Auction Basics for Wholesale Power Markets: Objectives and Pricing Rules

Leigh Tesfatsion, *Member, IEEE*

Abstract—Power systems have distinctive features that greatly complicate the development of auction designs. This study reviews the theory and practice of auction design as it relates specifically to U.S. restructured wholesale power markets, i.e., centrally-administered wholesale power markets with congestion managed by locational marginal prices. Basic auction concepts such as reservation value, net seller surplus, net buyer surplus, competitive market clearing, market efficiency, market pricing rules, supply offers, demand bids, strategic capacity withholding, and market power are explained and illustrated. Complicating factors specific to wholesale power markets are clarified, and recent advances in computational tools designed to address these complications are briefly noted.

Index Terms—Auction design, restructured wholesale power markets, objectives, pricing rules, bid/offer-based optimal power flow, locational marginal prices, congestion rents

I. INTRODUCTION

AUCTION theory is often touted, with justification, as an economic success story ([1], [2]). Auction theory has been used to facilitate the design of auction markets for a wide range of goods, services, and financial assets. Examples include carbon emission permits in Germany, 3G mobile-phone licenses in the UK, and U.S. Treasury bills.

The optimal design of auctions for real-time, day-ahead, and other short-term wholesale power markets became a hot topic in the late 1990s with the restructuring of many wholesale power markets around the world. For example, as Joskow [3] notes, over 50% of generating capacity in the U.S. is now operating within the footprint of a restructured wholesale power market whose core architecture includes a day-ahead market organized as a double auction.

Nevertheless, auctions for restructured wholesale power markets are difficult to justify in terms of existing auction theory. The main finding of auction theory to date is the famous *Revenue Equivalence Theorem (RET)*, roughly stated as follows [1, p. 17]: Consider a collection of potential buyers bidding for a single object being sold by a single seller. Then, given certain regularity conditions, any design for this auction yields the same expected revenue to the seller. The bulk of auction theory research to date has built on the RET foundation. The difficulty for power industry researchers is

that real-world power markets tend to involve large deviations from the RET assumptions.

Consider, in particular, the many ways in which a restructured wholesale power market violates the RET assumptions. A restructured wholesale power market is an open-ended dynamic game among traders, a market operator, and one or more regulatory agencies. Traders (electric power producers, load-serving entities, power brokers) are highly heterogeneous, each with their own distinct goals, constraints, and structural attributes. The traders repeatedly make multi-unit supply offers and demand bids within one or more interrelated submarkets that constitute the wholesale power market in the large. These submarkets include spot and forward energy markets supported by a variety of submarkets for such things as ancillary services, capacity installation, and financial hedging instruments. The reservation values of the traders – i.e., the minimum sale prices and maximum purchase prices at which they are willing to trade – can vary over time in response to changes in such factors as fuel costs, weather conditions, and regulatory policies. When grid congestion arises, the energy market can segment into separate energy submarkets (load pockets) within which power suppliers profit from higher-than-normal energy prices because electric power quantities offered for sale at different buses now effectively represent distinct goods supporting a distinct array of prices.

Moreover, the extraordinarily complicated nature of the rules and regulations governing the operations of restructured wholesale power markets can provide opportunities for market participants to game the system to their advantage through strategic behaviors, either individually or in tacit collusion. Such gaming can involve anticompetitive behaviors (e.g., exaggeration of costs, physical capacity withholding, and entry deterrence) that result in operating inefficiencies and that raise adequacy and security concerns. Interestingly, the complexities of these rules and regulations, including in particular auction design features, have in many cases arisen for the very purpose of preventing such gaming.

In summary, researchers interested in the development of auctions for restructured wholesale power markets need to push the envelope of received auction theory to handle many additional complicating factors. The purpose of this auction primer is therefore threefold: (1) to explain and illustrate basic auction concepts and terminology relevant for restructured wholesale power markets; (2) to clarify the types of complicating factors that specifically arise for double-auction implementations of day-ahead markets within restructured wholesale power markets; and (3) to briefly note promising developments in computational modeling tools specifically

Leigh Tesfatsion (tesfatsi@iastate.edu), Economics, Mathematics, and Electrical and Computer Engineering Departments, Iowa State University, Ames, IA 50011-1070 USA.

designed to address these complicating factors.

II. AUCTION BASICS

A. Competitive Market Clearing

A *market* is any public context in which the selling and buying of a good, service, or financial asset takes place. Hereafter, the symbol q will be used to denote any particular quantity of good, service, or financial asset, and p will be used to denote the (per-unit) price of q. For example, q could denote a quantity of apples measured in bushels, and p could denote the price per bushel.

An *auction market* is a market that is centrally managed by an *auctioneer* (or "central clearing house") that facilitates trades between sellers and buyers. Auction outcomes are determined by the particular pricing rules enforced by the auctioneer in pursuit of its objectives. As clarified in section III, day-ahead markets in U.S. restructured wholesale power markets are organized as *double auctions*. A double auction is a two-sided auction in which sellers submit supply offers and buyers submit demand bids.

An *(inverse) true supply schedule* for a seller i is a schedule giving the minimum price p that seller i would be willing to accept for each successive quantity unit supplied. These successive minimum sale prices are referred to as seller i's *(sale) reservation values*. Conversely, an *(inverse) true demand schedule* for a buyer j is a schedule giving the maximum price p that buyer j would be willing to pay for each successive quantity unit demanded. These maximum purchase prices are referred to as buyer j's *(purchase) reservation values*.[1]

A simple illustration is given in Table I for gourmet Honeycrisp apples. Supplies and demands at wholesale are measured in bushels. The table entries give sale and purchase reservation values measured in \$/bushel. For example, the table entries for seller S1 indicate that his minimum acceptable sale price is \$10 for a first bushel, \$50 for a second bushel, \$65 for a third bushel, and \$70 for a fourth bushel. Four bushels are his maximum supply capacity, so his effective reservation value for any additional bushels is infinite. Similarly, the table entries for buyer B2 indicate that his maximum acceptable purchase price is \$70 for a first bushel, \$50 for a second bushel, \$25 for a third bushel, and \$0 for any additional bushels.

The *true total supply schedule* for a particular market gives the minimum sale reservation value p for each successive quantity unit supplied. This schedule can be represented as a function $p = S(q)$, where p denotes the minimum sale reservation value for the last quantity unit supplied at q. The function $p = S(q)$ is constructed from the individual sellers' true supply schedules by placing the sellers' reservation values

[1] The qualifier "true" is used to indicate that a supply or demand schedule accurately reflects sale or purchase reservation values. In section II-C the concept of a "reported" supply or demand schedule will be introduced for strategic traders whose offers and bids can diverge from reservation values. *Inverse* supply and demand schedules give reservation values for each additional quantity unit, whereas *ordinary* supply and demand schedules give maximum quantities that sellers and buyers are willing to sell and purchase for each successively higher price. This study uses the inverse form because the stress is on sellers and buyers who actively make price offers and bids for different considered quantities. For expositional simplicity, the qualifier "inverse" for supply and demand schedules will hereafter be omitted.

TABLE I
ILLUSTRATION: TRUE SUPPLY AND DEMAND SCHEDULES FOR TWO
SELLERS AND THREE BUYERS OF HONEYCRISP APPLES

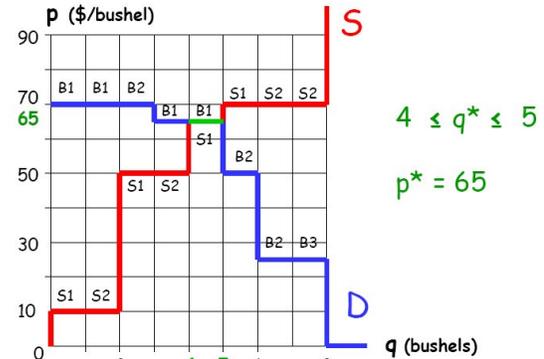| Bushels	S1	S2	B1	B2	B3
1	\$10.00	\$10.00	\$70.00	\$70.00	\$25.00
2	\$50.00	\$50.00	\$70.00	\$50.00	0
3	\$65.00	\$70.00	\$65.00	\$25.00	0
4	\$70.00	\$70.00	\$65.00	0	0
5	∞	∞	0	0	0
7	∞	∞	0	0	0
8	∞	∞	0	0	0

Fig. 1. Plots of the true total supply and demand schedules in Table I. The competitive market clearing points (q^*, p^*) are $4 \leq q^* \leq 5$ and $p^* = 65$.

in *ascending* order. Conversely, the *true total demand schedule* for a particular market gives the maximum purchase reservation value p for each successive quantity unit demanded. This schedule can be represented as a function $p = D(q)$, where p denotes the maximum purchase reservation value for the last quantity unit demanded at q. The function $p = D(q)$ is constructed from the individual buyers' true demand schedules by placing the buyers' reservation values in *descending* order.

Fig. 1 plots the true supply and demand schedules for the Table I market. A second example is given in Fig. 2.

The *net seller surplus on a quantity unit sold* is the difference between the actual sale price received by the seller for this unit and the seller's reservation value for this unit. The *net buyer surplus on a quantity unit bought* is the difference
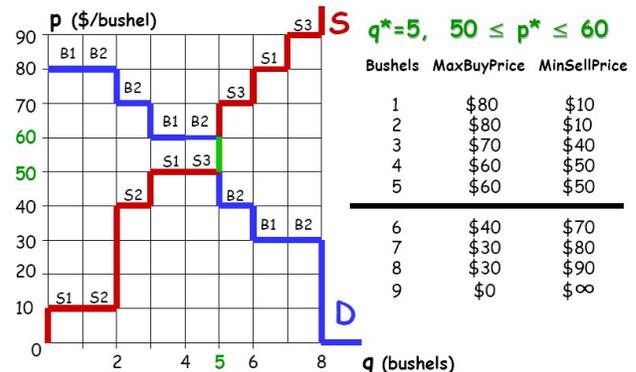

Fig. 2. A second illustration of true total supply and demand schedules. The competitive market clearing points (q^*, p^*) are $q^* = 5$ and $50 \leq p^* \leq 60$.
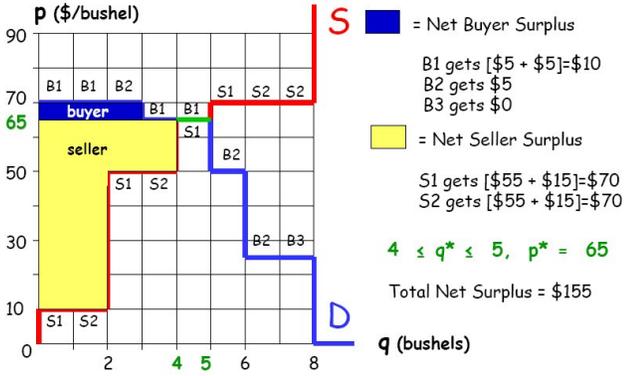
Fig. 3. Net seller surplus, net buyer surplus, and total net seller-buyer surplus calculations and depictions for the true total supply and demand schedules presented in Table I and plotted in Figure 1.

between the buyer's reservation value for this unit and the actual price paid by the buyer for this unit.

A quantity-price combination (q^*, p^*) is called a *Competitive Market Clearing (CMC) point* if plots of the true total supply schedule $S(q)$ and the true total demand schedule $D(q)$ (with vertical segments included) intersect at this point; that is, if $p^* = S(q^*) = D(q^*)$. The price p^* is called a *CMC price level*, and the quantity q^* is called a *CMC output level*. In general, CMC points can be multiple, unique, or non-existent. Multiplicity of CMC points is only possible for "step function" supply and demand schedules that have flat segments (either horizontally or vertically).[2] See, for example, Figs. 1 and 2.

What is the economic meaning of a CMC point? It can be shown that (q^*, p^*) is a CMC point if and only if the following two *no-more-surplus conditions* hold at this point: (a) Taking p^* as the given market price, no seller perceives an opportunity to *strictly* increase his extracted net seller surplus by changing his quantity offer; and (b) taking p^* as the given market price, no buyer perceives an opportunity to *strictly* increase his extracted net buyer surplus by changing his quantity bid.

Consequently, *if profit-seeking traders take prices as given at a CMC point*, they perceive no reason to unilaterally deviate from their current quantities supplied or demanded. In fact, however, traders in many real-world wholesale markets (e.g., restructured wholesale power markets) are not price takers. Rather, they are strategic participants in dynamic trading games requiring the repeated submission of supply offers and demand bids in the form of quantity-price schedules. This issue is further considered in section II-C below.

The *net seller surplus at a CMC point* (q^*, p^*) is the area between the horizontal line at the price level p^* and the true total supply schedule, up to q^*. The net seller surplus at (q^*, p^*) measures seller welfare at this point. It is the difference between what sellers would actually receive for the sale of q^* units at price p^*, i.e., $p^* \times q^*$, and the minimum payment they are willing to receive for the sale of q^* units.

[2]Discussions of markets in standard economic textbooks typically assume a strictly increasing true total supply schedule and a strictly decreasing true total demand schedule having a unique intersection point. On the other hand, supply offers and demand bids in U.S. restructured wholesale power markets (e.g., MISO [4]) are typically required to take the form of step functions.

The *net buyer surplus at a CMC point* (q^*, p^*) is the area between the true total demand curve and the horizontal line at the price level p^*, up to q^*. The net buyer surplus at (q^*, p^*) measures buyer welfare at this point. It is the difference between the maximum payment that buyers are willing to make for q^* units and the amount that buyers would actually pay for the purchase of q^* units at price p^*, i.e., $p^* \times q^*$.

The *total net seller-buyer surplus (TNSBS)* at any CMC point (q^*, p^*), i.e., the sum of net seller surplus and net buyer surplus at this point, is the *maximum* amount of net seller and buyer surplus that can be extracted from the market. That is, there is no way to arrange trades so that the total extracted net seller and buyer surplus is greater than at the CMC point. This follows from the arrangement of seller and buyer reservation values in ascending and descending orders in the construction of the true total supply and demand schedules.

The net seller surplus, net buyer surplus, and TNSBS corresponding to any one of the multiple CMC points depicted in Fig. 1 are calculated and plotted in Fig. 3. The invariance of TNSBS to the particular selection of a CMC point for a market with multiple CMC points follows from the definition of a CMC point.

B. Market Efficiency and Pricing Rules

Efficiency means that resources are not being wasted. Economists identify two types of wastage: (i) physical wastage, in the sense that some valued units of resource remain unused; and (ii) wastage of value, in the sense that some units of resource are not being used by those who value them most.

The concept of "market efficiency" addresses both (i) and (ii). Let *MaxTNSBS* denote the maximum possible total net surplus that sellers and buyers *could* extract from a particular market during some time period T for which sale and purchase reservation values remain unchanged. If a CMC point exists, MaxTNSBS coincides with the TNSBS extracted at this CMC point. Also, let *ActualTNSBS* denote the *actual* total net surplus that sellers and buyers manage to extract from the market during T. Then the *market efficiency (ME)* of this market during T is measured as follows:

$$\text{ME} = 100\% \times \left(\frac{\text{ActualTNSBS}}{\text{MaxTNSBS}} \right) \qquad (1)$$

By construction, $0\% \leq \text{ME} \leq 100\%$ if sellers never sell quantity units at prices below their reservation values and buyers never purchase quantity units at prices above their reservation values.

The quantity units lying to the left of the largest CMC quantity level are called *inframarginal*, and all other units are called *extramarginal*. ME will be 100% for a market during a given time period T if all inframarginal units trade, no extramarginal units trade, and the price paid to a seller equals the price paid by the buyer for each traded unit. An ME below 100% can only occur if some inframarginal unit fails to trade, some extramarginal unit succeeds in being traded, or the price received by some seller in a trade is not equal to the price paid by the buyer.

For example, for the market depicted in Fig. 1, suppose sellers S1 and S2 sell their five inframarginal bushels to buyers B1

and B2 with sale price equal to purchase price for each bushel. If these five bushels are inframarginal for B1 and B2, then the total net surplus extracted will be MaxTNSBS=$155 (i.e., the sum of the five purchase reservation values minus the sum of the five sale reservation values) regardless of the precise trading prices or the precise seller-buyer matchings. However, suppose S1 instead sells his first ($10) inframarginal bushel to the extramarginal buyer B3, whose purchase reservation value for this bushel is only $25. Then the actual total net surplus extracted in this market will be no greater than $115, regardless of remaining trades, implying a *deadweight loss* of at least $40 and a ME strictly below 100%.

Economists adhere to the market efficiency conception (1) because they are concerned that efficiency be achieved in both the short and longer run. Their short-run concern is to ensure maximum net surplus extraction using *existing* resources. Their longer-run concern is to provide proper incentives to encourage investment in *new* resources for producing and extracting more net surplus in the future. In particular, sellers with new ideas for supplying output at lower unit cost should have an incentive to introduce these ideas into a market to compete against current sellers for the extraction of seller surplus. Similarly, buyers with new ideas for gaining more benefit per unit of output purchased should have an incentive to introduce these ideas into a market to compete against current buyers for the extraction of buyer surplus.

It is instructive to consider how market efficiency is potentially enhanced in both the short run and longer run under a uniform CMC pricing rule.

As seen above, the total net seller and buyer surplus extracted at any CMC point is the maximum that can be extracted during a time period T for which sale and purchase reservation values remain unchanged. Thus, short-run efficiency is achieved if the market price for all quantity units is at a CMC price level.

However, the extraction of MaxTNSBS does not *require* a uniform CMC pricing rule. Rather, MaxTNSBS is extracted from the market during T if the following three conditions hold: (i) Traded output is at a CMC level q^*; (ii) only inframarginal units are traded; and (iii) for each traded unit, the price paid to the seller equals the price paid by the buyer.

For example, these three conditions are consistent with a k-discriminatory pricing rule ($k \in [0,1]$) defined as follows: For any matched seller-buyer pair with sale and purchase reservation values $p^s \leq p^b$, set their "strike" price at the level $p^k = k \cdot p^b + [1-k] \cdot p^s$ lying between their reservation values. This pricing rule simply distributes the net surplus generated by a trade between a seller-buyer pair in a ratio determined by k without affecting the amount of this net surplus. In particular, then, ignoring possible deleterious effects on incentives for truthful revelation of costs and capacities, as well as serious concerns related to recovery of non-variable costs, short-run efficiency could alternatively be achieved by using a 0-discriminatory pricing rule that awards all generated net surplus to buyers.[3]

[3]This is essentially the approach advocated for wholesale power markets in a series of studies on "payment cost minimization"; see, e.g., Yan et al. [5].

On the other hand, from a longer-run efficiency point of view, setting market price at a uniform CMC price level encourages sellers and buyers to innovate. In contrast, use of a k-discriminatory pricing rule, particularly with an extreme setting of $k = 0$ or $k = 1$, would tend to discourage innovation by reducing or even eliminating the gains to innovation.

More precisely, a CMC uniform pricing rule permits sellers of inframarginal units to keep any additional net seller surplus (net earnings) garnered from a lowering of their costs and hence a lowering of their sale reservation values. This provides sellers with an incentive to find new ways to supply output using fewer inputs or to supply more output from the same inputs. Similarly, setting market price at a uniform CMC price level permits buyers of inframarginal units to keep any additional net buyer surplus (net earnings) garnered from the development of new ways of achieving more benefit from each purchased unit of output and hence an increase in their purchase reservation values. This provides buyers with an incentive to pursue such developments.

All of the above analysis is conditioned on the presumption that the auctioneer knows (or can determine) the true reservation values of sellers and buyers. In fact, however, it is not necessarily to the advantage of sellers and buyers to reveal these private attributes in their publicly reported supplies and demands. This issue is taken up in the next section.

C. Strategic Trade Behavior

An offer to sell that is reported to an auctioneer is called a *supply offer*. A bid to buy that is reported to an auctioneer is called a *demand bid*. Hereafter it is assumed that supply offers take the form of *reported supply schedules* and demand bids take the form of *reported demand schedules*. Moreover, it is assumed that the objective of each seller and buyer in an auction is to attain the highest possible extraction of net surplus (net earnings) for itself. All else equal, then, sellers are better off the higher the prices they obtain for the quantity units they sell, and buyers are better off the lower the prices they pay for the quantity units they purchase.

Seller market power refers to the ability of a seller to profitably raise the sale price(s) he receives for the quantity units he sells, relative to CMC conditions. *Buyer market power* refers to the ability of a buyer to profitably lower the purchase price(s) he pays for quantity units he purchases, relative to CMC conditions. Sellers and buyers can attempt to exercise market power in two basic ways. First, sellers and buyers can misrepresent their true reservation values in their supply offers and demand bids. Alternatively, sellers and buyers can reduce the number of units they offer to sell or bid to buy in their supply offers and demand bids, even if the sale and purchase of these units would generate a positive net surplus.

For example, consider a market with sale and purchase reservation values as depicted in Fig. 2, and suppose the proposed CMC point for this market is $q^* = 5$ and $p^* = 50$. Sellers have two possible ways they might try to increase their extracted net seller surplus relative to this CMC point: *economic capacity withholding*, i.e., offering units for sale at prices higher than their sale reservation values; and *physical*

- **Traders**
  - LSEs (bulk-power buyers)
  - GenCos (bulk-power sellers with learning capabilities)
- **Two-settlement process**
  - Day-ahead market (double auction, financial contracts)
  - Real-time market (settlement of differences)
- **AC transmission grid**
  - LSEs and GenCos located at user-specified buses across the transmission grid
  - Congestion managed via locational marginal pricing

- **Independent System Operator (ISO)**
  - Day-ahead hourly scheduling via bid/offer-based DC optimal power flow (OPF)
  - System reliability assessments

Fig. 4. Simplified depiction of the core two-settlement system architecture of U.S. restructured wholesale power markets.

Fig. 5. Simplified depiction of ISO wholesale power market activities during a typical day D

capacity withholding, i.e., withholding units from sale. It is easily checked, for example, that seller S1 obtains a higher net seller surplus than at the proposed CMC point if he withdraws his second ($50) inframarginal unit from sale, thus ensuring a higher market price of $60 for his first ($10) inframarginal unit. Indeed, *all three* sellers obtain higher net seller surplus at this higher market price.

Recall from section II-A that sellers and buyers have no incentive to individually change their quantity sales or purchases at CMC points, assuming they take the market price as given. The important point above is that seller S1 does *not* take the market price as given. Rather, he correctly perceives his ability to increase the market price to his own advantage by withdrawing one of his quantity units from sale.

The detection of market power is particularly problematic for restructured wholesale power markets due to grid congestion effects on locational marginal prices. Before taking up this issue in section IV, below, we first set out in simplified form the flow of daily activities in a restructured wholesale power market encompassing an ISO-managed double-auction for a day-ahead energy market.

III. SIMPLIFIED FLOW OF DAILY ACTIVITIES IN U.S. RESTRUCTURED WHOLESALE POWER MARKETS

A. Overview

Fig. 4 depicts several core features of ISO-managed restructured wholesale power markets. Fig. 5 depicts ISO activities on a typical day D. A more detailed description of the logical flow of daily market events is as follows:[4]

- The wholesale power market operates over an *AC transmission grid* during successive days D, with each day D consisting of 24 successive hours H = 00,01, ...,23.
- The wholesale power market includes an *Independent System Operator (ISO)* and a collection of energy traders consisting of *Load-Serving Entities (LSEs)* and *Generation Companies (GenCos)* distributed across the buses of the transmission grid.

- The objective of the ISO is the reliable and efficient operation of the wholesale power market over time.
- In an attempt to attain this objective, the ISO undertakes the daily operation of a *day-ahead market* settled by means of *locational marginal pricing (LMP)*, i.e., the determination of prices for electric power in accordance with the location and timing of its injection into, or withdrawal from, the transmission grid.[5]
- The objective of each LSE is to secure for itself the highest possible net earnings each day through the purchase of electric power in the day-ahead market and the resale of this electric power to its downstream (retail) customers.
- During the morning of each day D, each LSE chooses a *demand bid* to submit to the ISO for the day-ahead market for day D+1. Each demand bid consists of two parts: a *fixed demand bid* (i.e., a 24-hour price-insensitive load profile); and 24 *price-sensitive demand bids* (one for each hour), each consisting of a demand function defined over a purchase capacity interval.
- The objective of each GenCo is to secure for itself the highest possible net earnings each day through the sale of electric power in the day-ahead market.
- During the morning of each day D, each GenCo chooses a *supply offer* to submit to the ISO for use in all 24 hours of the day-ahead market for day D+1. Each supply offer consists of a reported marginal cost function defined over a reported operating capacity interval.
- After receiving demand bids from LSEs and supply offers from GenCos during the morning of day D, the ISO determines and publicly reports hourly dispatch levels and LMPs for the day-ahead market for day D+1 as the solution to hourly bid/offer-based *DC optimal power flow (DC-OPF)* problems. *Transmission grid congestion* is managed by the inclusion of congestion cost components in LMPs.
- Any differences that arise during day D between actual

---

[4]This logical flow of events is captured in the AMES Wholesale Power Market Test Bed [6] developed by H. Li, J. Sun, and L. Tesfatsion for the study of ISO-managed restructured wholesale power markets with learning traders operating over AC transmission grids.
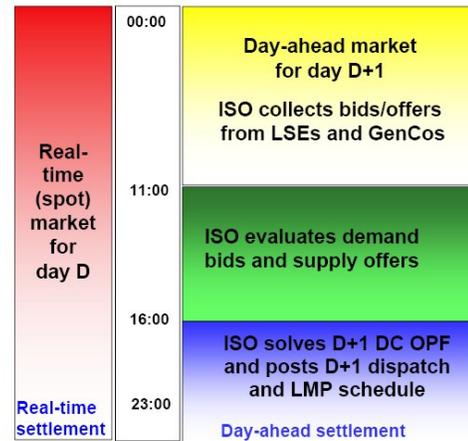
[5]Roughly stated, a *locational marginal price* at any particular transmission grid bus is the least cost to the system of servicing one additional megawatt (MW) of fixed demand (load) for electric power at that bus.
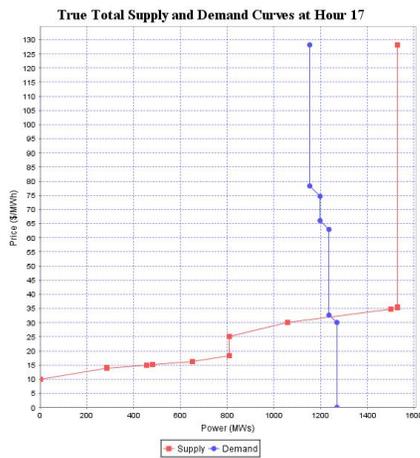
Fig. 6. True total supply and demand schedules during the peak-demand hour 17 of a day-ahead energy double-auction market for a benchmark (no-learning) 5-bus test case explored in [7] using the AMES Test Bed.

day-D demand and supplies and the demand bids and supply offers cleared during day D-1 in the day-ahead market for day D are resolved in the real-time market for day D.

- At the end of each day D, the ISO settles all of the demand and supply commitments for the day-ahead market for day D+1 on the basis of the LMPs determined for the day-ahead market for day D+1.
- At the end of each day D, the ISO settles all purchases and sales in the real-time market for day D on the basis of LMPs determined for the real-time market for day D.

IV. SELLER MARKET POWER AND MARKET EFFICIENCY IN DAY-AHEAD ENERGY AUCTION MARKETS

Fig. 6 illustrates true total supply and demand schedules for the peak-demand hour of a day-ahead energy double-auction market. The initial vertical portion of the true total demand schedule corresponds to LSE fixed demand bids and the remaining portion corresponds to LSE price-sensitive demand bids. The ratio of fixed to price-sensitive demand is calibrated to demand data from the MISO.

As indicated in Fig. 6, buyer market power is not considered problematic in current day-ahead energy auction markets since most bid-in demand takes the form of fixed demand. Consequently, only seller market power is considered below.

A. Detection of Seller Market Power

Although the exercise of seller market power in restructured wholesale power markets can have substantial adverse effects on the efficiency, reliability, and fairness of market operations, it is difficult to construct measures for its reliable detection ([8], [9]). Here we briefly review several critical issues.

On the one hand, the complexity of the rules and regulations governing market operations in restructured wholesale power markets creates opportunities for GenCos to exercise seller market power either individually or in tacit collusion. This seller market power takes two main forms: economic withholding of capacity through a reporting of higher-than-true

marginal costs; and physical withholding of capacity through a reporting of lower-than-true maximum operating capacity.

Economic withholding of capacity can induce higher prices for cleared supply as well as *out-of-merit-order dispatch*, i.e., more expensive generation dispatched in place of less expensive generation. This results in inefficient (and politically important) transfers of wealth away from LSEs and their downstream (retail) consumers and towards GenCos.

Physical withholding of capacity can induce higher prices for remaining offered capacity and hence higher net earnings for GenCos that withhold only a portion of their capacities. It can also result in out-of-merit-order dispatch. In addition, physical withholding of capacity increases the chances of *inadequacy events* in which offered capacity is insufficient to meet total fixed demand. Such events force ISOs to take special actions such as additional reserve commitments or load shedding to avoid the breakdown of power flow on the grid.

In short, seller market power in the form of capacity withholding results in distorted price signals as well as the possible need for non-market ISO interventions. This hinders the efficient and fair use of existing resources as well as the proper signalling of future transmission and generation investment needs.

On the other hand, the physical laws governing power flow on transmission grids mean that these grids are strongly connected networks. Injections or withdrawals of power at one location on the grid can have substantial effects on prices and power flows at distant locations, creating the appearance of seller market power even if no seller market power has been deliberately exercised.

Consider, in particular, the impact of grid network effects on prices. The *locational marginal price (LMP)* at each bus of the transmission grid is defined as a right-hand *system* marginal cost: namely, the least cost *to the system* of servicing an *additional* megawatt (MW) of fixed (price insensitive) demand at that bus. This additional MW of fixed demand can be serviced in three possible ways: (a) increasing generation somewhere in the system; (b) cutting back on cleared price-sensitive demand (if any) somewhere in the system to free up generation for servicing the additional fixed demand; or (c) some combination of the two. The particular option used to determine the LMP is the option resulting in the least additional system cost.[6]

It follows that only GenCos that are "marginal" at a system operating point can contribute to the determination of LMPs at that point. A GenCo is said to be *marginal* if it is capable of servicing additional demand because it is currently operating strictly within its operating capacity limits. If the injection of power at a particular grid location leads to grid congestion, this will cause at least some separation of LMPs across the grid. Moreover, this grid congestion will also result in a failure to fully extract all possible net surplus (i.e., a *deadweight loss* or *congestion loss*) if price-sensitive demand then has to be curtailed.

For example, the intersection of the true total demand and

---

[6]Note that LMPs are *not* conceptually equivalent to the CMC price levels defined in Section II-A.

supply curves depicted in Fig. 6 for the peak-demand hour 17 occurs at (1245MW, $32/MWh). Note that cleared demand at this CMC point would include both fixed and price-sensitive demand. However, as shown in [7], due to congestion on the branch between bus 1 and bus 2 only total fixed demand (1153.58MW) is actually cleared; no price-sensitive demand is actually cleared, indicating a deadweight loss. Moreover, the congestion causes strong LMP separation across the grid. The LMP at bus 1 is $14.02/MWh, the LMP at bus 2 is $78.24/MWh, the LMP at bus 3 is $66.07/MWh, the LMP at bus 4 is $32.61/MWh, and the LMP at bus 5 is $17.32/MWh. The average of these actual LMP values is $41.65/MWh, much higher than the CMC price level $32/MWh depicted in Fig. 6.

As is well understood, the LMP paid to each individual *non-marginal* (i.e., capacity-constrained) GenCo for each MW it sells at its operating point can strictly exceed its *left-hand* marginal cost.[7] The MWs supplied by these non-marginal GenCos constitute inframarginal quantity units, similar to the quantity units to the left of q*=5 in Fig. 2. What is not as well understood, however, is that an LMP at some bus k can strictly exceed the *right-hand* marginal cost of all *marginal* GenCos located at buses $k' \neq k$ if grid congestion requires out-of-merit-order dispatch.

For example, to service an additional MW of demand at some bus k for some hour H in the presence of grid congestion might require that less expensive generation at some second bus k' be backed down, e.g., by 2MWs at $20/MWh, and that more expensive generation at some third bus k'' be brought up, e.g., by 3MWs at $30/MWh, in order to avoid overloading an already constrained transmission grid branch. In this case the *system* marginal cost of servicing an additional MW of demand at bus k for hour H—i.e., the LMP at bus k for hour H—is $50/h = [3MWs·($30/MWh) - 2MWs·($20/MWh)]. If the GenCo at k'' supplying 3MWs at $30/MWh has additional operating capacity available at a marginal cost less than $50/MWh, then the LMP at bus k strictly exceeds the right-hand marginal cost of this marginal GenCo.

Indeed, this scenario occurs during the peak-demand hour 17 depicted in Fig. 6 for a benchmark case in which all Gen-Cos report their true marginal costs (reservation values). At the system operating point, the LMP at bus 2 strictly exceeds the true right-hand marginal costs of the three marginal GenCos 1, 4, and 5 located at buses 1, 4, and 5.

Consequently, care is needed in applying standard economic measures for the detection of seller market power that rely on comparisons between price and marginal cost, e.g., the *Lerner Index* $[p - MC(q)]/p$.

B. Attainment of Market Efficiency

This section expands upon a key point raised in [9]: the basic form of the bid/offer-based OPF objective function used by ISOs to determine LMPs and power commitments for day-ahead energy markets renders problematic the attainment of market efficiency.[8]

An LSE that has a positive fixed (price-insensitive) demand for electric power has a vertical demand schedule for these quantity units, implying an infinite maximum willingness to pay for them. In ISO-managed day-ahead markets, fixed demand must be met in all normal market circumstances. Consequently, this infinite benefit does not help to distinguish between the efficiency attained under distinct (normal) market scenarios characterized by the same fixed demand because in effect it cancels out when the benefits arising under any two of these scenarios are differenced.

For this reason, power market researchers routinely omit consideration of LSE fixed demand benefits when constructing measures for comparative market efficiency evaluations. A special case of this is when all LSE demand is fixed and attention is focused solely on minimization of the total avoidable costs incurred in satisfying this fixed demand. Here we consider the more general case, reflective of many actual ISO-managed day-ahead wholesale power markets, in which LSE demand bids consist of both fixed and price-sensitive parts.

Consider an ISO-managed day-ahead market participated in by I GenCos and J LSEs. For each GenCo i, let $k(i)$ denote its bus location, and let p_{Gi} denote the quantity of electric power that GenCo i is cleared to sell in the day-ahead market for hour H of day D+1. Also, let $LMP_{k(i)}$ denote the LMP for bus $k(i)$ in hour H of day D+1. Assuming GenCos have no avoidable fixed costs, the net seller surplus of GenCo i corresponding to its cleared supply p_{Gi} is

$$NSS_{Gi} = LMP_{k(i)} \cdot p_{Gi} - \int_0^{p_{Gi}} [MC_i(p)] \, dp \quad (2)$$

In (2), $MC_i(p)$ denotes GenCo i's true left-hand marginal cost (sale reservation value) for an increment dp of power, evaluated at p.

Similarly, for each LSE j, let $k(j)$ denote its bus location. Also, let p_{Lj}^S and p_{Lj}^F denote the quantities of electric power that LSE j is cleared to buy in the day-ahead market for hour H of day D+1 corresponding to its price-sensitive demand bid and its fixed demand bid, respectively. Also, let $LMP_{k(j)}$ denote the LMP for bus $k(j)$ in hour H of day D+1. The net buyer surplus of LSE j corresponding to its total cleared demand, adjusted to omit the infinite benefit corresponding to its fixed demand, is then

$$AdjNBS_{Lj} = \int_0^{p_{Lj}^S} [D_j(p)] \, dp - LMP_{k(j)} \cdot p_{Lj} \quad (3)$$

In (3), $p_{Lj} = [p_{Lj}^S + p_{Lj}^F]$, and $D_j(p)$ denotes LSE j's true left-hand maximum willingness to pay (purchase reservation value) for an increment dp of power, evaluated at p.

The total net seller-buyer surplus attained in the day-ahead market in hour H of day D+1, adjusted by omission of the infinite benefit corresponding to LSE fixed demand, thus takes

---

[7]The marginal cost curve of a capacity-constrained GenCo goes vertical at its maximum operating capacity, implying that the *right-hand* marginal cost of a GenCo at maximum operating capacity is effectively infinite.

[8]ISOs depend on closed-source proprietary software for OPF, so the exact form of these OPF objective functions can only be gleaned indirectly from summary descriptions appearing in manuals and other sources.

the following form:

$$AdjTNSBS = \sum_{i=1}^{I} NSS_{Gi} + \sum_{j=1}^{J} AdjNBS_{Lj} \quad (4)$$

The objective of the ISO is the constrained maximization of the area between the *reported* total supply schedule and the *reported* total price-sensitive demand schedule as constructed from the supply offers $MC_i^R(p)$ and price-sensitive demand bids $D_j^R(p)$ of the participant GenCos and LSEs.[9] More precisely, the ISO's objective for hour H of the day-ahead market in day D+1 is the constrained maximization of

$$B^R - C^R = \sum_{j=1}^{J} \int_0^{p_{Lj}^S} \left[D_j^R(p) \right] dp - \sum_{i=1}^{I} \int_0^{p_{Gi}} \left[MC_i^R(p) \right] dp$$
(5)

with respect to choice of the supply and demand commitment levels $\{p_{Gi}, p_{Lj}^S \mid i = 1, \ldots, I, j = 1, \ldots, J\}$.

The payments from LSEs and to GenCos for the day-ahead market in day D+1 are settled through the ISO at the end of day D. Let *NetISOSurplus* denote the total net payments collected by the ISO in the day-D settlement for hour H of the day-ahead market in day D+1. Using previously introduced terminology, NetISOSurplus can be expressed as follows:

$$\left[\sum_{j=1}^{J} LMP_{k(j)} \cdot p_{Lj} - \sum_{i=1}^{I} LMP_{k(i)} \cdot p_{Gi} \right] \quad (6)$$

Combining (2) through (6), it is seen that

$$B^R - C^R = \left[AdjTNSBS^R + NetISOSurplus \right], \quad (7)$$

where $AdjTNSBS^R$ denotes AdjTNSBS based on supply offers and demand bids rather than true reservation values.

Expression (7) shows that strategic reporting of supply offers and demand bids results in divergence of $B^R - C^R$ from AdjTNSBS. Also, a positive value for NetISOSurplus is a subtraction from net seller and buyer surplus. Either case prevents attainment of market efficiency as standardly defined.

With regard to the first difficulty, ISOs generally engage the services of independent agencies to monitor for the exercise of market power (i.e., strategic offers and bids). However, the experimental findings reported in ([7], [11]) suggest that seller market power will remain a serious concern as long as the bulk of LSE demand is passively bid in as fixed demand.

With regard to the second difficulty, in the absence of grid congestion and branch losses, all LMPs collapse to a single uniform level across the grid. In this case NetISOSurplus is zero since the total quantity of power sold is then constrained to equal the total quantity of power bought through standard OPF nodal balance constraints. On the other hand, Alderete [10, Prop. 2.1] proves for a general DC OPF formulation that NetISOSurplus is always strictly positive in the presence of grid congestion. Also, Li et al. [7] report consistently high positive NetISOSurplus outcomes in a suite of AMES

---

[9]Sometimes additional "unit commitment" costs are also considered, such as no-load and start-up costs, but this does not affect the essential point of this section.

---

experiments for a dynamic 5-bus test case even though grid congestion arises on only one branch.

In response to this problem, ISOs typically distribute at least part of their NetISOSurplus as "congestion rent" payments to participants in Financial Transmission Rights (FTR) markets. The basic rationale for FTRs is two-fold: to provide hedging instruments against day-ahead market price volatility caused in large part by congestion; and to encourage appropriate investment in new transmission capacity to alleviate the congestion causing the positive NetISOSurplus.

As discussed in [12], however, FTR markets as designed to date have not yet succeeded in meeting these two goals. This issue requires further study.

V. CONCLUDING REMARKS

As indicated in previous sections, restructured wholesale power markets encompass extraordinarily complex dynamic auction-market games taking place over transmission grids exhibiting strong network effects.

Recent advances in computational modeling tools are increasingly permitting researchers to address this level of complexity [13], offering hope that auction designs can be sufficiently pre-tested to prevent serious unintended consequences. Nevertheless, these tools must be complemented with increased transparency in rules of operation as well as with a fuller and more timely release of data if useful models are to be achieved.

REFERENCES

[1] P. Klemperer, *Auctions: Theory and Practice*, Princeton University Press, Princeton, NJ, 2004.
http://hicks.nuff.ox.ac.uk/users/klemperer/papers.html
[2] P. Milgrom, *Putting Auction Theory to Work*, Cambridge University Press, Cambridge, UK, 2004.
[3] P. Joskow, "Markets for power in the united states: An interim assessment," *The Energy Journal*, vol. 27, no. 1, pp. 1–36, 2006.
[4] MISO, Business Practices Manual 002 (Energy Markets), Midwest ISO, Inc., 2008. www.midwestmarket.org
[5] J. H. Yan, G. A. Stern, P. B. Luh, and F. Zhao, "Payment Versus Bid Cost Minimization in ISO Markets," *IEEE Power & Energy Magazine*, March/April 2008, pp. 24–36.
[6] L. Tesfatsion, "AMES Wholesale Power Market Test Bed Homepage," hosted by the Economics Department, Iowa State University, www.econ.iastate.edu/tesfatsi/AMESMarketHome.htm
[7] H. Li, J. Sun, and L. Tesfatsion, "Separation and volatility of locational marginal prices in restructured wholesale power markets," ISU Economics Working Paper #09009, June 2009.
[8] P. Twomey, R. Green, K. Neuhoff, and D. Newbery, "A review of the monitoring of market power: Possible roles of TSOs in monitoring for market power issues in congested transmission systems," 05-002 WP, Center for Energy and Environmental Policy Research, March 2005.
[9] A. Somani and L. Tesfatsion, "An agent-based test bed study of wholesale power market performance measures," *IEEE Computational Intelligence Magazine*, Vol. 3, No. 4, November 2008, pp. 56–72.
[10] G. B. Alderete, "Alternative models to analyze market power and financial transmission rights in electricity markets," PhD Thesis, Electrical and Computer Engineering, University of Waterloo, 2005.

[11] H. Li, J. Sun, and L. Tesfatsion, "Dynamic LMP response under alternative price-cap and price-sensitive demand scenarios," *Proceedings of the IEEE Power and Energy Society General Meeting*, Pittsburgh, PA, July 2008.

[12] California Independent System Operator (CAISO), 2004, "Transmission economic assessment methodology (TEAM)," www.caiso.com/docs/2003/03/18/2003031815303519270.html

[13] L. Tesfatsion, "Agent-Based Computational Economics Research on Restructured Electricity Markets," hosted by the Economics Department, Iowa State University, www.econ.iastate.edu/tesfatsi/aelect.htm

Leigh Tesfatsion received her Ph.D. degree in Economics from the University of Minnesota in 1975. She is currently Professor of Economics, Mathematics, and Electrical and Computer Engineering at Iowa State University. Her principal research area is Agent-based Computational Economics (ACE), the computational study of economic processes modeled as dynamic systems of interacting agents, with a particular focus on restructured electricity markets. She is an active participant in IEEE Power Engineering Society working groups and task forces focusing on power economics issues and a co-organizer of the ISU Electric Energy Economics (E3) Group. She serves as associate editor for a number of journals, including the *Journal of Energy Markets*.