Final Opinion on
Payment for Provision of Flexible Ramping

James Bushnell, Member
Scott M. Harvey, Member
Steven Stoft, Member
Benjamin F. Hobbs, Chairman
Market Surveillance Committee of the California ISO

August 16, 2011

1. Introduction

This opinion comments on the ISO's "Opportunity Cost of Flexible Ramping Constraint" proposal ("Flexiramp Proposal").[1] This proposal would implement a new set of constraints in the Real-Time Pre-Dispatch (RTPD) and Real-Time Dispatch (RTD) processes. The goal of these constraints is to reserve unloaded rampable capacity ("flexiramp") to meet ramping needs in the real-time market. The proposal would also provide for payments to reserved capacity based on calculated marginal opportunity costs from the RTPD market. The proposal is meant to be an interim measure, to be superseded by a ramp-related product to be defined as part of the ISO's Renewable Integration Market & Product Review, Phase 2 ("Renewable Integration Review").[2]

In preparing this opinion, we benefited from interaction with ISO staff and from the written comments made by stakeholders and the questions and concerns raised in the public calls on the proposal.

We support the proposal as a strictly interim measure to increase the supply of upward ramp in the real-time market, pending market reforms resulting from the Renewable Integration Review that better address the fundamental issues.

We have assessed the proposal against seven criteria that we propose for evaluating such interim market proposals. First, we anticipate that the proposal will generally be effective in committing more rampable capacity in real time in the RTPD, and that this flexiramp will be useful to the market. The ISO has performed market simulations of the flexiramp proposal that provide some indication of the likely cost and impact on the real-time availability of ramp. We have not have the opportunity to study these results, but have been informed by ISO staff that the simulations indicate that more ramping capability will result. Some adjustment of parameters used in implementing the flexiramp constraint in RTPD and RTD may be necessary in order to strike an appropriate balance between the risk of too little flexiramp and the excessive payments that may occur if much more than is needed is acquired. It is also possible that at some point in the future there will be insufficient commitable capacity in real-time to meet the need for flexiramp, and

[1] Draft Final Proposal, July 20, 2011, foliweb7.caiso.com/2bc1/2bc1e2b53ba90.pdf

[2] www.caiso.com/informed/Pages/StakeholderProcesses/RenewablesIntegrationMarketProductReviewPhase2.aspx

the possibility of acquiring flexiramp in the day-ahead market will need to be revised. Finally, flexiramp is to be acquired on a system-wide basis; therefore, the ISO will need to monitor where flexiramp is acquired to determine whether it is being obtained in, for instance, generation pockets where it cannot be accessed if needed due to transmission constraints.

Second, we are satisfied that the proposal is reversible. It does not tie the ISO's hands concerning development in the Renewable Integration Review of a ramp product that will better address the fundamental issues. Third, the proposal has the advantage of being simple to implement, and the opportunity cost-based settlement price is automatically calculated in RTPD. However, because of complicated interactions of energy, ancillary services, and flexiramp in different intervals, the causes that contribute to the flexiramp price in particular intervals will inevitably less than transparent. This is, however, a generic problem with energy and ancillary services pricing as well. Fourth, interim measures should avoid large shifts in costs, and ISO staff have stated that the market simulations indicate that this will likely not be a problem. If, however, large costs are incurred without corresponding benefit, then the flexiramp feature can be modified or turned off relatively easily.

Fifth, price signals should be efficient, bearing a relationship to the costs incurred in supplying flexiramp, with costs being allocated to responsible parties. In the interests of simplicity of implementation and reversibility, this criterion is somewhat sacrificed, which we believe is justified for an interim measure. We have two main concerns that lead to us to conclude that the settlement method used should not be made a permanent feature of the ISO markets. First of all, it is not clear what level and form of compensation is needed to incent investments in efficient amounts of upward ramp capability, or its short-run provision in the ISO's markets. Second, we believe that in most cases, the opportunity cost-based payment from RTPD is likely to be greater than the actual opportunity costs, as MW capacity that is designated as flexiramp by RTPD will be free to sell energy in the first (binding) interval of RTD (subject to the need to provide flexiramp in later intervals). Thus generators will receive a payment based on RTPD's estimate of opportunity cost, but in most cases actually incur little or no such cost. If flexiramp prices indeed turn out to be very low, as the market simulations indicate they are likely to be, then we believe that their interim nature makes this possible overpayment an acceptable risk in exchange for providing needed upward ramp capacity. We also note that this is also analogous to the situation with payments for non-contingent spinning reserves to the extent that such reserves can be dispatched for energy, and so the proposal has the advantage of treating those reserves and flexiramp in a broadly consistent manner. The issue of efficient and appropriate payment for ramping capability will be addressed, we hope, in the Renewable Integration Review.

Sixth, we see some potential for unexpected consequences. The potential is not great because of the finite life of this interim measure, and its restrictive scope relative to alternatives that would involve creation of a product that would be acquired both day-ahead and in real-time or which requires submission of bids. Further, we have been told that the ISO's market simulations of the flexiramp proposal indicate that flexiramp prices will be low. However, we believe that the inconsistency between how the RTPD process calculates opportunity costs for flexiramp and the likely smaller opportunity costs that will actually be experienced by generators in the RTPD and RTD processes may lead to unexpected changes in bidding behavior. For example, since spinning reserves will be paid for opportunity costs in RTPD based both on energy and flexiramp opportunities, while only energy opportunity costs are considered in the day-ahead market, this may alter day-ahead spin bidding behavior.

The seventh and last criterion is consistency in philosophy with existing market features. This is not achieved by this proposal because flexiramp is treated differently from ancillary services involving bidding and both day-ahead and real-time markets. We believe that this is acceptable for an interim proposal, given the need for simplicity, reversibility, and limited potential for unexpected consequences. We further believe that the Renewable Integration Review is a more appropriate forum for fundamental consideration of the relationship of any ramp product to other products in the market, and the appropriate design of the market.

The remainder of the Opinion is structured as follows. In Section 2, we characterize in general terms the problem that the flexiramp proposal is addressing, and possible fundamental solutions to the problem. Because this is an interim measure that may be greatly changed as a result of the Renewable Integration Review, it is not feasible or desirable to implement a fundamental solution. In Section 3, we summarize several criteria for evaluating interim measures to provide flexiramp, and we assess this proposal against those criteria. Section 4 presents conclusions.

2. Problem Description and Fundamental Solutions

2.1 The Physical Source of the Problem Addressed

Balancing load and generation is difficult because of unexpected events on both sides of the market. The day-ahead market (IFM) relies on a forecast of load that is implicitly assumed to contain no surprises and purchases generation based on an assumption that generators will perform as bid or self-scheduled. But the IFM does not consider only generation capacity when scheduling generation, it also considers the ramping ability of generators. So if the average hourly load is expected to increase by, say, four GW from one hour to the next, the IFM makes sure the committed generators can increase their output by that amount.

The IFM also purchases ancillary services (AS), regulation and operating reserves. The purchase of operating reserves are designed to take account of unexpected events—"contingencies."[3] But "contingencies" evaluated with regard to operating reserves pertain to the availability of enough capacity to meet forecast load, mainly considering the possibilities of generators and lines going out of service due to failures, not to whether there is enough ramping capability to enable generation to meet load in each dispatch interval. As noted in the flexiramp proposal, many other kinds of "non-contingent" events can disrupt the system's balance—eight are listed. These include fluctuations in load and in non-dispatchable supply, such as wind, solar and run-of-river.

Because operating reserves are by-and-large held in reserve for contingencies,[4] no product is specifically procured to cover the eight "non-contingencies," although regulation is used to meet them within 5-minute dispatch intervals. Fortunately, the generators providing energy have ramping capability available to handle non-contingent events adequately under most circums-

[3] Regulation, in contrast, balances supply and demand within five-minute intervals, since dispatch of generation by RTD is unable to do so.

[4]In the Eastern Interconnection, operating reserves are also times also used to meet ACE when there is so much dragging that regulation is unable to do so. Page 5 of the CAISO's flexiramp proposal suggests that the CAISO also uses its operating reserves to balance load and generation when it is ramp constrained or there is such dragging.

tances. However, particularly during the morning ramp, when many generators are at their maximum up-ramp limits, there are regular shortages of the extra ramping capacity needed to handle non-contingencies. Just under 1% of all five minute real-time intervals suffer power balance violations due to inadequate upward ramp capacity in the second quarter of 2011, according to the most recent DMM Quarterly Report.[5] That same report indicates, however, that the system power balance constraint is violated in RTD due to lack of upward ramp in many hours of the day. In the future, this problem is anticipated to get worse and to occur more often outside of the morning ramp period partly because of the increasing penetration of variable renewable energy sources.

2.2 The Consequences of the Problem Addressed

A shortage of ramping capability to handle non-contingencies results in a number of problems. However, mild shortages of rampable capacity in RTD do not endanger reliability as long as either sufficient regulation is available or, if regulation cannot maintain the desired CAISO Area Control Error (ACE), California can import more power. California is well connected to the Western Interconnection (WI) with uncontrollable AC lines. As a result, if the ISO generators provide 100 MW less than its load (including scheduled imports and the response of regulation), an extra 100 MW will flow into the ISO from the WI. This is a matter of physics and it happens instantly. Moreover generation remains equal to load for the entire WI. However, the WI will not be "in balance." Instead that 100 MW shortfall means that there will be a tiny reduction in voltage and frequency throughout the system. This will force a (slightly) undesirable reduction in load, although generation outside the ISO will quickly adjust to rebalance the WI as a whole. The frequency reduction is uniform throughout the WI, but voltage reductions will be more localized to the ISO. The main point here is that a short-term ACE imbalance in the ISO does not directly impose a high cost. No real problem may be noticeable.

However, unscheduled imports from the WI (leaning on the system) are undesirable and so there are NERC rules, limiting ACE, and there are penalties if ACE limits are violated too often by too much. (These rules are designed both to limit ACE and to maintain the frequency of the WI.) So the first cost to the ISO of imbalances comes from ACE-violation penalties.[6] A much less likely, but more dramatic cost is associated with reliability problems. If it happens that the extra imports due to an ISO imbalance overload a tie-line, that overload must be corrected. The choices are then to (1) use up operating reserves, (2) dispatch emergency demand response, or (3) shed load.

[5] CAISO Department of Market Monitoring, Quarterly Report on Market Issues and Performance, August 2011, Figure 1.7 (www.caiso.com/Documents/QuarterlyReport_MarketIssues-Performance-August 2011.pdf) shows that the power balance constraint is relaxed on average in about 1% of the hours, and that these violations are well dispersed (with 12 of the 24 hours equaling or exceeding that 1% frequency). However, regulation may have been used in these instances to bring CAISO supply and demand closer to balance. The CAISO market design sets the energy price at the power balance constraint penalty when it must use regulation to balance load and generation.

[6] Consistent ACE violations with excess imports when the price of power is above average would make the ISO deservedly unpopular in the WI. However this problem could easily be remedied by extra exports in the afternoon when the price is generally higher than in the morning. Even revenue-neutral ACE violations are thought to provoke disapproval in dispatching circles, but this might not by itself provide sufficient reason for spending much ratepayer money.

The first is undesirable and could endanger reliability, the second is costly, and the third is much more costly.

So there are five apparent sources of cost from the present lack of non-contingent operating reserves, use of regulation for imbalances, use of operating reserves, penalties for ACE violations, purchases of emergency demand response, and load shedding. We have not heard that any of the latter three have actually occurred as a result of this problem, perhaps because regulation has been sufficient. However, we understand from ISO staff that high prices have sometimes been paid—generally out of market—for generator performance that prevents the occurrence of these problems. Presumably, those payments are less than the costs that would otherwise be borne if these problems occur, though we have seen no information about either payments or avoided costs. Assuming the payments are reasonable, then flexiramp can be partially evaluated by asking if it will reduce total financial costs. Again, we have seen no estimates, but if flexiramp turns out not to be cheaper than other alternatives, it has been designed by the ISO to be easy to turn off—even in real time.

2.3 The Market-Design Sources of Problem Addressed

Unexpected balancing events (non-contingencies) are to be expected. In fact, they are certain to occur, but at unpredictable times. The fact that the market design does not sufficiently account for them now and therefore acquires too little ramping capacity therefore cannot be attributed to their unexpected nature.[7] It appears that there are two main flaws in the present market design that flexiramp is designed to address. Because this market patch is being made under considerable time pressure, flexiramp does not fix either flaw, but instead, implements a compensating measure.

The first market flaw is that non-contingent imbalance events are dealt with only by regulation, and are disregarded by the real-time commitment and dispatch market for energy and operating reserves. As wind penetration increases, using regulation alone will become an increasingly inefficient way of dealing with those imbalance events. That the energy market ignores non-contingent imbalance events is easily explained. Taking them into account explicitly in the market dispatch would require the software to look at not just one possible future—the forecast load—but at a great many possible futures. This sort of stochastic-programming approach, though theoretically ideal and the subject of significant research and development by universities and vendors, is simply out of reach at this time. It is quite possible that even if it could be implemented, it would provide only small savings compared with a heuristic approach similar to what is now used for ancillary services. So the market flaw of using a non-stochastic energy dispatch should likely remain, though a heuristic correction to increase the market's procurement of ramping capability might make sense.

That leaves the flaw in the operating reserves markets. Jumping ahead a bit, we note that flexiramp is simply another way of buying non-contingent reserves, a product that the ISO already purchases in its AS markets. The problem is that the ISO has no control over the amount of non-contingent reserves that it purchases. It simply buys reserves and lets the sellers designate them as contingent or non-contingent. The ISO could buy more reserves, but much or most of these

[7] Regulation is acquired for these contingencies within dispatch intervals.

would likely be contingent reserves, as is now the case. Fixing this flaw is non-trivial (but not particularly difficult) because, if the ISO required a specific quantity of non-contingent reserves, it would need to price them separately as they would at times be a more expensive product than the less-frequently used contingent reserves. This would occur when the contingency reserve requirement is binding, resulting in price separation between contingent and non-contingent reserves.

From an economic perspective, there is little difference between contingencies and non-contingent events that cause sudden imbalance problems. Standard contingencies tend to be quicker and so they require more regulation, but after regulation (and some very temporary leaning on interconnections[8]), both types of surprises are handled the same way by 10-minute spinning and non-spinning reserves.[9] Since the basic approach to AS is well established and largely implemented, it is likely to make more sense in the long run (probably in the Renewable Integration Review) to fix the AS market rather than to invent a new flexiramp market. However, in the short-run, flexiramp is the more practical approach.

2.4 The Flexiramp Design

Flexiramp is described by the proposal as a constraint, but as usual, the constraint simply determines the ISO's demand for the flexiramp product. The capacity providing flexiramp functions, in essence, as a form of non-contingent reserves procured at a different price than standard spin, and at a price that is not set by offers from the generators providing the service.[10] The amount of capacity scheduled to meet this constraint (the demand for flexiramp) is the Minimum online Required Upward ramping Capability requirement (MRUC).[11] This value will be determined by a simple heuristic formula involving recent ramping needs as modified by the system operator's input.

Flexiramp will be priced in RTPD. Generators providing flexiramp in the first (binding) interval of RTPD will be paid a price equal to the shadow price (or dual variable) of the RTPD flexiramp constraint (MRUC). This shadow price equals the opportunity cost of devoting capacity to provide flexiramp rather than to providing some other product.[12] Bids for flexiramp do not set the

[8] The response to an outage contingency generally involves such leaning in the seconds after a large outage, as allowed by reliability standards. These standards accept that neither the CAISO nor any other balancing authority area would carry enough regulation to cover a large outage by itself.

[9] Another difference is that the loss of a unit results in a persistent loss of capacity that can be replaced by starting a unit, whereas, variations in wind or load can be shorter in duration for which starting a unit with a minimum run time might not be appropriate.

[10] Note, however, that despite the similarity in function, flexiramp is defined as a separate product from noncontingent spin. Capacity designated as flexiramp cannot also be designated as spinning reserve in the same interval.

[11] No downward ramping capacity is being required at this time, although the DMM Quarterly Report on Market Issues and Performance for the second quarter of 2011 (op. cit.) indicates that more intervals suffer power balance relaxations (violations) due to inadequacies in downward ramp capacity than because of inadequate upward ramp capability (see Figs. 1.5 and 1.6 and the accompanying discussion).

[12] For instance, a generator might provide 20 MW of flexiramp in interval 1 of RTPD. If it could instead have provided 20 MW of spin, priced at $5/MWh and if it had bid $2/MWh to provide that spin, then it

price of flexiramp, as there are no bids. The RTPD market in essence assumes that all available rampable capacity can be used for flexiramp up to the limits imposed by capacity constraints and ramp constraints; generators cannot otherwise limit the amount of its capacity that can be scheduled in the RTPD process for this purpose.

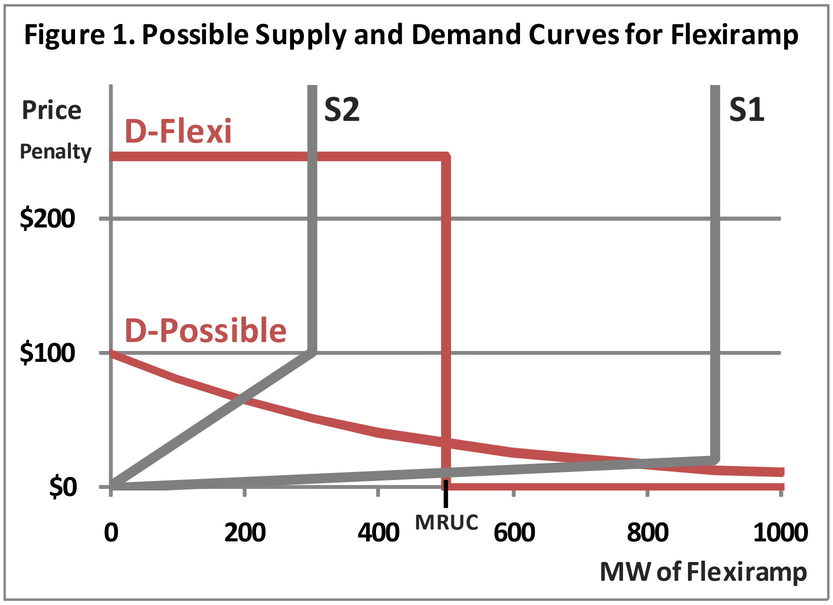
In each 15-minute real-time pre-dispatch (RTPD), generation will be scheduled to the end of the hour in such a way that MRUC MW of flexiramp would be available in each 5-minute interval if everything went as planned. This is accomplished by dispatching some off-line generators and by backing off some generators that are near or at maximum capacity. Backing off these generators will be costly if other generators need to be dispatched or if more expensive generators need to be ramped up to take up the slack. This cost is the market's opportunity cost of flexiramp, and for intervals in which it is non-zero this opportunity cost will be paid to all MRUC MW of flexiramp that is planned in the RTPD. Of course in the real time dispatch (RTD), some of this flexiramp will not be provided. This is because it may be used to cover non-contingent balancing events and because RTD relaxes some of the requirements for flexiramp relative to RTPD. In particular, no flexiramp is required in the first (binding) 5 minute interval, while reduced amounts (relative to RTPD's MRUC) are acquired in the non-binding intervals immediately after the binding interval.

The provision of flexiramp will in most intervals likely be supported by RTPD committing generation that would otherwise be off-line. After the initial (binding) 5-minute interval in RTD, flexiramp is maintained by RTD by positioning some generators below their maximum output. However the flexiramp constraint is reduced in the first few non-binding 5-minute interval in RTD, as just mentioned. Hence, generators that were projected to incur an opportunity cost of providing flexi-ramp in RTPD have a high likelihood of being dispatched to provide energy in RTD, thereby incurring no opportunity cost. For this reason, and because of thermal generation needs arising from non-contingencies, the actual opportunity cost in RTD will often turn out differently than anticipated in RTDP. The result is that the actual market cost of flexiramp can be quite difficult to predict. However, there can be no doubt that the flexiramp mechanism will procure an increased level of what is, in effect, non-contingent spin that can be used for solving the ramping problems that are currently plaguing the ISO.

The RTPD's process for acquiring flexiramp can be viewed as an interplay of a demand curve for flexiramp with supply. In particular, the MRUC requirement in combination with the penalty price on the flexiramp constraint defines the ISO's demand curve (shown as D-Flexi in Figure 1) for the flexiramp product.

---

bears a $3/MWh opportunity cost for providing flexiramp rather than spin. The RTPD optimization will then automatically yield a flexiramp shadow price of at least $3/MWh. Alternatively, the opportunity cost might instead result from diverting that unit's capacity from profitable energy production, in which case the opportunity cost would be the difference between the energy price and the unit's energy bid. (Note, however, energy is not settled in RTDP, so this is a surrogate for possible opportunity costs in the RTD market where energy is actually sold.) The opportunity cost calculation can become more complicated by intertemporal ramp constraints, which means that the shadow price for flexiramp in interval 1 might be determined by foregone opportunities for AS or energy in later RTPD intervals.



The higher, step-function demand curve, labeled D-Flexi is the demand curve determined by the flexiramp proposal, although MRUC will vary from hour to hour. The lower, smooth demand curve, labeled D-Possible, is one possibility for what an optimal demand curve might look like. (It is possible, of course, that the optimal curve might be higher than D-Flexi for some values of flexiramp to the left of 500 MW; we do not mean to prejudge the marginal value of flexiramp.) In one respect, D-Possible is more reflective of the value of flexiramp than D-Flexi because it is smooth. It is unrealistic to think that MRUC could be worth $247/MWh (its penalty price) when 499 MW is available and $0/MWh when 501 MW is available. Such an on-off demand curves makes sense when complying with a black-and-white regulatory requirement, such as NERC's requirement for operating reserves, but they make little sense (except as operational compromises) when they are meant to reflect the value of a physical hedge against random contingencies. The penalty level of the D-Flexi demand curve is also unlikely to reflect the economic value of flexiramp since it was determined with little regard for the value of flexiramp.

Fortunately, based on reports of the ISO's market simulations of flexiramp, the supply curve for flexiramp seems to be frequently most similar to S1. In other words, flexiramp is anticipated to be usually very inexpensive. In this case the D-Flexi curve will demand too little, but will pay a lower price, often zero, for a fairly generous supply of flexiramp. However if the system runs into conditions under which flexiramp is not cheap, perhaps as shown by supply curve S2, then using the wrong demand curve will likely cause significant over procurement at an uneconomic cost. Since the ISO's reported simulations indicate that a generous supply of flexiramp is currently available, we do not recommend a change in the current design. However, for the ISO's Renewable Integration Market & Product Review, Phase 2 ("Renewable Integration Review"), we suggest that the demand side of the market be seriously considered in a manner consistent with scarcity pricing for other market products.

3. Criteria for Evaluation of an Interim Solution

The fundamental solutions presented above are either impractical in general (as is an explicitly stochastic IFM) or they are practical for consideration in the Renewable Integration Review, but not so for a short-lived interim measure. Thus, if the problem is to be addressed in the interim, a temporary fix would need to be implemented. We suggest the following general criteria for evaluating such temporary market changes:

1. Cost effectiveness (economic efficiency) in addressing the problem.
2. Reversible, so as to avoid constraining the ISO from implementing a fundamentally different approach as a result of the Renewable Integration Review.
3. Simplicity and transparency.
4. Avoidance of large shifts in costs among market parties.
5. Efficient price signals: compensates market parties for costs incurred, and is paid for by market parties responsible for the costs
6. Limited potential for unexpected consequences.
7. Consistency in philosophy with existing market features.

Below we offer our assessment of the flexiramp proposal in light of these criteria.

3.1 Effectiveness

In those RTPD intervals when there would otherwise be insufficient ramp, the formulation of the flexiramp constraint in RTPD will increase the amount of unloaded capacity that could provide energy at short notice. This is anticipated to occur by increasing the amount of committed capacity. We have confirmed this possibility by undertaking our own simplified simulations (without data) of the impacts of imposing the flexiramp constraint in the RTPD and RTD processes, including extensions of the examples presented in the Technical Bulletin.[13]

There are three circumstances we identified in which the flexiramp constraint may be ineffective. First, it is possible that in some circumstances that load (net of wind generation) will be appreciably higher than anticipated in the first interval of RTD, using most or all of the scheduled flexiramp. In that case it may be difficult for RTD to meet flexiramp constraints in subsequent intervals, even though the amounts of flexiramp required in the advisory intervals immediately after the first interval are reduced relative to the RTPD requirement. In these cases, the flexiramp constraint may be unsatisfied. We understand that this did not occur often in the ISO's market simulations of the flexiramp constraint. However, if this situation occurs with some frequency, then adjustments to the amount of flexiramp acquired in RTPD and in advisory intervals of RTD are likely to correct the problem.

The second circumstance is where inadequate capacity is committed in the day-ahead (IFM) and residual unit commitment processes such that there is too little committable capacity available for RTPD, so that less flexiramp can be acquired than is desired by the ISO. Again, this did not appear to be a problem in the market simulations. However, it could arise under conditions other

---

[13]"Flexible Ramping Constraint", Technical Bulletin 2011-02-01, April 19, 2011, www.caiso.com/Documents/TechnicalBulletin-FlexibleRampingConstraint_UpdatedApr19_2011.pdf

than those considered in the simulations. It is also possible that the cost of commitment and energy might be higher than if additional commitment to accommodate flexiramp could take place day-ahead. If either circumstance occurs, it is possible that the problem could be partially solved if relatively high prices for flexiramp in RTPD encourage generators who otherwise would not be available to self-commit and make themselves available.

This possibility would have at least partially been dealt with by the ISO's original proposal, which proposed that flexiramp also be obtained in RUC. However, the ISO believes that this additional complicating feature is not needed to provide sufficient flexiramp at this time.[14] Therefore, we suggest that the sufficiency and cost of supply of flexiramp in RTPD be monitored and if there problems occur frequently that some day-ahead version of the flexiramp constraint be revisited. Meanwhile, we anticipate that the Renewable Integration Review will result in a proposal in which a product that provides flexiramp functionality would be acquired both day-ahead and in real-time, avoiding the latter problem.

The third possibility arises because flexiramp is acquired on a system-wide basis, unlike energy which is acquired nodally or AS which are acquired zonally. A difficulty may arise in that it might be acquired in places where it would be ineffective. As an extreme case, flexiramp might all be scheduled in a generation pocket which has lots of extra generation and ramp precisely because it cannot be dispatched. Locational flexiramp constraints have not been included in the flexiramp design for the sake of simplicity. Therefore, the ISO should monitor where flexiramp is acquired so that if inaccessibility of flexiramp occurs frequently, then appropriate fixes can be implemented relatively quickly.[15]

3.2 Reversibility

The concern here is that the particular flexiramp constraint formulation and settlement procedure would restrict the range of alternatives that would be considered in the Renewable Integration Review, perhaps raising an obstacle to a more fundamental solution. We understand that the Review has been considering several possibilities for providing flexiramp functionality in the ISO markets unrestrained by the interim flexiramp design, most involving definition of a product that

[14]Opportunity Cost of Flexible Ramping Constraint Issue Paper & Straw Proposal, June 24, 2011, www.caiso.com/Documents/Flexible%20ramping%20constraint%20-%20papers%20and%20proposals/OpportunityCost_FlexibleRampingConstraint_IssuePaper_StrawProposal.pdf

[15]A third situation does not involve inadequate amounts of flexiramp, but rather payment might be made in RTPD without any change in the amount of resources or their schedules in binding intervals of RTD. In particular, it is possible that in some intervals, the flexiramp constraint in the first interval of the RTPD constraint might be binding with a positive shadow price, but without changing unit commitment or the distribution of operating reserves among generators. As a result, generator schedules and prices might not be changed in the first interval of the RTD compared to the situation. In that situation, there will be opportunity cost-based payments to generators without changing actual dispatch or commitment. However there is no reliability concern in that situation as the system will have enough rampable capacity in this case. We note however that the ISO's market simulations indicate that in many or most cases when there is a significant payment, there will also be more capacity committed, resulting in more ramp capability..

would be bid for and settled in both day-ahead and real-time markets. Therefore, there does not seem to be a concern with respect to this criterion.

3.3 Simplicity and Transparency

The proposal has two basic aspects. First, it adds sets of constraints to the RTPD and RTD processes to reserve rampable capacity to accommodate unexpected changes in RTD load, and to ensure that ramp rate limitations for generation units are still satisfied should the flexiramp be called upon. Second, it uses the shadow price the binding interval's flexiramp supply constraint in RTPD to determine payments for sources of flexiramp in that interval. Payments are made by load. Market participants are not asked to submit bids, and no changes are made in the day-ahead markets. Therefore, from a mathematical perspective, the flexiramp constraints are the minimum necessary to ensure feasibility, and the payment rule has the appeal of being based upon a single price that is automatically provided by RTPD, which is RTPD's estimate of the marginal cost to the market of the provision of flexiramp.[16]

However, the simple use a shadow price does present transparency problems, in that it can be challenging to understanding how opportunity costs arise any particular interval. Because RTPD schedules energy (non-binding), ancillary services (binding), and flexiramp simultaneously, opportunity costs can arise in very complicated ways, even in simplistic cases. As can often happen in such cases, the price of any one of these commodities can reflect the complex movements, both up and down, of supply of more than one commodity from several generators that are necessary to ensure that capacity, ramp, and transmission constraints remain satisfied.[17] We support the use of opportunity cost calculated in this way not only because of its mathematical convenience, but also because it correctly reflects the physical interactions of how the energy, AS, and flexiramp are supplied.[18] Unfortunately, in complicated electricity markets, it will not always be transparent how shadow prices originate in a particular circumstance, and sometimes unexpected extremes can occur. The examples in the Technical Bulletin[19] are helpful, and additional ones reflecting, for instance, changes in commitment and significant changes in load between RTPD and RTD would be helpful in informing stakeholders.

3.4 Avoidance of Large Shifts in Costs

This is desirable in an interim measure because, in general, an interim measure will involve compromises in market design in the interest of simplicity and reversibility that may result in

---

[16]An alternative payment schemes that might be as easy or easier to understand is to instead have a fixed $/MW/hour payment. But such a payment would violate our other criteria by being arbitrary, unreflective of market conditions, and inconsistent with how energy and ancillary services are paid.

[17] See the Technical Bulletin, op. cit. In one simple three generator simulation we considered with just energy and flexiramp, the price of flexiramp was determined by a complex reshuffling of energy and flexiramp among two different generators that arose because of a binding ramp constraint.

[18] A partial exception is when a product or flexiramp is in shortage in the RTPD scheduling run, so that the prices in that run reflect constraint violation penalties, such that the pricing run of RTPD yield different prices than the scheduling run.

[19] Op. cit..

payments not satisfying the next market design criterion (cost compensation and responsibility). As a result, if large payments are made, they may not be made to those who bear the costs or in proportionate to those costs, and payments may be made by parties who are not responsible for the need for flexiramp. In theory, the amounts of flexiramp to be acquired (hundreds of MW) and the potential for prices in the tens of $/MW/hr means that there is potential for significant cash flows. However, the market simulations conducted by the ISO show zero or very low flexiramp prices in most intervals; if this turns out to be actually the case when flexiramp is implemented, then this criterion will be satisfied. Because the constraints for flexiramp are similar in many respects to those for spinning reserve, while the amount of MWs involved is much less, we anticipate that the aggregate payments over time will be much less than for operating reserves.

If, contrary to expectation, it turns out flexiramp settlement expenses are much higher than anticipated, the relationship of these costs to the actual opportunity costs (in terms of AS and energy rescheduling) should be verified.

3.5 Efficient Price Signals: Cost Compensation and Cost Responsibility

Efficient price signals impose two requirements. First, the prices paid should reflect benefits (price should equal marginal benefit). Second, prices should reflect the costs of what is purchased (price should equal marginal cost). The result in this case, if both requirements were satisfied, would be that the marginal cost of flexiramp would equal the marginal benefit of flexiramp. This would mean the procurement was cost effective (economically efficient). To our knowledge, no estimate has been made of benefits, although we agree with the ISO's view-that the benefits could be substantial. This means that efficient price signals, those equal to the marginal benefit of additional flexiramp to the system, are currently out of the question. However, it is still desirable to achieve supply-side efficiency (minimize the cost to the market) by ensuring that all providers face the same price, and the current proposal takes a useful step in that direction.

As a general market design principle, minimizing the cost to the market of acquiring flexiramp is desirable for an interim solution. However, it may be compromised because of the need for quick implementation, and the cost of implementing more complex designs that would send more efficient prices to the market parties that can provide or reduce the requirements for a commodity. This is the situation with flexiramp. On the supply side, a more efficient design would involve, as noted above, acquisition both day-ahead and in real-time. Also, to ensure cost reflectivity of prices, a bidding procedure would be preferable so that suppliers can reveal their opportunity costs of providing flexiramp. However, in interests of simplicity, day-ahead acquisition and a bidding procedure has been omitted.

One simple step that could contribute to flexiramp payments being more cost reflective is to use the the capacity bid for spinning reserve as the bid price for flexiramp in the objective functions of RTPD and RTD. This idea has appeal because of some similarities between spinning reserve and flexiramp from the generators' point of view. However, there is a key difference: flexiramp is more likely to be used to produce energy in RTD, and so opportunity costs should, on average, be less than for spin, especially non-discretionary spin. There are not obviously other significant costs associated with provision of flexiramp (although more frequent ramping itself may impose costs[20]), so we support the use of a zero bid price for the interim flexiramp process.

We note that there is an aspect of this proposal that, like ancillary services, can result in divergence of payments from actual costs. This largely arises from differences between the RTPD and RTD designs and prices. In particular, the opportunity costs that would be paid to flexiramp in RTPD are based, in part, upon energy prices that are not actually paid; actual energy payments occur in RTD. The lack of convergence, on average, of RTPD and RTD prices in practice means that the actual opportunity cost of flexiramp (if arising from foregone RTD energy revenues) is not paid. Because RTPD prices have tended to be less than RTD prices, this could imply a downward bias for payments. Note, however, that this is true also of other ancillary services acquired in RTPD. On the other hand, unlike contingency-only spinning reserve and regulation, capacity reserved for flexiramp can be dispatched for energy and paid in RTD, and indeed has a reasonably high likelihood of doing so. Therefore the opportunity cost of flexiramp calculated in RTPD may overstate the actual lost opportunity in the ISO's real-time markets.[21,22] This could affect bidding behavior for spin, because capacity scheduled for flexiramp would generally have more opportunity than spin to earn a gross margin from energy revenues in RTD.

However, we note that this possible overstatement of opportunity costs for flexiramp is also analogous to the situation with payments for non-contingent spinning reserves. Non-contingent spin, which is only acquired in the day-ahead market, is paid the shadow price of the spin constraint, but that may overstate opportunity costs since non-contingent spin can, under some conditions, be dispatched in RTD. Therefore the proposal has the advantage of treating flexiramp opportunity costs in a manner broadly consistent with non-contingent spin opportunity costs. The issue of the correct estimation of opportunity costs for ancillary services and how they are compensated should be addressed in the Renewable Integration Review.

On the demand side, identification and allocation of costs to responsible parties, which could include both suppliers and consumers of energy, would in theory be preferable to charging all flexiramp costs to load. One approach would be to charge flexiramp costs to parties with imbalances in RTD. However, this would be imperfect, because it is not imbalances that create the need for flexiramp but large variations of load and generation within the five minute RTD intervals. Another approach would be to identify the relative contribution of load and supply (especially

---

[20] C. Wang, S. M. Shahidehpour, "Effects of Ramp-Rate Limits on Unit Commitment and Economic Dispatch", *IEEE Trans. Power Systems*, Vol. 8, No. 3, 1993.

[21] Some opportunity cost can occur in RTD because the flexiramp constraint in later (nonbinding) intervals of RTD could result in changes in dispatch in the first interval in order to position generators who would be providing flexiramp in nonbinding RTD intervals.

[22] A fundamental fix would be to include an energy settlement in RTPD and then only settle imbalances from that position in RTD, but this would be a large and fundamental departure from the present market design. Another approach would be to estimate the actual opportunity costs incurred in RTD rather than use the RTPD shadow price. This might possibly be estimated based, for instance, on the shadow price of ramp rate constraints that limit energy production in the first (binding) RTD interval for units that provide flexiramp in later (nonbinding) intervals, considering the amount of energy provided in the first RTD interval relative to that scheduled in the corresponding RTPD interval. However, implementation of such a procedure would require significant study and would complicate the proposal.

variable uncontrollable supply) to those fluctuations, and to charge accordingly; however, this is a topic more suited for the Renewable Integration Review. Therefore, in the interest of simplicity, we support charging all flexiramp costs to load, while looking forward to the results of the Review for definition of a more complete and cost-reflective charging mechanism for ramping capability.

3.6 Limited Potential for Unexpected Consequences

Almost by definition, this criterion is difficult to assess, since it requires anticipation of the unanticipated. However, large changes to a market design in both philosophy and implementation details are more likely to yield unexpected problems than circumscribed changes.

The limited duration of the mechanism (until the results of the Renewable Integration Review are implemented) and its limited scope (no bidding mechanism, scheduling only in the real-time market) suggests that unintended consequences are less likely than they would be for more permanent measures involving broader changes to the market. However, we believe that two elements of the flexiramp design that appear at present to have the potential for troublesome unintended consequences. One is that setting payments based on an opportunity cost in RTPD that is due to a constraint that will not be enforced in the same way in RTD; as we pointed out, the actual opportunity costs experienced by resources devoted to flexiramp will differ from (and likely be less than) the payments made, even on average. The second is the impact of flexiramp opportunity cost payments on the day-ahead spinning reserve market; it is possible spinning reserve prices might be higher in RTPD than in the day-ahead market due to the presence of the flexiramp constraint in the former but not in the latter. Hence it would be desirable for the California ISO both to evaluate in testing the practical potential for such unintended consequences and to also carefully monitor this aspect of flexiramp performance when it is in operation.

3.7 Consistency in Philosophy with Existing Market Features

In the interest of simplicity, the flexiramp process has not been proposed as a full-fledged product with bidding and both day-ahead and real-time markets. Thus, it differs in philosophy with the processes for acquiring and paying for ancillary services, which we believe is undesirable in the long run. However, because it is interim, we accept that the need for simplicity in the proposal. We look forward to participating in the Renewable Integration Review where a more fundamental examination of the need for and alternatives for acquisition of ramp products is underway.

4. Conclusion

We conclude that the flexiramp proposal is likely to be an effective and reversible measure to address a need for more upward ramp capability in the real-time ISO market. The ability to tune the parameters of the flexiramp constraints, including the RTPD requirements and the amounts acquired in nonbinding intervals of RTD, give considerable flexibility to respond to unforeseen problems. We look forward to more detailed reporting of market simulations designed to confirm the anticipated effectiveness and price impacts of the flexiramp proposal.

However, the flexiramp proposal only partially addresses the fundamental reasons for the inadequate amount of upward ramping capacity. Furthermore, the payment based on opportunity costs in RTPD may overstate actual opportunity costs that providers of flexiramp will experience in the ISO's real-time markets. Therefore, we anticipate that any ramping product design that emerges from the ISO's Renewable Integration Review will likely have a very different structure and settlement procedure. That Review will need address the issues we have identified concerning appropriate incentives for short- and long-term provision of upward ramp and consistency with other products in the ISO's day-ahead and real-time markets. Among these issues are whether and when to pay capacity payments for ramp and (other) ancillary services; how they interact in the operational constraints in the day-ahead and real-time markets; assignment of responsibility for the costs of ancillary services; and the design of scarcity payments.