Western Power Trading Forum (WPTF) Comments on Parameter Tuning Issues

June 20, 2008

Contact: Ellen Wolfe 916 791-4533

WPTF appreciates the opportunity to offer comments on the CAISO's revised proposal for Parameter Tuning policies.

Please reference comments previously submitted by WTPF on the CAISO's initial Parameter Tuning white paper (attached below as Attachment 1), as those comments continue to be relevant This set of comments amplifies on WPTF's earlier remarks and provide additional clarification.

1. WPTF believes it is of the utmost importance to have transparent process around uneconomic adjustments, and we urge the CAISO to provide further details about the information that will be available for uneconomic adjustments. For example, WPTF believes a market participant should be able to determine when uneconomic adjustments have occurred, what the relevant constraints were, and what the outcome of the uneconomic adjustment was.

2. WPTF continues to believe that the uneconomic adjustments in both the scheduling runs and pricing runs will affect rates, terms and conditions of service, and therefore they should be part of the FERC-approved tariff. Certainly, the CAISO must, at a minimum, indicate in the tariff how the parameters will be set and changed from time to time, and the CAISO should include all the information and processes related to uneconomic adjustments (for both the scheduling runs and the pricing runs) in a BPM. As such, the CAISO proposal to include pricing run parameters in a BPM, but to include the scheduling run parameters only in an Operating Procedures should – at a minimum - be modified to have both sets of information included in a BPM. Simply, the CAISO should rely on the BPM change management process for which it so strongly advocated so that both the pricing run parameters and the scheduling run parameters are managed with equal rigor, transparency and market participant input.

3. The CAISO indicated that it planned to use a $5000 shadow price for constraints, and it further indicated during the last call that this parameter has been selected because it is equivalent to a 10% effectiveness factor that operators typically use to adjust dispatch to manage constraints, given that under MRTU a $500 bid cap will apply initially. WPTF requests that the CAISO provide further information about the operational practices

upon which the market design experts are basing this value, and we would like to understand how those past practices are relevant for this parameter setting activity for MRTU.

WPTF continues to believe a 5% effectiveness factor is more appropriate and believes this is similar to practices in both the MISO and PJM markets, as submitted in our initial comments. The CAISO challenged WPTF to support its comments that other markets use an effectiveness threshold lower than 10%. WPTF believes that MISO uses an effectiveness threshold of 4% to 6%, but we were unable to locate specific references to these levels. PJM does cite lower values in its manuals as follows.

> Generation Redispatch
>
> PJM, prior to initiating redispatch, reviews available controlling actions and the distribution factor effect on the overloaded facility. PJM also considers whether there are sufficient resources available to control transmission facilities within acceptable limits.
>
> 1. Contingency Operations
>
> PJM will initiate off-cost if reasonable controlling actions are available with an impact effect generally greater than 5%. Once off-cost is initiated, UDS tools will redispatch generation based on dollar per MW effect, considering all on-line flexible units with an impact of 1% or greater. PJM staff has the ability to adjust the controlling percentage on an individual constraint basis. PJM will initiate a Post Contingency Local Load Relief Warning/Action if post-contingency flows exceed designated ratings and insufficient resources are available to control the overloaded facilities.
>
> 2. Normal / Actual Overload
>
> In general PJM initiates off-cost and utilizes controlling actions greater than 5% impact, however, since an actual overload causes real-time loss-of-life on the affected facility, PJM will load generation with an impact effect less than 5%. Once off-cost is initiated, the UDS tool will redispatch generation based on dollar per MW effect, considering all on-line flexible units with an impact of 1% or greater. PJM staff has the ability to adjust the controlling percentage on an individual constraint basis.
>
> The UDS software continues to monitor projected flows on constrained facilities and sends ramp-limited set points to re-optimize redispatch for constraint control to the designated threshold. The eligibility of units to set Locational Marginal Price is determined by comparing the desired output as calculated by UDS to the actual output as calculated by the State Estimator. (Manual 12, Attachment B, page 78.)

Lastly, with respect to setting a shadow price as opposed to using an effectiveness, it seems fitting that since the dollar shadow price is primarily derived from an underlying effectiveness factor it should adjust in proportion to the increases in bid caps over time.

For instance, while the underlying bid cap is $500, the CAISO's proposed effectiveness factor of 10% equates to a $5000 shadow price and WPTF's proposed effectiveness factor of 5% equates to a $10,000 shadow price. However, at a $1000 bid cap the equivalent shadow price would be $10,000 and the equivalent WPTF shadow price would be $20,000. If the CAISO does not take this price, and specifies a fixed shadow price for uneconomic relaxation of transmission constraints, then the equivalent effectiveness factor would essentially increase as the bid cap goes up.

4. WPTF continues to encourage the CAISO to consider increasing the (negative) bid floor. The CAISO has not provided any specific examples or explained how its concerns about intrazonal congestion management in the pre-MRTU market design will continue to be relevant under MRTU. Adjusting the bid floor may eliminate the need for uneconomic adjustments in the decremental direction; this is the opportune time to align the markets and avoid/minimize minimize instances where negative penalty prices would be needed. The CAISO should take the time to fully explore this topic and its ramifications.

5. WPTF has recently become aware of a proposal in ERCOT that may avoid many unintended outcomes of uneconomic adjustments. In ERCOT stakeholders and the ERCOT market designers are contemplating an approach to penalty prices that seems to avoid the adverse impacts of the discontinuities that result from using discrete penalty prices by employing a formulaic approach to setting penalty prices. A paper presenting this approach was distributed to ERCOT stakeholders and is included as Attachment 2. Though the paper includes some mathematical formulations and is focused on the ERCOT market, the underlying concepts appear to be very applicable to the process the CAISO is designing through parameter tuning. The paper's author, Dr. Ross Baldick, points out that discrete penalty prices often produce adverse impacts given the discontinuities or "lumpiness" of the applied penalty prices. Using discrete prices means a one MW change in a constraint or schedule can produce price impacts in the hundreds of dollars, create inappropriate incentives and lead the CAISO to make difficult and often costly tradeoffs between violating constraints and paying more extreme prices. Baldick offers an alternative approach: "A natural solution to avoiding the compromise … is to use a quadratic or a piecewise linear penalty that enables the penalty function to be more closely tailored to the amount of violation", for example a formulaic penalty that increases in magnitude as the extensiveness of the violation increases. This seems to be an innovative and rational approach that should be given some focused consideration by the CAISO before "locking in its design" in a way that would likely produce adverse and unintended outcomes.

6. Finally, WPTF reiterates that the CAISO should focus more directly on the extent to which self-scheduling is occurring in its markets, what ramifications that self-scheduling has for its market outcomes, and what incentives and safeguards are in place to ensure that self-scheduling does not unduly or inappropriately affect market outcomes. In this regard, the CAISO proposal to cap the pricing run to the existing bid cap provides no incentive for entities to limit self-scheduling, because an entity that has self-scheduled faces no risk that it will pay more than the bid cap in the event that overall self scheduling creates dispatch and reliability issues. While self-scheduling may be a valuable tool for market participants to use to manage their portfolios, self-scheduling by entities that own significant amounts of resources, especially when those resources receive out of market revenues through regulated rate mechanisms, can have a significant impact on market outcomes. Thus, WPTF urges the CAISO to undertake a comprehensive review of how self scheduling is being utilized, the extent to which is being utilized, and whether the levels of self scheduling the CAISO markets are similar to those experienced in other organized markets. With such analysis in hand, the CAISO could determine how its proposal here could be modified to provide direct incentives to limit self scheduling so that uneconomic adjustments can be avoided, including modifications to the pricing run caps so that they are more in line with the scheduling run parameters.

While WPTF recognizes that the CAISO has initiated a very compressed process, the CAISO should develop a strategy to follow through on its goal of striving for “best in class” market designs.

Attachment 1

Western Power Trading Forum (WPTF) Comments on Parameter Tuning Issues

May 23, 2008

Contact: Ellen Wolfe 916 791-4533

WPTF offers a number of comments including some guiding principles followed by some more specific comments for creating and revising “penalty price” parameters.

- First and foremost, WPTF members believe that in order to obtain optimal outcomes from the CAISO’s markets, parties need to submit economic bids with their schedules. Self-Schedules frustrate the CAISO’s efforts to obtain optimal outcomes. They require the CAISO to administratively set “penalty prices” when otherwise inflexible Self-Schedules must be adjusted. The CAISO should ensure that its choice of penalty prices, the process it uses to set and revise them, and the impacts of penalty prices on the LMPs encourage parties, in the first instance, to submit economic bids and in the second, to avoid economic harm to parties that offer to be dispatched economically.
- Penalty prices should not be set in a way that would – or could -- allow the targeted Self-Schedule to clear at a price less than the cap for load, or above the floor for supply (WPTF provides added guidance below regarding the floor). In this manner, highly-effective Self-Schedules will be priced in a manner which will reduce the potential for very high LMPs while not unduly dampening LMP prices for economically bid units.
- WPTF feels strongly that the tuning of the parameters will have a direct and measurable impact on the price clearing process. We believe that FERC must, pursuant to its obligation to ensure just and reasonable prices, have the ability to review and approve the creation of, and any revisions to, penalty prices. Therefore, WPTF believes that penalty prices and any tuning thereof, must be a part of the FERC-approved tariff and

certainly should not be delegated to an Operating Procedure. And certainly specific rules should be possible to specify regarding the values[1] and place within the tariff.

- From a process perspective, determining how to clear the market when sufficient economic bids are not available is very complex, and involves a number of difficult issues. A meaningful stakeholder process that thoroughly examines all of these issues will require a commitment by the CAISO to freely share its view of the results, concerns and unintended consequences from the market simulation and its related testing of the tuning parameters. Finally, the final rules should promote transparency to the maximum extent possible.
- WPTF is concerned about the potential incentives any specific parameter values may have on bidding and scheduling behavior that would lead to adverse market outcomes. WPTF requests that any perverse incentives identified by the CAISO be shared with market participants for further consideration so that they can inform stakeholder positions on the parameter design. For example, the choice of parameters and thresholds should not allow parties that submit Self-Schedules to influence LMPs to the disadvantage of parties that submit economic bids.
- WPTF believes that the any proposal to modify the current tariff to allow out of merit, uneconomic adjustments before all economic bids have been exhausted creates a possibility of unwarranted market intervention and inappropriate market price suppression, and therefore the setting of the parameters should protect against this.
- WPTF believes a threshold of 10% for relaxation of a constraint is too high. Other markets have used thresholds on the order of 5% or 6%. Moreover, WPTF advocates that the CAISO employ a shadow price-based threshold for relaxation of transmission constraints as is done in some of the other ISO markets. This would allow both consideration of effectiveness and the cost. For example a shadow price threshold of $10,000 would reflect an effectiveness of 5% at $500/MWh or an effectiveness of 1% if

[1] For example, "penalty prices will be adjusted in a specified way if effectiveness factors are less than some threshold (e.g. 5%) and cause shadow prices of greater than some amount (e.g. $1500/MWh)."

bid prices were $100/MWh. We believe such a shadow price threshold is more appropriate than an effectiveness threshold, especially when price differential between the distant node and the local node is small.

- Once a shadow price threshold is agreed upon, WPTF recommends this value become the basis for the penalty prices used by the CAISO to relax self-schedules such that all the penalty prices have a common fundamental basis.
- WPTF is also very concerned about the proposed level of the bid floor, its perverse impact on SCs' willingness to submit economic bids during low load conditions, and the interaction with penalty prices that will result. There may be a significant pool of bidders interested in submitting economic bids that are below the minus $30/MWh floor but these bidders may instead submit fixed schedules to avoid the economic harm they would otherwise suffer if they were economically dispatched based on the bid floor.[2] Instead, with current rules, at a minus $30/MWh floor all such economic bids are deemed uneconomic, and the true economic bidders are not fully cleared before the CAISO makes its "uneconomic" adjustments. WPTF suggests that the CAISO consider setting the magnitude of the negative bid penalty price equal to the magnitude of the positive cap (but with the opposite sign), as this would allow the market to manage reliability most economically and provide appropriate incentives to submit economic bids that result in efficient dispatch. This approach would also allow demand to help manage over generation conditions by being paid to receive energy. Alternatively, maintaining a minus $30/MWh bid floor could significantly increase the number of "uneconomic" adjustments and it would - among other things - increase the impact of any distortions in the penalty prices.
- While WPTF is not proposing specific scheduling run penalty prices at this time, WPTF strongly believes pricing run values must produce pricing results and a system dispatch that are consistent with system dispatch results from the scheduling run. WPTF is

[2] For example, during the spring when a combination of low off-peak demand, high hydro conditions, and high wind energy production could cause generation to exceed demand by a sizable margin generators with commitments for the following day may not have an incentive to reduce their output.

concerned that the CAISO's dispatch will be inefficient and LMPs will be distorted if the penalty values used for the scheduling run are substantially different from the penalty prices that are used for pricing and settlement. In fact, LECG apparently indicated some concerns about this very issue when they recommended using fixed LDFs for LAP clearing. WPTF is concerned that the CAISO's proposal could otherwise result in distorted prices, excessive use of uneconomic adjustments, little incentive to reduce Self-Schedules and instead provide economic bids, and undermine confidence in and the robustness of the CAISO's markets.

- WPTF believes that Convergence Bidding offers a very complementary set of functionality that SCs can use to protect themselves from the effects of imperfectly chosen penalty prices and from attempts to manipulate prices using self-schedules. In the Day-Ahead market SCs who believe the parameters could create distortions would be free to submit convergence bids. Though the convergence bids will be "backed out" in real time, nevertheless in real time the CAISO's load forecast is used to set prices. WPTF continues to encourage the CAISO to implement convergence bidding as soon as possible.

Attachment 2

ERCOT Market Stakeholder-Submitted White Paper

Discussion and Proposal for Maximum Shadow Price Methodology

Ross Baldick
Updated Draft
May 29, 2008
Prepared for CPA

1 Introduction and summary

This paper discusses the proposal entitled "Maximum Shadow Price Methodology," by Resmi Surendran of ERCOT, May 2008. In that proposal, penalty prices are used to relax constraints in the SCED when conditions of high demand or tight supply would otherwise lead to very high prices or violated transmission constraints.

It is not explicitly stated in "Maximum Shadow Price Methodology," but there are at least three roles played by the penalty prices that are to be used to relax constraints in SCED:

1. Encode operator practice to not be greatly concerned about a few MW violation of, for example, a thermal contingency constraint that would not lead to cascading outages. That is, small violations of such constraints should not result in LMPs that are significantly above the highest offer prices in the market.
2. Discourage, to the extent possible, violation of constraints that would lead to cascading outages, even if that necessitates rolling blackouts. That is, even small violations of such constraints should be avoided.
3. Ensure that SCED can find a dispatch under all circumstances, even if that solution indeed violates some constraints. That is, all constraints can be violated *in extremis*.

Drawbacks of "Maximum Shadow Price Method" will be discussed from the perspective of these three roles using a simple example. The basic conclusion is that the proposed approach in "Maximum Shadow Price Method" cannot satisfactorily accomplish these three roles. Then an alternative is proposed that allows for these roles to be satisfactorily achieved through a more general form of penalty function.

2 Discussion of "Maximum Shadow Price Methodology"

In "Maximum Shadow Price Methodology," values of penalty prices, the "Maximum Shadow Prices," are proposed for:

- violating transmission constraints, and
- violating the energy balance constraint

in the SCED operation. The value of these penalty prices are keyed to observations about preferences for deploying generation, and are differentiated by voltage class.

As a threshold matter, it is not stated in "Maximum Shadow Price Methodology" as to whether the penalty prices are to be incorporated into the LMPs or whether the LMPs are to be "decontaminated" by not incorporating the penalty prices. The choice has a significant effect on incentives, but the choice is not specified and is not discussed in "Maximum Shadow Price Methodology." In Section 6.5.7.1.11(2) of the Nodal Protocols,[3] it says that:

> "ERCOT shall establish a maximum Shadow Price for each constraint as part of the definition of contingencies. The cost calculated by SCED to resolve an additional MW of congestion on the constraint is limited to the maximum Shadow Price for the constraint. ERCOT shall develop a policy for setting maximum Shadow Prices for approval through the PRR process."

Although Section 6.5.7.1.11(2) of the Protocols does not explicitly discuss the implications for LMPs, it is suggestive that the intent is to incorporate the penalty prices into the LMPs. Given that the penalty price for energy balance is very large, the implication is that prices might also be very large. Moreover, even though the intent is apparently to apply these penalty prices to the real-time market *only* through the operation of the SCED, it is important to realize that asymmetries between the real-time and day-ahead market design will potentially provide inefficient arbitrage opportunities. The discussion here will focus only on the SCED and not on arbitrage opportunities.

2.1 Basic assumptions and terminology

We assume that an offer is associated with each generator. As per Section 6.5.7.3(3) of the Nodal Protocols, some of the specification of a generator offer may be implicit if the explicit offer does not include the full range of generation from LSL to HSL. In that case, for offer quantities between the maximum explicitly offered quantity and the HSL, the offer is specified for SCED to be at or close to the system-wide offer cap. Consequently, in keeping with the context of high demand and/or violated transmission constraints, we will consider that some offer prices are high or equal to the system-wide offer cap, which will be denoted by P. For the purposes of concrete examples, we will assume that P = \$2500/MWh.

The value of lost load will be denoted by V and it is assumed that $P \leq V$. For purposes of concrete examples, we will assume that $V \approx$ \$5,000/MWh to \$10,000/MWh, so that $P < V$. From a welfare maximization perspective, it is inappropriate to serve generic load (that is, excluding particular high-value loads such as hospitals) if the marginal cost to meet demand exceeds V. That is, rolling blackouts should be instituted once LMPs would exceed V since, by definition, the cost of serving incremental load exceeds the value of its consumption. The connection between V and the choice to institute rolling blackouts is not discussed explicitly in "Maximum Shadow Price Methodology."

Conceptually, the "offer cost," or indefinite integral of each offer, is used as the term in the objective corresponding to the offer. We use the symbol c with an appropriate subscript to denote the offer costs.

[3] The Nodal Protocols are available from http://nodal.ercot.com/protocols/index.html.

We assume that demand is specified (or forecast) rather than bid, although the extension to demand bids is theoretically straightforward.

To model the proposed use of penalty prices by ERCOT, we will relax each constraint using a violation or "surplus" variable, with each such variable denoted by the symbol v having an appropriate subscript. When the surplus variable is non-zero, we say that the "unrelaxed constraint is violated."

To represent each relaxed constraint, the cost function in the model of SCED will include a penalty function that is equal to the product of:

- the penalty price for violation, β, with an appropriate subscript to refer to the constraint, multiplied by
- the corresponding violation variable.

(In the proposal in Section 3, we will generalize this approach to include more general penalty functions.)

The model of SCED is offer-based transmission-constrained economic dispatch, with constraints relaxed by the penalty functions. That is, SCED:

- minimizes an objective specified by the sum of:
  - offer costs, plus
  - penalty functions for all constraints,
- subject to the relaxed constraints.

2.2 Example

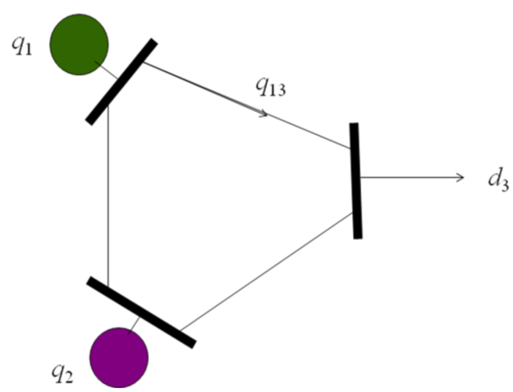

Figure 1. Three bus three line example system.

Consider the simple three bus three line example system shown in Figure 1. There are generators at buses 1 and 2 and demand at bus 3. The generations at buses 1 and 2 are, respectively, q_1 and q_2. The offer by the generator at bus 1 is \$300/MWh for a quantity q_1 of zero to 100MW and equal to P = \$2500/MWh for q_1 from 100 to 110MW. The corresponding offer cost $c_1(q_1)$ (in \$/h) is piecewise linear with:

- slope \$300/MWh for q_1 between zero and 100MW, and
- slope P = \$2500/MWh for q_1 between 100 to 110MW.

The offer by the generator at bus 2 is \$100/MWh for a quantity q_2 of zero to 100MW and equal to P = \$2500/MWh for q_2 100 to 110MW. The corresponding offer cost $c_2(q_2)$ (in \$/h) is piecewise linear with:

- slope \$100/MWh for q_2 between zero and 100MW, and
- slope P = \$2500/MWh for q_2 between 100 to 110MW.

The demand at bus 3 is d_3 = 205MW. Writing v_3 for violation of the power balance constraint and using a penalty cost of β_3 for violation of the power balance constraint, we formulate the power balance constraint as $q_1 + q_2 - d_3 + v_3 = 0$ and include a penalty function in the objective that is equal to $\beta_3 v_3$.

Assume that all line impedances are equal and that there is a flow limit on the line from bus 1 to bus 3 of Q_{13} = 100MW. We initially assume that violation of this constraint would not lead to cascading outages. Using the DC power flow approximation, the flow on the line joining buses 1 and 3 is $q_{13} = (2/3)q_1 + (1/3)q_2$. (We have implicitly made bus 3 the price reference bus.) Writing v_{13} for violation of the line flow constraint and using a penalty cost of β_{13} for violation of the flow limit constraint, we formulate the line limit constraint as $(2/3)q_1 + (1/3)q_2 - v_{13} \leq Q_{13}$ and include a penalty function in the objective that is equal to $\beta_{13} v_{13}$.

"Maximum Shadow Price Methodology" proposes that β_3 = \$100,000/MWh and we will initially consider this value.[4] We observe that β_3 is much larger than the value of lost load of $V \approx$ \$10,000/MWh. All other penalty costs for non-cascading outages proposed in "Maximum Shadow Price Methodology" are considerably lower than V. In particular, the range of values of penalty prices proposed in "Maximum Shadow Price Methodology" for constraints associated with non-cascading outages are all between:

- the value \$2,500/MWh that we have assumed for the system-wide offer cap, P, and
- the range of values \$5,000/MWh to \$10,000/MWh that we have assumed for the value of lost load, V.

For concreteness, we will consider β_{13} = \$3,000/MWh, so that, consistent with "Maximum Shadow Price Methodology, for the values we have specified, it is the case that $P < \beta_{13} < V$.

2.3 Formulation and solution of example

The example offer-based economic dispatch problem, including the relaxation of constraints with penalty functions is:

$$\min_{q1,\, q2,\, v13,\, v3}\{c_1(q_1) + c_2(q_2) + \beta_3 v_3 + \beta_{13} v_{13} \mid q_1 + q_2 - d_3 + v_3 = 0, \\ (2/3)q_1 + (1/3)q_2 - v_{13} \leq Q_{13}\}.$$

We use the symbol λ_3 for the Lagrange multiplier on power balance and the symbol μ_{13} for the Lagrange multiplier on the line constraint in the solution of this problem. Because of the use of penalty functions, the values of these Lagrange multipliers can never exceed the penalty prices β_3 and β_{13}, respectively. Moreover, whenever an unrelaxed constraint is violated then the Lagrange multiplier on that constraint equals the corresponding penalty price.

[4] "Maximum Shadow Price Methodology" actually proposes \$100,000/MW, but I interpret this as meaning \$100,000/MWh.

Putting aside any "decontamination" of the LMPs, the values of the Lagrange multipliers together with the shift factors determine the LMPs.[5] In particular, the LMPs at the three buses, p_1, p_2, p_3, are given by:

$$p_1 = \lambda_3 - (2/3)\mu_{13},$$
$$p_2 = \lambda_3 - (1/3)\mu_{13},$$
$$p_3 = \lambda_3.$$

For reference in Section 3, we observe that the LMP p_k at bus k is equal to the difference between:

- the Lagrange multiplier on the power balance constraint, minus
- a weighted sum of the Lagrange multipliers on the line constraints.

The weights in the weighted sum are the shift factors to the corresponding lines for injection at bus k and withdrawal at the reference bus. The shift factors have values that are less than one in absolute value. Moreover, unless there are two constrained lines "in series," the sum of the shift factors for any bus k will be less than one.

We now solve for the optimal transmission-constrained dispatch and for the Lagrange multipliers. We first consider the dispatch. Note that if it were the case that q_1 = 100MW and q_2 = 100MW in the optimal transmission-constrained dispatch, then the line flow would just be at its limit. Since this amount of generation does not quite meet the demand of d_3 = 205MW, the offer-based transmission-constrained economic dispatch solution will involve violating one or other of the unrelaxed constraints.

Using the very large penalty cost β_3 = \$100,000/MWh proposed in "Maximum Shadow Price Methodology" will ensure that another constraint besides power balance is violated before power balance is violated. This is presumably the explicit intent of using such a large value; however, it does not satisfy the principle articulated above that load should be curtailed if the marginal cost to serve the load exceeds the value of lost load.[6] We will consider the implications of this choice further in the context of pricing.

To serve the demand of d_3 = 205MW, the unrelaxed line flow constraint will be violated. In particular, given the offer costs, offer-based minimum cost dispatch would result in:

$$q_1 = 100\text{MW},$$
$$q_2 = 105\text{MW},$$

[5] For details see course notes for EE394V, Restructured Electricity Markets: Locational Marginal Pricing, available from http://users.ece.utexas.edu/~baldick/classes/394V/EE394V.html.

[6] It is possible that the intent of the high value of β_3 = \$100,000/MWh is to implicitly represent an operator action such as reserves being used to satisfy the power balance constraint. In that case, however, the cost of using reserves (and of, therefore, violating the reserve constraints) should be considered explicitly in the formulation along the lines of restoring section 6.5.7.3(6) of the Nodal Protocols that was removed under NPRR051. Since this is not explicitly stated in "Maximum Shadow Price Methodology" and, consequently, there are no stated penalty costs for violating reserve constraints, I will only consider here the power balance constraint and value of penalty cost as proposed in "Maximum Shadow Price Methodology." In the proposal in Section 2, however, there will be discussion of representation of use of reserves to satisfy power balance.

λ_3 = \$3,500/MWh,
$\mu_{13} = \beta_{13}$ = \$3,000/MWh,
$p_1 = \lambda_3 - (2/3)\mu_{13}$ = \$1,500/MWh,
$p_2 = \lambda_3 - (1/3)\mu_{13}$ = \$2,500/MWh,
$p_3 = \lambda_3$ = \$3,500/MWh.

It is important to recognize that the LMPs as specified provide incentives to each market participant that are consistent with the underlying economics. In particular, the generator at bus 2 is "marginal." It is paid at its offer price (for dispatch of q_2 = 105MW) of \$2500/MWh. Generator 1 is paid above its offer price (for dispatch of q_1 = 100MW), but the price is not enough to make it want to generate more than 100MW since its offer price for quantities greater than 100MW is \$2500/MWh. Finally, the demand pays \$3,500/MWh, which is the marginal price to serve additional demand at bus 3 based on the assumption that violating the line flow constraint in fact incurs a cost of β_{13} = \$3,000/MWh. (To see this, note that to deliver an additional 1MW of demand to bus 3 would require:

- increasing production at generator 1 by 1MW, costing \$2,500/h, and
- violating the line limit constraint by an additional (1/3)MW, costing (1/3)×\$3,000/h,

totaling \$3,500/h for the 1MW, or a marginal cost of \$3,500/MWh.)

The flow on the line is 101.67MW, just a little above the limit, so the violation variable is v_{13} = 1.67MW. Nevertheless, the LMP at demand bus 3 is significantly, by \$1,000/MWh, above the highest offer price of \$2,500/MWh. The choice of penalty prices therefore does not satisfy the first role for penalty prices as articulated in Section 1.

2.4 Decontaminating LMPs

"Decontamination" of LMPs is not discussed in "Maximum Shadow Price Methodology;" however, the potential for LMPs that are significantly higher than the highest offer price may prompt a desire by market participants to avoid reflecting the penalty prices into the LMPs. In this section, I will consider the most straightforward way of doing this, which is to:

- use the dispatch corresponding to the solution with penalty prices, but
- set μ_{13} = \$0/MWh in the LMP *calculation*.

I will show that such "decontamination" will result in incorrect incentives to market participants and is therefore inappropriate.

Setting μ_{13} = \$0/MWh in the LMP calculation and assuming no change to the other Lagrange multiplier of λ_3 = \$3,500/MWh would yield the "decontaminated LMPs" of:

$p_1 = \lambda_3 - (2/3)\mu_{13}$ = \$3,500/MWh,
$p_2 = \lambda_3 - (1/3)\mu_{13}$ = \$3,500/MWh,
$p_3 = \lambda_3$ = \$3,500/MWh.

In this case, the effect is to *increase* prices at all buses. Moreover, the incentives are now incorrect: both generators at buses 1 and 2 would prefer to generate more given the LMPs.

The particular change to prices with decontamination is due the choice of the price reference bus, which is bus 3 in the example. Different choices of price reference bus would result in different decontaminated prices:[7]

1. If bus 1 were the price reference bus then the decontaminated LMPs would all be $1,500/MWh. In this case, the incentives would again be incorrect since the generator at bus 2 would be unwilling to produce 105MW at this price.
2. If bus 2 were the price reference bus then the decontaminated LMPs would all be $2,500/MWh. In this case, the incentives would again be incorrect since the generator at bus 1 would want to generate more than 100MW at this price.

To summarize, decontaminating the LMPs will result in incorrect incentives to market participants. Although it may be tempting to carry out some decontamination of the LMPs, this would be an inappropriate policy since it will result in incorrect incentives to market participants.

2.5 LMPs including penalty prices.

The above discussion indicates that removing the effect of the penalty costs through decontamination is unsatisfactory. The reason is that the incentives for generators (and for demand) will then be incorrect with the decontaminated prices.

However, as mentioned in Section 2.3, pricing using LMPs that include penalty costs that are on the order of thousands of dollars per MWh is also somewhat unsatisfactory, at least for small violation of thermal constraints associated with non-cascading outages. The reason is that the actual cost of violating such a constraint by a *small* amount as in the example is relatively small because the exact rating of the line is not exactly certain. For example, for a pre-contingency thermal constraint on a transmission line, the cost is related to the *slightly* increased likelihood of the limiting element sagging into a tree or other structure. For a contingency thermal constraint, the cost is related to the slightly increased likelihood of the limiting element sagging into a tree or other structure multiplied by the probability of the contingency actually occurring. There may also some degradation of the lifetime of the elements. In both cases, these costs are relatively small for a small overload and, in any case, the limit is somewhat uncertain. In the example, there was a 1.67MW overload on a 100MW line. Given uncertainties in ambient temperature (unless dynamically measured) and given uncertainties in sags of transmission spans, it is unlikely that the thermal capacity is precisely known to within 1.67MW in 100MW. Consequently, the underlying economic considerations do not support a large increase in prices for a slight overload of this type.[8]

To summarize, there is uncertainty in the specification of thermal limits. This implies that there should not be a sudden jump from zero to a large penalty when a flow *just* exceeds a particular

[7] Of course, the choice of price reference bus does not affect the LMPs when decontamination is not carried out. The issue here is that decontamination is, of necessity, zeroing out some of the Lagrange multipliers, and the choice of price reference bus does affect the values of these Lagrange multipliers and consequently affects the decontaminated prices.

[8] For comparison, if the line capacity increased to 102MW then the prices would be $2,500/MWh at all buses.

established limit.   Using large penalties for even a small overload is not consistent with the underlying costs and violates the first role of penalty prices discussed in Section 1.

2.6 Constraints associated with cascading outages.

A different drawback of "Maximum Shadow Price Methodology" occurs for constraints associated with cascading outages. In particular, for such constraints it is imperative that there be *no* violation. That is, the operator should and must curtail demand in preference to violating such constraints. The proposed penalty price in "Maximum Shadow Price Methodology" for such cascading outage constraints is around $6,000/MWh. However, "Maximum Shadow Price Methodology" proposes an extremely high penalty price for power balance of β_3 = $100,000/MWh. This extremely high penalty price for power balance will typically result in the cascading outage constraint being violated before violation of power balance.

To summarize, the proposed penalty prices in "Maximum Shadow Price Methodology" do not satisfy the second role of penalty prices discussed in Section 1. Similarly, for large violations of a thermal constraint, it is appropriate to not violate the constraint. The penalty prices on such constraints should be appropriately above the penalty price on power balance so that such constraints are not violated, except *in extremis* as a computational technique that enables SCED to find *some* dispatch that is near to feasible.

3 Alternative proposal

3.1 Summary of observations

The selection of a *single* penalty price for constraint violation is a trade-off between:

- making the penalty price high enough to avoid violation of the constraints when generation offers are available to satisfy the constraint, but
- making the penalty price low enough to comport with the observation that for small violations of non-cascading constraints, there is only a very small cost incurred by violating the constraint.

Unfortunately, this trade-off is inherently difficult to make because prices in electricity markets would be volatile even in the absence of, for example, ramp-rate and transmission constraints. That is, "high" and "low" are relative concepts that vary over time.

Moreover, the trade-off forces a single parameter to play two distinct roles. To understand these roles, consider again a thermal transmission flow limit such as in the example system in Figure 1. As discussed above, typical thermal flow limits have some uncertainty and, furthermore, the limits can be exceeded by a small amount without significant damage to lines or significantly increased risk of conductors sagging. Consequently, for a small violation above the stated limit, a large penalty factor is not appropriate. On the other hand, a large violation of a limit is unlikely to be acceptable: in this case a small penalty factor is *not* appropriate. This alternative proposal is aimed at resolving this issue by the simple expedient of generalizing the notion of a penalty function to having more than one parameter.

3.2 A more general penalty function

A natural solution to avoiding the compromise described in Section 3.1 is to use a quadratic or a piecewise linear penalty that enables the penalty function to be more closely tailored to the amount of violation. The two parameters in a quadratic function (or potentially more than two parameters in a piecewise linear penalty function) can be targeted at each of the three specific roles of a penalty function described in Section 1.

For example, again consider a thermal flow limit associated with a non-cascading outage. Instead of a penalty function for violation that had a fixed slope as specified by the penalty price, a quadratic penalty could have a modest initial slope for small violations but the slope could then increase rapidly as violation became significant. Consider a transmission element that is rated as follows:

- 100 MW rating for steady-state operation,
- 110 MW for 30 minutes, and
- 120 MW for 10 minutes.

For this line, a small violation of a few MW is not of great concern; however, a violation of over 10 MW is of considerable concern.

Consider a pre-contingency limit associated with this line. Let the flow on the line be f. The violation of the flow limit constraint is $v = f - 100$. A possible choice of *quadratic* penalty function $c(v)$ associated with the violation v would be:

$$c(v) = \begin{cases} 0, \quad \text{if} \quad v \leq 0, \\ v \times 500 + (v)^2 \times 100, \quad \text{if} \quad v > 0. \end{cases}$$

(A piecewise linear function could also be used to approximate this behavior in a market that did not support quadratic costs.) There are two parameters in the specification of this penalty function:

1. linear term that specifies the initial slope of the penalty function, in this case \$500/MWh, and

2. the quadratic coefficient, in this case $100/(MW)^{2}h.

With this penalty function, flows below 100 MW contribute a zero penalty because such a flow is below the limit and so there is no violation of the unrelaxed flow limit constraint. Flows just above 100MW contribute costs that are specified by the linear term of the penalty function. That is, for small violations, the penalty function is similar to a *conventional* penalty price of $500/MWh as described in "Maximum Shadow Price Methodology." That is, the coefficient of the linear term in the penalty function can be tailored to the role of keeping the penalty relatively low when constraints are only slightly violated.

However, as flows rise significantly above 100MW, and the value of the violation variable rises above 10MW, the penalty rises rapidly: the penalty for a 10MW violation in this case would be $1,500/MWh on average (and a marginal penalty of $2,500/MWh). That is, the quadratic coefficient in the penalty can be tailored to discourage significant violation of the constraints.

To summarize, the choice of the two parameters in the quadratic penalty can be targeted at each of the first two roles of the penalty function described in Section 1 without the compromise inherent in using only one parameter. In particular, since the slope of the penalty function at $v = 0$ is $500/MWh, there would only be violation of the constraint if supply offers at correspondingly low prices were exhausted. That is, violation would not occur unless the supply was moderately tight. Moreover, the low "initial" penalty price reflects the intrinsic uncertainty in thermal flow limits. However, when flows are significantly above limits, the penalty becomes large, reflecting the imperative not to significantly overload a line.

In the case of a *contingency* constraint on the same line, a similar penalty function could be used but the unrelaxed limit flow limit would be 110MW. This penalty would allow, at modest penalty, violations of the 30 minute limit but would much more strongly penalize a violation of a 10 minute limit.

This same approach could be applied to most constraints. For example, in the context of a power balance constraint, which is sometimes violated in favor of using regulating reserves to provide energy, an appropriate penalty would involve a modest penalty for small violations that is based on the costs of deploying the regulating reserves. The penalty would rise to the value of lost load when the violation reached a significant fraction of the total capacity of available regulating reserves.

3.3 Systematic choices for penalty prices

In the following discussion, systematic choices for the parameters of the penalty function are discussed. The focus is on estimating the underlying costs of violating a constraint and of ordering the violation of constraints by SCED.

3.3.1 Power balance

From the discussion above, the "initial" slope of the quadratic penalty function on power balance should be at a level reflecting the relatively low cost of deploying regulating reserves but should then increase towards the value of lost load, V. A more careful assessment of this issue would involve an assessment of the costs due to the risk of cascading outages when using significant regulation reserves for energy.

For example, suppose that an assessment was made that the economic cost, in terms of risk to security, of deploying just a few MW of regulating reserves for energy was some value V' denominated in \$/MWh. That is, violating power balance by v_3 incurs a cost of $v_3 \times V'$, at least for small values of v_3. Presumably, $V' < V$ in order for this to be a sensible action for the operator to take: otherwise curtailment should undertaken in preference to deploying reserves for energy, since curtailment of v_3 of demand costs $v_3 \times V$, by definition.

Furthermore, it is presumably also the case that $V' > P$, since regulating reserves would not be deployed if there were any available offered generation capacity and generating capacity must be offered at no more than the offer cap P. For concreteness in an example, suppose that $V' = \$3,000 / \text{MWh}$.

Moreover, suppose that it was also assessed that if more than, for example, approximately $R' = 100\text{MW}$ of reserves were deployed then the risk to security would have become unacceptable so that curtailment should be carried out. A piece-wise quadratic penalty function on violation v_3 of power balance would then be:

$$c_3(v_3) = \begin{cases} 0, \quad \text{if } v_3 \le 0, \\ v_3 \times V' + (v_3)^2 \times (V - V') / R', \quad \text{if } 0 < v_3 \le R', \\ v_3 \times V, \quad \text{if } v_3 \ge R', \end{cases}$$

where it is understood that if $v_3 > R'$ then curtailment should be carried out by the operator.

3.3.2 Constraints associated with non-cascading outages

In the example in Section 3.2 of a quadratic penalty function for violation of a flow limit associated with a non-cascading outage, a value of $500/MWh was used for the penalty price associated with initial violation of the constraint. However, this value should be designed to comport with the penalty function on power balance and also with costs of violating flow limits on transmission lines and transformers estimated in terms of accelerated aging due to overloads.

In particular, if violating such constraints by a small amount imposes essentially no significant cost then such constraints should be violated in preference to using regulation to satisfy power balance, since using regulation to satisfy power balance was assumed to have a non-zero cost of V'. Consequently, constraints associated with non-cascading outages should have an initial penalty price that is low enough to guarantee that such constraints would be violated in preference to violating power balance. From the discussion in Section 3.3.1, power balance would be violated if the Lagrange multiplier on power balance were to rise above V', which is itself above the highest offer price of P in the system.

To enable a choice of the linear coefficient that determines the initial penalty price, we make the following assumptions:

1. For each bus *k*, the sum of the shift factors in the corresponding expression for the LMP is less than one.
2. For small overloads of line constraints, there is no cost due to transmission line and transformer degradation nor increased likelihood of elements sagging into trees or other structures.

If the first assumption holds, then, for any bus *k*, the difference between:

- the Lagrange multiplier on the power balance constraint, and
- the LMP at bus *k*,

is no more than one times the maximum value of a Lagrange multiplier on a line constraint. For any marginal generator at a bus *k*, the LMP at bus *k* cannot be more than the offer cap *P*.

Consequently, if the linear coefficient in each penalty function on line limits is equal to $(V'-P)$ then for small violation of the line constraints the Lagrange multiplier on power balance could never rise above $P+(V'-P)=V'$. That is, we will not use regulating reserves in preference to slightly violating these line limit constraints.

A more careful assessment of this initial penalty price would also involve assessing the cost of transmission line and transformer degradation due to thermal overloads and increased likelihood of lines sagging into trees and other structures. For example, manufacturer ampacity information together with thermal properties of the conductors could be used to refine the assessment of the linear coefficient. On the one hand, for a conductor that was relatively robust to overloads, the value $(V'-P)$ might be appropriate. On the other hand, for a conductor with greater potential for sagging or for a transformer, a somewhat higher value would be more appropriate.

At higher violations, however, above some violation limit *R*, the slope of the penalty function should rise rapidly. In order to ensure that power balance would be violated before a large overload occurs, the slope of the penalty function should rise to being on the order of:

- V', divided by
- the smallest typical shift factor impact on the constraint, σ, say.

Summarizing, an appropriate penalty function could be:

$$c(v)=\begin{cases} 0, \quad \text{if} \quad v\leq 0, \\ v\times(V'-P)+(v)^2\times V'/(R\sigma), \quad \text{if} \quad v>0. \end{cases}$$

3.3.3 Constraints associated with cascading outages.

These constraints should not be violated except *in extremis* and therefore should have high values of initial penalty price that are above *V*, the value of lost load. Having high, but finite, values would enable the penalty functions to satisfy the third role of penalty functions described in Section 1.

3.3.4 Constraints associated with other issues

Besides transmission limits there are a number of other constraints represented in SCED, including implicitly, for example, generator ramp rate constraints. Potentially these could also be relaxed, involving the consideration of issues such as generation wear-and-tear costs for violating ramp rate limits.

Acknowledgment: In preparing this document, I have benefitted from discussions with Lin Xu, Resmi Surendran, and during the CMWG meeting on May 19, 2008.