21 February 1997

ACTIVITY RULES FOR A POWER EXCHANGE

Robert Wilson*

This paper studies the design of activity rules for a power exchange (PX). The suggested rules are summarized in Section 2 after their motivation is described in Section 1. They are developed in a practical form in subsequent sections, and collected together in Appendix A, along with several variants in Appendix B. Appendix C describes an alternative offer format that eliminates withdrawals. Appendix D describes the uniform-price double auction that is the principal alternative to the auction design studied here. Section 10 concludes with a list of unresolved issues.

1. The Role of Activity Rules

Self-scheduling is the principal feature of a PX auction. Bids and offers are for energy only. Fixed components such as start-up and no-load costs are absorbed by suppliers. There are several market designs that provide a supplier with some assurance that these fixed costs are covered by the difference between total revenue and incremental energy costs. One type allows offers on a full-cost basis; this type includes bilateral bid-ask markets as in Appendix D, and auctions that allow combination tenders for multiple hours, as described in Appendix C. A second type is represented by the California PX Protocol, in which an iterative process enables suppliers to withdraw from hours with prices insufficient to cover their total costs.

The key role of withdrawals in the PX Protocol is due to an interaction between the tender format and the pricing rule. The tender format requires separate offers for each hour, while the uniform-price rule encourages each supplier to offer incremental quantities at prices close to incremental cost. The uniform-price rule stems from the CPUC order that in the PX all energy in a given hour is to be traded at the same price, exclusive of the zonal surcharges for transmission, and the policy decision that the PX takes no net position. There are ways to implement uniform pricing without withdrawals, using the offer format in Appendix C. Alternatively, one could forego the uniform pricing rule, in which case the natural auction process is a bid-ask market. In a dynamic version of such a market, during the week prior to dispatch, each trader can post bids or offers, or accept any posted bid or offer; each transaction is a binding bilateral contract immediately upon acceptance. Dynamic markets preclude a uniform price but they have the important advantage that they ensure impatience to trade. This is an impatience borne of fear that profitable opportunities will be missed: when a demander posts a good bid, each supplier is eager to accept it before a competing supplier grabs it first. In such markets the volume of trade rises fairly steadily as the dispatch time approaches, and the accuracy of traders' predictions about the best bid and ask prices that will

* Stanford Business School, Stanford CA 94305-5015. The preparation of this report was supported by the California Trust for Power Industry Restructuring, which does not necessarily endorse the conclusions herein, and in collaboration with Charles Plott.

prevail at the close improves correspondingly. Impatience to trade is one way to solve the fundamental problem of reliable price discovery.

Price discovery is more problematic in the PX Protocol, because no transactions occur until the close of the final iteration. Any dynamic or iterative process provides a sequence of price signals to traders. If these interim prices are good predictors of the final prices that will prevail at the close, then they enable suppliers to make accurate judgments about which plants to operate and in which hours. In turn, early resolution about which plants to operate in each hour ensures stable convergence, since later iterations focus on the simpler task of finding the clearing price for energy.

In the PX, activity rules are needed to ensure that price discovery is reliable. The issue is very simple: without activity rules, and with uniform pricing, no trader has any incentive to make serious bids or offers until the final iteration; and without serious bids and offers, the tentative clearing prices in early iterations are unreliable predictors of the final clearing prices. Indeed, any large trader has the opposite incentive: it withholds information about its own final offers in the early iterations, preferring instead to rely on others to provide such information contributing to price discovery. So in the absence of impatience of trade, activity rules are imposed in order to force all traders to reveal early some credible signal about the bids and offers they will tender in the final iteration.

In designing the activity rules, the guiding principle is that they should be the least restrictive rules that suffice to assure reliable price discovery. Ideally, they impose no limit on the efficiency attainable at the close of the market. In particular, they should impose no significant restrictions or disadvantages on suppliers who elect to offer their actual costs. The only effect of the activity rules is to limit gaming by imposing constraints on revisions of offers during the iterative process. During the iterative process, these constraints create increasingly strong incentives for cost-based offers: the net effect is about the same as rounding up wild horses by driving them into the chute at the vertex of a V-shaped fence. If the activity rules are successful, as preliminary experimental evidence indicates they are, then suppliers will learn that there is little to be gained by strategic bidding – it may delay convergence somewhat, but the final outcome is largely determined by cost-based offers in the closing iterations.

Realistically, costs must be interpreted here as opportunity costs rather than actual running costs, since each supplier also has opportunities to trade in other markets. In addition, opportunity costs must be interpreted in relation to market power. There is no activity rule that can prevent a supplier from realizing the profit obtained when it offers the running cost of the next plant in the merit order – which the supplier might know from long experience in which its plant is at or near the margin that determines the final clearing price.

To preserve self-scheduling, the activity rules cannot be invasive; e.g., they cannot rely on any additional solicitation of reports about traders' private information. On the other hand, we shall see that the activity rules can be designed using the principle of “revealed preference.” By interpreting previous offers as reliable indicators of what is feasible and profitable for the supplier, we can impose constraints on subsequent offers. As the auction progresses, these constraints narrow the supplier’s allowed strategies, until in the final iteration there is little room for offers that differ significantly from the supplier’s actual costs.

As a practical matter the activity rules must be easily understood by traders, and simple for the PX to implement. The activity rules should be applied automatically by the PX software: the portion of any submitted tender that violates the rules is discarded, without any "negotiation" with the trader.

Activity rules are generally of two kinds. One kind pertains to the opening and closing of the auction, and the other pertains to the ways in which tenders can be revised or withdrawn from one iteration to the next. The rules treat demanders and suppliers symmetrically: the rules for demanders differ only by interpreting price decrements as price increments. To avoid confusing the reader with separate phrasing regarding demanders and suppliers, I refer here only to the rules applicable to suppliers.

2. General Statement of the Activity Rules

The activity rules can be derived from a single formulation that is quite general in its application. To express this formulation succinctly, it is useful to interpret the tendered supply function as a bundle of contingent offers: each offer consists of a price for a particular increment of supplied energy. For example, one point on the tendered supply function might offer a price of $23 for the 107th mega-Watt (MW) of power delivered in the hour from 10 to 11 AM. Thus, I interpret a point (p,q) on the tender as offering the price p for the q-th increment of energy production.

The rule has three parts. In each iteration after the first, for each quantity increment included in the tender submitted in the first iteration:

1. The price cannot be increased.

2. The price can be decreased only if the new price is less than the clearing price in the previous iteration by at least a specified price decrement (e.g., $1.00 or $0.10 per MWh). We say in this case that the new price "improves" the previous clearing price.

3. The price cannot improve any previous clearing price not improved at the first opportunity.

Part 1 is a fundamental requirement for a competitive auction. Part 2's requirement that a price change improves the clearing price eliminates extraneous revisions. A minimum decrement is necessary to avoid stalling the auction. The number of iterations required for convergence is sensitive to the choice of this design parameter.

Part 3 is the key provision. To make it precise requires the following clarification: the "first opportunity" is the first iteration following an iteration in which the offered price exceeds the clearing price. For instance, if a supplier offers a price of $25 in iteration 1, in which the clearing price is $23, then iteration 2 is the first opportunity to improve this clearing price. If the supplier offers a price less than $23 in iteration 1 then for present purposes it has no obligation or "opportunity" in iteration 2 to improve the $23 clearing price obtained in iteration 1. Therefore, Part 3 imposes no restriction on suppliers who offer prices below the clearing price; in particular, these suppliers are not disadvantaged by refusing to improve the clearing price in the next iteration. However, among those suppliers who offer exactly the $23 clearing price there may be some whose offers are rejected according to the Rationing Rule. For these suppliers, iteration 2 is indeed the first opportunity to improve the previous clearing price.

With this clarification, Part 3 says the following, expressed via the example. Suppose the specified price decrement is $0.50. If in iteration 2 a supplier who offered $25 in iteration 1 does not improve iteration 1's clearing price of $23 then this is taken as de facto evidence that its cost increment for this quantity increment exceeds $22.50; consequently, the supplier is precluded from offering a price equal to or less than $22.50 in any subsequent iteration. However, if the clearing price later rises above $23, say to $24 in iteration 5, then the supplier can in the next iteration 6 improve this clearing price by offering any price between $22.50 and $23.50 – but if it fails to do so then thereafter it cannot offer any price equal to or less than $23.50. Similarly, a supplier who offers exactly the clearing price of $23 in iteration 1 and is rationed, and then declines to improve its offer to a price at or below $22.50 in iteration 2, cannot offer a price in this range later.

The effect of Part 3 is to "freeze" any part of a supplier's tendered supply function for which there is presumptive evidence that its cost exceeds a previous clearing price. It is only frozen, not rejected irrevocably, because there remains the possibility that it is “thawed” if the clearing price rises sufficiently in some later iteration. Part 3 prevents a supplier from profiting by withholding supply until the final iteration.

This general form of the activity rule is not in itself sufficient. The reason is that it allows suppliers to offer very high prices in the first iteration. If demanders similarly offer very low prices in the first iteration then the auction gets off to a slow start due to the resulting gap between supply and demand. This is an inherent problem in all auctions; the usual way of correcting this deficiency is an Opening Rule that governs the first iteration, typically by specifying reserve prices. Suggestions for fast-start provisions in the Opening Rule are included in Appendix B.

The Competitive Process

Activity rules of the form described above produce a characteristic process of competition among suppliers. After each iteration the offers are divided into those that are infra-marginal, because their offered prices are less than the clearing price, and those that are extra-marginal, because their offered prices are more than the clearing price. In the next iteration, each extra-marginal offer must improve the previous clearing price or forego all subsequent opportunities to offer lower prices – because it is frozen until later clearing prices rise above the previous clearing price. Thus, if the previous clearing price exceeds the supplier’s cost then the incentive to revise the offered price is quite strong, since this is the last opportunity. However, when the offer is revised, it ejects some previously infra-marginal offer, which now becomes extra-marginal, and that supplier now faces a similar problem. The resulting process resembles a tug-of-war between the suppliers to determine which one’s offer will be accepted at the clearing price. This battle is resolved when the clearing price is driven down to the cost of one of the contenders, who then prefers to let the offer be frozen. The characteristic pattern is that in each iteration there are many bids and offers near the previous clearing price; but if one side of the market must be rationed, say the suppliers, then those whose offers are excluded, and their costs are less, find it advantageous to reduce their prices.

3. The Auction Process and the Bid Format

The PX has 24 forward markets for delivery in the hours of the next day. A clearing price is computed separately for each hourly market. The auction is conducted in discrete iterations. After each iteration, a clearing price for each hourly market is computed independently from the current tenders. Each tender is specific to a particular hourly market, and consists of a step function that states the supply offered at each price. This function is interpreted as a bundle of contingent offers: each point (p,q) on the tender is an offer to deliver the quantity q in that hour at any price no less than p . Similarly, a step on the schedule offers a price p for any quantity within a corresponding min-max interval [m,M].

The activity rules apply separately to each point (p,q) on the tender. Thus, when checking the activity rules, no distinction is necessary regarding the exact form in which the tender is submitted: the same rules apply to tenders that are points, intervals, piecewise-linear, or step functions. For simplicity in the exposition, however, I assume that schedules are step functions. The rules are essentially unchanged if the alternative format for supply schedules described in Appendix C is used; nevertheless, for consistency I assume below that each supply schedule applies to a single hourly market.

I do not dwell here on the necessity that every tender is a binding bid or offer that remains in force until it is revised or ultimately rejected by the PX. A revised tender replaces all previous tenders for the same portfolio and hour. Except for those withdrawn or replaced, all tenders continue in force for the next iteration. If a withdrawal is revoked then the tender is automatically in force for the next iteration. At the close of the auction, those supply tenders with prices above the clearing price are rejected, with ties resolved by a Rationing Rule. The remaining tenders are accepted, and each becomes automatically a binding contract, with the PX as the counter-party, for the tendered quantity at the final clearing price. This contract is an obligation for energy; the supplier remains liable also for the transmission surcharge, calculated as the difference between the zonal price and the PX clearing price.

Sections 4 to 8 list the key ingredients of the basic activity rules. Less stringent variants are suggested in Appendix B. These weaker versions might be used in the iterations of a first stage that allows traders more flexibility, followed by iterations in a second and final stage that uses the stringent version. Alternatively, experimental testing or direct experience might indicate that relaxing some rules does no harm, especially if the number of traders is sufficiently large, and no one has any substantial market power, so that the vigor of the competitive process provides adequate protection against gaming.

4. The Opening Rule

The first part of the Opening Rule is simple:

- Opening Rule: A new tender can be submitted only in the first iteration.

In particular, in each later iteration the only tenders allowed are revisions of ones submitted in the first iteration. This rule ensures that the maximum supply in each hourly market is revealed in the first iteration. This rule is essential for effective price discovery, else a trader could wait until the final iteration to submit its first tenders.

One possible exception requires further study: a hydro portfolio might be restricted only by an initial declaration of its total energy available in all the hourly markets combined. For a thermal portfolio, the tender submitted in the first iteration essentially declares its maximum power capacity (MW) in each hour. In contrast, a hydro portfolio is typically constrained by the energy available in each day (MWh/day), or in a smaller interval if generation is limited to a few hours. Thus, a hydro portfolio differs in that supplies in different hours are substitutes. Efficiency therefore requires that a hydro portfolio can shift among the hourly markets in order to suppress the peak prices. The shifts allowed must be flexible, because flows from different reservoirs in the same drainage may be closely linked. The design of activity rules for total-energy suppliers has not yet been undertaken but an initial sketch is provided in Appendix B.

The second part of the Opening Rule is intended to get the auction off to a quick start. There are several choices available, so this part is relegated to Appendix B.

5. The Exclusion and Revision Rules

I first describe these rules along the lines of Section 2 and then elaborate their motivation.

All tenders that were not withdrawn after previous iterations are automatically carried over to the current iteration.   (So too are tenders whose previous withdrawals are revoked, if a less stringent version were to allow revocable withdrawals.) Based on the prior history of the auction, the steps on these tenders are divided into those that are frozen and those that are active: active steps can be revised, whereas frozen steps cannot. In the first and second iteration, all steps are active. In each iteration after the first:

- Exclusion Rule: A previously active step on a supply tender becomes frozen after the current iteration if its offered price was not revised to improve the previous clearing price, and in the previous iteration its offered price was above this clearing price – called its Activation Price. A frozen step cannot be revised. A frozen step becomes active again after an iteration in which the clearing price is higher than its Activation Price.

The Exclusion Rule operates as follows. If a tender's offered price for a particular quantity was less than the clearing price in the previous iteration then the supplier has no obligation to revise the offered price, but is not excluded from doing so. However, if its offered price exceeds the previous clearing price (or equal, and a portion of the step is rationed), then its offered price must be revised to something less than the previous clearing price, else it is frozen until the clearing price regains the previous level. For example, if the previous clearing price was $20 and the supplier now declines to offer a revised price less than $20 then this step cannot be revised again until after the clearing price rises above $20. As described in Section 2, the Exclusion Rule is based on the inference that refusal to improve the previous clearing price signals that the revised price would be insufficient to recover the supplier's cost.

The restriction that frozen steps cannot be revised is essential to reliable price discovery. Otherwise, a supplier could wait until the last iteration to revise, and in the meantime other traders would be getting no information about lower prices the supplier might be willing to offer. Thus, each tendered supply price that is above the clearing price in one iteration must be revised in the next iteration lest it thereafter be excluded from revisions until the clearing price rises again to comparable levels.

- <u>Revision Rule</u>: An active step can be divided into two active steps with the same offered price. An active step can be revised only by offering a lower price that improves the previous clearing price. That is, the revised step must offer a new price for the same quantity interval that is less than the previously offered price, and also less than the previous clearing price by at least the specified price decrement.

This particular phrasing of the Revision Rule is peculiar to the present supposition that each tender is represented as a step function. In this case, an active step corresponding to an offered price for an interval [m,M] of quantities can be revised by breaking it into two steps with intervals [m,k] and [k,M] . Then, one step is revised to offer a new price that improves the previous clearing price, and the second step is frozen. For the frozen step, the offered price is unchanged and its Activation Price is the previous clearing price.

Note that in each hourly market the clearing price is computed using all steps on the current tenders, both frozen and active. This reflects the fact that even frozen steps remain binding offers to the PX. However, those steps that offer a higher price for a smaller quantity than another step are excluded from the merit order used for the computation, so they have no effect on the clearing price obtained.

It is important to realize that the price decrement (and a comparable price increment for demanders) is an important design parameter that can substantially affect the rate of convergence of the iterative process. In a worst-case scenario the clearing price moves by no more than the price decrement from one iteration to the next. The appropriate magnitude cannot be determined a priori; rather, it must be based on judgment, experience, and predictions about current supply and demand conditions, especially the price elasticities and variances of supply and demand. A practical procedure might start in iteration 2 with a large value, say $1.00/MWh, and then decrease it steadily in later iterations to a final value, say $0.20/MWh. However, experimental evidence indicates that it is not evident that a small decrement will produce clearing prices closer to the theoretical clearing price. A large decrement has the advantage that it produces stronger pressure on suppliers to tender initial offers closer to actual costs: due to the large decrement, a price slightly above actual cost cannot be revised profitably, so a supplier must contend with the risk that a profitable opportunity will be missed. Experiments as well as subsequent experience will provide guidelines about how to set the price decrement to ensure timely conclusion of the auction.

Another important ingredient is the Rationing Rule. In a typical iteration there can be many offers at the clearing price, and if demand at that price is less than supply, then some of the supply steps must be rationed. Presently, the experimental evidence indicates that it best to reject entire steps rather than allocate the marginal demand *pro rata* among the supply steps at the margin. This avoids a proliferation of subdivided steps, and appears to accelerate convergence.

6. The Withdrawal Rule

The following formulation assumes that after withdrawals the clearing prices are re-computed before the next iteration. Re-computing the clearing prices is desirable to ensure that other traders can take account of this information when revising their tenders for the next iteration.

- Withdrawal Rule: After each iteration except the first and last, each supplier has the option to withdraw a tender entirely and irrevocably from any hourly market. The clearing prices are re-calculated after the withdrawal round. For the purposes of the Exclusion and Revision Rules and setting Activation Prices, these become the clearing prices for this iteration.

The purpose of withdrawals is to allow a supplier to exit one or more markets if prices are insufficient to recover fixed costs. It is clear that withdrawals cannot be revoked easily, else a supplier could withdraw until it re-enters in the final iteration. On the other hand, it might be argued that for efficiency it is necessary to revoke withdrawals if prices rise later. I have studied this problem in considerable detail, but find it very difficult to construct a simple revocation rule that is invulnerable to gaming. Within the strictures of the PX Protocol, my solution to the problem is the Revision Rule in Section 5, which is constructed explicitly to enable a supplier to offer tenders that assure coverage of its average costs. Appendix C offers an alternative solution that departs from the PX Protocol by using a different format for supply tenders. Both solutions enable a supplier to cover its average cost whenever the clearing price exceeds the supplier's minimum average cost. Consequently, my conclusion is that there is no need, and no easy prospect, to allow withdrawals to be revocable.

Note that after the final iteration, an accepted tender cannot be withdrawn and the supplier is financially liable for delivery. An alternative procedure based on combination tenders is suggested in Appendix B.

7. The Closing Rule

- All the hourly markets close simultaneously. They close automatically after any iteration in which no tender is revised. Otherwise, before time expires, the final iteration is announced, closing procedures are specified, and special rules may apply.

One important possibility for a special rule is described below. Its purpose is to reduce the risk inherent in the exclusion of withdrawals after the final iteration. Others are included in Section 9 regarding failsafe provisions.

Offers of Strips in Combined Tenders

In the final iteration only, it is possible to allow combination tenders for strips. A combination tender applies jointly to several hourly markets, typically a sequence of consecutive hours. Such a tender consists of a valid (according to the activity rules) tender for each hour in the sequence, plus an aggregate revenue requirement specified by the supplier. The evaluation process works as follows. After the clearing prices are computed, if there is a combination tender with a revenue deficiency then the one with the greatest deficiency is rejected irrevocably, the clearing prices are re-calculated, and the process repeats until there remains no combination tender with a revenue deficiency. Note that prices increase after each rejection of a combination supply tender, which is why the rejection must be irrevocable. If the final iteration allows combination tenders then it may be sufficient to exclude withdrawals after previous iterations, provided experience proves that this does not impair price discovery.

Allowing combined tenders in the final iteration requires a sufficiently slight violation of the principle of separate evaluation of the hourly markets that I think it can be considered a viable alternative to allowing withdrawals after every iteration except the last. It enables suppliers to avoid financial risks in the final iteration, and it averts gaming of the Withdrawal Rule of the sort described in Appendix B. Nevertheless, as I have emphasized in Sections 5 and 6, the current form of the Revision Rule is intended to diminish the role of withdrawals. Therefore, I do not recommend combination tenders unless this prediction proves wrong.

8. Reporting of Results

The minimum reporting requirement is that the clearing prices are broadcast after each iteration, and again after withdrawals. It is possible to hold each supplier responsible for inferring the status of its tenders. The sole exception is that the result of the Rationing Rule must be reported to those suppliers with steps whose offered prices are the same as the clearing price.

It is desirable that after each iteration the imbalance in each zone is broadcast. This information enables traders to obtain better predictions about the magnitude of the transmission charges imposed in the subsequent congestion management phase.

9. Failsafe Provisions

The motive for an iterative auction process is to allow suppliers to withdraw units whose fixed cost components are not covered by the clearing prices for energy. This brings the fundamental risk that prices could be affected substantially by withdrawals of large plants shortly before the final iteration. This section discusses some options that might be used to reduce this risk.

One option excludes withdrawals after the penultimate iteration, so that at least two iterations are available after the last withdrawal. A second option interprets the withdrawal or rejection of a tender as relevant only to transactions at the clearing prices.

In this case, the tenders remain binding offers that the PX can accept at the tendered prices if this is deemed necessary according to some prudent criterion. A similar option is that the PX takes a position that it then clears in later markets; again there must be prudent guidelines.

The above options intervene in the energy market. The principal alternative is to rely on trades in the market for Inc/Dec options that occurs in the hour after the ISO's advisory re-dispatch. Because this market for trades among all the competing market makers will function in any case to resolve price discrepancies among their markets, it is likely the best solution.

10. Conclusion

The purpose of the activity rules is to encourage convergence to an efficient outcome by suppressing gaming. The rules proposed here are based on the principle of "revealed preference." Essentially, a supplier's refusal to improve a previous clearing price is taken as evidence that such a lower price would not recover its cost. The resulting process forces suppliers at the margin to compete: each extra-marginal bidder that improves the previous clearing price ejects some infra-marginal bidder who is thereby forced to reduce the offered price or forego any profit it might obtain. Each refusal freezes a step of the tender until the clearing price rises that high again later.

These rules are complemented by procedures for opening and closing the auction, and allowance for withdrawals.   All tenders must be submitted at the opening to preclude a strategy of waiting until the final iteration that would impair price discovery. Withdrawals must be irrevocable and in any case withdrawals after the final iteration must be excluded unless explicit provision is made for combination tenders.

There are three main issues unresolved presently. One is the design of alternative activity rules for those suppliers affected by a total energy constraint (MWh/day) such as a hydro portfolio, in addition to the usual power capacity constraint (MW) that affects thermal units. The second is the issue of whether to exclude withdrawals. Withdrawals could pose problems in attaining efficiency if suppliers' offers are based on incremental costs, because then early withdrawals might not be justified by the higher final clearing prices. More likely, however, suppliers will recognize that a superior strategy is to base offers on full costs, in which case the outcome is efficient and convergence is quick and stable without withdrawals playing any major role.   Appendix C outlines an alternative bid format that directly ensures full-cost offers. The third issue is to determine how best to design the rationing rule and the price decrement to ensure convergence within the allowed time of two hours. This issue is being examined in the ongoing experimental tests, but a full resolution will require experience in the actual auction with 24 hourly markets and many traders.

The residual risks can be divided into those that are procedural and those that jeopardize efficiency.   I mention each with the likely remedies.

- One procedural risk is that this auction format might operate too slowly to provide reliable convergence within the two hours allowed. The possible remedies include a large price decrement parameter, continuous-time bidding, the alternative offer format for load slices described in Appendix C, or a bid-ask market (either a dynamic market for bilateral transactions, or a uniform-price market as described in Appendix D). Alternatively, the auction can simply conclude when time expires, with a final iteration that allows combination tenders; or earlier, if a criterion for adequate convergence is satisfied.

- The second procedural risk is that clearing prices will be ill-defined because of gaps in the supply function produced by minimum stable generation levels. The possible remedies include a revised bid evaluation rule for computing the clearing prices, allowing the PX to take a net position to clear the market, or a bid-ask market. The alternative offer format will not eliminate this risk.

- One risk to efficiency is that withdrawals might play an important role due to incremental-cost offers from suppliers with substantial fixed costs, and that price discovery occurs insufficiently early to ensure an efficient selection of withdrawn units. The possible remedies include informing suppliers about the likely advantages of full-cost offers, the alternative offer format for load slices, or a bid-ask market.

- The second risk to efficiency is that suppliers affected by total-energy constraints, such as apply to a hydro portfolio, might find the activity rules interfering with optimal allocation of their supplies to peak hours. The possible remedies include separate activity rules for total-energy suppliers as sketched in Appendix B, or a bid-ask market.

There are also risks from the market power of major suppliers, but most of these cannot be eliminated by any variant of the activity rules. To exclude predatory tactics of the kind described in Appendix B, the best remedy is prohibition of withdrawals except via combination tenders in the final iteration.

Appendix A

Activity Rules

The following stringent version of the activity rules is stated for supply tenders; symmetric rules apply to demand tenders. Variants are suggested in Appendix B, and some options for withdrawals are included here.

Tenders: Each step of each tender is a binding offer to trade at any price not less than the offered price. Each tender remains in force until it is withdrawn or validly revised by the trader, or rejected by the PX. A revised tender replaces the previous tender for the same segment. At the close of the auction, those steps with prices above the final clearing price are rejected; ties are resolved via the Rationing Rule. The remaining steps are accepted, and each becomes automatically a binding contract, with the PX as the counter-party, for the tendered or rationed quantity at the final clearing price – except a step at the margin, for which only a portion of the offered quantity might be accepted. This contract is an obligation for energy; the supplier remains liable for the transmission surcharge – the difference between the zonal price and the clearing price.

Opening Rule (First Part): A new tender can be submitted only in the first iteration. After the first iteration, the only valid tenders are those submitted in the first iteration or validly revised subsequently that have not been withdrawn.

Exclusion Rule: An active step on a supply tender becomes frozen after the current iteration if its offered price is not validly revised to improve the previous clearing price, and in the previous iteration its offered price was above this clearing price – called its Activation Price. A frozen step cannot be revised. A frozen step becomes active again after an iteration in which the clearing price is higher than its Activation Price.

Revision Rule: An active step can be divided into two active steps with the same offered price. An active step can be revised only by offering a lower price that improves the previous clearing price. That is, the revised step must offer a new price for the same quantity interval that is less than the previously offered price, and also less than the previous clearing price by at least the specified price decrement.

Withdrawal Rule: After each iteration except the last, each supplier has the option to withdraw a tender entirely and irrevocably from any hourly market. The clearing prices are re-calculated after the withdrawal round. For the purposes of the Exclusion and Revision Rules and setting Activation Prices, these become the clearing prices for this iteration. After the final iteration, an accepted tender cannot be withdrawn and the trader remains financially liable for delivery. Alternatively, withdrawals can be excluded until the final iteration, when they are based on submission of combination tenders; or they can be limited to complete withdrawal from all markets simultaneously; or they can be precluded whenever the allocated quantity exceeds the minimum stable generation.

Closing Rule: All hourly markets close simultaneously. They close automatically after an iteration in which no tender is revised. Otherwise, before time expires, the final iteration is announced, closing procedures are specified, and special rules may apply.

One such rule allows combination tenders, in which case withdrawals might be prohibited after previous iterations.

Failsafe Rule: After the final iteration, the PX can elect various options to ensure orderly markets. One such option allows it to accept any rejected or withdrawn tender at the tendered price; another allows it to swap trades with other scheduling coordinators and/or trade in the later market for Inc/Decs during the congestion management stage.

Appendix B

Variants of the Activity Rules

The Second Part of the Opening Rule

There are several options for ensuring that the auction gets off to a good start in the first iteration. Because the merits of these options require the judgment of persons more familiar with the power industry, I merely list them here rather than suggest one in particular. As before these are phrased only in terms of supply tenders. One might interpret each of the three specified percentages as 20% for purposes of discussion.

- The PX establishes reserve prices based on capacity options it acquires on a monthly basis.
- A specified percentage of the highest-priced supply steps offering prices above the clearing price in the first iteration are immediately frozen. Their Activation Price is set midway between the clearing price and the offered price.
- A step cannot be offered at a price more than a specified percentage above the best price at which it has been offered in recent weeks.
- A step cannot be offered at a price more than a specified percentage above the predicted clearing price based on current supply and demand forecasts.

The first option, which relies on reserve prices, is the standard method in similar auctions in other contexts. The second option exerts genuine competitive pressure to offer a price close to the predicted clearing price to avoid the risk of being frozen. The other options are problematic because they do not exclude the possibility that all the initial supply tenders are at the upper limit. In the worst case, all supply tenders exceed all demand tenders, so the initial clearing price is the highest demand bid; then in the second iteration the suppliers must improve this clearing price, which is what gets the auction started – but in the meantime an iteration has been essentially wasted.

The incentives required to get an auction started are ordinarily not a major concern, but in the PX the short time frame and the small number of iterations make it imperative that progress towards convergence not be delayed. The large number of traders in the PX will likely suffice to prevent implicit coordination on extreme offers in the first iteration. This is because it takes only a few traders making serious offers to establish a meaningful clearing price in the first iteration; then in the second iteration all the others are forced by the Revision Rule to improve this clearing price, or their tenders are frozen. The pressure on suppliers is stronger, since the PX Protocol sets the clearing price at the highest unserved demand bid whenever there is a gap between the supply offers and the demand bids.

If this prediction proves to be wrong then an option such as one of the four above should be invoked to ensure a fast start. Further analysis and experiments are needed to determine the effectiveness of these options. Discussion is also needed because I mentioned only the first option in my oral presentation.

Relaxing the Exclusion Rule

The Exclusion Rule prevents an idled tender from being revised before the clearing price exceeds its Activation Price. On the other hand there can be advantages from allowing a supplier to signal that it is willing to offer a price that is not far above its Activation Price. Although further analysis is required, it appears that allowing a frozen tender to offer any revised price between its current offer price and an increment above its Activation Price does not introduce significant gaming possibilities, while allowing some degree of signaling.

If the clearing prices are updated continually, after the arrival of each revised tender, then the Exclusion Rule must be specified in terms of the time *duration* allowed for revision of a supply step whose offered price is above the clearing price. A typical form of this rule is as follows:

> At any time t that the offered price for an active step exceeds the current clearing price $P(t)$, or this clearing price exceeds the Activation Price of a frozen step, the supplier is notified that it has until a specified time $T > t$ to improve that clearing price, else the step is frozen at time T with the new Activation Price $P(t)$. Note that if the clearing price $P(T)$ at time T exceeds $P(t)$ then the frozen step is immediately re-activated and the rule is applied again, allowing a new duration of length $T-t$ in which to improve $P(T)$. Also, if the step's offered price is revised to $P(t)-d$ at time $s < T$ and this price exceeds the current clearing price $P(s)$, then again the rule is re-applied to require a revised offer that improves $P(s)$ within the next time interval of length $T-t$.

The experimental implementation at Caltech includes this rule as an option. It is one way to include an Exclusion Rule in the uniform-price double auction described in Appendix D.

Relaxing the Revision Rule

As described in Sections 2 and 5 the Revision Rule is quite stringent. A step whose offered price is above the clearing price in one iteration must improve that clearing price in the very next iteration or it is frozen, and it cannot be revised thereafter until the clearing price rises that high again. I specify this stringent form in order to accelerate the iterative process, since the allowed time for the auction is short. It might suffice, however, to adopt relaxed versions that are not so Draconian. One possibility is to allow two iterations for the offered price to improve the clearing price. Another is to allow that the price improved could be midway between the offered price and the clearing price. A third is to re-activate a frozen supply tender at some lower price, such as the clearing price in the next iteration – the one in which it failed to improve the previous clearing price.

Each option brings a risk that the added flexibility will encourage gaming and/or slow convergence. On the other hand, re-activating a tender at a lower price enables a previously frozen tender to be participate in the competitive fray. A potential compromise is to divide the auction into, say, two stages (as in the FCC spectrum auctions): during iterations in the first stage a relaxed Revision Rule applies, but the stringent Revision Rule applies in the iterations of the second stage.

Altering the Withdrawal Rule

The Withdrawal Rule could be strengthened to exclude withdrawals from hourly markets in which the portfolio is allocated a quantity exceeding the minimum quantity tendered, since this can be taken as presumptive evidence that hour-specific fixed costs are covered at the clearing price. In this case, a supplier cannot withdraw from such an hourly market unless the portfolio is withdrawn from *all* hourly markets simultaneously. It is also plausible to consider weakening the Withdrawal Rule to allow shifting a sequence of consecutive hours to an earlier or later start time to take better account of the pattern of hourly prices. Shifting the start time can only be allowed if it is restricted to prevent its use to defer entry into some hourly markets until the last iteration. The exact form of such a rule is likely to be an analog of an opening rule applicable to a total energy declaration for a hydro portfolio.

In general, allowing withdrawals creates a potential vulnerability in the activity rules. Steven Stoft emphasizes that they might be used in a predatory way if some firm has significant market power, as in the following scenario. A supplier could offer a large but uneconomic portfolio at low prices, thereby inducing marginal suppliers to withdraw irrevocably, after which the uneconomic portfolio could be withdrawn, resulting in higher prices for the supplier's other portfolios that are economic. That is, the irrevocability of withdrawals creates the possibility that uneconomic portfolios are used to stalk and threaten others; indeed, in the worst case one can envision that a large supplier uses a non-existent portfolio, or one whose maximum capacity is exaggerated, to predate in this way. This is one reason that I offer the alternative in which withdrawals are excluded except in the final iteration when they are based on combination tenders, and the scheme in Appendix C that enables exclusion of withdrawals altogether. In any case, the possibility that irrevocability of withdrawals could be exploited in a predatory way points to one of the implications of the activity rules for FERC's monitoring of adherence to its market power mitigation requirements.

Partial Clearing

Because thermal units typically have a minimum stable generation (MSG) level, the first step of a portfolio might specify a minimum quantity that is positive. In this case there can be any number of market clearing prices, including zero, due to the resulting gaps in the aggregate supply function. Apart from incentive considerations, the efficient solution in this case is to use as the clearing price the least price that provides supply as large as demand, provided the surplus supply can be sold in later markets, as the ISO presently allows for energy obtained from spinning reserves. If markets for spinning reserve are included directly in the PX then it is also possible to assign portions of more fully loaded plants to spinning reserve. On the other hand, this problem can also be solved by the supplier individually, using a bilateral contract to cover the MSG, or by a financial contract that hedges the price. If marginal units at their MSGs are typically rejected then suppliers presumably will learn to use these solutions, or to bid flexibly, planning to complete their MSGs in later markets. When incentive considerations are included, these individual solutions are superior, since otherwise a supplier has incentives to exaggerate the MSG of a unit likely to be at the margin, as a way of forcing the PX to absorb the difference or to reduce the operating rate of a more efficient unit.

Swaps

A central requirement in the overall design of the energy markets is enabling sufficient arbitrage among the scheduling coordinators to ensure that their energy prices are similar. In the long term this arbitrage will be accomplished by traders moving among the various markets. In the short run the arbitrage can be done during congestion management after the ISO's advisory re-dispatch by enabling a market for trading Inc/Dec options among the various scheduling coordinators, including the PX. Alternatively, it might be done during their parallel energy auctions by allowing swaps. If the PX were to engage in swaps with another scheduling coordinator then after each iteration the market clearing prices could, for example, be computed using both its own tenders and those rejected in the market of the other scheduling coordinator, and vice versa. This requires either coordinated market clearing or an iterative process, but even a partially effective procedure might be helpful in avoiding large disparities in energy prices.

Alternative Activity Rules for Total-Energy Portfolios

I interpret a total-energy portfolio here as representing a hydro complex and sketch one version of alternative activity rules that could apply to such a portfolio. I assume that the technology can be described by an opportunity cost for spilled water, together with a total energy constraint, a maximum power rate, and possibly linkages among the hours, as in the case that the water spills through a series of reservoirs. The key feature is the uniform opportunity cost. Based on this feature, the activity rule can be cast in terms of a requirement that the supplier’s tenders must offer the same price in each feasible hour for all quantities up to the maximum power rate. For each portfolio, the set of feasible hours (which may be limited for COB importers, for instance), the total energy, and the maximum power rate(s) must be declared in the initial iteration. Thereafter, the portfolio’s offered price must improve the clearing price in at least one hour (or the entire portfolio is frozen), and the total energy tendered at the offered price cannot increase from one iteration to the next. In each iteration, the revised tenders can specify the allocation of the total energy among the hourly markets – or at the option of the supplier this can be done automatically by the PX, which allocates the total energy to the hours with the highest prices.

I caution that these alternative rules for hydro portfolios have not been thoroughly studied nor tested experimentally. Nevertheless, preliminary analysis indicates that these rules are sufficient. Their main effect is to offer a tradeoff: a supplier can designate a hydro portfolio as a total-energy portfolio, and thereby gain greater flexibility in allocating its supply to the peak hours, in exchange for two restrictions – its offered prices must be uniform across the feasible hours, and its total supply cannot increase.

Appendix C

An Alternative Bid Evaluation Procedure

This appendix describes an alternative procedure for evaluating bids and offers to establish clearing prices in the hourly markets. Unlike the text of this report, which adheres closely to the PX Protocol, the evaluation procedure presented here departs from the current protocol – but in only the one respect that it alters the format in which suppliers submit offers. Its key feature is that it enables suppliers to offer their full cost for a sequence of consecutive running hours, while at the same time ensuring that a uniform market clearing price is established in each hourly market. This brings the advantage that withdrawals can be prohibited; or even if they are allowed, they are unlikely to play an important role. Thus, this modified evaluation procedure is an important backstop that can used if it turns out that withdrawals are a serious problem when the current protocol is used.

I begin by explaining the motivation for the alternative offer format and its corresponding evaluation procedure and then I outline how it is implemented in practice.

Motivation for the Alternative Bid Evaluation Procedure

As described in Section 1 of the text, the current PX Protocol assumes that several iterations are conducted to arrive at a uniform clearing price in each hourly market, each evaluated independently. It is an inherent feature of this design that withdrawals are an important part of the iterative process. Because each supplier competes with others for whom the price is sufficient to cover fixed costs, there is a strong tendency for revised supply offers to be driven close to the incremental energy costs of those suppliers at the margin who, in effect, establish the clearing price. If the resulting clearing price is insufficient to cover its fixed costs then a supplier’s main recourse is to withdraw from the hourly market, or to withdraw from the entire day. This produces a risk that prices in the PX will be volatile in the last few iterations as some suppliers withdraw, and that premature withdrawals might produce inefficiencies.

The purpose of this appendix is to describe a procedure that allows suppliers to bid their full costs, not just incremental costs; in turn, this allows the PX to prohibit withdrawals. This procedure can also be used if the time window for the auction is insufficient, since the procedure works effectively even if only a single iteration is allowed. A single iteration would also simplify software development.

Bid and Offer Formats

For buyers the auction is exactly as in the PX Protocol; in particular, each buyer submits a separate demand schedule to each of the 24 hourly markets. A demand schedule has the usual form, stating the quantity demanded in that hour at each price. The power quantity, say q, is measured in MW to be received continuously over the hour, and the power price, say p, is measured in \$/MWh accumulated continuously over the single hour.

Suppliers also submit several different schedules showing the power quantity offered as a function of the price, but in this case the price R is the total to be paid for power supplied continuously over a stated number of consecutive hours of operation [I use capital letters to designate that this price is a total amount]. Thus, if the offer is for $q = 100$ MW supplied continuously for $h = 10$ hours, then the offered price $R(100, 10) = \$20,000$ corresponds to the average power price $p = \$20/\text{MWh}$ over each of the 10 hours. In general, then, each supplier submits a schedule $R(q,h)$ as a function of the power rate q for each number h of consecutive hours of operation.

A supply tender in this format can be interpreted as an offer to serve a horizontal slice of the load-duration curve. In contrast, a buyer's schedule is based on time-of-day prices and the quantities represent a vertical slice of the load-duration curve. In California, supply offers for load-slices are feasible since there are few days in which the system has more than a single peak, and even then the mid-day trough is small.
Thus, there is usually a well-defined relationship between the ordering of the hours used to construct the load-duration curve and the temporal ordering in real time. For instance, it is largely predictable that the 10-hour slice will be used to serve the load between 8 AM and 6 PM, and the 12-hour slice, between 7 AM and 7 PM. Consequently, each supplier can be fairly sure of the start and stop times associated with each length of load slice. Further, each offer can include all the costs associated with serving that load slice, including fixed costs of start-up and no-load, as well as the incremental costs of increasing the power level above the minimum stable load level.

If a supplier prefers to submit time-of-day schedules then they can be submitted as negative demands. Similarly, a buyer can submit load-slice schedules as negative supplies.

Determination of the Hourly Clearing Prices

There is a simple accounting formula that determines the relationship between the load-slice prices and the time-of-day prices. This formula can be used to convert the load-slice offers into equivalent time-of-day offers, and then from these and the demand bids, the clearing prices in each hour are calculated in the usual way. For simplicity of exposition, I assume here that the demand load is symmetric in time on either side of a peak-load time.

The key formula recognizes that the payment R(q,h) for the power rate q over h hours must be the sum of the time-of-day payments p(t)q over the temporal hours spanned by the duration h . Here p(t) represents the time-of-day price in temporal hour t . This implies that the average price P(q,h) = R(q,h)/q in $/MW must be the sum of the time-of-day prices over the h hours. To apply this formula, one proceeds as follows, assuming for simplicity that the load durations h are denominated in even numbers of hours.

First, for each number h = 2, 4, … 24 of consecutive hours, stack the suppliers' load-slice offers in merit order in the usual way. This yields the aggregate supply curve S(P,h) showing the total power supplied continuously over the h hours as a function of the average price P , measured as $/MW for h hours of continuous operation.

The second step observes that, if P(h) is to be the market price for the load-slice of length h , then in terms of time-of-day prices it must be that p(t) = [P(h) - P(h-2)]/2 , where p(t) is the time-of-day price in each of the 2 temporal hours at either end of the interval of h consecutive hours for that slice. (Obviously this is more accurate if the time slices are short: rather than hours it is better to use fifteen-minute or even five-minute intervals. The formula invokes the symmetry assumption.) This formula provides the equivalent time-of-day price.

The third and final step is to use this time-of-day price to calculate the market clearing price. In particular, if the aggregate incremental demand function is D(p,t) for temporal hour t , then market clearing is obtained by finding the equilibrium time-of-day price p(t) for that hour such that D(p(t),t) = S(P(h),h) , where t is the first or last hour in the sequence of h hours, and p(t) = [P(h) - P(h-2)]/2 .

Unlike time-of-day offers, all of which might be accepted, load-slice offers are typically mutually exclusive, so only one can be accepted from each portfolio. Thus, if a supplier's h-hours offer is accepted then its offers for each other number of hours are rejected. The iterative calculation is easiest if the load-slice prices P(24), P(22) , et cetera, are determined sequentially in that order. For instance, if a supplier's first accepted offer is for, say, a 12-hour slice then its offers for load slices of duration 10 hours or less are deleted in the continuation of the calculation – on the assumption that each portfolio will run for the maximum duration accepted – except for any portion of its capacity that is unused by the assigned 12-hour slice, which is carried over to the market for 10-hour slices.

As set forth above, this procedure requires one to predict beforehand the two temporal hours t that correspond to the start and stop hours for each duration h . In practice, these temporal hours for start and stop cannot be known exactly beforehand. Consequently, the calculation actually involves several iterative steps to arrive at the exact correspondence between the predicted and actual start and stop times to which each load-slice applies.

One way to see the essential structure is to interpret time as continuous, in which case the conditions for market clearing (in the symmetric case assumed here) are the two equations

$$dD(p(t),t)/dt = S(P(h),h) \quad \text{and} \quad P'(h) + p(t) = 0 .$$

These two conditions can be interpreted as determining the two price schedules. The third condition that determines the duration $h(t)$ associated with each time t before the peak-load time $t = T$ (for which $h(T) = 0$) is an intertemporal optimization condition: this condition specifies that $P'(h)$ matches the marginal cost of an additional hour for those suppliers selected to begin operation at time t for which the duration is $h = h(t)$. Thus, the proper formulation expresses the supply q offered for each duration h by each supplier by two functions Q and H , each of which depends on both P and P' . In particular, the proper statement of the supply function has the form $S(P(h),P'(h),h)$.

Summary

The advantage of this alternative procedure for evaluating bids and offers is that it allows suppliers to express their offers in terms of their total "going forward" costs, including start-up, no-load, and incremental energy costs. As required by the CPUC directive, a uniform market clearing price is obtained for every hour. For suppliers, there is the advantage that the hourly prices suffice to cover all their costs, both fixed and incremental. In principle, the results should be the same as those resulting from the current PX Protocol, but with no price volatility in the final iterations. This alternative procedure, and its load-slice format for supply offers, is feasible mainly because of the convenient feature of the California demand pattern that only rarely are there multiple peaks during the day, and in any case the trough is small. Consequently, there usually exists a well-defined correspondence between each load slice and the sequence of temporal hours in which it serves demand.

A recent draft report from London Economics shows that this alternative bid format for full-cost offers is not inherently necessary. In the PX a practical, and perhaps optimal, strategy for each supplier is to offer its tenders initially on a full-cost basis. That is, the tender submitted in the first iteration already includes an allocation to each hour of the no-load cost, as well as the start-up cost, on the assumption that this might be the only hour in which the supplier's tender is accepted. Subsequent iterations provide increasingly accurate predictions about the final clearing prices, and on the basis of these improved predictions, the supplier can extend the range of consecutive hours over which the fixed cost is allocated – which is what enables the supplier to improve successively lower clearing prices down to the level of its actual full cost for the hours in which it expects to operate. As their study establishes, strategies of this sort produce an efficient outcome, and convergence is quick and stable. Withdrawals play a minor role because the selection process (who will operate when) is accomplished mostly via the freezing of offers above the clearing prices. This avoids the risk of inefficiency from an inaccurate selection process based on withdrawals, which is susceptible to errors if a portfolio that is cost effective at the final clearing prices is withdrawn early when prices are lower.

Appendix D

Relationship to Other Designs

The activity rules suggested here have been designed to conform to the format in the PX Protocol for California, which specifies a sequence of iterations, each of which concludes with a single clearing price for each hourly market, but only the last iteration produces binding transactions. The principle of "revealed preference" that characterizes these activity rules has not previously been used in any existing market. Consequently, it brings the risks inherent in any innovation. The principle alternative is also novel but it has had a longer history of experimental studies conducted at the University of Arizona. The principal reference to this work is by McCabe, Rassenti, and Smith, "Designing a Uniform-Price Double Auction," Chapter 11 in Friedman and Rust (eds.), *The Double Auction Market*, Addison-Wesley, 1993.

The uniform-price double auction (UPDA) devised by these authors is a call market. At any time before the hourly markets are "called" each trader can submit binding orders in the form of bid or ask prices for single-unit trades in each hourly market. The orders in each market are recorded in a book, which is divided between those orders tentatively accepted and rejected. Whenever those tentatively rejected include a bid no less than an ask, this pair is tentatively accepted. Moreover, a new or revised order, say an ask, can displace one that is tentatively accepted if it offers a lower price, provided it satisfies an "update rule." When the market is called, the tentatively accepted orders are accepted, and all transactions are made at the same clearing price – which in conformity with the PX Protocol is the greater of the highest accepted ask and the highest rejected bid.

During the auction, each trader can submit a new order or revise a previous order, provided the revised price improves the price asked previously. In addition, a key element of their construction is the update rule: for a new or revised order to displace one that has been tentatively accepted, it must meet the best terms on the other side of the market. Thus, if it is an ask then it must be as low as the highest rejected bid. The effect of this rule is to simulate the effect of a dynamic bid-ask market in which binding bilateral transactions occur continuously.

This auction design is a viable alternative, especially if too many iterations are required for the current design to converge. Its principal advantage is that it assures that the market closes at a specified time. This might entail an intense flurry of offers in the closing minutes, however, so software must cope with a large volume of orders in a short time. Whether it would enable suppliers to choose their unit commitments efficiently has not been studied, but it is clear that rapid price movements in the closing minutes pose the chief risk to operating efficiency. Because this design includes no activity rules, there is no assurance that reliable price discovery occurs early. In effect, impatience to submit orders is induced entirely by the risk of being unable to submit a profitable order amidst the flurry in the final minutes. In the Arizona experiments, the market is typically called after four or five minutes and the number of subjects is small, which differs considerably from the context of a PX.

There are also other auction designs closer to those studied here, but none of these includes an analogue of the Exclusion Rule, which in my judgment is essential to reliable price discovery.