June 29, 2004

The Honorable Magalie Roman Salas
Secretary
Federal Energy Regulatory Commission
888 First Street, N.E.
Washington, D.C. 20426

Re: California Independent System Operator Corporation, ER02-1656

Dear Secretary Salas,

The California Independent System Operator Corporation ("ISO") hereby respectfully submits for filing an original and fourteen copies of two reports on the performance of the Automated Mitigation Procedures, one for the three months ended September 30, 2003, and one for the three months ended December 31, 2003, as directed by Commission's July 17, 2002 Order, *California Independent System Operator Corporation*, 100 FERC ¶ 61,060 (2002). Please return one file-stamped copy of each report to the messenger.

These reports also will be posted on the ISO's web site (http://www.caiso.com).

Thank you for your assistance in this matter.

Respectfully submitted,

Anthony J. Ivancovich
Senior Regulatory Counsel
California Independent System Operator Corporation
151 Blue Ravine Road
Folsom, CA 95630
(916) 608-7135

Third Automated Mitigation Procedure Report
July 1, 2003 through September 30, 2003
California Independent System Operator Corporation
June 29, 2004

Executive Summary

As directed by the Federal Energy Regulatory Commission ("Commission") in its July 17, 2002 Order[1], the ISO has prepared this third Quarterly Report on the Performance of the Automated Mitigation Procedure ("AMP"). AMP, proposed by the ISO in its May 1, 2002 Market Redesign 2002 filing, was approved by the Commission with modifications in the July 17 Order. This report provides observations and analysis of trends pertaining to the effectiveness of AMP mitigation for the period covering July 1, 2003 through September 30, 2003.

Because the real-time market for incremental balancing energy has become increasingly thin in recent months, it remains uncertain whether AMP would provide effective market power mitigation during times of market stress. Through September 2003, AMP has yet to mitigate a single bid, and fundamental market conditions, particularly in the incremental balancing market, have been relatively mild. While market conditions observed during this period do not provide a stress test equivalent to another energy crisis, they do provide some insight into the potential effectiveness of the mitigation measures.

This report takes a new direction in monitoring the performance of AMP. Whereas previous Quarterly Reports looked generally at AMP as it affected bidding behavior on an overall basis, the present analysis focuses on hours in which bidding behavior that is consistent with the exercise of market power is most extreme. This will enable us to draw conclusions pertaining to AMP's efficacy in mitigating those bids in particular.

One of the major concerns with AMP is that the rolling-average bid characteristic of reference levels introduces the perverse incentive to raise bids in order to increase future reference levels. As a result, sellers may have the opportunity to bid in a manner that is consistent with the exercise of market power. Alternatively, the reference level may drift sufficiently high that bids are actually constrained not by the reference level, but by the market-wide $250/MWh soft price cap.

Another problematic element of AMP has been the "price screen." This refers to the requirement that bids will be subject to the AMP Conduct and Impact Tests only in the case that the real-time market-clearing price is expected to exceed $91.87/MWh in any zone in any interval during the hour of operation. Due to operational system limitations, this price screen effectively is applied 53 minutes prior to the hour of operation based on the projected imbalance energy dispatch for that hour of operation. This means that if AMP is not triggered due to an expected price greater than $91.87/MWh in the next hour at 53 minutes before that hour, AMP will not be triggered at all for the next hour, even if a contingency occurs after 53 minutes before the beginning of the hour that causes the actual price to be greater than $91.87/MWh (an interval of time of one hour and 53 minutes). Furthermore, the price prediction algorithm, which necessarily is imperfect, itself suffers from both "false-positive" and "false-negative" type errors: the algorithm may predict that a price will exceed $91.87/MWh when in fact the actual price is below that level; or it predicts that a price will be below $91.87/MWh when in fact the actual price is above that level.

[1] *California Independent System Operator Corporation*, 100 FERC ¶ 61,060 (2002) ("July 17 Order").

We have observed the following:

- Among the most extreme thirteen hours in which the estimated market impact from price markup due to bidding behavior consistent with the exercise of market power exceeded $100,000, bidders failed the Conduct Test in only two hours, and would have failed the Conduct Test in only one hour had the $91.87/MWh price screen had not prevented AMP from being applied at that time.
- The Conduct Test identified units that bid in a manner consistent with the exercise of market power in nine of 75 hours (12% of the time) during which the measured price-to-cost markup was greater than 40 percent between October 30, 2002, and September 30, 2003.
- Awarded bids were from units whose reference levels were sufficiently high such that their bids would have to be above the $250/MWh price cap to fail the Conduct Test in 11 of the same 75 hours (14.7% of the time). [2]
- On days with the most extreme markup (markup in excess of 100 percent of estimated costs, or 50 percent of the price) when units failed the AMP Conduct Test, mitigation would likely have had a minimal effect on prices (reducing the price by less than six percent).
- Reference levels in general have not increased when considering the entire portfolio of units in the ISO Control Area.
- However, reference levels among units that actually are awarded dispatch instructions have increased noticeably, even when controlling for natural gas price adjustments to reference levels.

The following analysis looks at each of these issues in further detail. The Department of Market Analysis has developed several new indicators to aid in answering these questions. Many of the charts on the following pages address several of the above questions simultaneously, and will be used again in future AMP Quarterly Reports.

Effectiveness of the AMP Price Screen and Conduct Test

The $91.87/MWh price screen significantly reduces the likelihood that AMP will be implemented during a price spike. In the case that the predicted price exceeds $91.87/MWh and bids are then subject to the Conduct Test, bids fail whenever a single segment's price exceeds the unit's corresponding reference level by the lesser of at least $100/MWh or 200 percent. We categorize Conduct Test failures according to the following definitions:

- A bid whose price exceeds its reference level by the lesser of 200% or $100/MWh in any bid step, regardless of whether the Price Screen would have prevented the application of the Conduct Test, is a *Potential Conduct Test Failure*. A bid that officially failed the Conduct Test, by bidding sufficiently above its reference level threshold in an hour in which

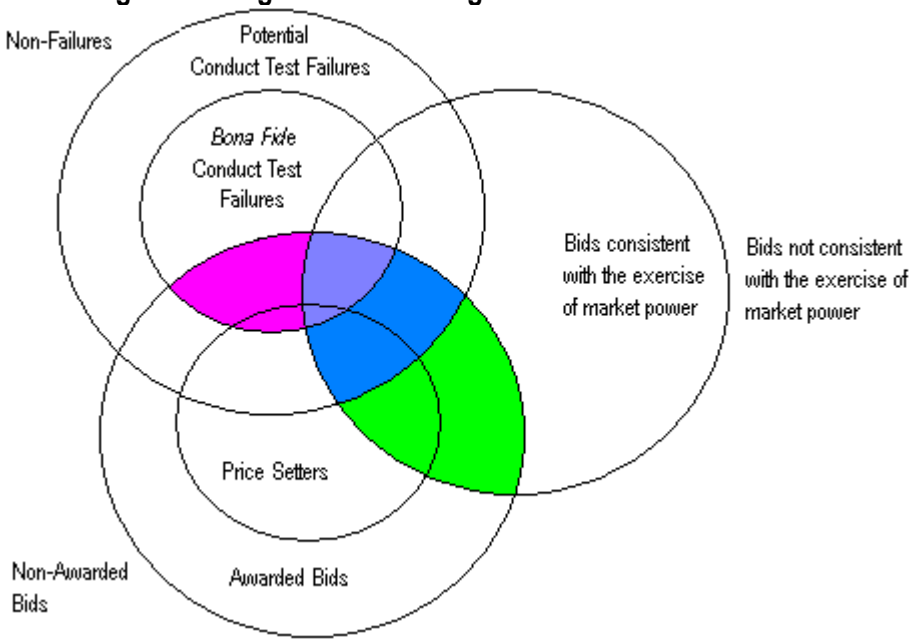
[2] These include but are not limited to two municipal units that bid exclusively above the price cap, by specific order of the South Coast Air Quality Management District, and are dispatched only when the bid stack is otherwise fully exhausted (usually during contingencies or price spikes).

the Price Screen exceeded $91.87/MWh and thus AMP was activated, is a *Bona Fide Conduct Test Failure.*

- A bid at a price in excess of $100/MWh and at least 20 percent above its estimated marginal cost, in an hour in which total price-to-cost markup is estimated to be above 40 percent, is considered to be *Consistent with the Exercise of Market Power.*
- A bid either is or is not awarded an in-sequence dispatch. If it is, it is noted as an *Awarded Bid.* In each interval with a dispatch, a single awarded bid sets the market-clearing price; this bid is known as a *Price Setter.*

None of these categories is either necessary or sufficient for any other category, with the exceptions that a B*ona Fide* Conduct Test Failure necessarily is a Potential Conduct Test Failure, and Price Setters comprise a subset of Awarded Bids. That is to say, a bid can fall into the category of Awarded Bids Consistent with the Exercise of Market Power, that are not Potential Conduct Test Failures, or non-awarded B*ona Fide* Conduct Test Failures, etc. The following Venn diagram depicts the potential categories of bids:

Figure 1. Diagram of Bid Categories for the Conduct Test



In this diagram,

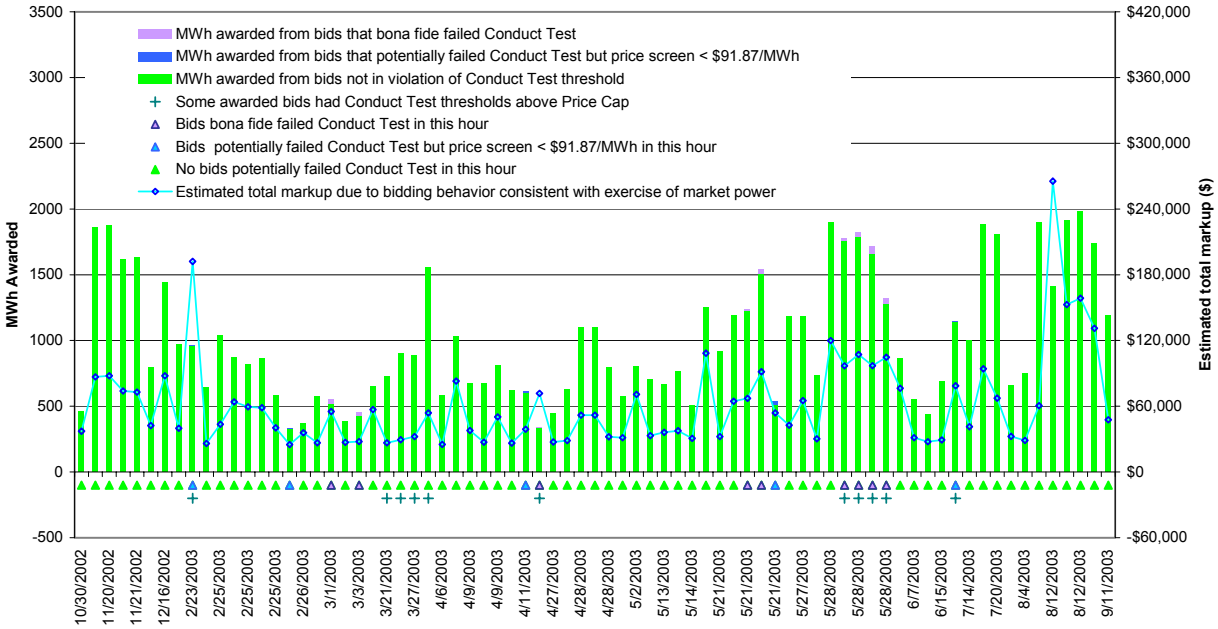
- The **pink** and **purple** areas depict the Awarded Bids that are also *Bona Fide* Conduct Test Failures. These are the units that would be subject to the Impact Test and potentially could be mitigated. However, only bids in the **purple** area are those that are of concern, since they are consistent with the exercise of market power. The pink area is an example of "Type II Error" – test failures that are of no risk to the market. There is a small risk that the low-priced, non-harmful bids in this region will be subject to mitigation if they also fail the Impact Test.

- The **blue** area depicts those bids that are Consistent with the Exercise of Market Power, but to which the Conduct Test was never applied, due to the fact that the predicted price 53 minutes ahead of the hour was below $91.87/MWh. This is an example of "Type I Error" – non-failures that are of concern, because the price screen prevents AMP from being applied to those bids.
- The **green** area depicts those Awarded Bids Consistent with the Exercise of Market Power that pass the Conduct Test, since they are within the permitted Conduct Test thresholds. This is also an example of "Type I Error," because those bids have the potential to exercise market power while bidding within the allowed parameters of AMP.

While the purple region represents the set of bids that are the desired potential candidates for mitigation and should be as inclusive as possible, the test itself may create incentives to bid such that few bids will actually fall into this region in the long run.

The following chart compares average volume for the purple, blue, and green regions during the most extreme price spike hours where prices exceeded $100/MWh, and the estimated price-to-cost markup was at least 40 percent, between October 30, 2002, and September 30, 2003.

Figure 2. Awarded Bids above $100/MWh in Hours with Markup of at least 40 Percent above Cost and at least $25,000: Conduct Test Failures and Non-Failures



The above instances where the market costs related to the markup were greater than $100,000 has occurred in 13 hours since the implementation of AMP in October 2002 through the end of the quarter. Among these hours, dispatched units failed the Conduct Test in only two hours (both on May 28, 2003). Dispatched units also bid in a manner that would have failed the Conduct Test, had the price screen not been below $91.87/MWh, in one hour on February 23 (the all-time second highest estimated markup hour). In the other four of the top five markup hours, all on August 12,

2003, no awarded units bid in a manner that would have failed the Conduct Test. Therefore, even during the highest price-to-cost markup hours, the conduct test was seldom violated given the $100/MWh or 200 percent conduct test threshold.

Awarded Bids resulted in *Bona Fide* Conduct Test Failures in nine of the 75 hours during which incremental balancing energy was procured for at least $100/MWh in any interval. Hourly markup was at least 40 percent of estimated marginal cost and exceeded $25,000 in the 11 months through September 2003. All of those hours occurred between March 2003 and May 2003; none occurred during the quarter that is the subject of this report.

Awarded bids above $100/MWh potentially would have failed the Conduct Test, had the predicted price exceeded $91.87/MWh, in another five of the 75 hours during which incremental energy was sold for at least $100/MWh, and markup exceeded 40 percent of cost and was at least $25,000. Of these five hours, only one hour occurred during the subject quarter. This hour was during a price spike on July 2, during which the market price reached $250/MWh. The total cost in that hour that can be attributed to markup was approximately $79,000, or 51.4 percent of cost.

In 11 hours since October 30, 2002, awarded bids had corresponding reference levels for some portion of the bid above $150/MWh. If such a unit had bid that portion above its reference level threshold and had bid been awarded, it would have been required to cost-justify the bid to the Commission, or be subject to refund.

The following table summarizes the bids with *Bona Fide* and/or Potential Conduct Test Failures in the 12 hours in which markup exceeded 40% of estimated marginal cost and the market-clearing price was above $100/MWh. Except in the cases noted, the bids that set the prices in these hours were the units that failed the Conduct Test.

Table 1. Awarded Bona Fide and Potential Conduct Test Failures, And Hours in which Awarded Bids had Conduct Test thresholds above the Price Cap, During which Market Power may have been Exercised

Date of Operation	Hour of Operation	Estimated total markup due to bidding behavior consistent with exercise of market power	Estimated Markup Index (markup as a percentage of marginal cost)	Bids bona fide failed Conduct Test in this hour ("Purple")	Bids potentially failed Conduct Test but price screen < $91.87/MWh in this hour ("Blue")	Some awarded bids had Conduct Test thresholds above Price Cap
2/23/2003	19	$ 192,163.50	166.5%		*	*
2/25/2003	19	$ 25,011.58	111.6%		*	
3/1/2003	23	$ 55,183.81	104.4%	*		
3/3/2003	20	$ 27,887.99	82.1%	*		
3/21/2003	20	$ 26,604.27	42.2%			*
3/25/2003	19	$ 29,329.05	42.8%			*
3/27/2003	19	$ 32,375.26	42.9%			*
3/30/2003	19	$ 53,885.00	41.3%			*
4/11/2003	23	$ 38,979.66	107.9%		*	
4/25/2003	23	$ 71,743.00	355.7%	*		*
5/21/2003	16	$ 67,259.31	45.5%	*		
5/21/2003	17	$ 91,393.14	51.6%	*		
5/21/2003	18	$ 53,899.36	133.7%		*	
5/28/2003	15	$ 96,735.36	44.6%	*		*
5/28/2003	16	$ 107,268.70	45.7%	*		*
5/28/2003	17	$ 97,087.74	43.8%	*		*
5/28/2003	18	$ 104,799.79	78.4%	*		*
7/2/2003	20	$ 78,722.05	51.4%		*	*

In all other hours in which balancing energy was procured for at least $100/MWh and hourly markup exceeded the minimum of (1) 40 percent of cost or (2) $25,000, no unit that was awarded energy bid in a manner that could have failed the Conduct Test. In particular, four of the five highest total estimated hourly markups since the implementation of AMP occurred on August 12, during which the incremental price stood at $189/MWh. These markups ranged from $131,000 (73.5 percent above cost, in Hour Ending 17) to $265,500 (243.1 percent above cost, in Hour Ending 14). The following table shows high-markup hours during the subject quarter in which awarded units did not bid in a manner that would have failed the Conduct Test.

Table 2. Hours in which no units potentially failed the Conduct Test, yet Markup Index Exceeded 40% and MCP Exceeded $100/MWh, July-September 2003

Date	Hour	MWh Awarded from Bids that Did Not Potentially Fail Conduct Test ("Green")	Est. Total Market Impact of Markup	Markup Index	Notes
7/20/2003	15	1885	$ 94,049.63	42.1%	
7/24/2003	16	655	$ 32,414.96	44.3%	
8/12/2003	14	1409	$ 265,500.82	70.9%	
8/12/2003	15	1909	$ 152,684.11	42.3%	3
8/12/2003	16	1983	$ 158,631.37	42.3%	2
8/12/2003	17	1741	$ 131,043.56	42.3%	2

On August 12, the SP15 Zone reached its all-time record Demand, and limited transmission into the Zone caused the market price for this Zone to be set at $189/MWh for approximately three hours. Due to the very high dispatch volume during this spike, its total market impact was approximately $800,000,[4] of which at least $576,000 can be attributed to bidding behavior that is Consistent with the Exercise of Market Power, as shown in the preceding table. These spikes are described in further detail in the Market Analysis Reports for July-August and September 2003.[5]

In five (5) of 54 high-market-impact hours, the price-setting bid's corresponding threshold above which it would fail the Conduct Test exceeded the soft price cap of $250/MWh, since the reference level corresponding to the price-setting bid was at least $150/MWh. That is, a unit that failed the Conduct Test would have set a $250/MWh price (even if its bid was higher), would have been paid as bid, and also would have had to provide cost justification for the bid to FERC.

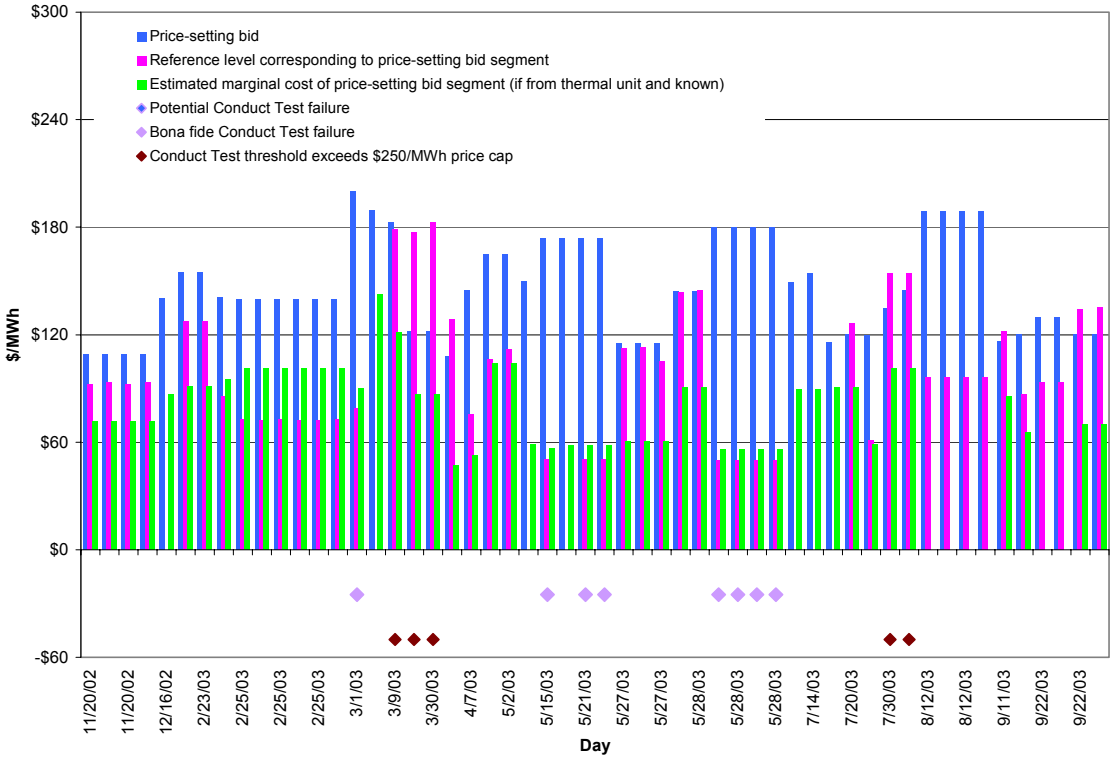
The following chart shows individual price-setting bids with market impacts above $60,000, with corresponding reference levels, marginal costs, and Conduct Test information.

[3] Hours ending 15-17 on August 12, 2003, show the same markup index because the marginal unit was the same in each of these hours.

[4] " Market Impact" is defined as the difference between the total cost of energy during the spike and the cost of an equal quantity of energy at the weekly average price.

[5] These reports are available on the ISO web site at http://www.caiso.com/docs/2000/07/27/2000072710233117407.html.

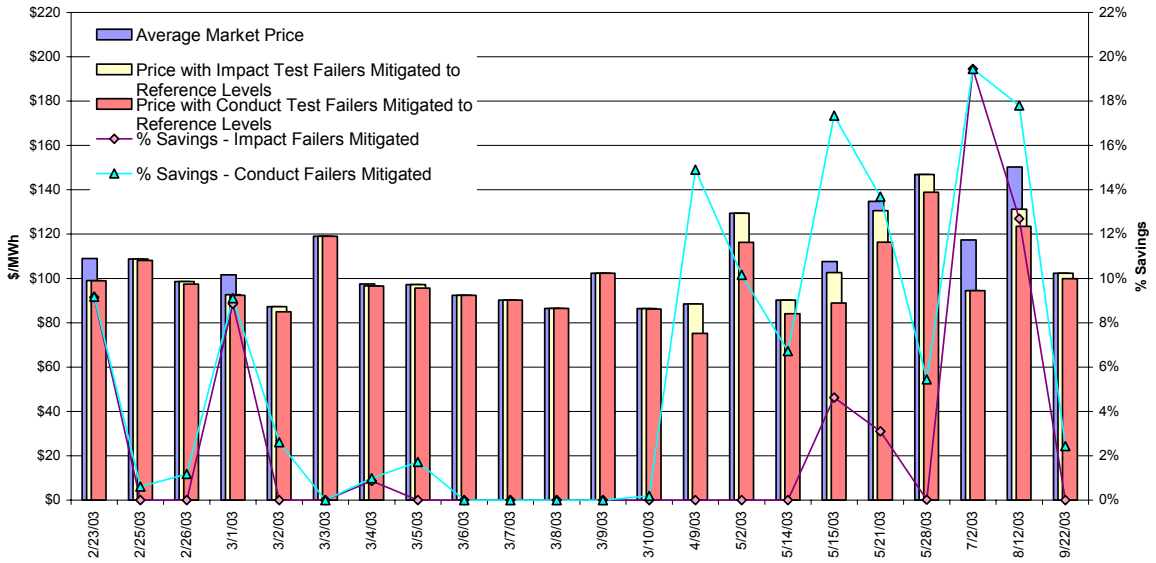
Figure 3. Bid Segments that Set Prices Causing Market Impact to Exceed $60,000, with Corresponding Reference Levels, Estimated Marginal Costs, and Conduct Test Status



Estimates of Market Savings due to Mitigation

To review the effects of the $91.87/MWh price screen and Impact Test thresholds on the overall market impact of AMP, the ISO's Department of Market Analysis has developed estimates of the real-time market prices under the scenarios in which units that failed the Conduct Test are mitigated to their reference levels and then re-dispatched. When these units' bids are mitigated, the set of Awarded Bids changes, since high-priced and out-of-merit order bids, which were in excess of their corresponding Conduct Test thresholds, may become in-merit when reference levels are substituted for the original bids. The following chart compares daily average actual prices (denoted in blue) on days in which the daily average price exceeded $85/MWh to average prices that would have occurred had the price screen not been required (denoted in yellow), and to prices that would have occurred had the price screen and Impact Test not been required for mitigation -- that is, if all bids from units that failed the Conduct Test had been mitigated (denoted in orange).

Figure 4. Daily Average Price compared to Estimated Prices When Bids are Mitigated to Reference Levels, through September 2003, when Daily Average Price exceeds $85/MWh



The largest difference between the actual price and the estimated price that would have occurred had all the units failing the Conduct Test been mitigated occurred on April 9, May 15, July 2, and August 12. The mid-day spike on April 9 had a total estimated market cost impact of approximately $134,000, and lasted approximately five hours; no actual dispatched volume in this hour either potentially or *bona fide* failed the Conduct Test (blue or purple). The May 15 spike was intermittent over three hours, and had a cost impact of approximately $160,000. The July 2 spike, whose price in several intervals was set by a participating load bid, had a market impact of approximately $170,000 (of which nearly $79,000 is due to markup in hour 20). The August 12 spike, discussed above, was set by a unit that bid $189/MWh, a price that would not potentially have caused that particular unit to fail the Conduct Test, for that one day only. However, other units did fail in this hour. Interestingly, the spike on May 28, during which the Price Setter bid was also a *Bona Fide* Conduct Test Failure, would have had an average price six percent lower with the Conduct Test-failing unit's bids mitigated to its reference levels. Under the existing structure, the price-setting bid still would have been the marginal bid setting the market clearing price regardless of the price screen. Thus, even though the price screen resulted in AMP not being implemented during many price spikes and the Impact Test resulted in no bid mitigation, had these barriers been removed and AMP triggered and applied, there still would have been very little difference in market costs due to mitigation.

Trends in Reference Levels

Overall, reference levels among *all* units – including the vast majority of units that do not actively participate in real-time markets -- increased in July 2003, and decreased again in the third quarter both on an absolute basis, and on a relative basis when adjusting for changes in natural gas prices, due primarily to an extraordinary prolonged price spike on July 2 and residuals of high prices in May in the 90-day rolling average composition of reference levels. The effect of this spike

largely dissipated by September, as spikes became few and far between, and prices for balancing energy decreased. The charts below show average peak-hour reference levels among all awarded and non-awarded bids from all units, system wide, grouped by generation type, and for a small portfolio of units that set the market-clearing price at least eight times in the 90 days ending September 30 (the "MCP Setters"); both on an absolute basis, and normalized to October 2002 gas prices.

Figure 5a. Peak-Hour Non-Normalized Reference Level Trends by Generation Type and MCP Setters Portfolio through September

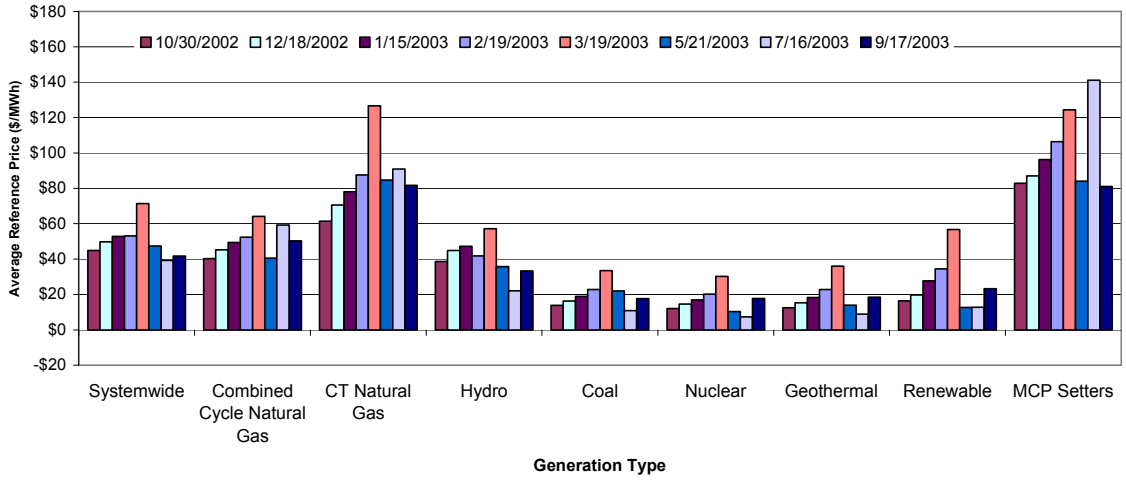
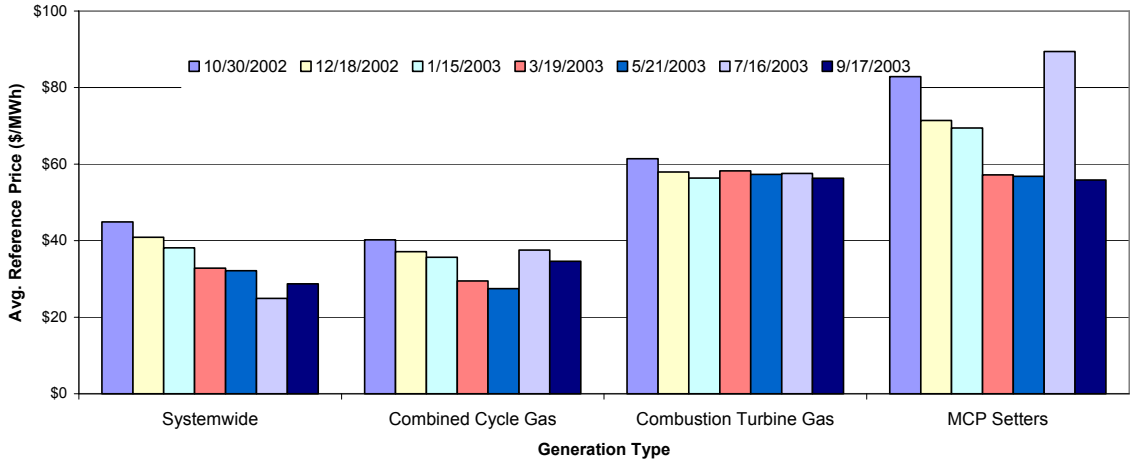


Figure 5b. Peak-Hour Reference Level Trends for Thermal Units by Generation Type and MCP Setters Portfolio through September, Normalized to Oct 2002 Gas Price



Reference levels decreased in September on an overall basis because a primary contributing factor in the calculation of reference levels is average market prices, which on average were lower in the summer than in the spring. Because these indices consider all units whether or not they are dispatched, they give a flavor of overall system wide trends but may be somewhat misleading.

A better indicator is to look at trends of reference levels of specific individual units, and particularly those units that set prices frequently. When adjusted for changes in the Natural Gas deflator index, reference levels of individual price-setting units have tended to trend upward. While some units appear to cycle through "booms" and "busts," by selling alternately during periods in which price are high and low, others appear to be able to maintain high prices by only selling when prices are relatively high (primarily by only bidding high prices). While gas-normalized reference levels corresponding to price-setting bids were in the range of $50 to $60/MWh in early 2003, they generally increased to the range of $65 to $75/MWh in the second quarter, and some frequent price setters were in the range of $70 to $90/MWh in the third quarter.[6] The following chart shows gas-normalized reference levels corresponding to awarded bids with prices above $100/MWh, specific to individual thermal units, between October 30, 2002, and September 30, 2003.

[6] The apparent thinness of price-setting reference levels in the third quarter is due to the fact that most real-time balancing in this period was in the decremental direction due to schedules in excess of actual load. On many days in the quarter, incremental energy was awarded in only a few hours.

Figure 6. Unit-Specific Reference Levels for Individual Awarded High-Priced Bids From Thermal Units in Peak Hours, Normalized to Gas Index[7]

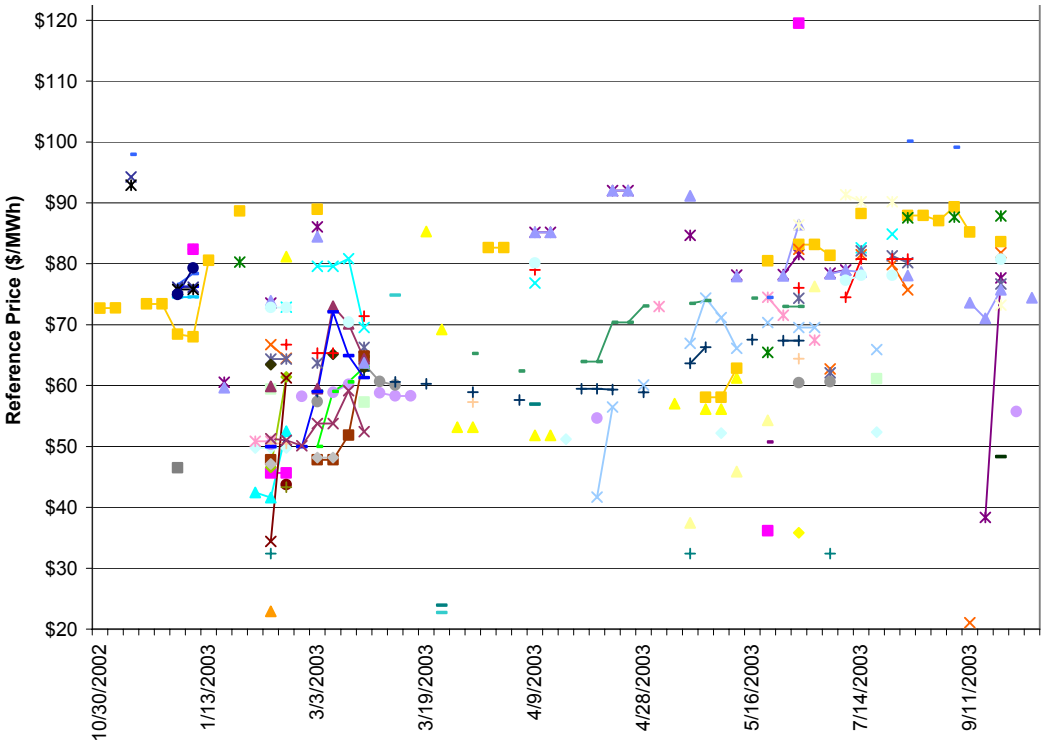


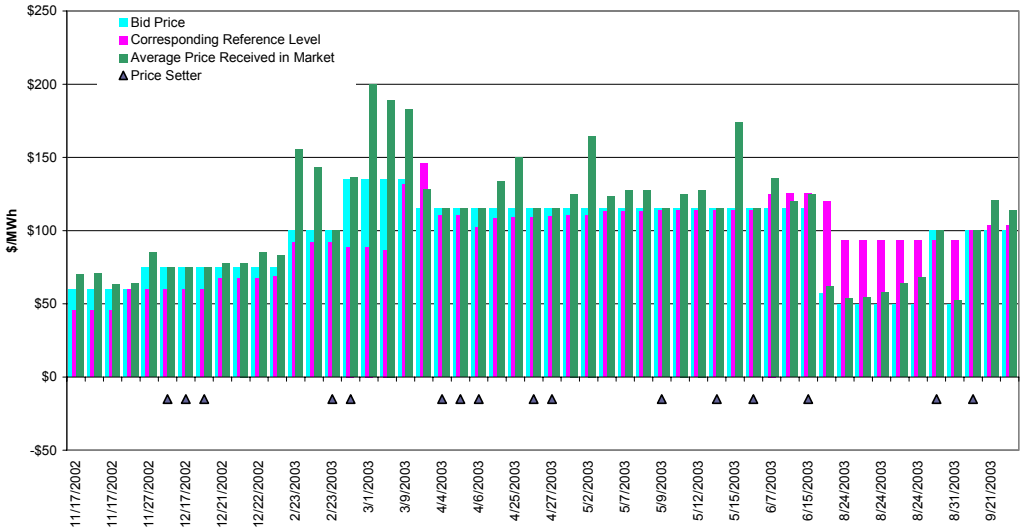
Figure 6 shows that reference levels have trended upward, and that certain unit's reference levels can be quite volatile, increasing nearly 100 percent in a short period of time. Certain units in particular have been able to maintain high reference levels, and to collect high prices for energy, without failing the Conduct Test. The following chart presents an example of a particular thermal resource of approximately average heat rate. This unit has set high prices more frequently in off-peak hours than any other unit, but has never potentially failed the Conduct Test. By systematically increasing bids over time since AMP have been in effect, this unit has effectively increased its reference level and now usually receives prices in excess of $100/MWh for its energy.[8] The chart below depicts the unit's bid price, corresponding reference level, and average price received in off-peak hours in which the unit was awarded energy. The chart also identifies whether the unit set the market-clearing price in at least one of the six dispatch intervals within the hour.

---

[7] Maximum average reference levels corresponding to awarded bids less the $6.00/MWh O&M adder, divided by monthly gas index used in deflating reference levels, multiplied by October 2002 gas index of $3.34/MMBtu, plus $6 O&M adder. Gas index ranges from $3.34 to $7.27/MMBtu (high in March 2003). Peak-hour reference levels are computed independently of off-peak-hour reference levels.

[8] The exceptions were July 31, August 24, and August 31, during which the unit offered energy below its reference level.

Figure 7. Bid Prices, Corresponding Reference Levels, and Average Prices In Off-Peak Hours for a Particular Unit



Conclusions

We can draw the following conclusions concerning system wide incremental AMP performance:

- Reference levels on an overall basis have declined. However, reference levels of certain units that often set prices have increased. In fact, some reference levels have increased to the point that a unit would need to bid above the $250/MWh price cap to fail the Conduct Test.
- The $91.87/MWh price screen has proven to be problematic. Due to the difficulties in forecasting real-time prices, AMP was not applied during many of the price spikes observed since implementation.
- The exercise of market power is neither necessary nor sufficient for a unit to fail the Conduct Test. In fact, the Conduct Test's ability to identify bidding behavior that is Consistent with the Exercise of Market Power appears to be somewhat disappointing.
- Even if the bids for units failing the Conduct Test were to be mitigated, it appears as though mitigation would have had a minimal effect on real-time prices and price-to-cost markup.

Although it is difficult to analyze what the effectiveness of AMP would be under tighter supply to demand conditions, analysis of the few hours when prices spiked and significant price to cost markup was present shows that AMP may have limited effectiveness in mitigating market power due its sporadic implementation as a result of the price screen, increasing reference levels, and wide Conduct and Impact Test thresholds.