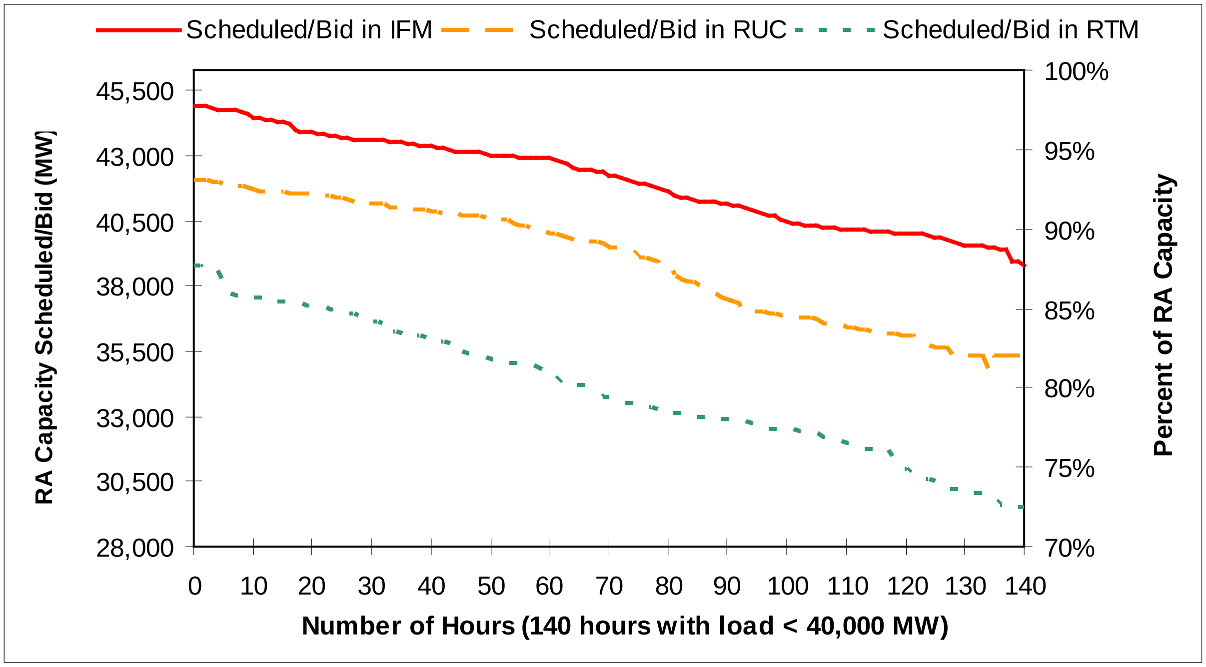


Table 6.1 provides a more disaggregated summary of the analysis depicted in Figure 6.2 in terms of a variety of different types of generation resources, including sub-totals for two categories of resource types: (1) resources for which the ISO creates bids if a bid or self-schedule is not submitted for RA capacity, and (2) resources for which no bids are created by the ISO to ensure that resources adhere to RA bidding requirements.

- ***RA Capacity After Reported Outages and De-rates*.** The first three numerical columns of Table 6.1 list the approximately 46,000 MW of capacity used to meet RA requirements in Q3 that were examined in this analysis and the actual capacity after adjusting for reported outages and de-rates during the 140 highest load hours (in megawatts and as a percent of total RA capacity). As shown in Table 6.1, the total availability for the over 23,000 MW of non-use-limited gas-fired generation was about 92 percent, representing an outage rate of about 8 percent during the highest 140 load hours. The overall RA capacity, including all resource types, after adjusting for reported outages and de-rates was about 94 percent of the overall RA capacity, representing an outage rate of about 6 percent during the highest 140 load hours.

- ***IFM Availability*.** Table 6.1 then lists the average amounts of bids and self-schedules actually scheduled or bid in the IFM (in megawatts and as a percent of total RA capacity). For the 23,000 MW of thermal RA resources for which the ISO submits bids based on their reported availability, IFM availability was unchanged at an average of 92 percent. For the more than 22,000 MW of RA capacity for which bids are not automatically submitted in the IFM, the total amount of capacity that was scheduled or bid in the IFM averaged only 90 percent, bringing the total average availability of all RA capacity examined in this analysis in the IFM to about 91 percent. This is somewhat less than the 94 percent of RA capacity that was available after adjusting for reported outages and de-rates.[52]

- ***RUC Availability*.** Table 6.1 then lists the average amounts of bids and self-schedules actually scheduled or bid in the RUC process. The overall percentage of RA capacity made available in the RUC process drops to 88 percent compared to 91 percent in the IFM. As shown in Table 6.1, the major reason for this is that RUC bids are not submitted for all RA imports.[53] As previously noted, DMM understands that the ISO is currently implementing software modifications that will address this issue by automatically inserting RUC bids for any RA import capacity that is not bid into RUC.

- ***RTM Availability*.** The last three columns of Table 6.1 compare the total RA capacity from these resources that were obligated to be available in RTD with the actual schedules and bids for these resources in RTD during the 140 hours examined in this report. The RA capacity that should have been available to the RTM is calculated as the remaining RA capacity from resources with an IFM or RUC schedule plus RA capacity from uncommitted short-start units. On average, about 92 percent of the RA capacity that was potentially available to the RTM was made available. This is slightly more than the percentage of RA capacity for which bids and self-schedules were submitted to the IFM.

[52] Some of this difference may also have been due to outages of import resources, for which market participants cannot currently report outages or de-rates through the ISO's SLIC system.

[53] These shortfalls are most likely attributed to some market participants failing to submit RUC bids rather than due to resources not being physically available. If a resource is available for a given day and a bid or self-schedule is submitted to the IFM, then that resource should presumably be available for the same day in RUC.

Table 6.1 Average RA Capacity and Availability to IFM, RUC, and RTM July-Sept. 2009

Resource Type	Total RA Capacity (MW)	Net Outage Adjusted RA Capacity		IFM Bids and Self-Schedules		RUC Bids		Total RTM RA Capacity (MW)	RTM Bids and Self-Schedules	
		MW	% of Total	MW	% of Total	MW	% of Total		MW	% of RTM
ISO Creates Bids:										
Gas-Fired Generators	23,020	21,205	92%	21,182	92%	21,182	92%	17,364	16,522	95%
Other Generators	990	913	92%	913	92%	912	92%	986	908	92%
Subtotal	24,010	22,118	92%	22,095	92%	22,094	92%	18,350	17,430	95%
ISO Does Not Create Bids:										
Use-Limited Gas Units	913	902	99%	850	93%	743	81%	887	802	90%
Hydro Generators	6,406	6,220	97%	5,788	90%	5,774	90%	6,406	5,420	85%
Nuclear Generators	4,870	4,716	97%	4,645	95%	4,645	95%	4,870	4,694	96%
QF Generators	4,505	4,352	97%	3,910	87%	3,787	84%	4,483	3,875	86%
Wind Generators	659	657	100%	388	59%	388	59%	659	503	76%
Other (Non-Dispachable)	744	568	76%	527	71%	527	71%	744	529	71%
Imports	4,194	4,194	100%	3,866	92%	2,921	70%	3,818	3,762	99%
Subtotal	22,291	21,609	97%	19,974	90%	18,785	84%	21,867	19,585	90%
Total	46,301	43,727	94%	42,069	91%	40,879	88%	40,217	37,015	92%

6.3 Conclusion and Recommendations

During the peak hours examined in this analysis, the overall average availability of RA resources was relatively high: about 91 percent in the IFM and 88 percent in RUC. This represents an overall availability just slightly below the 92 percent level that is implicitly incorporated in RA program requirements.[54] DMM notes that under higher loads that equal or exceed the 1-in-2 year peak load conditions used in setting RA requirements, this difference could have a significant impact on ISO market performance and system reliability. DMM also believes these findings reinforce the need to maintain or improve overall availability of RA resources, and for the ISO to continue to consider future refinements to the RA process and the ISO's recent RA Standard Capacity Product (SCP) tariff provisions.

Performance incentives for internal RA resources to be implemented under the SCP only address forced outage rates.[55] Refinements to the SCP to measure the amounts of all RA capacity actually made available by RA resources to the ISO markets through bids or self-schedules may help ensure that the required overall level of availability of RA resources can be maintained. DMM believes the following findings should be considered in developing future refinements to the RA and SCP provisions:

- DMM has calculated that under the SCP provisions, the penalty assessed for non-availability during any individual critical peak hour would be no more than about \$34/MW.[56] Under critical system or peak load conditions, this may not provide a high incentive to ensure that RA capacity is available to the ISO markets. From a longer-term perspective, this may also not provide an efficient price signal for investment in new capacity that would actually be available to the ISO system on these highest load hours.

- For RA imports, which are not required to be backed by specific generating resources under current RA program guidelines, it could often be more profitable under peak load conditions to simply incur the maximum potential \$34/MW charge under the SCP provisions rather than procure the energy and transmission needed to fulfill this RA obligation (or perhaps sell any available energy to other buyers). For example, a supplier without a physical generating resource that is supplying RA import capacity by procuring energy in the bilateral market may find the cost to procure energy and transmission exceeds the ISO price paid during a critical peak hour. In practice, under new SCP provisions, if these RA importers met their full RA bid obligation during other non-critical hours, there would actually be no cost to these importers of not making RA imports available during the few critical peak hours.[57]

---

[54] 115 percent RA requirements less 7 percent operating reserve = 108 percent. Thus, after accounting for operating reserve, just over 92 percent of remaining RA resources would be necessary to meet the 1-in-2 year peak load used in setting the RA requirement.

[55] *Comments on Updated Proposal for Standard Capacity Product*, Department of Market Monitoring, December 19, 2008, http://www.caiso.com/20a2/20a2e7b12ae60.pdf.

[56] Also, it is important to note that a resource would not pay any penalty as long as the resource did not fall below the minimum threshold for availability over all hours in the month used to calculate the penalty provisions of the SCP (i.e., being available about 93 percent of 100 peak hours per month).

[57] For instance, if an entity with an RA import obligation bids its RA capacity at the price cap during about 93.5 percent of hours (with the lowest expected prices and loads), the RA importer could fulfill its RA obligation without incurring any penalties during the 6.5 percent of hours with the highest expected prices and loads.

- During the 140 highest load hours examined in this analysis, the overall availability of QFs and other renewable and non-dispatchable resources was less than 90 percent of the amount of these resources' capacity that was counted to meet RA requirements. The qualifying capacity of these resources that may be used to meet RA requirements is generally based on each unit's historical average output during the hours of noon to six of the month the resource is to be counted as RA capacity.[58] Thus, the qualifying capacity of these resources has already been adjusted to reflect their average availability during peak hours. However, as shown by these results, the actual availability of these resources can be significantly lower during the highest load peak hours when this capacity is most critical to reliability and market performance. Moreover, these resources will initially be exempt from SCP provisions.

- For the over 900 MW of use-limited gas-fired resources (most of which are limited to 360 hours of operation per year under air permitting regulations), these SCP performance incentives are based on reported outages, rather than the capacity actually made available to the ISO markets. As shown in Table 6.1, while these resources were reported to be available an average of 99 percent of their RA capacity during Q3, an average of only about 93 percent of the reported capacity of these resources was actually made available in the IFM, and an average of only 90 percent of this capacity was made available in the RTM. This underscores the potential usefulness of a performance incentive based on submitted bids and schedules, rather than being based on only reported unit availability based on de-rates and outages. For example, this would provide an incentive for use-limited gas resource to schedule or bid a unit as contingency-only non-spinning reserve during most or all hours, in order to make this capacity available in the market while limiting the actual hours the resource is dispatched.

- Finally, the fact that the availability of use-limited resources during peak hours was generally less than the planned RA capacity reinforces the need for the ISO to thoroughly review the use-plans submitted for use-limited resources. The initial operation of the new market during summer conditions provides historical data that can be used to evaluate these use-plans in the future.

[58] The methodology to calculate the RA capacity that wind resources can provide to entities under the jurisdiction of the CPUC has been revised for next year and will be based on the capacity a resource can provide in more than 70 percent of the peak hours.