Design of Ancillary Service Markets[1]

Shmuel S. Oren,
University of California at Berkeley
Berkeley, CA 94720
Oren@IEOR.Berkeley.edu

Abstract

We examine the design of bid selection protocols and settlement rules in ancillary service markets. Such markets are typically operated by an independent system operator (ISO) for competitive procurement of reserves that are needed to ensure the secure operation of a competitive electric power system. Reserve types are characterized *in terms of response time and they are downward substitutable (faster responding reserves can replace slower ones). We explore how this substitutability is accounted for in alternative market protocols and we analyze the efficiency, distributional aspects and incentive compatibility of such protocols.*

1. Introduction

A common aspect shared by the various designs of restructured electricity markets in the US and around the world is the designation of a system operator that is responsible for the reliable real time control of the transmission system that enables operation of a competitive energy market. In California as in PJM, NYPP, New England, and the proposed ERCOT market this role is performed by an Independent System Operator (ISO) that is responsible for real time load balancing, congestion management and provision of ancillary services. These services include voltage support, black start capability, AGC and reserves with varying levels of response time. The precise definition of ancillary services varies across different restructured systems and so is the design of markets for competitive procurement and provision of such services. In many markets such as California, New England, NYPP and ERCOT separate ancillary service products have been defined which are procured by the ISO through an auction on behalf of market participants so as to meet NERC reliability standards. The proper product definition and design of ancillary service markets are the primary determinants of efficiency and liquidity in these markets that in turn influence system reliability. This lesson was learned in California through the poor performance and low liquidity of its ancillary service market that necessitated price caps and other forms of direct intervention by the ISO.

In New England, according to the New England ISO market report for June 1999 "During June and July the ISO identified design flaws in the operating reserve markets of Ten Minute Spinning Reserve (TMSR), Ten Minute Non-Spinning Reserve (TMNSR), and Thirty Minute Operating Reserve (TMOR). The operating reserve design flaws resulted in markets that are not workably competitive." The report identifies *price reversals* as the primary symptom of the design flaws. It States: "In a competitive market, higher quality goods should command higher prices. In the operating reserves markets prices did not appear to be reflective of costs and lesser quality reserves received higher prices. For example, Thirty Minute Operating Reserve ("TMOR") is less useful than Ten Minute Non-Spinning Reserve ("TMNSR"), Ten Minute Spinning Reserve ("TMSR"), or Energy. However, at times, TMOR prices exceeded the prices for all three of the other products." These flaws resulted in numerous administrative interventions overriding the market prices and eventually in the temporary suspension of the ancillary service market operations in New England. Likewise in New York, NYPP filed on March 27, 2000 an emergency request to FERC to suspend market based pricing of 10-Minutes Reserves due to market failure that was attributed to exercise of market power.

Market reforms aimed at repairing the ancillary service markets are under way in California, New England and New York. A key feature of these various markets is the cascading nature and downward substitutability among the various reserve types. (i.e., faster response reserves are regarded as higher quality and can substitute for lower response reserves). Hence, price reversals where lower quality reserves get a higher price are generally considered undesirable consequences of a market design that may create perverse incentives and misrepresentation of bids. However, the remedies proposed or adopted by the different systems vary with emphasis on the specific problematic issues in each system. In New York the main focus of the reform is on reducing market power of the ancillary service suppliers by recognizing the congestion problems between east and west that limit competition in the 10-minutes reserve market. In New England the focus of the reform is on introducing price elasticity in the reserve markets and modifying the dispatch protocol in order to neutralize the gaming that plagued those markets. In California the

[1] The work described in this paper was coordinated by the Consortium for Electric Reliability Technology Solutions on behalf of the Department of Energy.

0-7695-0981-9/01 $10.00 (c) 2001 IEEE

primary objectives of the reform have been to improve liquidity and efficiency in the ancillary service markets and reduce procurement cost. Key elements of the California reform have been to move from a sequential to a simultaneous procurement auction for the various services and to allow the ISO to exploit the substitutability among the different reserve types.

Specifically, the California ISO (CAISO) needs to procure four types of reserves: Regulation (RG) for AGC, Spinning reserves (SP) that are synchronized and available within 10 minutes, Non-Spinning reserves (NS) that are not synchronized but can be made available within 10 minutes and Replacement reserves (RS) that can be made available within 60 minutes. These products are, for practical purposes, hierarchically substitutable. Regulation resources can also provide SP, NS and RS. Likewise spinning reserves can provide NS and RS whereas non-spinning reserves can provide RS. In the initial market design the auction for the four reserve products was conducted sequentially from the highest to the lowest quality. Generators were allowed to rebid in each round their uncommitted resources at new prices that could exploit thin markets in the lower quality reserves. As will be illustrated below, even without market power, such sequential markets are susceptible to price reversals and create perverse incentives. The CAISO reform attempts to stabilize prices by permitting generator to submit a single bid specifying reserve type, quantity, and price whereas the ISO can use any of the procured resources to meet demand for any of the reserve products that a resource can provide (this is often referred to as cascading bids). So for instance the ISO may procure spinning reserves and use them to meet the need for replacement reserves.

Within the framework of a simultaneous multiproduct auction with possible cascading of bids, there are several alternatives for organizing the auction, that define how winning bids are selected and how payoffs are determined. These alternatives have diverse efficiency and distributional implications, which have been recognized and debated by the ISO board. The main objective of this paper is to analyze the implications of these design alternatives with respect to efficiency, distribution of gains between buyers and sellers and incentive compatibility. The scheme that has been adopted in California employs a "rational buyer" bid selection criterion that minimizes total procurement cost and pays to winning bids of each reserve type (as declared by the bidder) a uniform market clearing price set by the highest accepted bid of that type. In New York, on the other hand the bid selection is based on minimizing social cost as reflected by the bids and market clearing prices for each reserve type are set to the *marginal value* of that reserve (that equals the highest accepted bid for all reserves of equal or lower quality). We will reexamine the implications of these designs and of other design options considering bidders' potential strategic behavior that is affected by the auction design and by the opportunities available to the generation owners.

In addition to the auction design on the procurement side we also address the question of how to price the ancillary service products to the users. Again several options are available with diverse efficiency and revenue adequacy implications. Efficiency of the price signals will affect proper use of resources and the potential for inefficient bypass of the ISO market through self-provision. Revenue adequacy, on the other hand, will determine whether the total payment by customers match the procurement cost or whether a shortfall or surplus is possible. For instance, the current pricing rule adopted by the CAISO board does not meet the revenue adequacy criterion and inefficient price signals due to price reversals are possible.

2. Review of Relevant Literature

The bidding literature addressing multipart and multi-product auctions is quite young. A comprehensive survey of that literature is provided by Klemperer [1]. A significant body of work in this area was spurred by the FCC spectrum auction and the multitude of interesting questions they have raised. The two major camps in this area are those advocating combinatorial auctions that employ combinatorial optimization algorithms vs. those favoring simultaneous multiround auctions with activity rules for the individual components. The latter approach, which has been adopted in the early phases of the FCC spectrum auction, is justified on the grounds of the computational complexity and lack of transparency involved in a combinatorial auction. However, as illustrated by Rothkoph, Pekee and Harstad [2] these objectionable aspects of combinatorial auctions may be irrelevant when their special structure can be exploited. Background information and discussions as well as theoretical and empirical studies related to the FCC auction have been reported in Milgrom [3] MacMillan [4], Crampton [5], McAfee and McMillan [6] Ausubel and Crampton [7] Issac and James [8].

Literature focusing on bidding in the context of electricity markets initially addressed PURPA auctions for IPP contracts. Work examining electric power auctions by Kahn, Rothkopf, Eto and Nataf [9] examined efficiency aspects of the acquisition process but did not treat cases where bids were curtailable. Stoft and Kahn [10] have examined questions concerning "bias" in scoring of curtailable bids but there has not been an explicit treatment of the trade-off between fixed and variable price component. Bushnell and Oren [11], [12] addressed the question of whether bidders in the Biennial Resource Planning Update (BRPU) auction mandated by the California Public Utility Commission (CPUC) for procurement of IPP capacity by the utilities, will be induced to reveal their "true" fixed and marginal cost of generation. Such information was essential for efficient procurement and dispatch of procured resources. Their analysis predicted that the proposed CPUC scoring rule is likely to induce understatement of marginal generation cost (which indeed happened). The outcome of that auction, which was eventually voided by FERC due to "gaming behavior" by the bidders are discussed in an article by Gribik [13]. Bushnell and Oren [11] also propose a Vickery auction type scheme for which true revelation of marginal costs is a dominant strategy for all bidders.

Von der Fehr and Harbord [14] is the first paper we are aware of that analyzed a multi unit auction for electricity in the context of the UK power pool. The UK old design like the later market designs adopted in PJM, NYPP and New England have preserved the central unit commitment aspect of the vertically integrated utilities by means of a multipart auction in which bidders provide the necessary inputs for central unit commitment optimization. Unfortunately multipart auctions are not well understood and with few exceptions that we will mention below have limited theoretical foundation that would enable a practical incentive compatible design. Johnson, Oren and Svoboda [15] illustrated some inherent difficulties arising from central unit commitment in a competitive environment with dispersed generation ownership. The difficulties are due to indeterminacies of the optimal unit commitment solution and basic incompatibility between competitive behavior and truth telling due to integer affects and non-convexities. Wilson [16] and Elmaghraby and Oren [17] have proposed alternative designs of one-part auctions that enable bidders to internalize non-convexities due to fixed cost. Hobbs et al. [18] studied an incentive compatible multi-part electricity auction based on the Vickery-Clarke-Grove mechanism. Unfortunately, as noted by the authors, this approach has limited practical value due to revenue deficiency and the discriminatory nature of the payment scheme.

Reserve markets can be viewed as special types of multi-part auctions were generators compete for the provision of reserves by submitting two-part bids consisting of capacity and energy prices. For that special case, Chao and Wilson [19] were able to develop an incentive compatible design. In their proposed scheme generators submit separate capacity and energy bids. The energy bids are used in case it is necessary to call the reserves to supply energy and all dispatched energy is remunerated uniformly at a price set by the highest dispatched energy bid. They proved that under this setup a scoring rule that ranks reserves bids based only on their capacity component and pays all accepted bids a capacity price set to the highest accepted capacity bid is incentive compatible, i.e., bidders will be induced to reveal their true marginal energy cost as well as their capacity cost (that may be an opportunity cost for forgone profit from sale of energy.)

The problems associated with the California ancillary service market that led to the "rational buyer" reform are documented in the Annual report of the California ISO market surveillance committee [20] and discussed by Laura Brien [21]. Some of the deliberations concerning the ISO ancillary service market reform have been documented in unpublished memoranda. An analysis of alternative market designs for ancillary services is contained in the Griffes [22] report to the California PX that is considering establishing its own internal ancillary service market for the purpose of self-provision. Non of these discussion explicitly address bidders strategic response to the market design consequences.

3. Design Options

As discussed earlier, an important aspects of ancillary services is their hierarchical nature that allows substitution of a high quality reserve for a lower quality one. Both social efficiency and rational procurement behavior dictate that such substitution should be allowed. In a perfectly competitive market such substitution would occur naturally in a sequential auction (from high to low quality) since bidders would rebid their rejected bids in the subsequent auctions for which their resources are eligible. In principle, assuming bidders bid and rebid their true cost, such a sequential auction would lead to socially efficient procurement. In the absence of market power uniform market clearing prices in each auction will indeed induce bidders to bid their true cost. Unfortunately, a sequential auction with independent uniform market clearing prices in each round may result in price reversal, i.e., the market clearing price for a high quality resource (e.g. regulation) may be lower than that for a lower quality resource (e.g. spinning reserves). Indeed such price reversals have been observed in California and in New England ancillary service markets.

Price reversals pose serious incentive compatibility problems since even price taking generators anticipating such reversal may be induced to understate their capability and wait for a later round of the sequential auction that is expected to fetch a higher market clearing price. With market power the situation is exacerbated as losing bids in early rounds may raise their bids in subsequent round when they perceive potential scarcity of bids. That type of low liquidity and high prices for low quality reserves has been observed in California and has motivated the proposed reforms. The following example is used to illustrate the possible price reversal described above. For simplicity we will only consider two type of reserves (RG and SP) with demand of 500MW for each reserve type and assume that bidders have no market power so that in a uniform price auction they bid their true cost. Let us assume that the following bids have been submitted in a sequential uniform price auction and the rejected RG bids in the first round are rebid in the second round at the same price:

RG	600 MW at $10/MW	100MW at $15/MW
SP	200MW at $5/MW,	300MW at $20/MW

In the first round (RG auction) the ISO procures 500MW of RG at $10/MW which sets the market clearing price for the RG auction at $10/MW. The remaining 100MW RG at $10/MW and 100MW RG at $15/MW are rebid for SP at the same price.

In the second round (SP auction) the ISO take 200MW of SP resources at $5/MW then takes the rebid 100MW RG at $10/MW and 100MW RG at $15/MW and finally

100MW SP at $20/MW. The market-clearing price for SP is therefore $20/MW which is paid to all the reserve capacity procured in the second round. The resulting total procurement cost is $15,000 and the social cost (assuming bid reveal true cost) is $10,500. While the procurement resulting from this approach is socially efficient we notice that the inferior reserves (SP) are paid twice as much as the superior reserves (RG). Furthermore, the losing bids in the first round end up being paid more than the winning bids. Clearly, this price reversal creates perverse incentives for withholding bids or raising prices in the RG auction. This is true even for resources that have no market power to affect market clearing prices but simply try to take advantage of intermarket arbitrage opportunity by selling their product in the auction that will fetch the best price. In reality bidders in the California ancillary service market also exploited market power due to low liquidity and raised their bids in the subsequent auctions.

The proposed reform to the California ancillary service markets replaces the sequential auction with a simultaneous auction were each resource submits a single bid specifying reserve type, capacity bid and energy price if called. The ISO is allowed to substitute demand for a lower quality reserve with a higher quality and thus, use higher quality resources to meet demand for lower quality reserves. Similar schemes are used in NYPP and were recommended for New England by Crampton and Lien [23].

Within the above simultaneous auction framework there are several degrees of freedom in the market design:

- The bid selection objective function
  - Minimum social cost
  - Minimum procurement cost
- Settlement rule
  - Pay uniform price based on bid type (demand substitution)
  - Pay uniform price based on usage (product substitution)
  - Marginal value pricing
  - Pay as bid
- Pricing of the products to buyers
  - Set product price to highest accepted bid of that type
  - Set product price to highest price paid to meet product demand
  - Marginal value pricing

Much of the deliberations in the California ISO focused on the choice of objective function for implementing the rational buyer approach. The implication of that choice depends on the settlement rule. Paying resources based on usage (product substitution) is a natural choice in sequential auctions where rejected resources in one auction can rebid or are carried over to the next. In such a case the payments are based on the clearing price of the auction in which a resource wins. Under such a settlement rule, however, if a rejected bid in an early auction is accepted in a later one (assuming no bid change) it follows that the clearing price in the later auction (for lower quality reserves) will be higher than the predecessor auction in which the bid was rejected. This implies price reversal where lower quality products are paid more than the higher quality ones and as discussed above such reversal creates perverse incentives for misrepresentation of quality. Avoiding such situation is one of the main arguments in favor of a simultaneous auction. However, a simultaneous auction in which resources are paid based on usage is de facto equivalent to a sequential auction in which rejected bids are carried forward. Therefore, we will not consider that option any further.

The remaining settlement rules are uniform payment based on bid type, Uniform payment based on marginal cost and pay-as-bid (PAB). To date, a PAB approach has been adopted only in the UK under the New Electricity Trading Arrangements (NETA). Some of the implications of that approach with respect to the balancing market are explored by Green and McDaniel [24], [25] and the advantages of PAB auctions in avoiding collusive behavior are discussed by Klemperer [26]. Under the uniform pricing options the two possible bid selection objectives with demand substitution can lead to different dispatch results. In particular, the procurement cost minimization option adopted by the CAISO may result in a socially inefficient dispatch and in price reversal as will be shown below. We will also show that even if bids are selected so as to minimize social cost price reversals could still occur if procurement prices are not set to marginal value.

4. Simultaneous Auction with Uniform Pricing and Social Cost Minimization.

To simplify notation we will assume four hierarchically substitutable reserve type (regulation, spinning reserves, non spinning reserves and replacement reserves) and denote them by an index i=1,2,3,4, respectively. The subsequent analysis, however, generalizes to any number of hierarchically substitutable products. We assume that each resource is bid as a quantity and price pair and designated for a particular reserve type. Thus, it is possible to rank all bids for a particular resource type i according to price and obtain a supply function $p_i(q_i)$ specifying the market clearing price (or marginal price) that will enlist capacity q_i of resource type i. This function will have the form of an increasing staircase rising vertically at the maximum capacity level bid for that resource type.

We introduce the additional following notation:

$$Q_i = \sum_{j=1}^{i} q_j - \left\{ \begin{array}{l} \text{Cumulative acquired capacity} \\ \text{of resource type i or better} \end{array} \right.$$

d_i - Demand for reserve type i

$$D_i = \sum_{j=1}^{i} d_j - \left\{ \begin{array}{l} \text{Cumulative demand for} \\ \text{reserve type i or better} \end{array} \right.$$

The uniform price rule incents bidders with no market power to bid true cost. Hence, in the absence of market power social cost is minimized by minimizing the cost (as bid) of the dispatched resources. Thus, the minimum social cost problem can be written in the mathematical form:

$$\underset{q_1,q_2,q_3,q_4\geq 0}{Min}\sum_{i=1}^{4}\int_0^{q_i} p_i(\varphi_i)d\varphi_i$$
$$s.t\ \ Q_i \geq D_i \ \text{ for } i=1,2,3,4$$

Furthermore, since $p_i(q_i)$ is a staircase function the above can be formulated as a simple Linear Program. Let (p_{ik}, q_{ik}) denote the price and quantity of the *k-th* block of type i where $k=1,....,N(i)$ (sorted from the lowest to highest price) and let x_{ik} denote the quantity of block *ik* selected. Then the bid selection problem is

$$\underset{x_{ik}}{Min}\sum_{i=1}^{4}\sum_{k=1}^{N(i)} p_{ik}x_{ik}$$
$$s.t\quad \sum_{j=1}^{i}\sum_{k=1}^{N(j)} x_{jk} \geq \sum_{j=1}^{i} d_j \ \text{ for}$$
$$0 \leq x_{ik} \leq q_{ik} \ \text{ for all } i,k$$

If there exist a feasible solution to the above LP then it can be solved by a simple "greedy algorithm" that successively fills the demand for each service from highest to the lowest quality using bids in order of increasing price and pushing any unused bids to the next level. So bids of type 1 are selected in increasing order of bid price until the cumulative quantity reaches the demand for service type 1. The remaining bids of type 1 are pushed forward and mixed with the type 2 bids and the selection out of the combined pool is again done in increasing order of bid prices until the cumulative quantity reaches the demand for type 2 reserves. It is assumed that fractional bid quantities can be selected or split among different reserve types and that price ties at the margin are resolved by prorating the selected quantities of tied bids.

The dual of the above LP problem is:

$$\underset{\{y_i,w_{ik}\}}{Max}\sum_{i=1}^{4} y_i\sum_{j=1}^{i} d_j - \sum_{i=1}^{4}\sum_{k=1}^{N(i)} w_{ik}q_{ik}$$
$$s.t\ \sum_{j=i}^{4} y_j - w_{ik} \leq p_{ik} \ \text{ for } k=1,...,N(i),\ i=1,...,4$$
$$y_i \geq 0,\ w_{ik} \geq 0 \text{ for all } i,k$$

In this dual formulation, the variables $y_1,...,y_4$ represent the shadow prices corresponding to the cumulative demand constraints in the primal LP and the variables w_{ik} can be interpreted as scarcity costs corresponding to the respective bid quantity limits q_{ik}. From the dual objective it can be easily seen that the marginal value of an incremental unit of reserve type i is $MC_i = \sum_{j=i}^{4} y_j$ and each incremental unit of bid ik will reduce total cost by w_{ik}. One immediate implication of the above observations is that the marginal value of the different reserve types are monotonically increasing in quality, i.e., $MC_1 \geq MC_2 \geq MC_3 \geq MC_4$ so pricing each reserve at its marginal value will avoid price reversal. If the primal LP problem has no feasible solution, i.e., there are not enough bids to meet the demand for at least one reserve type, then the dual problem is unbounded and the marginal value of some reserve types may be infinite. Such a situation calls for some form of caps which can be formally introduced into the above formulation as an artificial bid of type 1 (highest quality) with unlimited quantity and a bid price that equals the desired cap. The artificial bid could be interpreted as a demand side bid in which case it would make sense to interpret the cap as shortage cost. With this interpretation one may wish to introduce an artificial bid for each of the reserve types with prices reflecting shortage costs or hypothetical demand side bids that vary by reserve type. The above modification will result in back propagation of scarcity rents due to shortages of any reserve type so that if the marginal cost of a reserve type reaches the cap so will the marginal costs of all the higher quality reserves.

Given the optimal bid selection (obtain by the roll forward algorithm) the corresponding marginal costs of each reserve type can be determined as follows. For each reserve type i let $\overline{p}_i$ be the price of the most expensive bid selected of that type (in case of shortage that price is set to the price cap for that reserve type). The Marginal value of any reserve type is the largest value $\overline{p}_i$ of all the reserve types of equal or lower quality i.e., $MC_i = Max_{j\geq i}\overline{p}_j$ (this rule is used for pricing reserves in NYPP).

As discussed above, there are three uniform pricing alternatives that can be implemented with the social cost minimizing bid selection approach. The product substitution approach that sets the clearing prices paid to each bid according to use is equivalent to the sequential auction discussed earlier which has incentive compatibility problems that could result in misrepresentation of costs and capability. The second alternative is to set the uniform clearing price for each bid type to the price of the highest bid selected of that bid type (demand substitution). The third approach is to set the uniform clearing price of each bid type to the marginal value of the corresponding reserve type (i.e., the marginal value of the highest quality reserve that a bid can provide). We illustrate the difference between the last two pricing schemes using the example described earlier.

The social cost minimizing bid selection for that case is:

RG - 600MW at $10/MW +100MW at $15/MW
(200MW used to meet SP demand)
SP - 200MW at $5/MW + 100MW at $20/MW

The social cost of that selection is \$10,500, however, the procurement cost will depend on the pricing rule. If market clearing prices for each bid type are set to the highest accepted bid of that type then the market clearing price for RG is \$15/MW whereas the market clearing price for SP is \$20/MW. The corresponding procurement cost is \$16,500. On the other hand, the marginal value for both RG an SP in this case is \$20 (since meeting an incremental MW of either reserve type will increase social cost by \$20). Thus, the procurement cost under marginal value pricing is \$20,000.

We observe that procurement cost has gone up under both pricing rules as compared with the sequential auction whereas the social cost is still the same. We also note that although the bid selection is socially efficient, price reversal can still occur so there are still perverse incentives for bidders to understate the quality of their product[2]. In order to avoid the price reversal and guarantee incentive compatibility one would have to set the uniform price of each reserve type to its marginal value. (as shown above those are guaranteed to be monotone). At first sight, however, such a remedy appears expensive from the consumer perspective and it cannot be justified without explicitly analyzing the potential consequences of strategic misrepresentation of bid types. As in classic Vickery auctions the efficiency and incentive compatibility is achieved at the expense of a transfer from buyers to sellers. Critics of that approach would argue that this policy achieves incentive compatibility by giving away to the sellers all they could hope to get through cheating.

The above considerations have led the CAISO to adopt a bid selection approach that explicitly focuses on minimizing procurement cost (without considering strategic responses by the bidders)[3]. As we will demonstrate below this goal is achieved at the expense of both social efficiency and incentive compatibility.

5. Simultaneous Auction with Uniform Pricing and Procurement Cost Minimization (Rational Buyer).

We introduce the additional notation:

$$P_i(q_i) = q_i p_i(q_i) - \left\{ \begin{array}{l} \text{Total payment to resource type i as a} \\ \text{function of acquired capacity of that type} \end{array} \right.$$

The least procurement cost problem can then be written in the mathematical form:

$$\underset{q_1,q_2,q_3,q_4 \geq 0}{Min} \sum_{i=1}^{4} P_i(q_i)$$
$$s.t \;\; Q_i \geq D_i \;\; \text{for } i = 1,2,3,4$$

The objective function in this problem is nonconvex. Furthermore, because of the discrete nature of the supply functions, this problem is combinatorial in nature and may have multiple local solutions. Hence, the problem is not easily solved by standard linear or nonlinear programming algorithms. Fortunately, however, the special structure of the problem allows us to formulate it as a simple Dynamic Programming problem whose solution guarantees a global minimum.

$$\underset{D_4 \geq Q_3 \geq D_3}{Min} \left[\begin{array}{l} P_4(D_4 - Q_3) \\ + \underset{Q_3 \geq Q_2 \geq D_2}{Min} \left[\begin{array}{l} P_3(Q_3 - Q_2) \\ + \underset{Q_2 \geq Q_1 \geq D_1}{Min} \left[\begin{array}{l} P_2(Q_2 - Q_1) \\ + P_1(Q_1) \end{array} \right] \end{array} \right] \end{array} \right]$$

This is a four stage deterministic dynamic program (number of stages equals the number of product types) that will determine the optimal quantities of each of the resource types that will minimize total acquisition cost to the ISO for meeting the demands for the four reserve types. The stages in this DP formulation represent the subsequent resource types in the hierarchy while the states are the cumulative amount of resources acquired at each stage. The resource quantities are discretized using an appropriate increment for the required precision. The computation time will depend on that discretization. The solution involves one forward and one backward pass. In the forward pass we start with the most inner minimization first computing the cost of acquiring any feasible quantity of regulation capacity in the range between the demand for regulation and the combined demand for all services. Next we compute the least cost feasible mix of regulation and spinning reserves resources for any total amount of the two in the range between the combined demand for regulation and spinning reserves and the demand for all services combined, subject to the constraint that the regulation capacity exceeds the demand for regulation. Similarly we then compute the least cost feasible mix of the first two resources and non-spinning reserves for each possible total amount of the three and so on. In the backward pass we start with the total amount of the four ancillary services and trace back the least cost path from which we can extract the optimal procured quantity of each resource type. Once we have these quantities we can use the supply functions to determine the corresponding market clearing prices.[4]

[2] Under social cost minimization, price reversals are due to the steps in the supply function and the limited quantities of each bid. Such reversals would not occur if the supply functions for the different reserve types were continues and unbounded.

[3] Some have rationalized this bid selection criteria on the grounds that in enables the ISO to function as a countervailing force to suppress the exercise of market power by the suppliers.

[4] The solution to the above dynamic program can be implemented in a simple spreadsheet even for realistic size problems. For example, if the combined demand for spinning, non-spinning and replacement reserves is 6000MW and we measure quantities in 1 MW increments, the calculations can be contained in a 6000 rows by 3 column spreadsheet which is well within the capability of EXCEL (the amount of regulation capacity required does not affect the computational complexity of the problem). The fill of the spreadsheet depends on the demands for the individual services. Suppose for example that the requirement for each of the reserves is 2000MW then the solution involves the computation

We will again illustrate the implication of this approach using the simple numerical example introduced earlier. The least procurement cost bid selection is:

RG - 600MW at $10/MW

SP - 200MW at $5/MW + 200MW at $20/MW

The resulting market clearing price for RG is $10/MW whereas the market-clearing price for SP is $20/MW. The corresponding social cost is $11,000 while the procurement cost is $14,000. We note that the outcome is less efficient than the sequential auction and the social cost-minimizing auction but the procurement cost is lower then in the other two designs. There is a clear tradeoff here between efficiency and procurement cost. While it would be more efficient to use 100MW RG at $15/MW to replace 100MW of SP at $20/MW, such substitution would raise the market-clearing price of RG from $10/MW to $15/MW resulting in a net increase of $2,500 in total procurement cost. Therefore, the rational buyer protocol forgoes such efficiency improvement in order to lower procurement cost. We also note that this approach may lead to price reversal which again may incent sellers to understate the quality of their resources and that may lead to shortages of high quality reserves[5]. Kamat and Oren [27] present a systematic analysis of equilibrium bidding strategies for hierarchically substitutable products under demand uncertainty and a Rational Buyer protocol. They show, for a stylized model, that in equilibrium some bids will understate their resource quality (declare themselves as SP rather than RG). When high quality resources are limited, such misrepresentation will result in random shortages of the high quality reserves and price spikes that wouldn't occur had the bids revealed their true quality.

6. Simultaneous Auction with Pay-As-Bid Settlement.

The efficiency difference between the minimum procurement cost criterion and minimum social cost criterion is a consequence of the payment rule that sets a uniform price for each bid reserve type to the highest accepted bid of that type or to marginal cost. That discrepancy would disappear under a pay-as-bid (PAB) rule. Under such a rule bidders will no longer have an incentive to misrepresent their reserve type but do have an incentive to bid above marginal cost. However, under a rational expectations assumption and common knowledge of the total supply function (which is not unreasonable in a daily repeated market[6]) rational bids in a PAB auction are monotonically increasing in true costs. Hence the least social cost objective and the least procurement cost objective are aligned in a PAB auction. Bid selection in the PAB auction is based on solving the LP problem presented in the social cost minimization case, which in fact minimizes bid cost. Although the value of the objective function no longer represents social cost, the minimization of bid cost (that equals procurement cost) will produce the socially optimal bid selection (under certain assumptions on information structure)

The idea of using a PAB rule in the ancillary service market has not been considered in any of the US markets but it is investigated by Kamat and Oren [27]. The New Electricity Trading Arrangements (NETA) in the UK adopted a PAB approach in the balancing market (starting 3.5 hours before delivery) as a way of dealing with the heterogeneity of generation resources close to real time. A PAB approach may be advantageous as we get closer to real time since it eliminates the need to differentiate the energy commodity (by location, ramp rate, flexibility, etc.) and create separate (and often illiquid) markets that are settled at uniform market clearing prices. Instead, each resource is dealt with as unique but partially substitutable and dispatched so as to minimize total cost subject to operational constraints. It is up to the bidders to figure out whom they are competing with and how much they can charge (which is not an unreasonable task in a market that is repeated daily).

PAB auctions have the advantage of reducing opportunities for collusive behavior as articulated by Klemperer [26]. Furthermore a PAB settlement rule eliminates the incentive for bidders in a multiunit auction to employ a "sacrificial lamb" strategy or "last man out" strategy. These strategies risk high bids on marginal units hoping to set a high clearing price for the inframarginal bids. In a PAB auction each bid is bearing its own risk and rewards. Finally, since in a PAB auction there is an incentive for bidders to absorb some of the quantity risk, the average prices will exhibit lower volatility (with uniform clearing prices all the volatility due to demand uncertainty is reflected in the clearing price).

7. Pricing Ancillary Services to the Buyers.

While the cost of certain ancillary services such as reactive power and black start capability are recovered through an uplift charge, reserve requirements are typically attributed to the loads that must either self-provide their

---

of 12,000 quantities involving simple comparison operations. The worse case fill occurs when there is no demand for spinning and non-spinning reserves so all the resources compete to provide replacement reserves. In that case the spreadsheet is full and 18,000 entries need to be computed.

[5] Price reversals have indeed been reported in the CAISO ancillary service market: On March 20, 2000 in [hour]HE 19, there was a spike in the Replacement Reserve price. The following are published prices for [location] NP15:
Reg-Up: P1 = 18.44 $/MW
Spin: P2 = 35.97 $/MW
Non-Spin: P3 = 18.00 $/MW
Replacement: P4 = 198.98 $/MW
Note that P4 > P3, P2, P1. Also P2 > P1.
A similar pattern in AS prices was observed on Feb 29,2000 in [hour] HE18 when replacement reserve prices were 122$/MW.

[6]Further implications of the daily repetition are discussed by Rothkopf [28]

reserves or purchase them from the ISO. Thus, the ISO must set selling prices for the reserves it procurers. There are several, often conflicting, objectives that guide the determination of the reserves selling price including efficiency, revenue sufficiency and equity. The procurement auction design also affects the options and choices of product pricing approaches. We will examine the various objectives and pricing options in the context of the various procurement auctions discussed above.

Efficient price signals are necessary in order to induce efficient utilization of resources and prevent inefficient bypass of the ISO reserve market through self-provision. The ideal pricing scheme in term of efficiency is to price each reserve at the marginal social value as determined by the LP formulation described in Section 4. In principle such pricing is possible regardless of how resources are procured. However, such pricing makes sense only in the context of the social cost minimization bid selection with marginal value based purchase prices. In that case LP duality guarantees revenue sufficiency (sales revenue equals purchase cost) and there are no bypass incentives.

In general there are two other viable alternatives for product pricing in conjunction with uniform price procurement auctions. The ISO can sell each reserve type at the price of the most expensive bid used to provide that reserve type (regardless of resource type) or sell each reserve type at the market-clearing price for the corresponding resource type. The latter approach has been adopted by the CAISO in conjunction with the Rational Buyer procurement protocol. Clearly the two pricing methods converge to marginal cost pricing when procurement is based on social cost minimization with marginal value payments. However, they differ in unpredictable ways when other procurement mechanisms are used. Since quantities purchased and sold at each price will differ, revenue adequacy cannot be assured. Furthermore, price reversals that incent suppliers to understate the quality of their resource will also incent inefficient bypass of the ISO market. For instance if regulation clearing price is lower than that of spinning reserves as in our example, some load will find it attractive to employ a resource that could provide regulation in order to self-provide its spinning reserves obligation.

In our numerical example, under the least social cost auction, with procurement prices for each bid type set to the last accepted bid price, the cost of the highest bid selected to provide RG is $10/MW and the highest bid selected to provide SP is $20/MW. Hence setting the selling prices to these levels will produce $15,000 in revenue, which is $1500 short of the procurement cost of $16,500. On the other hand the market-clearing price for RG is $15/MW and for SP is $20/MW. Setting the selling prices to these market-clearing prices will produce $17,500 in revenue which creates a $1000 surplus.[7] Revenue shortfalls may also occur under the Rational Buyer protocol and the CAISO pricing rule that set the selling price for each reserve type to the market clearing price of the corresponding resource type. Suppose that there is no price reversal and that the marginal regulation resources are selected under the Rational Buyer protocol to meet demand for spinning reserves (to avoid a shortage or replace more expansive spinning reserve bids). Then under the CAISO pricing rule the selling price for spinning reserves will not cover the procurement cost of the regulation resources that are used as spinning reserves but paid the higher regulation clearing price.

With a PAB auction, The natural pricing scheme is to price each reserve type at the marginal procurement cost of that type. These marginal procurement costs can be determined analogously to the marginal social cost computation described in Section 4. Thus, given the bid selection, the marginal procurement cost for each reserve type is given by the highest priced bid selected of that bid type or any lower quality bid type. This rule ensures price monotonicity. Furthermore, The LP duality theorem guarantees that sales revenues under these scheme will match procurement cost. However, because of the discriminatory nature of a PAB auction it is impossible to eliminate profitable bypass of the ISO market reserve market through bilateral arrangements for self-provision. In order to insure liquidity in the ISO reserve market and avoid "cherry picking" of cheap resources by influential buyers, it may be necessary, to prohibit self-provision when reserves are procured through a PAB auction.

8. Conclusions.

Practical experience and theoretical analysis demonstrate the prevalence of price reversals in ancillary service markets where lower quality reserves clear at prices higher then the higher quality reserves. Such phenomena create perverse incentives that will induce rational bidders to misrepresent their reserves and potentially create shortages in of high quality reserves. Price reversals can be avoided by a market design in which bids are selected to minimize bid cost and selected bids are paid the marginal value corresponding to their bid "quality". This approach is socially efficient and incentive compatible but has undesirable distributional implications that raise procurement cost. Consequently, some systems such as CAISO have opted for market designs that are susceptible to price reversals in order to reduce procurement cost. Pay as bid (PAB) auctions eliminate the discrepancy between social cost and procurement cost minimization in bid selection and have several attractive features that have led the reformed UK system to adopt that approach. However, more theoretical and experimental

[7] Under the Rational Buyer protocol this example produces the same revenue, which happens to match the procurement cost with both pricing rules, but this is not true in general.

analysis are needed to asses the merit of the PAB approach in the context of ancillary service markets in the US.

Finally, it should be emphasized that even a good market design can be defeated by market concentration (i.e., market power) and lack of demand response, which will result in noncompetitive pricing and inefficient allocation of resources

9. References.

[1] Klemperer, Paul, "Auction Theory: A Guide to the Literature", Journal of Economic Surveys, Vol 13(3), July 1999, pp. 227-286.

[2] Rothkopf, M. A. Pekee, and R. Harstad, "Computationally Managable Combinatorial Auctions," Management Science, Vol. 44, No. 8, pp. 1131-1147, (1998)

[3] Milgrom, P. "Putting Auction Theory to Work: The Simultaneous Ascending Auction," Maryland Auction Conference, (May 29-31, 1998)

[4] McMillan, J., "Selling Spectrum Rights," Journal of Economic Perspective, Vol. 8, No. 3, pp. 145-162, (1994)

[5] Crampton, P. "Money Out of Thin Air: The Nationwide Narrowband PCS Auction," Jornal of Economics and Management Strategy, Vol. 4, No, 2 pp. 267-343 (1995)

[6] McAfee, P. and J. McMillan, "Analyzing the airwave Auction," Journal of Economic Perspective, Vol. 10, No. 1, pp. 159-175, (1996)

[7] Ausubel, L., and P. Crampton, "Demand Reduction and Inefficiency in Multi-unit Auctions", Working Paper No. 96-07, University of Maryland (1996).

[8] Isaac, R., D. James, "Robustness of the Incentive Compatible Combinatorial Auction," Maryland Auction Conference, (May 29-31, 1998).

[9] Kahn, E., Rothkopf, M., Eto, J., and Nataf, J.M., "Auction for PURPA Purchases: a Simulation Study," Journal of Regulatory Economics," Vol. 2, pp. 129-149, (1990).

[10] Stoft, S., and Kahn, E., "Auction Markets for Dispatchable Power: How to Score the Bids," Journal of Regulatory Economics, Vol. 3, pp.. 275-286 (1991)

[11] Bushnell, J. and S. S. Oren, "Bidder Cost Revelation in Electric Power Auctions," Journal of Regulatory Economics, V. 6, No 1 (1994), pp. 1-22.

[12] Bushnell, J. and S. S. Oren, "Two-dimensional Auctions for Efficient Franchising of Independent Power Generation" Proceedings of the 29th Annual Conference of the Operational Research Society of New Zealand, , Auckland University, New Zealand, (August 1993).

[13] Gribik, Paul: "Learning from California's QF Auction," Public Utilities Fortnightly, (April 15, 1995)

[14] Von der Fehr and Harbord, "Spot Market Competition in the UK Electricity Industry", *The Economic Journal*, 103, pp 531-546 (1993)

[15] Johnson B. Raymond, Shmuel S. Oren, Alva J. Svoboda, "Equity and Efficiency of Unit Commitment in Competitive Electricity Markets.", Utilities Policy, Vol. 6. No. 1, (1997) pp.9-19

[16] Wilson , R. "Activity Rules for The California PX Electricity Auction", Presented at the UCEI POWER Conference, Berkeley, California, (March 1998).

[17] Elmaghrabi Wedad and Shmuel S. Oren, "Efficiency of Multi-Unit Electricity Auctions", Energy Journal, 20, 4, pp. 89-115 (1999).

[18] Hobbs, B., M. Rothkopf, L. Hyde and R. O'Neill, "Evaluation of Incentive Compatible Auction for Energy Markets with Nonconcave Benefits," Working Paper, Presented at the INFORMS meeting, Seattle, Washington (October 1998).

[19] Chao, H. P., and R. B. Wilson, "Multi-Dimensional Procurement Auction for Power Reserves: Incentive-Compatibility Evaluation and Settlement Rules",Presented at the UCEI POWER Conference, Berkeley California, March 1999 (Revised October 1999)

[20] California ISO Market Surveillance Committee Annual Report (June 1999)

[21] Brien, Laura, Why the Ancillary Services Markets in California Don't Work and What to do About It," The Electricity Journal, 12(5): 38-49. (1999)

[22] Griffes, Peter, "A report to the California Power Exchange: The benefits of a simultaneous vs. sequential PX market for energy and ancillary services", (March 1999)

[23] Crampton Peter and Jeffrey Lien, "Eliminating Flaws in New England's Reserve Markets", Presented at the UCEI POWER Conference, Berkeley, California, (March 2000).

[24] Green Richard and Tanga McDaniel, " Expected Revenues in the Balancing Market: Equivalence between Pay-as-Bid and SMP", Working Paper, (February 24, 1999)

[25] Green Richard and Tanga McDaniel, " Modeling RETA: A model of forward trading and the balancing mechanism ", Working Paper, (April 16, 1999)

[26] Klemperer Paul, "What Really Matter in Auction Design", Working Paper, (June 21, 2000) www.nu!.ox.ac.uk/economics/people/klemperer.htm

[27] Kamat Rajnish and Shmuel Oren, "Rational Buyer Meets Rational Seller: Reserve Market Equilibrium Under Alternative Auction Designs", Working Paper, UC Berkeley, (June 2000)

[28] Michael H. Rothkopf, "Daily Repetition: A Neglected Factor in the Analysis of Electricity Auctions," The Electricity Journal 11, pp. 60-70, April 1999.