



Memorandum

To: The ISO Board of Governors
From: Frank A. Wolak, Chairman, Market Surveillance Committee of ISO
cc: ISO Officers
Date: October 12, 2005
Re: ***MRTU Conceptual Filing***

The Market Surveillance Committee has been following and providing periodic comments on the comprehensive market redesign proposed under the Market Redesign and Technology Upgrade (MRTU) project since the ISO initiated this project. The MSC recently completed an opinion commenting on seven of the more than twenty aspects of the MRTU design that have been under discussion with stakeholders since the ISO's May 13 conceptual filing. The MSC has long-supported a different approach from the ISO on several of these seven issues, so it is not unexpected that there are remaining points of disagreement between the MSC and ISO staff. The MSC continues to support the transition to a locational marginal pricing (LMP) market for California. For the reasons outlined in the opinion, the MSC believes that its recommendations increase the likelihood that consumers, producers and the ISO operators will realize the maximum benefits from a LMP market in California. However, we do not believe any of the concerns expressed in the MSC opinion are reasons to delay the MRTU process. We urge the ISO Board to give the ISO management its approval to move forward with MRTU.

There are four major points of disagreement between the MSC recommendations and the current MRTU proposal. Some of these are fundamental while others can be addressed during Release 2 of MRTU. The MSC has consistently advocated against the inclusion of bid adders in the mitigated bids of units subject to local market power mitigation because these bid adders can significantly degrade overall market efficiency. The MSC has also been a consistent advocate for allocating as opposed to auctioning Congestion Revenue Rights (CRRs) and for simplified mechanisms for allocating CRRs on the grounds that they are more transparent and lower cost than the more complex multi-round nomination process proposed by the ISO. The MSC has also been a persistent supporter of a prospective local market power mitigation mechanism for both energy and ancillary services, but is concerned that it is too risky to make competitive path assessments without the benefit of actual market outcomes under a locational marginal pricing (LMP) market. An issue on which an ISO Board decision is not being requested at this time is the simplified Hour Ahead Scheduling Process (HASP). Ever since it was introduced by the ISO, the MSC has been concerned that there may be adverse unintended consequences associated with eliminating a formal hour-ahead price-setting and settlement mechanism. The experience in late 2004 with the Real Time Market Application (RTMA) with the bid-or-better pre-dispatch process for imports and exports has only increased our concerns with the HASP mechanism.

Despite the concerns with the MRTU conceptual filing expressed in our opinion, we still strongly support moving forward with a LMP market for California. We continue to endorse the following paragraph taken from an MSC opinion completed on April 7, 2003.¹

"LMP is a small, but important, part of a well-functioning wholesale market for electricity. Demand-responsiveness to both locational and temporal price differences is another important source of benefits from a wholesale electricity market. LMP is a necessary step towards achieving this long-term goal. In the short-run, the phased implementation of LMP (as proposed by CAISO) carries little potential costs and provides several short-term benefits. These benefits include: (1) the ability to secure effective local market power mitigation tools from FERC, (2) reduction in undesirable trading strategies (e.g., the "dec game"), (3) greater transparency, efficiency, and reliability in system operation, (4) improved demand responsiveness (given the ability of dispatchable loads to bid and respond as generation and receive the LMP), and (5) greater granularity in the costs of transmission congestion to aid the transmission planning process."

We continue to believe that an effective local market power mitigation mechanism is essential to the success of an LMP market in California. This is why we support a cautious approach to including transmission paths in the set of competitive paths and would like the ISO to have one year of experience with only the existing three zonal interfaces and interties into California as the only competitive transmission paths. Following this year, the ISO would be free to apply its competitive path methodology to all transmission paths to determine which ones will be deemed competitive during subsequent years.

Although we are confident that a LMP market can deliver benefits to California consumers and ISO operators, the market design must be tailored to capture these benefits. If properly designed, a LMP market can improve wholesale market efficiency and system reliability in California relative to the existing zonal market design. The motivation for our comments on the seven aspects of the MRTU conceptual filing is to increase the likelihood that a LMP market would benefit system reliability and market efficiency in California. We look forward to working with the ISO to address the concerns expressed in this opinion as it moves forward with a LMP market for California. All of these concerns can be addressed within the context of the existing ISO MRTU proposal. Therefore, none of the changes we are recommending should be interpreted as a reason for the ISO Board not to move forward with MRTU as soon as possible.

¹ Comments on Locational Marginal Pricing and the California ISO's MD02 Proposals, April 7, 2003, (available from <http://www.caiso.com/docs/2003/04/07/2003040713192323878.pdf>).

Opinion on Aspects of the California ISO's Market Redesign and Technology Upgrade (MRTU) Conceptual Filing

by

**Frank A. Wolak, Chairman; Brad Barber, Member;
James Bushnell, Member; Benjamin F. Hobbs, Member
Market Surveillance Committee of the California ISO**

September 30, 2005

Introduction

We have been asked by the ISO to comment on a number of aspects of the California ISO's Market Redesign and Technology Upgrade (MRTU). The specific issues addressed in this opinion are: (1) the use of bid adders for frequently mitigated units, (2) competitive path assessment to implement the ISO's local market power mitigation (LMPM) mechanism, (3) the formulation of the Hour-Ahead Scheduling Process (HASP), (4) the formation of trading hubs, (5) the rules for allocating Congestion Revenue Rights (CRRs), (6) rules for allocating CRRs to loads located outside of the ISO control area, and (7) rules for allocating CRRs to merchant transmission owners. Because the ISO has not formulated concrete proposals on a number of these issues, at this point we only provide general comments on the issues and a discussion of the proposals receiving the most attention in the stakeholder process.

In preparing this opinion, MSC members have discussed each of these issues with ISO staff on a number of occasions. MSC members have attended several of the MRTU stakeholder meetings held over the past two months on these topics. The MSC has also discussed these issues with ISO staff and management at both the July 7 and September 22 MSC meetings. We received written comments from Calpine, Energy User's Forum, Sacramento Municipal Utility District (SMUD), and Strategic Energy on various aspects of the MRTU. At the September 22 MSC meeting, public comments on MRTU were provided by Ellen Wolfe of the Western Power Trading Forum (WPTF), Mike Warner of the California State Water Project, Sue Mara on behalf of Strategic Energy, and Jeffrey Nelson of Southern California Edison. We are grateful to all of these individuals and entities for the input they provided.

1. Bid Adders for Frequently Mitigated Units

We believe that a shortcoming of the ISO's current LMPM proposal is the use of an ad hoc bid adder on top of the ISO's best estimate of the variable cost of providing energy from that generation unit to compute its mitigated bid level. The ISO's proposal to set a substantially higher ad hoc bid adder for frequently mitigated units—defined as units that are mitigated for more than 80% of their run hours—has the potential to introduce substantial costs on both California consumers and generation unit owners facing significant competition for their output, with no counter-balancing market-efficiency benefit. While we understand the ISO's desire to be consistent with market rules that the Federal Regulatory Commission (FERC) has approved in other markets, we believe that the market inefficiencies introduced by the use of bid adders for frequently mitigated units are substantial enough to justify deviating from FERC precedent. Although we strongly support providing revenue adequacy for generation units needed to meet local reliability needs that may not earn sufficient revenues from the energy and ancillary

services market, providing this revenue adequacy need not come at the cost of substantial inefficiencies in system operation.

The goal of bid mitigation is to replace the supplier's bid with the bid that the market participant would submit if it faced effective competition. If the supplier did face effective competition it would bid its minimum variable cost of supplying energy. By this logic, the economically efficient mitigated bid level is the ISO's best estimate of the generation unit's minimum variable cost of supplying energy. Using a bid adder the ISO knows is larger than this minimum variable cost contradicts the primary goal of locational marginal pricing to obtain the most efficient dispatch possible. A scheme that systematically biases the bids of mitigated generation units upward relative to the ISO's best estimate of the unit's minimum variable cost of supplying electricity does not achieve this goal.

Generation units that face sufficient competition will bid close to their minimum variable cost. Mixing bids from such units with mitigated bids from other units set significantly above their variable cost is likely to result in those units facing significant competition being overused relative to what they should operate if the mitigated suppliers faced sufficient competition and bid their minimum variable cost of supplying energy. Suppliers bidding in a manner consistent with price-taking, competitive behavior have their units overused because of the ISO's decision to use an ad hoc bid adder for frequently mitigated units. The unnecessarily high mitigated bid level implies that mitigated units are dispatched for less energy than they would be if their mitigated bid was set equal to the ISO's best estimate of the unit's minimum variable cost. Generation units facing sufficient competition and bidding their minimum variable cost of supplying energy must therefore produce more output to make up the difference.

Including ad hoc bid adders in the computation of mitigated bid levels increases the incentives for unmitigated suppliers to distort their bids above their minimum variable cost. Nearby generation unit owners recognize that the mitigated bid must be dispatched so they face little risk of a reduced amount of energy sold but a substantial likelihood of achieving a higher price for their energy by bidding higher than their minimum variable cost of supplying energy. This bidding behavior enabled by the use of ad hoc bid adders results in additional market inefficiencies.

However, as noted in our opinion on the MRTU conceptual filing, these inefficiencies can be largely eliminated by requiring the ISO to use its best estimate of the unit's minimum variable cost of supplying energy as the mitigated bid level, regardless of how frequently a unit is mitigated.² We suggested a procedure (which we reproduce below) that the ISO could use to estimate of the unit's minimum variable cost. This proposal also provides strong incentives for suppliers to reduce their input fuel procurement costs and other operating costs. If a generation unit owner is unable to recover the unit's annual costs from short-term market sales, the unit owner should either make a cost-of-service filing to recover these costs or sign a long-term supply agreement with a load-serving entity (LSE) to provide the necessary energy in return for recovering its full costs on an annual basis. This LSE should be willing to sign such a contract if the ISO identifies this unit as needed for the LSE to meet its annual load obligations as part of the resource adequacy process.

² "Opinion on California ISO's Market Redesign and Technology Upgrade (MRTU) Conceptual Filing," April 26, 2005 (available from <http://www.caiso.com/docs/2005/04/26/2005042611125729395.pdf>).

The use of ad hoc bid adders for mitigated units introduces unnecessary potential cost to consumers to purchase this forward contract because the revenues a supplier expects to obtain from the spot market is the opportunity cost of signing a fixed price forward contract to supply energy. Suppose a supplier that owns a 100 MW unit with a variable cost of \$50/MWh expects to be mitigated more than 80% of its 2000 run hours and therefore will receive a \$40/MWh bid adder during these hours. This \$40/MWh bid adder translates into an additional annual payment of \$6.4 million, assuming that this bid always set the price the unit receives when it is mitigated. Suppose that the unit's remaining annual fixed-cost requirements after selling according to unmitigated bids during the remaining 20% of its run hours are \$4 million. Despite only needing an additional \$4 million to cover its annual fixed costs, unless the supplier receives a guarantee of \$6.4 million above its variable cost, it will not sign a fixed price forward contract with an LSE to supply this locational energy. Consequently, the presence of this \$40/MWh adder implies that consumers must pay \$2.4 million more on annual basis for the energy this supplier provides, despite the fact that there is no shortage of generation capacity in this location. This additional \$2.4 million payment is due to the exercise of local market power sanctioned by the ISO's requirement of a \$40/MWh bid adder on the unit's variable cost when it is mitigated.

Allowing these units to distort locational marginal prices (LMPs) throughout the control area and produce LMPs high enough at one location in the network to allow the unit at that location to recover its annual fixed cost requirements can also distort the LMPs at many other locations in the network. Particularly at the locations in the network with frequently mitigated generation units, the decision to include substantial adders in mitigated bid prices will set prices at these locations that provide signals for new generation to enter, even if there is no need for generation units at these locations. It is important to emphasize that there is sufficient generation capacity available to meet this local energy need. There is just not enough competition among those suppliers able to meet this local energy need to rely on a market mechanism to set the price paid for this local energy. Paying higher prices at this location and distorting the operating decisions of both mitigated and unmitigated generation units unnecessarily increases the prices that consumers must pay for energy throughout the transmission network because the higher LMPs at certain locations will be included in the Load Aggregation Point (LAP) prices paid by all consumers located in that LAP, with no increase in system reliability or market efficiency in either the short-term or long-term. We strongly urge the ISO to avoid setting LMPs above competitive levels by including ad hoc bid adders in an attempt to provide adequate revenues to owners of mitigated generation units.

MSC Bid Mitigation Proposal

Because the ISO has deemed these units necessary to operate the network reliably, these units should have the opportunity to recover their production costs including a reasonable return on capital invested. However, this cost recovery mechanism should not distort market outcomes and would ideally preserve incentives for efficient production. Except in rare circumstances (*i.e.* scarcity), there is no justification for prices in a perfectly competitive wholesale electricity market to rise above the incremental cost of the highest cost unit operating at that location. If the goal of the MRTU process is to achieve competitive market outcomes during as many hours of the year as possible, then it makes little sense for the ISO to apply ad hoc bid adders to a number of generation units during many hours when prices would otherwise achieve these competitive levels.

To limit the market inefficiencies associated with bid mitigation, we recommend that the ISO set mitigated bids for each unit equal to its best estimate of the variable cost of supplying electricity across all units of its type. This mechanism should also provide strong incentives for suppliers to minimize the cost of providing energy during the hours their bids are mitigated. We recommend a mechanism that would utilize a benchmark heat-rate, representative for units of similar technology. This heat rate would then be multiplied by the daily price of natural gas delivered to Henry Hub in Louisiana plus the regulated cost of transporting natural gas to the generation unit, including the relevant intrastate gas transmission and distribution charges. The heat rate times this benchmark delivered price of natural gas plus a typical variable operating and maintenance charge for generation units of this technology and vintage would constitute the mitigated bid level for this generation unit. This procedure provides a defensible estimate of the unit's minimum variable cost of supplying electricity. If the unit owner believes that it can procure its natural gas at lower cost, it would be able to keep the difference between this natural gas price benchmark and its actual natural gas costs during all of the hours it is dispatched in addition to the difference between the market-clearing price of energy at this location and its mitigated bid level. The same logic applies to the benchmark heat rate and to the variable operating and maintenance cost estimates. This scheme for setting mitigated bids would provide strong incentives for least cost procurement of natural gas and operation by mitigated generation units while limiting the distortions introduced into the spot market mechanisms as a result of bid mitigation.

General Comments on the Use of Bid Adders

The FERC has articulated the belief that it is appropriate that some portion of the fixed costs of mitigated units be allowed to set market prices. In other words, such units should not just be allowed to recover their fixed costs for themselves, but those costs should be reflected in the prices earned by other non-mitigated units. The FERC is essentially arguing that prices should be set at long-run average cost, as they would in the long run in a competitive market. There are two problems with this view. The first is that the FERC would set prices to recover at least these average costs during *all hours* the unit operates. In a competitive market the high prices during certain periods would offset prices at incremental costs during the majority of hours with abundant supply. The average of all these resulting prices would trend toward long-run average cost. The adder approach gets prices wrong all the time, producing the problems described above.

If the ISO identifies a unit as necessary to meet its local reliability requirements, the bilateral negotiation process between this unit owner and the local LSE that wants to ensure a reliable supply of electricity to its customers seems far more likely to balance the competing goals of cost recovery for the unit owner and protection against local market power for final consumers. As discussed earlier, including an ad hoc bid adder on top of the ISO's best estimate of the unit's minimum variable cost of supplying electricity unnecessarily increases the cost and complexity of this bilateral negotiation process. We strongly urge the ISO to eliminate all ad hoc adders from the mitigated bids that it sets.

2. Methodology for Competitive Path Assessment

Objectives and Effects of Competitive Path Assessment

The goal of Local Market Power Mitigation (LMPM) mechanisms, and of wholesale electricity markets in general, is to produce locational prices that accurately reflect the incremental cost of withdrawing power at each location in the network. An efficient price should reflect the incremental cost (or benefit) to the system of additional consumption (or supply) at that location in the transmission network. Unless there is a shortage, a price that is above the short term incremental cost is inefficient because it can deter consumption whose value is greater than the cost of production, but below the price. Further, when an individual generation unit sets its price above its incremental cost, other more expensive units may be chosen to supply energy in its place. In the absence of shortages, prices that deviate from incremental costs cause inefficient consumption and inefficient production. In perfectly competitive markets, firms will choose to produce as long as the price is above their incremental costs. The only time the economically efficient price should be above the incremental cost of withdrawing energy at that location is when supply at that location is capacity constrained (*i.e.*, there is a scarcity of supply).³

The general idea of local market power mitigation is to induce an offer price from a generation unit owner with local market power equal to the bid the unit owner would submit if its unit faced significant competition. As we noted in the above discussion of bid adders, a unit that faces sufficient competition would offer a price equal to its minimum variable cost of supplying additional energy. When the LMPM mechanism is triggered, the offer price of such a unit is set to a regulated level. By the above logic, this regulated level should be equal to the ISO's best estimate of the unit's variable cost of supplying energy, assuming that a scarcity pricing mechanism is in place that would raise prices above bid levels in case of a capacity shortfall.⁴

For local market power mitigation (LMPM) to be effective, a methodology must be devised to identify generation units that can significantly raise the price they are paid through their own unilateral actions. These suppliers possess substantial local market power because they know that there are no economically viable substitutes for the production of their generation units, either because this supplier owns a significant fraction of the generation capacity needed to meet local demand or because transmission constraints limit the number suppliers that are able to compete against this supplier.⁵ The lack of competitive substitutes means that prices must rise if a supplier chooses to withhold even small amounts of capacity from the market.

³ In this case, the efficient locational price is above the variable costs of all generation units. Ideally, it is set by the willingness of demand at that location to curtail its consumption. In practice, it is usually set equal to bid cap on the energy market.

⁴ It is important to recognize that a form of scarcity pricing will be in effect under the initial release of MRTU. Although some aspects of this approach to scarcity pricing may not be ideal, it does in general allow for prices to rise above bid levels in periods of overall shortage of reserve and energy needs.

⁵ This sort of market power is associated with areas that rely on distant generation units to meet a significant fraction of their energy needs, and is the primary focus of the LMPM mechanism. However local market power can also be exercised in regions where there is significantly more local generation than there is load to serve. In these regions,

The ISO's approach to market power mitigation is to delineate clearly the regions and conditions when generation units are subject to mitigation. If these conditions do not apply, the ISO should not intervene to reset the bids unit owners submit. This approach identifies regions that are sufficiently competitive and operates under the assumption that if there is no congestion within these regions, no mitigation is necessary.⁶ If transmission congestion does occur within these regions on transmission paths that are not sufficiently competitive, then mitigation will apply for the purposes of altering the dispatch to deal with congestion on this transmission path. The process is implemented by using the Integrated Forward Market (IFM) to identify generation units that possess local market power. The ISO process utilizes a sequence of pre-IFM runs in which, first, only "competitive" transmission constraints are imposed on the generation scheduling process, and then a second run in which all constraints are imposed. Units whose output is increased in the second run relative to the first run are then subject to mitigation.

This two-step process for determining whether to mitigate the bid of a generation unit accounts for the fact that there are substantial uncertainties in the ISO's local reliability needs. In particular, system conditions can arise when virtually any generation unit can possess substantial market power to resolve these local reliability constraints despite the best efforts of the LSE to purchase the appropriate mix of local reliability energy in fixed-price forward contracts far in advance of real-time system operation. Consequently, the second stage of the process imposes mitigation for the remaining, "uncompetitive" set of transmission constraints that can cause certain generation units to face insufficient competition for their output, implying that they possess local market power requiring mitigation. This mechanism for determining whether a generation unit possesses substantial local market power requires a procedure for dividing network constraints into competitive and non-competitive paths.

We have been, and continue to be, supportive of this approach to local market power mitigation. The identification of uncompetitive paths is obviously a critical aspect of this approach. There are two elements of this process of identification of uncompetitive paths that merit discussion: the general concept for determining a competitive path and its technical implementation, and the specific criterion used to declare a path uncompetitive. Before discussing these elements, it is worth reviewing the consequences of these decisions.

There are consequences to either a screen that is too conservative or too lax in its detection of local market power. The consequences of a screen that is too lax are the most obvious--firms will be able to raise prices substantially above competitive levels. The

generation unit owners may be able to raise local price by restraining output and decongesting constraints that limit exports from the region. The ISO's LMPM mechanism is not presently designed to mitigate this exercise of local market power. As the ISO points out in Appendix A of its Competitive Path Assessment proposal, the LMPM mechanism also cannot mitigate other, more sophisticated types of market power exercise involving deliberate creation of congestion by one generation plant in order to raise prices elsewhere for others.

⁶ An important assumption implicit in this approach is that LSEs have forward contracted for their expected energy needs in such a way that significantly reduces the incentives of suppliers to exercise market power over these larger geographic areas. Forward energy contracts are a critical element in determining the overall competitiveness of these broad regions. We emphasize that without adequate fixed-price forward contracting for energy between suppliers and LSEs, market power is likely to be a concern even at a system-wide level during high demand periods of the year.

consequences of too lax of a screen will become more serious as the overall bid cap in the market rises from \$250/MWh along its current expected trajectory to \$1000/MWh.⁷ In an ideal world, the consequences of a screen that is too conservative would be minor. The mitigated bids are intended to mimic those of competitive generators, and if that intent is successful, prices resembling competitive market outcomes would result. However, as we noted in the first section of this opinion, several aspects of all existing LMPM mechanisms, including the CAISO's proposed mechanism, bias the offer price upwards to guarantee that mitigated offer prices will be noticeably higher than those from units facing substantial competition. In particular, we are concerned with the impact of an ad hoc bid adder for mitigated units. As we expressed earlier, this adder has the potential to significantly distort market prices. A competitive path screen that is too conservative increases the chance that units will be frequently mitigated and therefore increases the potential distortions of the bid adder.

Last, it is important to remember that one of the goals of electricity restructuring is to remove the distortions to the incentive to produce in a least-cost manner that regulation can create. Local market power mitigation is a form of cost-based regulation. Instead of basing revenues upon the average costs of operation, LMPM sets bids equal to regulatory estimates of the unit's incremental cost, or some other proxy meant to represent these costs. It is highly probable that firms with units constantly subject to such mitigation will not find it worthwhile to work to reduce those incremental costs.⁸ If there is no prospect of real competition for the output of a plant, then the inefficiencies caused by regulation are almost certainly outweighed by the importance of mitigating the plant's market power, but we should be cognizant of the fact that such mitigation comes at a cost when considering the potential scope of mitigation.

Competitive Path Analysis

The MRTU November 2005 filing will include a proposed methodology for identifying and designating competitive constraints, or "paths." We believe that the CAISO proposed methodology, termed the Feasibility Index (FI) method, improves on the methods used or proposed for other ISOs. The FI method is consistent with the fundamental principle that local market power arises from constraints that cause a steep local residual demand curve—the difference between the local demand for energy and the willingness to supply (as function of the local price) this energy of all entities that can feasibly provide it. The FI method designates as uncompetitive any constraint (or, more generally, any set of constraints) whose imposition would result in an infeasible dispatch, should an entity controlling generation withdraw its output. This is equivalent to that constraint causing the supplier's residual demand curve to become vertical at some positive quantity of output.

Procedures for determining competitive transmission paths used by other ISOs unnecessarily complicate the designation process by defining "supply" and "demand" for congestion relief for particular transmission facilities using complex swing-factor based

⁷ The overall severity of local market power will also be influenced by the extent of RMR contracts, or their equivalents, and by the form of the resource adequacy obligation that is adopted by the California Public Utilities Commission (CPUC).

⁸ Our proposal in Section 1 to base mitigated bid levels on benchmark variable cost estimates by generation unit type is intended to provide incentives to reduce these incremental costs.

methods. This is inconsistent with how a LMP-based market works: generation unit owners bid to sell at their location and are paid the LMP. The residual demand curve describes the unit's LMP as a function of its output. LMPs and residual demand functions result from the interplay of demand, network constraints, and bids from other generation unit owners. Generation unit owners do not buy and sell the capacity of individual transmission lines. Splitting out a separate market for each transmission constraint using swing-factor methods is an unnecessary distortion of how the market actually functions. For instance, other ISOs use swing factors to determine how much congestion relief on a given line a generator can supply by increasing its output. However, these swing factors may not capture the actual relief a supplier can provide. This relief depends on where the matching decrease in energy output occurs to make room for this incremental supply, which in turn depends on system conditions and the bids of other generation unit owners.

Any method for mitigating the bids of generation units creates market inefficiencies, and the FI method is no exception. First of all, its focus on feasibility means that transmission constraints that significantly raise the costs of importing competitive power, but do not make it impossible, will not impact the screen. Furthermore, the method focuses on unilateral market power. Second, it does not explicitly account for how a collection of suppliers might exercise market power in a coordinated manner, although the use of a three-pivotal suppliers or two-pivotal suppliers test recognizes that coordinated actions may be more likely if there are a small number of jointly pivotal suppliers. Because coordinated actions among market participants to raise or lower market prices is illegal under US antitrust law and punishable by triple damages, detecting collusive behavior has not been the focus of LMPM, or other ISO market-power mitigation mechanisms.⁹ The goal of LMPM mechanisms is, rather, to detect and mitigate cases where the unilateral exercise of market power would be nearly as damaging as coordinated behavior among market participants.

Criterion for Determining Non-Competitive Transmission Paths

The CAISO proposal is that a path should be declared non-competitive if three suppliers are simultaneously pivotal to resolve a transmission constraint on this path. This procedure uses a prospective analysis of market outcomes for a range of system conditions to identify sets of transmission constraints where three suppliers acting jointly could produce an infeasibility, meaning that some of their generation capacity is needed to resolve the transmission constraint. These transmission paths are then designated non-competitive. Unfortunately, there is very little information at this time about the implications of the three-pivotal-supplier test, in terms of the number of non-competitive paths and units potentially subject to mitigation. We do not know whether this criterion would result in the appropriate number of non-competitive transmission paths. We would prefer to have more information on market participant behavior under a LMP market in California before making a definitive recommendation on this question.

⁹ The US Department of Justice and relevant state Attorneys General are best suited to detect and punish these modes of exercising market power. Further, sufficient fixed-price forward contracts for energy between suppliers and LSEs limit the incentives for these actions to occur.

We acknowledge that the three-pivotal-supplier approach is unlikely to be too lenient (*i.e.* it is unlikely to falsely designate transmission paths as competitive if they truly are not). However, it may prove too conservative and designate potentially competitive transmission paths as non-competitive. If this outcome does occur, then LMPs should not be significantly impacted if mitigated bids are set equal to ISO's best estimate of the unit's variable cost of supplying electricity as we recommend in the previous section. Suppliers facing significant competition find it unilaterally profit-maximizing to bid very close to their minimum variable cost of supplying energy. Therefore, we do not think that using this test during the initial year of operation of the LMP market in California has the potential to greatly harm market efficiency or system reliability if the ISO does not implement bid adders for mitigated units.

One additional complication with applying the three-pivotal-supplier test is that it is likely that during peak periods, the entire system will be in deficit if the available supply of the three largest suppliers is omitted. However, it is extremely unlikely that the entire ISO control area will operate in an unconstrained manner during peak system conditions. Historically, a number of local transmission constraints are binding during peak system conditions, so that LMPs should differ across locations in the ISO control area. This implies that local three-pivotal-supplier constraints will require higher levels of output from certain units at a number of locations in the ISO control area than is required from these same units by the system-wide three-pivotal-supplier test. In addition, given the quantity of available generation capacity supplied with mitigated bids by imposing the three-pivotal-supplier constraints locally, it is unlikely that the three-pivotal-supplier test will be violated on a system-wide basis. Further study is required to determine the number of hours of the year this logic holds. To the extent that the three-pivotal-supplier system-wide constraint is binding a substantial number of hours of the year, the ISO may wish to consider a two-pivotal-supplier or single-pivotal supplier test for its LMPM mechanism. We recommend that the proposed ISO study of its LMPM mechanism address these issues.

Although the LMP studies completed by the ISO have been informative about the geographic distribution of wholesale electricity prices in California under MRTU, these studies have not addressed a major source of uncertainty associated with an LMP market—market participant bidding behavior. Market participants are likely to change their bidding behavior under an LMP market. For this reason, we support designating the existing zonal transmission paths as the only competitive transmission paths during the initial year of operation of the market. This seems prudent given the substantial uncertainties about bidding behavior under an LMP market in California.

Rather than attempt to designate any new competitive transmission paths before the end of the first year of operation of the LMP market, we recommend that the ISO analyze market outcomes during the first year of the LMP market to determine the appropriateness of the three-pivotal-supplier test, two-pivotal-supplier, or a single-pivotal-supplier test with a price movement test (described below) for determining competitive transmission paths during subsequent years of the LMP market. During the early phase of the operation of the MRTU market, the risks of under-mitigation outweigh the consequences of over-mitigation if ad hoc adders are not applied to mitigated bids. Over the long-term, the regulatory distortions created by over-mitigation become a greater concern, and we suspect that a less stringent pivotal supplier

criterion, combined with a price movement test would produce a more appropriate level of mitigation.

Price Movement Test

The CAISO has also suggested a “Price Movement” test, which examines whether unilateral withdrawal of generation from the market will result in a significant price increase.¹⁰ The CAISO recommends further study of this approach, and if a transparent and practical test can be developed, that it be implemented after Year 1 of MRTU accompanied by a relaxation of the three-pivotal-supplier test criterion to a two-pivotal-supplier test. Because local market power mitigation, like all forms of regulatory intervention, is an imperfect mechanism for limiting unilateral market power, it is important for the ISO to consider alternative approaches to determining competitive transmission paths in its analysis of the results of the first year operation of the LMP market.

Major Stakeholder Issues

We conclude our discussion of the competitive path assessment approach by considering some specific issues that have been raised by stakeholders and the CAISO about the FI method, beyond the issues described above.

First, the ISO has offered three options regarding grandfathering (or pre-designation) of competitive paths. As noted above, we believe it would be imprudent to move beyond designating any transmission paths but the existing zonal interfaces as competitive during the first year of the LMP market. Specifically, the ISO should not designate paths that are minimally congested in the ISO LMP studies as competitive during the first year of market operation because these studies are not based on bids submitted to maximize the profits of generation unit owners under an LMP market. If the first year of operation of the LMP market finds that the magnitude of congestion is invariant to large but credible changes in the bidding behavior of all market participants able to relieve this constraint, treating this constraint as competitive in subsequent years should not have adverse consequences. Once the ISO settles on a mechanism for designating competitive transmission paths, it should review the competitiveness of all transmission paths on an annual basis.

A second important issue with competitive path assessment is the treatment of fixed-price forward contracts. Forward contracts, which can be interpreted broadly to include the load obligations of Utility Distribution Companies (UDCs) and RMR contracts, limit the incentive of suppliers to withhold output to raise market prices. To the extent that contracts are long-lived, verifiable, and bestow operational control of a generation unit on another entity, they should be considered. The CAISO proposal to consider the long term contracts reported to FERC is a reasonable start. We also think that consideration should be given to augment this set with UDC obligations, RMR contracts, and other publicly disclosed long-term contracts for output from particular supply sources.

¹⁰ The “price” could either be the shadow price of a transmission constraint, as proposed by the CAISO, or nodal prices. Of course, as is well known, the two are closely linked; one cannot increase without causing increases in the others, the exact linkage depending on swing factors.

A third issue concerns whether designations of competitive paths could be differentiated by season or time of day. If indeed a particular constraint is unlikely to be binding during a given period, or does not fail a three-pivotal-supplier test during that time, we see little reason to designate it as non-competitive. Assuming that the number of cases, types of load conditions, etc. to be considered in the path assessment can be agreed upon the CAISO and stakeholders, we would find such differentiation acceptable. However, system congestion should be monitored in case patterns of congestion change.

3. Hour-Ahead Scheduling Process

The ISO's current MRTU proposal will set market-clearing prices only in the day-ahead integrated forward market and the real-time market. There is no explicit price-setting process in the hour-ahead time frame, but market participants will have the opportunity to adjust their day-ahead schedules through the Hour-Ahead Scheduling Process (HASP) for real-time system operation. In this hourly scheduling process, the ISO also proposes to pre-dispatch bids from the interties to supply energy in real-time and set the price that importers will receive and exporters will pay during the subsequent hour.

We are worried that there are a number of unintended consequences associated with HASP mechanism because it is attempting to allow schedule changes, but not run a formal hour-ahead market. As noted by Calpine in its written comments submitted to the MSC, the proposed HASP mechanism requires internal resources to compete against importers to supply energy in the HASP, but does allow these internal resources to receive the hourly price or implement the changes in their day-ahead schedules implied by the pre-dispatch process.

Our preference is for the ISO to operate an hour-ahead market using the same network model and other operating constraints used in the day-ahead and real-time markets. Specifically, this hour-ahead market should re-run the integrated forward market in the hour-ahead time frame using its best estimate of real-time system conditions at this time and compute hourly prices for energy and ancillary services. This would ensure internal consistency of market and system operation across the day-ahead, hour-ahead, and real-time markets. At each time horizon, the ISO would operate the same pricing algorithm and network model, although the inputs to this pricing and dispatch model would be updated to reflect changes in bids, schedules, characteristics of transmission network, and other network reliability requirements as this information becomes available to the ISO.

Both importers and exports and internal resources should be allowed to compete in all three markets. Only those importers that are willing to schedule dynamically or are willing to manage real-time price risk within the hour should be able to sell in the real-time market. In the hourly market, both internal and external resources would compete to supply hourly blocks of energy and ancillary services at a fixed hourly price. There would be no separate pre-dispatch process for imports and exports. There would be no need for a pay as-bid mechanism for imports and exports or a bid-or-better mechanism. The ISO would run three identical markets and settlements, all of which honor the same set of reliability constraints and pay the same price to internal and external resources and subject them to same operating requirements.

Rather than allow the ISO operators to specify a demand into the hour-ahead market, the operators should instead specify minimum scheduling requirements that are necessary to operate the system in real-time. For example, if the ISO operators feel that they need Y percent of their forecast of locational energy and ancillary services needs scheduled in the hour-ahead time frame, then this constraint should be specified in the full network model and priced in the hour-ahead market. For the same reason, if the ISO operators believe that they need X percent of their day-ahead forecast of energy and reserves at some location in transmission network, this constraint should be imposed in the day-ahead market and reflected in market prices if it is binding. Rather than pretend that an operating constraint is not relevant in the pricing process and then purchase the required energy or ancillary services through an out-of-market or pre-dispatch process, the ISO should honor all relevant operating constraints in the pricing process in each market. One such operating constraint is that certain percentages of the ISO's locational load forecast and ancillary services needs are purchased in the day-ahead and hour-ahead time frames.

The ISO operators should also purchase the amount of ancillary services needed to operate the system at the level of granularity needed to maintain mandated system reliability standards. These locational ancillary services requirements should be built into all three markets as local reliability constraints and explicitly priced in the LMP mechanism. Rather than purchasing additional ancillary services outside of the formal market mechanisms or through a must-offer waiver denial process that requires unloaded generation capacity to be paid according to a regulated price, the ISO should ensure that the energy and ancillary services it needs to operate the system in real-time are purchased and scheduled in the day-ahead and hour-ahead markets. Rather than adhere to operating reserve requirements that were valid for the vertically-integrated regime and rely on out-of-market mechanisms such as the must-offer waiver-denial process to obtain the necessary operating reserves, the ISO operators should purchase the amount of reserve capacity needed to operate the system in a reliable manner.

A major reason we favor a formal day-ahead market is because California is import dependent and likely to become even more so. Consequently, given the increasing availability of imports, the ISO should facilitate participation by importers in a manner that faces them with the maximum amount of competition for their energy and ancillary services and maximizes their ability to provide effective competition for internal resources. Day-ahead, hour-ahead, and real-time markets where internal and external resources compete to supply energy and ancillary services on a level playing field is the best way to achieve this goal.

All costs of the energy purchased in the day-ahead or hour-ahead time frames that is not explicitly purchased by a load-serving entity (LSE) but is purchased to meet an ISO reliability requirement should be charged to load that consumes more than their final schedules in the real-time market. Capping the extent of this liability only provides incentives for load to continue to purchase in the ISO's real-time market. In this sense, the prospect of a very large cost to a small purchase from the real-time market can provide very strong incentives for loads and suppliers to schedule accurately in the real-time market. This approach generally accords with cost-responsibility principles in that the extra costs were incurred because load was under-scheduled.

We also encourage the ISO to specify its ancillary services requirements in a manner that reflects its actual needs to operate the market in a reliable manner, rather than adhere to a hard-and-fast rule that was designed for the former vertically-integrated utility regime. For example, if the ISO operators feel they need more than 7 percent of demand held in unloaded generation capacity or quick start generation capacity, the ISO operators should have the freedom to purchase the necessary ancillary services, rather than rely on a must-offer requirement which pays certain suppliers a regulated price for remaining on-line to provide reserves while other suppliers are receiving a market price to provide.

An additional benefit of three identical markets is that market participants could buy and sell ancillary services between the day-ahead and hour-ahead time frame and between the hour-ahead and real-time market in a multi-settlement process similar to what exists with energy. In this way, the ISO would only need to purchase the ancillary services actually used to operate the system during that hour.

One concern expressed with a formal hour-ahead market is that market participants could not receive feedback from the hour-ahead market quickly enough to be able to re-bid into the real-time market. There are a number of ways to address this concern. For instance, suppliers could be allowed to submit a different, but related, set of bids for the hour-ahead market versus the real-time market. This could be accomplished by requiring the price points of the bid curves to be the same across the two markets, but allowing the real-time bid quantity associated with each price point to be larger or smaller than the hour-ahead bid quantity for that price point, subject to the constraint that the total amount of energy bid into the real-time market is greater than or equal to the total amount bid into the hour-ahead market. Consider the example of an hour-ahead bid curve of 50 MWh at price of \$20/MWh, an additional 40 MWh at a price of \$30/MWh and finally an additional 30 MWh at a price of \$50/MWh. The market participant would be required to keep the price bids the same, but they could bid more or less output from each price step subject to the requirement that the sum of the steps of the bid curve is greater than or equal to $120 \text{ MWh} = (50 + 40 + 30)$. Both of the hour-ahead and real-time quantity bids and price bids (that are the same for both markets) would have to be submitted before the hour-ahead market, because suppliers do not have the opportunity to change their bids in response to hour-ahead market outcomes. A more straightforward, but less flexible alternative would be to require the same bid curve to be submitted to the hour-ahead and real-time markets. At the other extreme, suppliers could submit completely different bid curves to hour-ahead and real-time markets. In all of these cases, both bid curves would be submitted before the hour-ahead market closes.

Eliminating the pre-dispatch process and having a formal hour-ahead market would have the following benefits. First, it would encourage importers to schedule more energy in the day-ahead and hour-ahead markets. It would reduce the thinness of the day-ahead, hour-ahead, and real-time energy markets, because importers would be required to compete on equal footing with internal resources in these three markets. Finally, it is consistent with the market design goal of limiting the ISO's role as an energy market participant to just the real-time energy market, when it buys and sells energy on behalf of market participants that are out of balance with respect to their final schedule. Under the current market design, the ISO is the sole buyer of energy in both the pre-dispatch process and the real-time market. As a consequence, it is difficult for market

participants to arbitrage price differences between the pre-dispatch process and the real-time market. In particular, by purchasing a significant quantity of imports in the pre-dispatch process, the ISO operators can significantly reduce the demand for energy in the real-time market and thereby depress the real-time price. If the ISO is the single buyer in both pre-dispatch process and real-time market, there are few options available to market participants who would like to profit from attempts by the ISO operators to reduce the demand for energy in the real-time market. By eliminating the pre-dispatch process, market participants no longer need to be concerned with the ISO operators' attempts to reduce the demand for energy in the real-time market. Given the existence of the residual unit commitment (RUC) under MRTU, the reliability argument for pre-dispatch process for imports becomes less compelling because the ISO operators can purchase the necessary capacity through the RUC process.

4. Trading Hub Formation

The ISO is considering the development of trading hubs to: (1) facilitate the settlement of the existing State of California Department of Water Resources (DWR) forward contracts and (2) facilitate long-term contracting in an LMP environment. While we believe that a satisfactory resolution of the seller's choice of delivery location problem with the DWR contracts is necessary to move forward with an LMP market in California, one should not lose sight of the fact that most of these contracts will expire within a year or two following the proposed implementation date of MRTU. The ISO should therefore not compromise the long-term value of establishing trading hubs to solve this legacy problem.

It is also important to emphasize that there is potential downside to market participants from the ISO pre-specifying certain locations or combinations of locations as trading hubs. This choice by the ISO could limit the ability of market participants to choose the locations at which they would like to purchase their energy in the forward market. Consequently, unless having the ISO specify the location and composition of trading hubs facilitates the resolution of the DWR contracts seller's choice problem, the ISO should allow trading hubs to develop through the voluntary decisions of buyers and sellers to transact at certain locations in the transmission network. This will allow market participants the greatest flexibility to choose those locations or combinations of locations for trading energy and ancillary services that best serve their interests.

Background on Issue

On August 5, 2004, the ISO issued a white paper describing trading hubs being considered in an LMP setting. In fall 2004, there were a series of stakeholder meetings to discuss the development of trading hubs. The results of this stakeholder process are summarized in an October 26, 2004, ISO white paper. There was a general consensus that trading hubs should be based on existing zones (NP15, SP15, and ZP26), though there was no consensus about how the hub prices should be determined.

During 2005, Ellen Wolfe of the WPTF led a group of stakeholders that hoped to reach consensus on the trading hub issues. She described these efforts at the August 16-18, 2005, stakeholder meeting. Stakeholders agreed that they wanted a pricing algorithm that was transparent and created a deep liquid market, but did not reach consensus on what algorithm

would be preferred. As a result, the ISO continues to have several pricing options under consideration, which are described in an August 10, 2005, white paper:

1. Simple Average Of All Generation LMPs. There is no differential weighting of the LMPs in this option, merely a simple average of all generation nodes, so each generation node has the same weighting regardless of the capacity or average output of generation units at that location. The obvious disadvantage of this option is that a generator with a 10 MW schedule receives the same weighting as a generator with a 600 MW schedule. A variation of this approach is:
1b: Simple Average of a Subset of Nodes. LMPs in this option are not differentially weighted either, every LMP is equally weighted all the time, however the nodes are carefully chosen and the formulation is statistically verified to conform to the average price paid to generation in the zone.
2. P_Max (Maximum Capacity of Unit) Weighted Average Of All Generation LMPs. For this option, weights are fixed for a year and change once a year based on capacity additions and retirements. This approach will produce better representation of the average price paid to generation in the zone, but will likely bias the results up as peakers that run for short periods will receive the same weighting as a similarly sized base load unit despite the vast difference in their output.
3. Annual Average Output of All Generation Node LMPs. This approach will create a single set of weights for the entire year based on annual average output of all generator LMPs. This set of weights would change once a year and would be coordinated with the CRR auction process. It will most likely produce the best representation of the average price paid to generation in the zone, but will tend to dampen the oscillations around the mean, i.e., on peak days it will underestimate prices as peakers will have a minor representation, and during off-peak hours it will bias the results up slightly, as peaker prices will be included even though they are not running.

In a recent conference call, several market participants suggested a fourth option.

4. Dynamic Weighted Average Output Of All Generation LMPs. Under this operation the EZ Gen Hub price would be an output weighted average price of all generation LMPs, but the weighting would vary by hour depending on market outcomes. The benefit of this approach is that it will accurately capture the true average price paid to generation. However, it has some drawbacks in that the dynamic weights may affect the ability to attain perfect hedges through CRRs (since CRRs to or from a hub will have to be based on a fixed set of weights).

Although we believe there are many feasible trading hub definitions, there are a number of general guidelines that should be followed in specifying trading hubs. First, the definition of any trading hub must be clear to all market participants in terms of geographic region covered, specific nodes included, and method used to compute the trading price. Uncertainty in any of these dimensions will severely limit the usefulness of trading hubs.

The ISO should define the trading hubs as soon as possible in order to facilitate fixed-price forward contracting. It is unlikely that the trading hubs will be used for long-term forward contracting purposes until these definitions are finalized. There is also broad consensus to create hubs based on existing congestion zones, which seems reasonable. This consensus and the need to begin the forward contracting process as soon as possible implies that the ISO should quickly work to clarify the remaining details of specifying these trading hubs.

The only outstanding issue appears to be the price determination process. Here the major issues appear to be: (1) what choice will resolve the DWR seller's choice contract problem, (2) what mechanism will attract the greatest trading volume in the future, and (3) what will limit the scope for potential disputes among market participants. The simple unweighted average of LMPs within a geographic area is less attractive in terms of these criteria because there can be significant differences between the generation location and load withdrawal point in the amount of energy that is injected or withdrawn at that location, a fact that is not captured by the unweighted average price.

A number of stakeholders have argued for averaging prices to reduce price variation within in a region. However, we believe it is important to bear in mind that a less volatile trading hub price may fail to reflect the price variation a market participant actually faces. The trading hub price should reflect prices within the zone. If prices are volatile, the trading hub price should be volatile.

Our preferred solution is an hourly price index that is the hourly output-weighted average of all generation LMPs within a given geographic region. The generation nodes making up this hub price could be fixed for one year or changed as generation units exit or enter during the year. However, the weight applied to each LMP during the hour is based on the amount of energy actually injected at that node during that hour. The weight for each nodal price used to construct the hourly hub price is the quantity of hourly injections at that node divided by total quantity of hourly injections at all of the nodes included in the hub price. These weights would be the same for day-ahead, hour-ahead, and real-time markets. Because the amount injected at a node is not known until real-time, all of these hub prices can only be computed after the real-time market has operated.

We prefer this approach to one in which the quantity-weights are fixed for all hours of the day and only periodically adjusted. The use of hourly injections rather than sales in the day-ahead, hour-ahead or real-time market to construct the weights eliminates index price volatility associated with how generation unit owner schedules the energy it ultimately injects into the network.. A crucial pre-condition for setting this trading hub price is that the ISO releases the hourly weights and nodal prices entering into the calculation each hour. Without this transparency in the construction of the trading hub price, there may be a reduction in trading volume. We recognize that this approach may leave some residual unhedged congestion for CRR holders to bear unless the ISO issues CRRs that are sourced from this trading hub. However, if the actual hourly injections at each generation node are used to compute these weights then a very high LMP at a generation node injecting a small amount of energy will not exert undue influence on the hub price. In contrast, a high LMP at a generation node injecting a substantial amount of energy will increase the hub price.

In closing this section, we emphasize that the primary goal of the trading hub design process is facilitating a settlement to the seller's choice contract problem. Although we believe our preferred approach best balances the competing goals of designing a hub price, the ISO should pick a definition that most easily resolves the seller's choice contracts issues subject to the constraint that it does not reduce overall market efficiency. None of the trading hub price definitions described above, if adopted, should significantly reduce overall market efficiency. Consequently, the ISO should pick its preferred option as soon as possible so that the process of negotiating fixed-price long-term contracts for energy between suppliers and LSEs that clear against short-term prices from the MRTU market can begin.

5. Allocation of CRRs

The allocation of Congestion Revenue Rights (CRRs) is one of the most contentious remaining issues of the MRTU process. This topic received the most attention in the written stakeholder comments submitted to the MSC and the largest amount of discussion among stakeholders at the September 22 MSC meeting. The CRR allocation process will impact the distribution of hundreds of millions of dollars in congestion costs and revenues. This is also one of the issues that is *least* likely to have a major impact on the efficiency of market and system operation under an LMP market if the ISO respects a few important guidelines in allocating CRRs.

One of the major goals behind market redesign, and even the creation of an Independent System Operator under the original restructuring process, was to decouple the dispatch and operations of the grid from the ownership of transmission rights on the network. The previous paradigm of physical transmission rights has always been problematic for restructured markets because of concerns that physical rights holders could use their ownership to withhold transmission capacity and benefit their market positions. A rigid regime of pre-specified physical rights is also a poor match for the constantly changing nature of transmission congestion and the protocols needed to manage it.

Thus the initiatives that created the ISO and which have been developed under MRTU have strived to allow operators to manage the real-time operations of the network in the most efficient and reliable fashion without having to be constrained by the details of transmission ownership and contracting. These details are financially important, but should not be allowed to dictate real-time physical operations. Instead, financial transmission rights or similar financial instruments have been developed to create an avenue for ISOs to return congestion revenues to transmission owners and to facilitate the hedging of congestion costs by market participants. A second important goal of MRTU is to improve the information flow and management of transmission congestion in the California market by providing market participants with a more accurate representation of those constraints as well as an incentive to work to reduce the cost of those constraints. This is one purpose of applying locational marginal pricing (LMP) to the market.

Having established the two fundamental tenets of separation of transmission ownership from grid management and the provision of accurate information and incentives about congestion the distribution of the CRRs must not erode these goals. To us, this is the primary

concern about CRR allocation: that it should not interfere with the efficient operation of the market. The simplicity of the allocation process is an element of this. More complex allocation procedures raise transactions costs to all involved in the process. All else being equal, a simple allocation process that protects against the use of CRRs to degrade system reliability and market efficiency is preferred to more complex approaches that do not respect these concerns. As we discuss below, auction mechanisms to allocate CRRs can result in them being used to degrade system reliability and market efficiency.

The ISO has also identified several other principles to guide the allocation process, such as the “full payback” of congestion revenues to market participants, the revenue adequacy of the CRRs, and the fact that allocations should be “reasonably consistent with each LSE’s actual or expected use of the grid.”¹¹ These concerns, either directly or indirectly, address the equity of the allocation process. We address each of these topics, market efficiency and equity, in turn.

Allocating CRRs to Enhance Market Efficiency

Because financial transmission rights do not guarantee preferential access to the network, or directly reduce the marginal cost of congestion for using the network, there is great latitude in means for distributing these rights and not eroding these goals. However, there are still concerns to keep in mind.

First, there is a danger in linking CRRs to ongoing market transactions. To do so works to mute or even reverse the incentives for which LMP is being developed to provide. Consider the decision to build generation either on the Mexican border with California or within some of the constraints that separate that border region with the load centers in southern California. Results of the ISO’s LMP studies indicate that energy generated at the Mexican border will be less valuable, potentially much less so, than power generated closer to Los Angeles or San Diego. This difference may or not be enough to dictate the decision about where to site a new generation unit, as costs will obviously also differ by location. But if proper price signals are sent to producers, that decision will be based upon a proper accounting of costs and benefits that unit owners would receive. However, if upon constructing a new power plant on the Mexican border a firm is also awarded CRRs that offset its costs of bringing energy from this unit to the major load centers, then the allocation process has largely offset the incentives provided by the application of LMP. In effect, firms would be rewarded for siting power plants in locations with high congestion costs, where electricity is far less valuable.

Second, there is the danger that firms, upon purchasing or being awarded a financial CRR, would behave less efficiently than if they did not own that CRR. This second effect arises if firms are not behaving in a perfectly competitive manner. The scenario of greatest concern is a situation where a firm with generation into a load center, such as San Francisco, purchases CRRs that sink, rather than source, in that load center. Thus rather than acquiring the CRR for hedging purposes, as would be the case if the CRR were sourced in that location, the firm has in effect leveraged its position within the load pocket. The firm would profit twice from high local prices,

¹¹ These goals are summarized on page 22 of the CRR study 2, by Scott Harvey and Susan Pope, (available at <http://www.caiso.com/docs/2005/08/24/2005082417481216533.pdf>)

first by selling energy and second from raising the value of its CRR. The incentive to withhold output and raise prices is obviously increased.

This is one concern with an auction mechanism being used to allocate CRRs. The market participant able to cause the most congestion is willing to pay the most for CRRs that refund these congestion charges. Consider the following simple example of a CRR auction. Suppose that LSE1 expects to receive \$5 million in congestion revenues if it owns the CRRs being sold at auction. Generation unit owner 1 (GEN1) expects to earn \$15 million in congestion revenues if it owns these same CRRs because of how it bids and operates its generation units. Suppose these are the only two entities bidding for the CRRs and that each bidder only knows their own valuation of the CRRs. GEN1 only has to pay slightly above \$5 million (LSE1's value) to purchase CRRs that it expects to earn \$15 million in revenues from owning. By allocating the CRRs to LSE1, GEN1 no longer has the incentive to cause the level of congestion that makes the CRRs worth \$15 million. By this logic, a carefully considered allocation mechanism has the potential to limit the market inefficiencies associated with the ISO issuing CRRs relative to an auction mechanism for allocating CRRs. This example also illustrates the superiority of allocating CRRs rather than allocating Auction Revenue Rights (ARRs) and then running a CRR auction.¹² In a CRR auction, the market participant able to cause the most congestion is the one most likely to win the CRR auction. This logic could explain why ISOs, such as PJM, that went from allocating CRRs to allocating ARRs and then running a CRR auction have seen a significant increase in congestion.

Finally, there is the concern that firms would use the control of CRRs as a barrier to entry. If CRRs are the only available means to hedge congestion risk, and if that risk is great enough that financing for construction cannot be obtained without it, then new entrants will need access to CRRs. However, it is important to emphasize other entities besides the ISO can issue CRRs, so that a potential new entrant could obtain this hedge from a third-party. A third-party is unlikely to issue the necessary CRRs at a reasonable price if it believes market participants could influence its ex post congestion obligation.

Equity in CRR Allocation Processes

The reason why CRR allocation has been so contentious is that it is fundamentally about which firms, and in what proportion, will receive the congestion revenues collected by the ISO. Put simply, how much of the difference between what loads pay for the energy they consume and what generation unit owners are paid for the energy they produce will be redistributed, at no cost, to each market participant. At the root of the issue are two competing paradigms. One states that congestion revenues should flow in proportion to the share of fixed transmission costs paid by market participants. This interpretation is consistent with an allocation of revenues proportional to the fixed Transmission Access Charges (TAC) paid by market participants. The other states that congestion revenues should flow in proportion to the congestion costs paid by market participants. This interpretation is consistent with an allocation that attempts to mirror

¹² The ARR scheme gives rights holders auction revenues in proportion to their ARR allocation.

actual usage of the grid. If reduction of congestion risks through the holding of CRR hedges has economic value, then an allocation that matches use of the grid would enhance efficiency.¹³

The decision of how to allocate CRRs comes down to which paradigm is embraced. Unfortunately, neither paradigm has any economic efficiency properties to recommend it over the other paradigm. The transmission network has already been built and it must be paid for. This implies that the ISO has considerable latitude to design a CRR allocation mechanism that does not degrade market efficiency. Consequently, the primary concerns of the CRR allocation process should be its equity as perceived by the relevant stakeholder groups and the cost of the CRR allocation process chosen. If the ISO implements an allocation mechanism involving high set-up costs and ongoing operating costs, these costs must be recovered through a higher Grid Management Charge (GMC) paid by all loads. Those entities able to obtain more of the CRRs they want through an expensive allocation process may argue in favor of it, not because it improves the efficiency of energy market, but because they can pass a significant portion of the increased expense on to other market participants in a higher GMC. Thus, subject to the constraint that the CRR allocation process does not degrade overall market efficiency, the ISO should implement the allocation mechanism that costs market participants the least amount of money subject to the constraint that the majority of them perceive it as equitable.

California ISO Proposals

The ISO has discussed two sets of approaches to be used in some combination. One is an initial allocation process backed by a one-shot verification that the CRRs requested by each LSE can be credibly linked to actual historic usage of the network. The other is an auction of CRRs, either to supplement the original allocation or as the primary means of distributing all CRRs. For the reasons discussed above, we favor allocating CRRs to LSEs.

We strongly oppose any efforts to turn the CRR verification process into an ongoing effort to link grid usage to CRR distribution. We understand the desire of market participants to obtain congestion hedges for historic uses of the transmission network. However, we also understand the difficulty in defining historical uses of the transmission network. How the generation units that an LSE owns have been operated provides a straightforward method for determining the LSE's historical use of the transmission network. Long-term supply arrangements linked to specific generation units can also provide valuable input to the extent that the seller of the long-term supply agreement does not have the option to substitute alternative generation sources to meet its contractual obligations. In short, an allocation process based on credible historical use of the transmission network is likely to allocate a significant fraction of the available CRRs. For market efficiency reasons we also support a very conservative definition of the historical use of the transmission network.

The remaining CRRs should be allocated in the most cost effective manner possible. These CRRs are what remain after all credible historical uses of the network have been taken in account. Consequently, this stage of the allocation process should focus on refunding the

¹³ If trading in the secondary market is low cost then initial ownership of CRRs should not matter, because those who desire such hedges will be able to buy them.

difference between the amount paid by loads and the amount paid to generators in a least cost manner using a mechanism that is perceived to be equitable by the majority of stakeholders.

A final issue is whether to grandfather initial CRR allocations. As we stated earlier, we feel strongly that allocations should not be tied to ongoing market decisions. A grandfathering of CRRs is one, but not the only, way to accomplish that. We are concerned that the current uncertainty in the magnitude and geographic distribution of congestion charges may require that the ISO revisit its initial CRR allocations. For this reason, we recommend limiting the amount of rights that qualify for grandfathering in the initial CRR allocation. Those CRRs allocated based upon credible historical use should be given a higher priority in the allocation process in subsequent years. A prudent strategy for transitioning to an LMP market in California is for the ISO to preserve flexibility in its annual CRR allocations to ensure that no LSE experiences significant harm in the transition to an LMP market as a result of, for instance, the use of CRRs to magnify market power. The transition to LMP in the PJM market required significant revisions of the CRR holdings of several market participants. We expect similar unanticipated increases in congestion that require adjustments in the initial CRR allocation to also arise in California. The presence of this uncertainty about the effects of CRRs argues for an allocation design that is more conservative, rather than aggressive, in determining the amount of CRRs both available for allocation and eligible for grandfathering.

6. Out-of-Control-Area Load CRR Allocation

The ISO proposes to allocate CRRs to LSEs located outside of the ISO control area that pay the ISO's access charge. For each MW of access charge paid on a monthly or annual basis, this LSE will receive the right to request a 1 MW CRR. The ISO argues that this mechanism is consistent with its fundamental principle of eligibility for CRR allocation "that parties who support the embedded costs of the CAISO transmission grid are entitled to an allocation of CRRs in accordance with the nature and extent of their support for these costs." However, as was mentioned above, this principle has no economic efficiency properties to recommend it. The transmission network has already been constructed and the transmission owner is entitled to receive recovery of all prudently incurred costs under the rate-setting process.

In addition, if an LSE located outside of the ISO control area uses the ISO grid to serve its customers then it must pay the Transmission Access Charge (TAC) for each MWh it withdraws from the ISO control area. Consequently, the ISO's principle for allocating CRRs is based solely on equity considerations. Viewed from this perspective there is another approach to allocating CRRs, discussed below, that takes a different perspective on the equity consideration that may also have market efficiency consequences.

In any zonal or nodal locational pricing market, the ISO collects more revenues from load than it pays to generation unit owners if there are locational price differences. This occurs because the large load centers have higher locational prices than the outlying areas where major generation units are located. This merchandizing surplus is what allows the ISO to fund its CRR obligations. Consequently, one view of CRRs is that they are a mechanism for refunding congestion revenues to market participants. Loads will pay the delivered cost of electricity, including congestion charges and losses, under the MRTU design; in order to give the

appropriate locational price signals to generation unit owners to locate where LMPs are highest, the ISO should refund these congestion revenues to loads.

The logic that CRRs are a mechanism for refunding congestion revenues to loads implies an alternative principle for “equitable” allocation of CRRs: Congestion revenues should be refunded to those market participants that paid them in proportion to how much of total congestion charges they paid. Under this scheme, an LSE located outside of the ISO control area would pay the TAC only for MWh withdrawn from the network, but it would receive a CRR allocation on the expected amount of congestion charges that the ISO believes this market participant would pay over the duration of the CRR.

This guiding principle for CRR allocation also has market efficiency consequences if the ISO recognizes what generation resources the LSE owns or has long-term supply arrangements with in making a determination of how much congestion charges the LSE would have to pay on a prospective basis. In particular, if the ISO allowed LSEs outside of the ISO control area to purchase CRRs if they are willing to pay the TAC, then an LSE with significant generation outside of the ISO control area may be tempted to purchase a substantial volume of CRRs and use its external generation to cause a high locational price at its point of interconnection with the ISO network in order to profit from its CRR holding.¹⁴ Instead, if the ISO bases its allocation of CRRs on prospective levels of congestion, it can limit its allocation of CRRs to this market participant in such a way as to prevent it from using its own generation to cause congestion.

This principle of refunding congestion charges paid is consistent with the ISO’s desire to allocate CRRs based on historical use of the transmission network. If an LSE located outside of the ISO control area owns generation or has a generation unit-specific long-term supply arrangement that makes use of the ISO network, then it is hard to see why this LSE should not be granted a CRR allocation based this historical use of the transmission network, as long as the LSE also pays the TAC for all MWhs withdrawn from the ISO network.

We do not believe the guiding principle given above is superior to the ISO’s guiding principle. However, to the extent that the ISO provides justification for allocating CRRs to internal LSEs based on historical use of the network, it must apply this same principle in a non-discriminatory manner to the process of allocating CRRs to LSEs located outside of the ISO control area. We do not see any complications with applying the two-step CRR allocation process recommended in the previous section—first based on historical use of the network with the remaining CRRs allocated using a low-cost allocation rule—to the case of LSE located outside of the ISO control area.

7. CRRs for Merchant Transmission

The ISO proposes to allocate CRRs to merchant transmission owners for the life of their project and that these CRRs would be the sole cost recovery mechanism for the asset. We have supported this use of CRRs, subject to the pre-condition that this CRR allocation to the merchant transmission owner ceases to exist as soon as the transmission owner accepts one dollar of regulated revenues from services provided using these transmission facilities. It is essential that

¹⁴ This strategy can be profitable if the quantity of CRRs exceeds the amount of external generation.

there is a bright line between merchant and regulated transmission projects. In a number of markets around the world, there have been attempts to construct merchant transmission projects. A typical pattern is that an asset begins life as a merchant project but when the congestion revenues the project owner earns are insufficient to cover the cost of the project, the owner asks for the project to be put in the regulated transmission owner's rate base. If this sequence of events occurs in California, then the ISO should immediately revoke the CRRs issued to the merchant transmission owner and treat this project the same as any other project owned by the regulated transmission owner.

The timing of the assignment of the CRRs to a merchant transmission project can dramatically impact the likelihood that these projects will occur. Waiting until the project is in place before assigning CRRs, as appears to be the case under the ISO's proposal, will virtually guarantee that no merchant transmission projects will be constructed. It is very hard to imagine that a private investor would spend the money necessary to build a transmission facility without any commitment from the ISO in terms of the magnitude of CRRs the investors is entitled to. If the ISO would like to support merchant transmission investments, we urge it consider implementing a CRR allocation process that commits to a lower bound on the amount of CRRs awarded at the start of the project, with the option to increase this amount after the project is completed. Without this sort of up-front minimum CRR commitment a merchant transmission investor faces an unacceptable level of risk of ex post opportunistic behavior by the ISO or other entities overseeing the California transmission network once the investor has incurred the sunk cost of constructing the transmission facility. The CRRs that are allocated to merchant transmission should reflect the full contribution of the transmission addition to the ability of the grid to transmit energy and ancillary services. Restricting sinks of CRRs to LAPs could greatly undervalue the contribution of merchant transmission, as the major effect of many additions would be to reduce within-zone congestion.

Conclusion

The ISO has undertaken an extensive stakeholder process to arrive at its proposals. From our participation in these meetings and the written and oral comments that we received, it is clear that their efforts to understand and improve the market design have made important contributions to the final ISO proposals. However, in a number of instances this process has resulted in compromises that may significantly degrade market efficiency and system reliability. In closing we note these sources of potential market inefficiencies and our recommendations for addressing them.

The first such compromise is the use of ad hoc bid adders in constructing mitigated bids. As discussed in Section 1, we see no economic efficiency rationale for implementing bid adders for mitigated units despite the precedent for their use in a number of other ISOs. For competitive path assessment (Section 2), we encourage the ISO to take a cautious approach to deeming transmission paths competitive until it has a year of experience with LMP market outcomes. The existing zonal interfaces should be the only competitive interfaces during the first year of the market. This number is broadly consistent with the number of competitive transmission interfaces in PJM before the geographic boundaries of the PJM market expanded.

The HASP is another compromise solution that has potential adverse market efficiency consequences (Section 3). We hope that the ISO will continue to make provisions in the system software for a full hour-ahead market. The market efficiency benefits from an hour-ahead market consistent with the day-ahead and real-time markets appear to be sufficient to justify its existence. Except to resolve the seller's choice contract issue, the ISO should limit its intervention into the choice of trading hub prices (Section 4). Once this issue has been resolved, the ISO should define the details of trading hub prices as soon as possible so that California's LSEs can begin the process of forward contracting for energy and ancillary services that will delivery under MRTU.

The CRR allocation mechanism should be simplified to the greatest extent possible. As emphasized in Section 5, the CRR allocation process is primarily concerned with refunding the difference between what loads pay for electricity they consume and what suppliers are paid for the electricity they produce. Designing a more costly allocation mechanism means that consumers will ultimately receive less of the congestion refunds they are due because the refunds they receive are net of the cost of the design and implementation of the CRR allocation process. The same principles should be applied to allocating CRRs to LSEs located inside and outside of the ISO control area (Section 6). This allocation should first be based on historical use and then on some equitable allocation of any CRRs that remain. Finally, merchant transmission project should be encouraged and supported by CRR revenues as long as they remain merchant transmission projects (Section 7).