

NBER WORKING PAPER SERIES

UNDERSTANDING STRATEGIC BIDDING
IN RESTRUCTURED ELECTRICITY MARKETS:
A CASE STUDY OF ERCOTAli Hortacsu
Steven L. PullerWorking Paper 11123
<http://www.nber.org/papers/w11123>NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
February 2005

We thank seminar participants at various universities and conferences. We are grateful for assistance with data and institutional knowledge from Parviz Adib, Tony Grasso, and Danielle Jaussaud at the Public Utility Commission of Texas. Severin Borenstein, Jim Bushnell, Stephen Holland, Marc Ivaldi, Julie Holland Mortimer, Shmuel Oren, Peter Reiss, Steve Wiggins, Joaquim Winter and Frank Wolak provided very helpful comments. Hailing Zang, Anirban Sengupta, Jeremy Shapiro, and Joseph Wood provided capable research assistance. Hortacsu was a visitor at Harvard University and the Northwestern University Center for the Study of Industrial Organization during the course of this research, and gratefully acknowledges both institutions' hospitality and financial support. Puller was a visitor at the University of California Energy Institute's Center for the Study of Energy Markets, for whose hospitality he is grateful. This research was supported by the Texas Advanced Research Program Grant No. 010366-0202. The views expressed herein are those of the author(s) and do not necessarily reflect the views of the National Bureau of Economic Research.

© 2005 by Ali Hortacsu and Steven L. Puller. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

Understanding Strategic Bidding in Restructured Electricity Markets: A Case Study of ERCOT
Ali Hortacsu and Steven L. Puller
NBER Working Paper No. 11123
February 2005
JEL No. L1, L2, L5, L9

ABSTRACT

We examine the bidding behavior of firms competing on ERCOT, the hourly electricity balancing market in Texas. We characterize an equilibrium model of bidding into this uniform-price divisible-good auction market. Using detailed firm-level data on bids and marginal costs of generation, we find that firms with large stakes in the market performed close to theoretical benchmarks of static, profit-maximizing bidding derived from our model. However, several smaller firms utilized excessively steep bid schedules that deviated significantly from our theoretical benchmarks, in a manner that could not be empirically accounted for by the presence of technological adjustment costs, transmission constraints, or collusive behavior. Our results suggest that payoff scale matters in firms' willingness and ability to participate in complex, strategic market environments. Finally, although smaller firms moved closer to theoretical bidding benchmarks over time, their bidding patterns contributed to productive inefficiency in this newly restructured market, along with efficiency losses due to the close-to optimal exercise of market power by larger firms.

Ali Hortacsu
Department of Economics
University of Chicago
1126 East 59th Street
Chicago, IL 60637
and NBER
hortacsu@uchicago.edu

Steven L. Puller
Texas A&M University
puller@econweb.tamu.edu

1 Introduction

Many recent empirical models of oligopoly competition, including the analysis of bidding in auction markets, rely crucially on equilibrium assumptions.¹ By assuming that firms behave according to a particular strategic equilibrium model, the researcher can map firms' observed pricing or bidding decisions into their unobserved costs of production or their valuations for the auctioned object. The inferences drawn from such approaches rely heavily on the assumed strategic behavior. In most instances, testing the validity of a particular equilibrium model is left to the laboratory, where the researcher assigns costs/valuations to subjects and compares the subject behavior to the behavior predicted by the equilibrium model of competition. Outside of the laboratory, it is difficult to test equilibrium models because data usually are not available on bidder costs/valuations.

In this paper, we analyze the recently restructured electricity market in Texas, where we have the unique advantage of having very detailed data on firms' bidding strategies and their costs of producing electricity.² Most of the electricity in this market is traded through bilateral forward contracts between generators and users of electricity. To meet last-minute changes in aggregate electricity demand that fall beyond or below contracted quantities, generation firms submit bids to adjust their production levels into an hourly "balancing market" administered by ERCOT (Electricity Reliability Council of Texas). Firms of various shapes and sizes participate in these auctions, including large formerly regulated utilities, merchant generating firms, and small municipal utilities and power cooperatives. The auction mechanism is a multi-unit, uniform-price auction – firms bid supply functions and winning sellers earn the price at which aggregate supply bids equal demand.

We model competition in the hourly balancing market using Wilson's (1979) "share auction" formulation.³ In our model, firms choose bid functions to maximize expected profits under uncertainty. Firms face two main sources of uncertainty. First, total demand for balancing power is determined by events such as weather shocks, so it is stochastic from the perspective of the bidder. Second, firms possess private information on their forward contracts to supply power. These obligations determine the firms' net buy or net sell positions in the balancing market, and therefore affect bidding incentives. Because they are private information, these contract obligations generate uncertainty from the perspective of other bidders.⁴

¹Klemperer (2003) and Einav (2004) clarify the connections between oligopoly competition models and auction models.

²Our work is similar in spirit to papers that use "outside" estimates of marginal cost to measure price-cost margins and test strategic oligopoly models. For example, Genesove and Mullin (1998) analyzes the sugar industry in early 20th century. Wolfram (1999) and Sweeting (2004) analyze the electricity market of England and Wales. Bajari and Hortaçsu (2003) use experimental data with assigned bidder valuations to gauge the performance of structural econometric models of auctions.

³Ausubel and Cramton (2002), Wang and Zender (2000), Hortaçsu (2002b), Viswanathan, Wang and Witelski (2002) develop this theoretical framework further. See Cramton (2003) for a particularly accessible account of these models.

⁴Our model is closely related to, but more general than Klemperer and Meyer's (1989) supply-function equilibrium (SFE) model. The SFE has been influential in the modelling and analysis of electricity markets, as in Green and Newberry (1990), Green (1992), Rudkevich (1999), Baldick, Grant and Kahn (2004), Crawford, Crespo and Tauchen (2003). Sweeting (2004) uses data from the England and Wales market to test for static Nash equilibrium bidding in a full information environment. Like the SFE, our model assumes that generation costs are common knowledge

The Bayesian-Nash equilibrium characterization of bidding strategies on ERCOT’s balancing market yields three theoretical benchmarks of firm behavior. The first theoretical benchmark is “ex-post optimal” bidding, which is attained under a functional form restriction placed on Bayesian-Nash bidding strategies, namely that an equilibrium supply function has to be additively separable in its price and private information contract quantity components. Figure 1 explains the economics of “ex-post optimal” bidding in ERCOT.⁵ Firm i ’s marginal cost curve is given by $MC_i(q)$, and its forward contract position is labelled at QC_i . Suppose firm i is observed to submit the supply schedule $S_i^o(p, QC_i)$ in the balancing market. The market clearing price in the balancing market, and the actual amount of electricity that firm i will be called upon to generate will be determined by the intersection of $S_i^o(p, QC_i)$ and the residual demand (RD) curve faced by firm i . The RD curve is the sum of the supply schedules submitted by firms other than i , subtracted from the total demand for electricity in the market. The RD curve is uncertain from the perspective of firm i (since it depends on the realization of aggregate demand and the contract positions (QC_i s) of the other bidders). Given a particular realization, RD_1 , and given its actual supply schedule, $S_i^o(p)$, firm i would supply the quantity at the price given at point D in the figure. At this quantity, the firm would supply more electricity than it was previously contracted to sell (D is to the right of A), and its profits can be calculated as the profits from meeting its contract position, plus the profit made from providing additional power to the market.

However, firm i could do better. For the residual demand curve RD_1 , firm i could calculate the marginal revenue curve given by MR_1 .⁶ By equating marginal revenue and marginal cost, the firm could then select point B , which maximizes its “ex-post” profits.

This, of course, is what would happen if firm i knew that residual demand would be RD_1 . What if the residual demand curve were instead RD_2 ? Firm i could then calculate the marginal revenue curve corresponding to this realization of residual demand, and find the “ex-post” profit-maximizing point C . Note also that the “ex-post” profit maximizing point could be *below* marginal cost given another realization of residual demand sufficiently to the left of RD_2 , so that the firm is called upon to generate less electricity than QC_i . In this case, the firm has to “purchase” power from the spot market to cover its short position, i.e. it is in the position of a monopsonist so it bids below marginal cost.

As is apparent from the figure, firm i ’s profit maximization problem would be greatly simplified if the set of ex-post profit maximizing points (two elements of this set are points B and C) corresponding to *every* realization of residual demand could be connected by a monotonic function such

across bidders, since a lot of engineering information is publicly available about each firm’s generation technology and the spot price of fuel. However, the SFE does not allow bidders to possess private information, and takes only aggregate demand uncertainty into account.

⁵This explanation is based on Klemperer and Meyer’s (1989) construction of ex-post optimal supply function equilibria. Their construction is also utilized in Wolak (2003b), especially in Figure 1 of his paper. Our conceptual contribution here is to extend Klemperer and Meyer’s intuition to the more complex context of the balancing market, where private information about contracts, along with the aggregate uncertainty in total demand, requires the use of a Bayesian Nash equilibrium formulation.

⁶The calculation of marginal revenue is net of the contract position, hence it crosses RD_1 in the positive quadrant. Further details are given in Section 3.

as $S_i^{xpo}(p)$. In that case, firm i could submit $S_i^{xpo}(p)$ as its supply schedule, and be guaranteed a profit-maximizing result under every possible realization of uncertainty.

In section 3, we show that restricting supply schedules to be “additively separable in private information” allows one to characterize the Bayesian-Nash equilibrium of the game in terms of ex-post optimal supply schedules. The intuition under this result is that if supply strategies $S_i(p, QC_i)$ are additively separable in QC_i , then so is the residual demand function for each firm – i.e. any shifts in residual demand are going to be *parallel*. This allows the set of ex-post optimal points to lie on a monotonic supply schedule.

We acknowledge that additive separability is a rather strong a priori restriction to place on firms’ strategies. However, it yields ex-post optimal bidding as a theoretical benchmark: given information on marginal costs ($MC_i(q)$) and the contract position of a firm (QC_i)⁷, it is computationally trivial to find the ex-post optimal bid curve S_i^{xpo} by tracing the marginal revenue curves corresponding to parallel shifts of the residual demand curve observed in our data. And, as we discuss in detail in Section 4.2, this benchmark performs admirably well in explaining the behavior of the largest firm in this market. Figure 2 plots the “data-equivalent” of Figure 1 for Reliant Energy on June 4, 2002 when the ex-post optimal bid curve is very close to the actual bid curve. Note that this was not the only day in which Reliant performed very close to the ex-post optimal bidding benchmark. We show that Reliant’s bidding strategy allows the firm to capture 79% of its ex-post optimal profits on the balancing market.

However, not all firms perform as well as Reliant. In fact, as reported in Section 4.2, many do a lot worse. This result might be driven by several factors, which are considered in the rest of paper.

The first possibility we consider is that the set of assumptions under which “ex-post optimal bidding” is obtained may not be appropriate for this market. Therefore, in section 4.3, we provide two methods to test whether observed bidding behavior is driven by “ex-ante” expected profit maximizing behavior.⁸ The first method is to compare the profitability of actual bids to a bidding strategy that conditions only on information that is “ex-ante” observable to the bidders. Our alternative bidding strategy is to take the most recent realization of residual demand that bidders observe before choosing their bids, and to find the optimal schedule for this realization.⁹ We find that this “adaptive” strategy yields bid schedules that are very close to ex-post optimal, which reflects the fact that residual demand realizations are very persistent (up to a parallel shift) in the data. Thus, not surprisingly, we find that this adaptive strategy can systematically outperform all bidders in this market. The second method, described in section 4.3.2, is a GMM test of the first-order optimality conditions of the ex-ante profit maximization problem.

All three theoretical benchmarks show that, aside from the largest sellers in this market, many of the smaller firms submit bid schedules that are excessively steep from the perspective of static profit

⁷We discuss how we get this information in detail in Section 4.

⁸Bidding strategies that maximize “ex-ante” expected profits may not be ex-post optimal, in the sense that based on the actual realization of residual demand, the firm could have done better by bidding a point that is not on his supply schedule. In other words, an “ex-ante” optimal bidding strategy is subject to regret, whereas an “ex-post” optimal strategy is not.

⁹On ERCOT, firms observed the recent history of aggregate bids with a 2 day lag.

maximization. In Section 5.1, we examine whether technological adjustment costs may account for the excessive steepness of bid functions, but do not find evidence confirming this hypothesis. In section 5.2, we discuss whether transmission constraints may drive such bidding behavior. Although we find evidence that (expected) transmission constraints may drive a portion of the excess markups, it is unlikely to explain all of the deviations. Section 5.3 also discusses the possibility of collusion, though we do not find this to be a likely factor.

As discussed in Section 5.5, one empirical pattern that appears very robustly in our data is that the observed deviations from benchmarks are declining in the size of the firm.¹⁰ This suggests the presence of scale economies in setting up and maintaining a successful bidding operation – an intuition confirmed by our interviews with traders in the market. It also suggests specialization. Indeed, we find that some firms specialize in bidding, while others outsource their bidding operations. In Section 5.4, we provide more direct evidence for the heterogeneity in “sophistication” across bidders – which motivates why investments into bidding operations are important. In particular, firms do not make full use of the strategy space available to them. Each firm can submit supply schedules with up to 40 price-quantity points, but they use far fewer points. Moreover, our conversations with some of the smaller traders revealed the frequent use of bidding heuristics echoing the sunk-cost fallacy. Finally, we do find some evidence of learning over our sample period for the small firms in our sample. The learning rate is estimated to be 10% of performance improvement per year.

Whatever the sources of the observed behavioral deviations from the theoretical benchmarks, they might be economically insignificant if they do not lead to sizeable efficiency losses. In Section 6, we describe the two sources of efficiency losses in ERCOT. The first is the efficiency loss due to the (optimal) exercise of market power by profit maximizing firms. The second is the efficiency loss due to the “excessive steepness” of small firms’ bid schedules that we can not reconcile with expected profit maximizing behavior. When we decompose the total efficiency losses in this market into these two components, we find, somewhat surprisingly, that the latter source of inefficiency is larger.

The outline of the rest of the paper is as follows: in section 2, we describe the institutional setting of the Texas electricity balancing market and present summary evidence that market prices deviate from marginal cost pricing. In section 3, we model strategic bidding in this market as a uniform price share auction. We discuss the empirical implications of our model. In section 4, we compare our theoretical benchmarks with the actual bids in the data. Section 5 discusses these results and explores possible explanations for deviations from ex-post optimal bidding. Section 6 calculates the welfare losses and section 7 concludes.

2 How Bidding Occurs in ERCOT’s Balancing Energy Market

We analyze electricity transactions that occur through spot market auctions. In the Texas wholesale electricity market, most trades occur via bilateral agreements. In addition to this bilateral market,

¹⁰Interestingly, “governance” based explanations that might predict municipal-owned utilities to perform worse than investor-owned utilities appear not to be supported in the balancing market context.

ERCOT, the system operator, conducts an auction to balance supply and demand in real-time. Approximately 2-5% of energy is traded in this “spot market”, called the Balancing Energy Services auction, and we analyze the bidding into this auction.

The mechanics of electricity transactions on this market can be summarized as follows.¹¹ One day before production and consumption occurs, ERCOT accepts schedules of quantities of electricity to inject and withdraw at specific locations on the transmission grid. Those supply (“generation”) and demand (“load”) schedules may differ from the actual production and consumption in real-time for a variety of reasons such as an unpredictably hot day or an outage at a powerplant. The balancing market is a real-time market to balance actual load and generation. Depending upon whether more or less power is needed than the day-ahead schedule, the balancing demand can be positive or negative. As the time of production and consumption nears, ERCOT estimates how much balancing energy is required. Because there are virtually no sources of demand that can respond to prices in real-time, balancing demand is perfectly inelastic.

Bidders offer to increase (“INC”) and decrease (“DEC”) the amount of power supplied relative to their day-ahead schedule. Firms submit *hourly* INC and DEC bid schedules that must be increasing monotonic functions with up to 40 “elbow” points (20 INC and 20 DEC bids). The bid schedules apply to each of the four 15-minute intervals of the hour. Figure 3 displays a sample bidder’s interface for one participant’s daily bids to DEC its generation.¹² The bidder enters up to 20 elbow points for each hour, and these bids may be changed up until one hour prior to the operating hour. In addition, the bidder observes real-time information on its units’ generation, the load it is obligated to serve, and its net short or long position in the balancing market. A sample of this operations interface is shown in Figure 4.

Procurement occurs using a uniform-price, multi-unit auction. ERCOT clears the balancing market four times every hour by intersecting the hourly aggregate bid function with the 15-minute perfectly inelastic demand function. A generator called to INC is *paid* the market clearing price for all INC sales (i.e. production beyond the day-ahead schedule). Likewise, a generator called to DEC *pays* the market clearing price for the quantity of output reduced. A generator that DEC reduces output and purchases power from ERCOT at the market clearing price to satisfy existing contract obligations.

Bidders appear to have a great deal of information on the competitive environment when they choose their bid functions. Our conversations with several market participants suggest that traders have good information on their rivals’ marginal costs. The powerplants in Texas have very similar production technologies, and there are publicly available data on the fuel efficiency of each generating unit. Traders appear to know the major generating units that are on and off-line at any point in time. In addition, some market participants purchase real-time data on the generation of large

¹¹See Wilson (2002) and Joskow (2000) for more detailed discussions of transactions in restructured electricity markets.

¹²INC and DEC bids are submitted separately in this interface. We thank the real-time trading desk at Bryan Texas Utilities for providing Figures 3 and 4 and allowing us to observe their bidding operations. We also thank Reliant/TexasGenco for allowing us to visit and observe their trading desk and allowing us to interview their traders. A variety of other market participants provided very helpful insights into trading in the balancing market.

rival plants. An energy information company named Genscape has developed a technology that measures real-time output with remote sensors installed near the transmission lines out of a plant. Generators and traders can purchase subscription service to access the output data in real-time. Figure 5 displays a sample Genscape interface.¹³ This can be useful strategic information not only when initial bids are submitted but also if the trader wants to change the bids an hour before the market clears.

This information may allow the trader to accurately estimate her residual demand in an upcoming auction. The residual demand is the perfectly inelastic total balancing demand minus bids by all other firms. Total demand is stochastic, but shocks to total demand (e.g. weather) only shift residual demand left and right in a parallel fashion. Rival bids can be inferred in two ways. A trader equipped with knowledge of rivals' marginal costs and an equilibrium mapping of costs to bids, will be able to calculate the aggregate bids of rival firms. Alternatively, every trader has access to aggregate bid data with a two day lag. By knowing the recent aggregate bid stack as well as her own recent bids, the trader can infer the recent aggregate bids by all rivals. To the extent that rival bids several days before are similar to current rival bids, the trader can infer the shape of residual demand before placing their bids.

Congestion of the transmission grid poses a slight complication for our analysis. ERCOT is geographically divided into several zones. If transmission lines between zones are not congested, the balancing market is a single unified market across all Texas. However when lines become congested, ERCOT divides the state into separate markets with different market clearing prices. During congested hours, bids by some firms are technically feasible while bids by others are not. Therefore, our analysis uses only uncongested hours and we treat all bids as technically feasible. During our sample period, 74% of the time intervals are uncongested.¹⁴

We analyze weekdays from the beginning of the market in September 2001 to January 2003. Although we can analyze any period of the day, we focus on 6:00-6:15pm because the most flexible type of generators that can respond to balancing calls without large adjustment costs are likely to be online during this peak hour of the day. The average number of megawatts (MW) traded (both positive and negative) during this time interval is 915 MW.¹⁵ Figure 6 shows the probability density of the (perfectly inelastic) demand for balancing energy. Demand for balancing power is both negative and positive in many hours although is it close to zero on average.

The Texas market consists of a variety of investor-owned utilities, independent power producers,

¹³Also, the traders have interfaces with the ERCOT system that provide access to real-time charts of the grid's electrical frequency. If a major unit goes off-line, the frequency briefly dips and indicates that an outage may have occurred.

¹⁴Even with this sample restriction, there is another complication that firms may have different bidding strategies during congested and uncongested hours, and that congestion is not perfectly predictable. We investigate the possibility that firms anticipate congestion in section 5.2.

¹⁵To compare this with other intervals, the average MW transacted across all intervals varies only slightly over the day with an average volume of 890. Demand can vary substantially during the 15-minute intervals within a bidding hour. The average difference between the minimum and maximum balancing demand is 617 MW for the 6:00-7:00pm bidding hour. Because firms can submit only one bid function for each hour, optimal bids must cover a wide range of possible demand quantities.

municipal utilities and power cooperatives. Table 1 breaks down the generation capacity by owner. The two largest players are the two large former incumbent utilities: Reliant and TXU.¹⁶ The other major investor-owned utilities are Central Power & Light and West Texas Utilities. Municipal utilities (e.g. City of San Antonio Public Service and City of Austin) and power cooperatives (e.g. Lower Colorado River Authority) also sell to the balancing market. Finally, a large number of merchant generation firms of various sizes participate in the wholesale market. The generation technology is primarily natural gas and coal with small amounts of nuclear, hydroelectric and wind generation.

Each firm has some level of contract obligation to directly supply retail customers or provide power to utilities that serve retail load. If the day-ahead generation schedule differs from the contract obligation, the firm is either long or short entering into the balancing market. This residual contract position affects bidding incentives in the auction (Wolak, 2003a). We propose a method to estimate the long or short position in section 3.

Table 2 shows the average sales into the balancing market auction for each 6:00-6:15pm time interval by firm. Reliant is by far the largest participant selling an average of 431 MWh. The other large incumbent utility, TXU, averages substantially less relative to its capacity at 133 MWh. We suggest below that this relatively low quantity of balancing sales results from an apparent reluctance of TXU to supply “DEC” output to the balancing market, even when it appears profitable to do so. The third and fourth largest sellers, Extex Laporte and Calpine, are merchant generators.

2.1 Is There Evidence of Any Market Power?

A precondition for testing strategic bidding is evidence of non-price-taking behavior. In this market, strategic behavior can yield prices above marginal cost when balancing demand is positive and below marginal cost when demand is negative. Interestingly, market power can lead to prices that are too high or too *low*.¹⁷

Consider a firm that has submitted a day-ahead schedule to generate its contract obligations and therefore has no residual contract position in the balancing market. Firms that bid competitively will bid up the marginal cost curve to increase output and down marginal cost to decrease output. However if they exercise market power, the bid function will be steeper than the marginal cost function.

If demand is positive, a firm selling multiple units has incentives to markup the bid price above marginal cost, sacrificing additional sales to raise the price earned on inframarginal units. This is just the standard oligopoly result that a firm acts as a monopolist on residual demand.

However, firms will *mark down* bids below marginal costs if total balancing demand is negative. The logic is very similar but in reverse. Suppose Firm A has contract obligations to serve its customer 100 MW and has submitted a day-ahead schedule to generate 100 MW to cover that contract position. Additionally suppose that total electricity demand (but not demand from Firm

¹⁶Reliant’s powerplants were transferred to a subsidiary named Texas Genco.

¹⁷Kuhn and Machado (2004) also report this phenomenon in the Spanish electricity market.

A’s customers) is 10 MW lower than anticipated. ERCOT holds an auction to DEC generation by 10 MW and firms bid a willingness to pay schedule to reduce generation. If it is called to reduce generation by 10 MW, Firm A still must satisfy the 100 MW contract obligation so it pays the balancing market price on the 10 MW short position. Firm A exercises market power by bidding below marginal cost to sell itself into a short position but to lower the price at which it buys back that position. Or, put differently, a firm bids above marginal cost to exercise “monopoly power” if it is a net seller, and bids below marginal cost to exercise “monopsony power” if it is a net buyer. Therefore, positive price-cost margins in hours with positive balancing demand and negative margins in hours with negative demand are consistent with the exercise of some form of market power.

As a preliminary test, we calculate average prices during INC and DEC hours. For 6:00-6:15pm, prices averaged \$32.61/MWh during intervals of positive balancing demand and \$21.24/MWh during intervals of negative demand. In contrast, total ERCOT loads were not substantially different – load averaged 37,968 MW during the positive demand intervals and 33,937 MW during the negative demand intervals. Given the technology mix of the ERCOT system, these price differences are larger than the difference in system marginal costs for only about 4,000 MW additional production. More formally, we measure price-cost margins for each firm as the difference between the auction price and the marginal cost of the generating unit that would be called upon to supply more or less power. Across all firms, margins average \$4.70/MWh during INC intervals and -\$6.37 during DEC intervals.

The source of these margins and the efficiency consequences are the subject of the remaining sections of the paper. We model this strategic behavior, and test the predictions of this model using observed bidding and marginal cost data.

3 An Equilibrium Model of Bidding on the ERCOT Balancing Market

We now provide a model of strategic behavior in this market that incorporates the uncertainty faced by each firm when making their bidding decisions. To do this, we follow the uniform price share auction setup of Wilson (1979) to model competition in the balancing market. Specifically, let us index the costs of generation (at time t) of the N firms in this market by $\{C_{it}(q), i = 1, \dots, N\}$. We will take total demand $\tilde{D}_t(p) = D_t(p) + \varepsilon_t$ to be the sum of a deterministic price-elastic component and a stochastic constant term. Prior to the balancing market auction, each firm has signed fixed price contracts to deliver certain quantities of power each hour, given by QC_{it} . The analysis here will assume that these contracts have been written a long enough time ago that they are taken as “sunk” decisions from the purposes of a bidder making real time decisions on the balancing market.¹⁸

¹⁸Although the analysis of the two-stage game with endogenous contract quantity choices adds a further and interesting strategic dimension, we found that traders who negotiate forward contracts and traders who bid in the balancing (real-time) market are distinct, and often belong to different trading desks. Thus our model can be thought of as applying to the real-time trader.

In each time period t , each firm simultaneously submits a supply schedule, $S_{it}(p)$.¹⁹ In the balancing market setup of ERCOT, $S_{it}(p)$ can be thought of as the sum of a “day ahead schedule” component, and a balancing schedule component. (Notice that the lumping of these two components does not affect the strategic nature of the game if bidders are not provided any information about each others actions until the market clears.) Given the supply schedules of each firm, ERCOT computes the market clearing price, p_t^c , which satisfies the below market clearing condition, in the absence of any transmission constraints:

$$\sum_{i=1}^N S_{it}(p_t^c, QC_{it}) = \tilde{D}_t(p) \quad (1)$$

Each firm gets paid $S_{it}(p_t^c)p_t^c$ due to the uniform pricing rule. Hence, firm i 's ex-post profit, upon the realization of market clearing price p_t^c is:

$$\pi_{it} = S_{it}(p_t^c)p_t^c - C_{it}(S_{it}(p_t^c)) - (p_t^c - PC_{it})QC_{it}$$

The firm's payoff from its contract position is $-(p_t^c - PC_{it})QC_{it}$, because it has to refund its customers any differential between the contract and market prices for the contracted sales. Wolak (2003a) has shown this is identical to a contract for differences.

Clearly, the most important source of uncertainty in the profit equation above is p_t^c , the market clearing price at time t . In a strategic equilibrium, the uncertainty in p_t^c , from the perspective of firm i , is due to two factors: the uncertainty in market demand, \tilde{D}_t , and the unobserved components of i 's competitors' profit maximization problems, i.e. the contract positions and prices of rival firms, $\{(QC_{jt}, PC_{jt}), j \in -i\}$.

Following the discussion in Section 2, each firm bidding into the market is assumed to know its rivals' total cost functions. This knowledge would be sufficient to calculate equilibrium bids by rival firms, except the firm does not know its rivals' contract positions. Thus, following Wilson (1979), we look for a Bayesian-Nash equilibrium characterization of the game, in which firms' strategies are of the form $S_{it}(p, QC_{it})$.

To characterize a Bayesian-Nash equilibrium, we define a probability measure over the realizations of the market clearing price, from the perspective of firm i , conditional on firm i 's private information contract quantity, QC_{it} , and the fact that firm i submits the supply schedule, $\hat{S}_{it}(p)$, while his competitors are playing their equilibrium bidding strategies, $\{S_{jt}(p, QC_{jt}), j \in -i\}$:

$$H_{it}(p, \hat{S}_{it}(p); QC_{it}) \equiv \Pr(p_t^c \leq p | QC_{it}, \hat{S}_{it}(p))$$

¹⁹von der Fehr and Harbord (1993) note that the existence of pure strategy equilibria can depend upon whether bids are smooth or step functions for certain ranges of demand distributions. We follow the assumption in the supply-function equilibrium literature and allow bids to be smooth.

Utilizing the definition of the market clearing price in equation (1), we can rewrite this probability distribution as:

$$\begin{aligned} H_{it}(p, \hat{S}_{it}(p); QC_{it}) &= \Pr\left(\sum_{j \in -i}^N S_{jt}(p, QC_{jt}) + \hat{S}_{it}(p) \geq \tilde{D}_t(p) \mid QC_i, \hat{S}_{it}(p)\right) \\ &= \int_{QC_{-it} \times \varepsilon_t} \mathbf{1}\left\{\sum_{j \in -i} S_{jt}(p, QC_{jt}) + \hat{S}_{it}(p) \geq D_t(p) + \varepsilon_t\right\} dF(QC_{-it}, \varepsilon_t \mid QC_{it}) \end{aligned}$$

where the first line follows from the fact that the event “ $p_t^c \leq p$ ” is equivalent to there being excess supply at price p . The second line, $\mathbf{1}\{\cdot\}$ is the indicator function for the enclosed event, and $F(QC_{-it}, \tilde{D}_t \mid QC_{it})$ denotes the joint distribution of the vector of contract quantities, $\{QC_{jt}, j \in -i\}$, and the demand noise, ε_t , conditional on the contract position of bidder i , QC_{it} .²⁰

Now, rewrite the bidder’s expected utility maximization problem, where $U(\pi)$ is the utility enjoyed by the bidder from making π dollars of profit. This general utility formulation allows for both risk averse and risk neutral firms.

$$\max_{\hat{S}_{it}(p)} \int_{\mathbb{P}}^{\bar{p}} U\left(p\hat{S}_{it}(p) - C_{it}(\hat{S}_{it}(p)) - (p - PC_{it})QC_{it}\right) dH_{it}(p, \hat{S}_{it}(p); QC_{it})$$

where the expectation is taken over all possible realizations of the market clearing price, weighted by the probability density, $dH_{it}(p, \hat{S}_{it}(p); QC_{it})$. As shown in the Appendix B, Euler-Lagrange necessary condition for the (pointwise) optimality of the supply schedule $S_{it}^*(p)$ is given by:

$$p - C'_{it}(S_{it}^*(p)) = (S_{it}^*(p) - QC_{it}) \frac{H_S(p, S_{it}^*(p); QC_{it})}{H_p(p, S_{it}^*(p); QC_{it})} \quad (2)$$

where

$$\begin{aligned} H_p(p, S_{it}^*(p); QC_{it}) &= \frac{\partial}{\partial p} \Pr(p_t^c \leq p \mid QC_{it}, S_{it}^*(p)) \\ H_S(p, S_{it}^*(p); QC_{it}) &= \frac{\partial}{\partial S} \Pr(p_t^c \leq p \mid QC_{it}, S_{it}^*(p)) \end{aligned}$$

$H_p(p, S_{it}^*(p); QC_{it})$ is the “density” of market clearing price when firm i bids $S_{it}^*(p)$. $H_S(p, S_{it}^*(p); QC_{it})$ can be interpreted as the “shift” in the probability distribution of the market clearing price, due to a change in $S_{it}^*(p)$, i.e. this is the term that captures the “market power” of firm i . Notice that this derivative is always positive, because an increase in supply lowers the market clearing price, making the probability that the market clearing price is lower than a given price p higher.

Observe that the above first-order condition can be seen as a “markup” expression, where the markup in price above the marginal cost depends on how much market power firm i can exercise by shifting the distribution of the market clearing price through its own supply function $S_{it}^*(p)$.

²⁰Observe that we have not imposed independence on this joint distribution; contract quantities and demand noise can be correlated. However, as will be evident from the bidder’s objective function below, this is not a common value environment, as other bidders’ contract quantities does not enter into the bidder’s ex-post utility.

As an intuitive consequence, observe that if $H_S \rightarrow 0$, i.e. there is no market power, price equals marginal cost. Also observe that where $S_{it}^*(p) - QC_{it} = 0$, $p = C'_{it}(S_{it}^*(p))$. This has an interesting consequence for the empirical exercise to follow:

Proposition 1 *If $C'_{it}(S_{it}^*(p))$ is observed, one can calculate the contract position QC_{it} , by finding the quantity where the supply function of the firm intersects its marginal cost function.*

Observe also that:

Proposition 2 *Marginal cost schedules $C'_{it}(S_{it}^*(p))$ are nonparametrically identified through knowledge of contract quantities, QC_{it} and $H_{it}(p, S_{it}^*(p); QC_{it})$.*

The empirical implementation of the above identification lemma requires the estimation of $H_{it}(p, S_{it}^*(p); QC_{it})$ (and its partial derivatives), for each bidder i , in every period t . This estimate is the econometrician's estimate of the equilibrium belief of bidder i regarding the distribution of the market clearing price in auction t , conditional on his bidding strategy, $S_{it}^*(p)$. This might require strong parametric assumptions regarding the specification of bidder-specific beliefs, and especially the role played by economic unobservables entering into bidders' beliefs across different time periods.²¹ The latter concern is potentially the most troublesome: if bidders condition their beliefs on factors unobservable to the economist, estimating $H_{it}(p, S_{it}^*(p); QC_{it})$ by pooling data from a series of auctions without taking these unobservable factors into account may lead to incorrect estimates.

Observe also that the set of first-order conditions (2) for each firm, when written as a system of equations, characterizes equilibrium strategies, $S_{it}(p_t, QC_{it})$ for given primitives of the game.²² The computation of equilibrium strategies is not a trivial task, however, since $H_{it}(p, S_{it}^*(p); QC_{it})$ is determined endogenously through the market-clearing condition (1), and depends on the joint distribution of contract positions and the distribution of demand noise.

However, we now show that the characterization of equilibrium strategies in this game (2), along with econometric inference regarding unknown marginal costs, is greatly simplified when the functional form of the supply function strategies, $S_i(p, QC_i)$ is restricted to a class of strategies that are *additively separable in the private information* possessed by bidders:

Proposition 3 *If supply function strategies $S_i(p, QC_i)$ are restricted to the class of strategies: $S_i(p, QC_i) = \alpha_i(p) + \beta_i(QC_i)$, the markup relation (2) is given by the familiar "inverse-elasticity" markup rule: $p - C'_i(S_i(p, QC_i)) = \frac{S_i(p, QC_i) - QC_i}{-RD'_i(p)}$ where $RD'_i(p)$ is the price derivative of the ex-post realization of the residual demand curve faced by bidder i .*

Proof. See Appendix C. ■

²¹Some of these assumptions and possible estimation strategies based on such assumptions are discussed in the discriminatory (pay-as-bid) share auction context by Hortaçsu (2002).

²²The primitives of this game are the set of firms who are participating, N , their cost curves $C_{it}(q), i = 1, \dots, N$, the joint distribution of contract quantities and the distribution of the uncertain demand component.

The intuition underlying this result is tightly connected to the discussion in the introduction. Observe that the additive separability restriction implies that, in equilibrium, the residual demand function faced by bidders is also additively separable in its random component – i.e. all uncertainty (from the perspective of bidder i) *shifts* the residual demand curve but does not *rotate* it. Given this, it is easily seen that (subject to the concavity of the profit function), the bid function $S_i(p)$ provides a *pointwise best-response* to every possible realization of the residual demand curve. This also means that the class of additively separable equilibrium strategies, when they exist, are *ex-post optimal*, in the sense that seeing other bidders' supply functions would not change bidder i 's choice of supply function.²³

As an important caveat, however, note that the additive separability restriction is an *a priori* restriction on bidding strategies. It is not necessarily true that every specification of marginal cost functions, $C'_i(q)$ and joint distribution of contract quantities QC_i , will lead to equilibrium strategies of this form.²⁴ However, Appendix D works through an example in which firms possess linear marginal cost curves (these can be asymmetric across firms), under which equilibrium strategies are analytically characterizable, and satisfy the additive separability restriction. Based on our inspection of actual marginal cost curves, linearity appears to be a reasonable approximation.

This result has two immediate practical applications:

Proposition 4 *Suppose supply function strategies $S_i(p, QC_i)$ are restricted to the additively separable class of strategies: $S_i(p, QC_i) = \alpha_i(p) + \beta_i(QC_i)$. Then:*

1. *Given data on the marginal cost function, one can compute the ex-post optimal supply curve $S_i^{xpo}(p)$, which is the ex-post best response to the observed realization of the residual demand curve.*
2. *Equivalently, given QC_i , one can compute the entire marginal cost curve rationalizing a supply curve $S_i(p)$ observed in the data, using a single realization of the residual demand curve.*

We start with part (1). Observe that under the above restriction, a single realization of the residual demand curve, $RD_i(p, \varepsilon, QC_{-i})$ is enough to compute $\frac{d}{dp}RD_i(p, \varepsilon, QC_{-i}) = RD'_i(p)$ for all realizations. Then, for a range of prices, $p \in [\underline{p}, \bar{p}]$, one can solve the equation for S , in terms of p and QC_i :

$$p - MC_i(S) = \frac{S - QC_i}{-RD'_i(p)} \quad (3)$$

to trace out $S^{xpo}(p, QC_i)$, which constitutes an ex-post best-response to all possible realizations of residual demand.

²³The additive-separability restriction appears to be crucial. Note that without this linearity restriction, we can not, in general, collapse the stochastic terms (from the perspective of bidder i) into a single scalar random variable. See Appendix C for further discussion of the case where private information leads to rotations in residual demand, as opposed to shifts.

²⁴In particular, for certain specifications of marginal costs, a bidder's best-response to additively-separable bidding strategies by her opponents may not be additively separable.

Part (2) follows using the same logic. This time we solve the equation:

$$p - MC_i = \frac{S_i(p) - QC_i}{-RD'_i(p)} \quad (4)$$

for MC_i , in terms of p , $S_i(p)$ and QC_i . This yields MC_i as a function of $S_i(p)$, which exactly traces out the marginal cost function for all points on the support of $S_i(p)$.

Thus, although additive separability is a restrictive assumption, its imposition aids us greatly in solving for optimal supply schedules. Perhaps more importantly, ex-post optimality aids us in dealing with economic unobservables. Observe once again that the ex-post optimal supply schedule corresponding to a given marginal cost schedule (or equivalently, the marginal cost schedule rationalizing an observed bidding schedule) is derived using a *single*, ex-post observation of residual demand. The proposed empirical procedure does not pool data across auctions, and thus avoids the problem of making strong assumptions as to how to model the role of unobservables.

4 Analysis of Observed Bid Schedules

In our empirical application, we implement part (1) of Proposition 4. Specifically, we use data on the marginal cost functions of electricity generators to calculate the ex-post optimal supply curve, which is also an equilibrium bid function under the restrictions imposed in Proposition 3. Then we compare ex-post optimal bid schedules to actual bids.²⁵

The calculation largely follows the discussion of Figure 1 in the Introduction. In order to solve equation (3) for the ex-post optimal bid function, $S_i^{xpo}(p, QC_i)$, we require measures of QC_i , $MC_i(S)$, and $RD'_i(p)$. As we describe below, we collect data on each firm's hourly marginal cost function and bid schedule as well as information on the hourly realization of total balancing demand.

In order to measure the contract position, QC_i , we use Proposition 1 which allows us to identify the contract position under the assumption of equilibrium bidding. QC_i is measured as the quantity at which the actual bid schedule intersects the marginal cost function (point A in Figure 1). Note that we also can identify the contract position under certain forms on suboptimal or non-equilibrium bidding. This identification approach is valid as long as the firm is sufficiently sophisticated to bid above (below) marginal cost when it is a net seller (buyer), even if it errs in the size of the markup.²⁶ Our conversations with various market participants lead us to believe that traders clearly recognize the rationale for markuping up bids for quantities greater than the contract position (and vice versa), but have different heuristics for choosing the size of the markup.

Each firm's residual demand $RD_i(p)$ is the realized total demand minus the bids by all rival firms. Suppose RD_1 in Figure 1 is the actual realization of residual demand for firm i . We

²⁵Because only balancing sales in ERCOT are sold through the auction, we have to modify the notation of the model in fit our application. Demand \tilde{D}_t is now defined as total ERCOT demand minus the total quantity scheduled day-ahead. MC_{it} is now the marginal cost of changing output from the firm's day-ahead schedule. And QC_{it} is now the amount that the firm is long or short on its contracted sales after the day-ahead schedule and upon entering the balancing market.

²⁶More formally, we can identify QC_{it} in equation (2) even if firms have incorrectly calculated $\frac{H_S(p, S_{it}^*(p); QC_{it})}{H_p(p, S_{it}^*(p); QC_{it})}$.

calculate $RD'_1(p)$ and find the ex-post optimal (price, quantity) bid to be Point B, where the marginal revenue curve corresponding to $RD_1(p)$, MR_1 , intersects the marginal cost curve MC . Also, we can calculate the ex-post optimal bid under other possible realizations of uncertainty (i.e. other realizations of total balancing demand, ε , and rivals' private information, QC_{-i}). Because each form of uncertainty acts to shift residual demand in a parallel fashion, we can consider another possible realization of residual demand as RD_2 – the realized residual demand shifted parallel to the left. Under this realization of uncertainty, the optimal bidpoint is given by Point C, where $MR_2 = MC_i(q)$. We repeat this operation by adding parallel shifts to the actual realization of the residual demand curve to find the set of ex-post optimal points for various realizations of uncertainty. The ex-post optimal bid *function* is traced out by the set of ex-post optimal bid points to generate $S_i^{xpo}(p, QC_i)$. It is important to note that our assumption that the slope of residual demand is independent of uncertainty is necessary for the set of ex-post optimal points to be on a monotonic bid function.

We should also note that since the residual demand function is a step function whose derivatives are either zero or infinity, we follow Wolak (2003a) to obtain a “smoothed” version of the residual demand function to calculate the marginal revenue curve.²⁷

4.1 Data

We use data on hourly balancing demand and firm-level bids and marginal costs. To construct each firm's marginal cost function, we utilize data on the generating units operating on a given hour and the hourly declared capacity of each unit. The marginal cost of operating units represents the variable costs – fuel, operating and maintenance, and SO2 permit costs – of coal and natural gas fired units. A growing body of literature has developed on measuring the marginal cost of electricity production (for example, see Wolfram[1999], Borenstein, Bushnell and Wolak [2002], Mansur [2002], Puller [2004], Joskow and Kahn [2002], Bushnell and Saravia [2002]) and we use the same approach. Details of the data and additional institutional considerations are discussed in Appendix A.

Our “marginal cost of balancing power” function is the costs of providing more power (INCing) or the cost savings of reducing production (DECing) from the day-ahead scheduled quantity. We construct this function by first calculating the total marginal cost function of producing in a given hour, and then subtracting the quantity that has already been schedule one-day ahead. Total marginal cost is the marginal production cost of units that are reported by ERCOT to be operating and available in period t . It is reasonable to assume that firms produce in a least cost manner, so we stack up the marginal cost of each generating unit from cheapest to most expensive and construct the total marginal cost function.²⁸ We have data on how much generation has been scheduled

²⁷Let $\{(p_1, q_1), \dots, (p_K, q_K)\}$ represent the price and incremental quantities that form the residual demand curve seen in the data. The smoothed version of this function is $RD(p) = \sum_{k=1}^K q_k K(\frac{p-p_k}{h})$ where $K(\cdot)$ is a kernel function. With this representation, the derivative of residual demand is $RD'(p) = \sum_{k=1}^K q_k \frac{1}{h} K'(\frac{p-p_k}{h})$. We used a normal kernel and the smoothing parameter, $h = 10MW$ throughout our analysis.

²⁸Our conversations with several traders in the market suggest that they spend significant resources coming

day-ahead by the firm. Thus the marginal cost of supplying balancing power is characterized as the total marginal cost function with the origin re-centered at the day-ahead schedule. Certain types of generating units that cannot supply power on short notice are then excluded from this marginal cost stack. Natural gas-fired units, and to a lesser extent coal units, can be adjusted on relatively short notice to other production levels. Other types of units such as nuclear, wind, and hydroelectric typically cannot respond to balancing market calls. Therefore, we exclude nuclear, wind, and hydroelectric generating units from the marginal cost portfolio for providing balancing energy.²⁹ For a sample marginal cost function, see Figure 2 where Reliant has the capacity to increase production by 900 megawatts (MW) at a cost up to about \$40, and decrease production by over 3000 MW at a cost savings between \$30 and \$35.

We use firm (portfolio) bid schedules for each firm for the 6:00-7:00pm hour of each non-congested weekday.³⁰ To determine the sales into the balancing market, we intersect the actual bid with the realization of residual demand to determine sales. We assume a step function which is a good approximation to the actual dispatch algorithm used. This simulation of the actual auction predicts prices with only a 5% error.

4.2 Comparison of Actual Bids to Ex-Post Optimal Bids

Figure 7 displays representative actual and ex post optimal bid functions for three large suppliers – Reliant, TXU, and Calpine – and for one small seller, Guadalupe. In each panel, competitive bidding is offering the “MC curve” and optimal bidding is offering the “Ex-post optimal bid.” The intersection of the actual bids and marginal cost schedules is the contract position. For quantities above (below) the contract position, the ex-post optimal bid function is above (below) marginal cost. Visually, Reliant’s bids appear much closer to the optimal bids than to the marginal cost function. In addition, Reliant’s performance for incremental supply (Balancing MW > 0) is relatively similar to its performance for decremental supply.

TXU is close to the ex-post optimal bid function on the INC side, but bids below ex-post optimal prices on the DEC side. We see this tendency for TXU to offer DEC bids at only very low prices throughout our sample period.

Calpine offers some DEC bids but does not offer to INC supply. Although Calpine does offer INC bids in some periods, much of Calpine’s bids are to DEC supply. Those DEC offers are often at prices substantially below ex post optimal bids.

up with the “least cost” arrangement. For example, a specialized computer program called PCI GenTrader (www.powercosts.com) is utilized by some traders to compute the least cost dispatch arrangement.

²⁹Some hydroelectric units may be able to respond to balancing calls, however these units represent less than 1% of total capacity and are primarily owned by Lower Colorado River Authority.

³⁰A few firms submit their bids through other firms which makes it impossible to separate the bids by firm. We address this complication by using all bids when constructing residual demand curves, but only interpreting the bids as “firm-level behavior” if the bidder primarily submits bids for a single firm. If a single generation owner supplies at least 90% of the total generation supplied by the bidder, we define that bidder as a firm. This partially limits our ability to analyze behavior over parts of our sample. For example, the largest supplier of balancing power, Reliant, cannot be separated from City of San Antonio Public Service in 2001, so we cannot analyze Reliant’s behavior until 2002.

Guadalupe submits bids that are much steeper than ex-post optimal. Because it is a small seller, Guadalupe’s residual demand function is relatively flat as compared to the residual demand of the larger players. Nevertheless, small sellers have some potential to bid strategically and exercise market power. However, the actual bid functions are significantly above the optimal bids in INC periods and below optimal bids in DEC periods. This suggests that bids significantly different from marginal cost are not intended as a means to exercise market power, but rather to avoid being called upon to change production from day-ahead schedules. Many small sellers show similar bidding patterns. We will discuss several reasons underlying these patterns in Section 5.

Finally, note that the firms use coarse-grained strategies that utilize a small number of bidpoints. All three bidders use far fewer bidpoints than the maximum permitted, and it appears that using more bidpoints would allow the firms to better fit the relatively smooth optimal bid function. This observation is discussed in detail in Section 5.4.

The visual inspection of the figures are suggestive of the closeness of actual bidding behavior to ex-post optimal behavior, but a more meaningful metric to evaluate bidders’ performance is to measure how much profit they have foregone ex-post by deviating from the ex-post optimal bidding schedule.

To calculate the profit deviation, we calculate the difference of the producer surplus obtained at the actual submitted price/quantity point (point D in figure 1), and the surplus obtained at the ex-post optimal point (point B in figure 1). We calculate this difference each firm-period for 20 simulations of residual demand which we construct by adding uniformly distributed noise to the actual demand. These simulations allow us to evaluate the optimality of several points on the bid function that would be the market clearing price under possible realizations of residual demand. The results are generally robust to the scale of the noise added.

We also calculate the producer surplus achieved relative to a benchmark of “suboptimal” behavior to compare how much distance is closed between the benchmark of suboptimal pricing and optimal pricing.³¹ One possible benchmark is behaving non-strategically and bidding marginal cost. However, it appears that the “default” behavior is to bid to avoid being called to supply balancing power. As shown in the sample figures, smaller firms choose to bid only small quantities relative to both competitive and optimal bidding. Therefore, we measure performance as the fraction of (dollar) distance between “no bidding” and ex post optimal bidding that is realized by the actual bids. Producer surplus is:

$$\pi = P^{BAL}(Q_i^{BAL}) - TC(Q_i^{BAL}) - (P^{BAL} - PC)QC$$

We calculate π for three scenarios: (1) ex post optimal bidding, (2) actual bidding, and (3) bidding to avoid the balancing market (i.e. not bidding at all and buying/selling at the market clearing price any net short or long position on contracts). Our primary measure of performance is:

³¹We cannot use a measure such as the fraction of possible profits achieved because some firms tend to be short on their contract positions entering the balancing market, and we would have to make an assumption about the contract price. The measure we construct differences out the contract price and avoids this complication.

$$PercentAchieved = \frac{\pi^{Actual} - \pi^{Avoid}}{\pi^{XPO} - \pi^{Avoid}}$$

which we calculate for each firm in the market.

One should keep in mind that *PercentAchieved* is only a metric of the generator’s performance in the *balancing market*. It does not account for the profitability of the vast majority of output that is sold through bilateral transactions, and therefore may understate the overall profitability of electricity sales. In order to measure the profits of bilateral sales we would require data on contract prices, but such data are not available. However, we can construct an upper bound for overall profitability by assuming that each generator is maximizing profits in the bilateral market but may be sacrificing profits in the smaller balancing market:

$$Upper\ Bound\ Total\ Percent\ Profitability = PercentAchieved * \%Sales\ in\ Balancing \\ + 100\% * \%Sales\ in\ Bilaterals$$

4.2.1 Ex-post Profitability

Table 3 compares output and producer surplus in the balancing market under actual and ex-post optimal bidding. Potential profits relative to not bidding are shown in column 5. For the largest firms with average sales of at least 250 MWh under ex-post optimal bidding, the potential profits average about \$2300/hour for each 6:00-6:15pm interval, with Reliant having the most “money on the table” of \$4,333 each hour. Firms that would sell less than 250 MWh under ex-post optimal bidding would average about \$750/hour in potential profits ranging from \$2,176/hour for Air Liquide America to \$31/hour for Denton Municipal Electric.

The first column displays the percent of potential profits achieved under the actual bid schedules. Reliant achieves 79% of ex-post optimal profits, or \$3,422/hour of the \$4,333/hour of potential profits. The next two firms closest to ex-post optimal profits are two relatively small municipal utilities – Brownsville (50%) and Bryan (45%). TXU, the second largest incumbent utility, achieves 39% of potential profits while the largest independent power producers, Tenaska Gateway and Calpine, achieve 41% and 37%, respectively. The two other major incumbent utilities – West Texas Utilities and Central Power and Light – capture only 8% of potential ex-post profits. The other large municipal utilities earn a moderate fraction of potential profits: Austin (30%) and San Antonio (23%). The largest electricity cooperatives earn lower profits: LCRA (25%), Brazos (15%), and South Texas (3%).³²

Several characteristics of bid schedules appear to drive the foregone ex-post profits. We exam-

³²Note that two firms, Extex Laporte and Air Liquide, earn lower profits under actual bidding than bidding to “avoid the market.” Both firms, which are infrequent participants in the balancing market, have positive contract positions and relatively high cost units available. Both bid so they are called to INC despite the fact that it would be more profitable to not participate and buy its contract position from the market at a price lower than marginal cost.

ined many sets of actual and ex-post optimal bid functions similar to Figure 7. The key characteristic of underperforming bidders is that the actual bid functions tend to be “too steep” relative to optimal bids. By bidding too high during INC hours and too low during DEC hours, firms sell less in the balancing market than under ex-post optimal bidding and forego individually profitable sales. As shown in columns 6 and 7 of Table 3, firms sell less than the ex-post optimal quantity on average. Interestingly, the cause of foregone profits is not that firms are bidding more competitively than individually optimal, but rather that bid prices are too high. Also, firms use coarse-grained strategies that utilize far fewer bidpoints than the maximum permitted, as we discuss in more detail below.

It is important to keep in mind that these profitability measures do not necessarily represent generator’s overall performance in trading electricity. Many generators may focus their strategic efforts in the bilateral market where the vast majority of transactions occur.³³ If firms focus strategic efforts on the bilateral markets, the overall performance is substantially higher. The upper bound on total profitability, shown in column 8 of Table 3, ranges from 41% to 98% with a mean of 80%. These metrics are more likely to reflect the firms’ overall performance.

4.3 Tests of “Ex-Ante” Expected Profit Maximization

It is important to note that the calculation of foregone profits in the previous section relies on restrictions that we place on the economic environment. In particular, we assume that bidders’ equilibrium perceptions of the uncertainty regarding residual demand results in “parallel shifts” rather than “pivots” in residual demand. If this assumption is violated, it is possible that the set of ex-post optimal price-quantity points is a “cloud” of points that cannot be connected by a monotonic supply function. If this were the case, a profit-maximizing firm’s solution to the (ex ante) *expected* profit maximization problem may not be the set of *ex-post* optimal price-quantity points. That is, testing whether firms maximize “ex-post” profits as we did would not be informative about whether firms maximize “ex-ante” profits. Since “ex-ante” profitability is a more accurate measure of bidder performance in the presence of more general sources of uncertainty, we employ two additional testing strategies.

4.3.1 Best-Response to Previous Rival Bids

First, we test if firms could significantly increase profits by using a simple bidding rule that utilizes only information available to traders at the time of bidding. The bidding rule we will employ is a naive best response to recent rival bidding. As discussed in section 2, traders have access to the aggregate bid function with a two day lag. Using their own bids from the past, traders can calculate the aggregate bids by rivals in the recent past. To construct the naive best response of firm i for upcoming day t , we:

³³Although we have no formal evidence, it appears that the skill set of traders who negotiate bilateral trades is different from that of traders who sell through electronic auctions.

1. Use aggregate bids and own bids for day $t - 3$ and calculate aggregate rival bids on day $t - 3$
2. Assume rivals use the $t - 3$ bid schedule on upcoming day t
3. Calculate ex-post optimal bid function for various realizations of day t total balancing demand

Note that this algorithm uses only information available to firms when bids are submitted. We view this bidding rule as fairly unsophisticated – it uses only a small fraction of the information available to traders and it would be simple to program as an add-in to the trading interface used by the generators. We calculate producer surplus under “naive best response” bidding and compare to producer surplus under actual bidding. If uncertainty causes the expected profit maximizing bid to differ from ex-post optimal bids, the actual profitability should be much closer to the naive best response benchmark.

Results are shown in column 4 of Table 3. Across all firms, naive best response profits are substantially higher than actual profits and very close to ex-post optimal profits. The performance of actual bidding is significantly below the naive best reply benchmark for all firms except Reliant. In fact, most bidders’ performance measure of “Percent Achieved” rises very little when compared against the naive best reply benchmark. Moreover, naive best response profits are very close to ex-post optimal profits – the former average \$1,193 and the latter average \$1,204. This suggests that our findings of “money left on the table” do not arise from the restrictions placed on the economic environment.

4.3.2 Moments-Based Test for Expected Profit Maximization

A second test for expected profit maximization uses a GMM based approach inspired by Wolak (2003a). Under the null of expected profit maximization, each bid price and quantity is chosen optimally so that incremental changes in either the price or quantity should not increase profits *on average*. For example, if a firm uses 3 bidpoints, then the effect of increasing the price and holding the quantity constant (or vice versa) of any of the three bidpoints should not locally increase profits on average across the sample.

Assume a firm bids k bidpoints each period and each point is given by (p_k, q_k) . Let the bid function be given by $\Omega \equiv (p_1, q_1, \dots, p_k, q_k)$. Suppose any uncertainty in residual demand is determined by a random variable ν , and we do not restrict the relationship between ν and residual demand (i.e. it need not be additive). Let profits be given by:

$$\Pi(\Omega, \nu) = RD(p(\nu, \Omega), \nu) \cdot p(\nu, \Omega) - C(S(p(\nu, \Omega), \Omega)) - (p(\nu, \Omega) - PC)QC$$

The first-order conditions under expected profit maximization with respect to each price and quantity bid generate the following moment conditions:

$$E_\nu \left(\frac{\partial \Pi(\Omega, \nu)}{\partial p_k} \right) = 0 \quad E_\nu \left(\frac{\partial \Pi(\Omega, \nu)}{\partial q_k} \right) = 0$$

Wolak derives sample analogs to the first-order conditions and shows how the moments can be used to test expected profit maximization in a GMM setting. The sample analogs of the moments take the form:

$$\frac{\partial \Pi}{\partial p_k} = \left[\frac{\partial RD(\cdot)}{\partial p} p(\Omega, \nu) + RD(p, \nu) - QC - \frac{\partial C(\cdot)}{\partial S} \frac{\partial S}{\partial p} \right] \frac{\partial p(\Omega, \nu)}{\partial p_k} - \frac{\partial C(\cdot)}{\partial S} \frac{\partial S}{\partial p_k} \quad (5)$$

$$\frac{\partial \Pi}{\partial q_k} = \left[\frac{\partial RD(\cdot)}{\partial p} p(\Omega, \nu) + RD(p, \nu) - QC - \frac{\partial C(\cdot)}{\partial S} \frac{\partial S}{\partial p} \right] \frac{\partial p(\Omega, \nu)}{\partial q_k} - \frac{\partial C(\cdot)}{\partial S} \frac{\partial S}{\partial q_k} \quad (6)$$

The first part of each moment is “marginal revenue minus marginal cost” at the market-clearing price multiplied by the marginal effect of changing the price/quantity on the market-clearing price. The second part is the effect on production costs of changing the price/quantity and being called upon to change output.

Wolak (2003a) uses the above price moments to estimate the marginal costs that (in expectation) rationalize observed bids using GMM methodology. In our context, we know (or assume we know) the marginal costs for each period, hence we only need to evaluate the above first-order conditions. For each firm, this yields a vector of residuals for each period, $u_t, t = 1, \dots, T$. Under the null hypothesis of expected profit maximizing behavior, the moments evaluated at the actual bidding outcomes are zero in expectation. To test whether this is indeed the case, we evaluate the Hansen J-statistic, given by:

$$J_T = \left(\left[\frac{1}{T} \sum_{t=1}^T u_t \right]' S_T^{-1} \left[\frac{1}{T} \sum_{t=1}^T u_t \right] \right) \quad (7)$$

where $S_T = \frac{1}{T} \sum_{t=1}^T [(u_t - \bar{u})(u_t - \bar{u})']$ is the sample covariance matrix of the residuals.³⁴ Hansen (1982) shows, within the context of GMM estimation, that $TJ_T \sim \chi_{p-k}^2$, where p is the number of moments, and k is the number of parameters. In our application, there are no unknown parameters, so we set $k = 0$.

In our calculations of the J-statistic for the firms in our sample, we found that the test using quantity moments had much higher power than tests using price moments. One intuition for this is the following. Consider a hypothetical firm that does not maximize expected profits. Suppose the firm could profitably participate in the balancing market, but chooses not to participate except in the rare instance when the INC price is very high. The firm bids to produce zero output unless price is \$100 in which case it INCs 1 MW. Under almost all possible realizations of residual demand, the market clearing price is lower than \$100 so the firm is not called upon to INC. Although this firm is not maximizing expected profits, the price moments are essentially zero because the bid price is far enough away from typical market clearing prices that a small increase in the bid price is unlikely to affect the market clearing price ($\frac{\partial p}{\partial p_k} |_{p_k=100} \approx 0$) and increasing the bid price slightly above \$100 is unlikely to affect the quantity sold ($\frac{\partial S}{\partial p_k} |_{p_k=100} \approx 0$). Thus, tests using price moments could easily fail to reject the null of expected profit maximization for this hypothetical (non-expected

³⁴Inspection of the residual autocorrelation function did not yield significant temporal dependence. Accordingly, we found that the Newey-West estimator for the covariance matrix did not yield different J values.

profit-maximizing) firm.³⁵

However, quantity moment tests are likely to be higher power for firms that submit very steep bid functions as is the case for many firms in ERCOT with large bid-ask spreads. The quantity moment test asks whether our hypothetical firm could marginally increase profits by bidding some positive quantity at all prices up to \$100. Moments for the hypothetical firm are likely to be non-zero for two reasons – bidding to INC positive quantities at low prices is likely to increase the quantity called ($\frac{\partial S}{\partial q_k} > 0$) and the increased quantity bids will decrease the overall market clearing price ($\frac{\partial p}{\partial q_k} < 0$) at prices where marginal revenue exceeds marginal cost. Our choice of quantity over price moments stems from the fact that firms with large bid-ask spreads (e.g. Guadalupe) are more likely making “quantity mistakes” than “price mistakes”.

Table 4 reports the results of our statistical test of expected profit maximization using quantity moments. We report the results for the top 6 performers in Table 3, who realized more than 35% of their ex-post profits. To perform this test, we evaluated the quantity derivative on 11 equally spaced price points between 16 dollars and 36 dollars (the interquartile range of spot prices). The magnitudes of the J-statistic roughly follow the ordering in Table 3. However, none of these top 6 firms except Reliant and Brownsville are close to passing the test of expected profit maximization (the 5% critical value of the χ^2 distribution with 11 degrees of freedom is 19.7). Reliant only passes this test if we exclude sample observations after November 2002. This cutoff point in time is not spurious – when we inspected Reliant’s performance in terms of achieving its ex-post profits, we found a marked drop off starting November 2002 (Reliant achieves 84% of ex-post profits on average until the end of October 2002, but achieves only 31% of ex-post profits subsequently). Although we do not know the exact cause for the dip in Reliant’s performance, one of our interviewees indicated that there was an organizational change in the bidding operations of Reliant when a new subsidiary began to bid for some of Reliant’s generating units.

These results indicate that, except for Reliant, firms in this market violate the first-order optimality conditions that need to hold (on average) for expected profit maximization. This is broadly in agreement with our ex-post profitability measures and the “naive best-response” benchmark, and suggests that firm behavior in this market displays consistent departures from static, unilateral profit maximization.

5 Explaining Deviations from Optimal Bidding

This section investigates several explanations for the observed deviations from static profit maximization and the considerable heterogeneity across firms in terms of performance.³⁶ In particular, we consider the possibility that the observed deviations are driven by unmeasured costs of production. We then consider whether expectations regarding transmission congestion may drive some of the deviations from the optimal benchmarks. We also assess whether a dynamic pricing game

³⁵One could view the moments as only *local* tests of optimality.

³⁶Note that explanations based on risk-preferences are ruled out by the fact that the first-order condition of optimality (equation(2)) does not depend at all on the curvature of the bidder’s utility function.

may explain these patterns. Finally, we explore whether the observed deviations are driven by characteristics of the firms such as the firm size, the firm type (e.g. investor-owned utility vs. municipal utility), and the generation technology. We find that the most significant determinant of performance is the size of the stakes that each firm has in the balancing market.

5.1 Adjustment Costs

We consider the possibility that there are costs of supplying balancing power that we fail to measure. A first observation is that even if our measures of marginal costs are biased up to a *level* shift, this bias would not affect our finding regarding the slope of observed bidding schedules – i.e. the fact they are too steep. However, we also could be erring in measuring the slope of the marginal cost curves. Recall that our measure of marginal cost is the marginal fuel and operating/maintenance cost of a generator that is constantly operating; we do not explicitly measure the cost of adjusting the level of output from the day-ahead schedule.

If firms faced large adjustment costs, we would expect to see large steps in the bid schedules at a balancing quantity of zero.³⁷ We find that many firms forego profits because their bid functions have a large range of prices at which quantity offered is zero. To a varying degree, we see such phenomena in bids. For example, in Figure 7, Guadalupe is not offering to supply additional INC energy until the price reaches \$33 nor offering to reduce production until price is \$11 despite the fact that one of its units that can both INC and DEC has marginal cost of \$28. We call this difference in prices the “bid-ask spread.”

Although some firms’ bids are consistent with the presence of adjustment costs, there is only weak evidence that real adjustment costs exist. One source of these adjustment costs might be the “transaction cost” associated with the actual implementation of a balancing market adjustment in their generation – i.e. the cost of coordinating and communicating with the engineering staff to adjust production levels on short notice. However, our interviews suggest that most firms in ERCOT have invested into software and hardware that electronically receives the balancing call from ERCOT and adjusts the physical output from the generators. Adjustment costs could also arise if certain generating technologies are less flexible in adjusting output from the day-ahead schedule. In particular, coal-burning generators tend to be less flexible than natural gas-fired units. In order to test this explanation, we identify the fuel type of the marginal production unit for each firm-interval for those firms that serve as their own bidders. For all firms with the exception of one, a gas-fired unit is marginal in at least 70% of the time periods with several firms having gas units on the margin over 90% of the time.³⁸ In addition, the periods when there are no gas units available to INC or DEC are uncorrelated with lower percentages of profits achieved. Therefore,

³⁷Another means firms can utilize to indicate adjustment costs is the ramprates submitted with the bid schedule. The ramprates indicate the maximum rate at which a given generating unit can increase or decrease production. If the quantity a firm would be called under the bids violated a ramprate, the unit would be skipped over in favor of the next unit in the bidstack. When this occurs, the balancing price reported by ERCOT would differ from the price we calculate using the bid stack. However, as we discuss above, the two prices are very close for the periods we analyze.

³⁸The firm that does not frequently have a gas unit on the margin is South Texas Electricity Cooperative. We do not explore the possibility that some gas-fired technologies are more flexible than others.

we do not believe that the limited flexibility of coal units is causing us to understate the marginal cost and bias upwards our measure of foregone profits.

Although we cannot rule out the possibility that some unmeasured engineering adjustment costs may exist, it appears unlikely that those costs explain the large variation in bid-ask spreads which we observe across firms with similar generation technologies. Table 5 displays the average bid-ask spread in column 3. Reliant averages a relatively small spread (\$2.06) suggesting relatively smooth bid functions as depicted in Figure 2. The other generators' bids have bid-ask spreads averaging approximately \$20-\$30. If there were engineering adjustment costs, one would expect the costs to affect the bids of all firms with similar technologies.

Furthermore, bid-ask spreads tend to decrease over time for those firms that achieve the greatest fraction of ex-post optimal profits. This is seen in Table 8 in which we analyze the time trend in a regression setting. We regress the firm's daily bid-ask spread on a time trend and a dummy variable for the off-peak season when firms tend to have fewer units operating. The top six firms exhibit statistically smaller bid-ask spreads over time. In contrast, the remaining firms do not exhibit a significant downward trend.

There are two possible explanations for these observed trends. Perhaps some firms found means to reduce adjustment costs and those firms were able to earn higher profits in the balancing market. Such an explanation would be surprising because the production technology is fairly uniform. A more plausible explanation is that any adjustment costs are more "perceived" than real and that the best performing firms recognize the potential profits from participating in the balancing market, a possibility we will discuss in more detail in section 5.5.

5.2 Transmission Constraints

Another economic factor that may affect our analysis of static profit maximization are transmission constraints. Transmission constraints pose two complications for our analysis. First, when transmission lines between zones reach capacity, the market is divided into separate zones with different market clearing prices. To address this complication, our sample includes only the 74% of time intervals with no zonal transmission congestion when the entire ERCOT market is integrated into a single market.

Second, even if we restrict our sample to ex post uncongested hours, the ex ante possibility of congestion may affect the expected profit-maximizing bid function. Consider a generator in a zone that could be import constrained with positive probability. If the transmission line were to be congested, the generator would face no competition from outside the zone after the line reaches capacity. However, if the line were not congested, the generator faces competition from the entire state. This uncertainty introduces the possibility of kinks and convexities in the residual demand function that change the profit-maximization problem. If firms cannot perfectly forecast congestion before submitting their bid schedules, the expected profit-maximizing bid during intervals that are ex-post uncongested reflect the optimal bid if there were congestion. Therefore, the strategic effect of congestion may spillover to the uncongested hours we analyze.

We do not find evidence that the deviations from ex-post optimal bidding in uncongested intervals are significantly caused by uncertainty regarding the state of congestion. We test whether our metric of profitability is systematically related to the frequency of congestion. If the bidding strategy under congestion spills over to uncongested periods, the ex-post profitability should be lower in months with more frequent congestion. We regress the monthly measure of profitability, *PercentAchieved*, on the percent of 6:00-7:00pm intervals with transmission congestion. We also control for the percent of intervals when balancing demand is negative, firm fixed effects, and the quantity of sales under optimal bidding. Results are shown in Table 6 for all firms and for only those firms that act as their own bidders. After controlling for firm effects, we find that a high prevalence of congestion reduces profitability, but the effect is not significant for own bidders and only marginally significant for all firms. The coefficient of the congestion variable is very similar and robust to other (unreported) specifications. Despite the lack of statistical significance, we can assess the effect of congestion using the point estimates. It does not suggest the effect of congestion is large. A one standard deviation increase in the monthly percent of congested hours reduces *PercentAchieved* by 3.0%. Even moving from the month with the smallest to the largest frequency of congestion only reduces profitability by 9.4%. These reductions do not account for the measures of *PercentAchieved* substantially below 100%. Therefore, it does not appear that the possibility of congestion accounts for a significant fraction of the foregone ex-post profits measured in section 4.

5.3 Collusion

Another possibility is that the excessive steepness of observed bids (compared to *unilateral* ex-post optimal bids) results from collusion. However, the patterns of inelastic bids in ERCOT make an explanation based on dynamic pricing problematic.

First, many of the firms that sacrifice unilateral profits by submitting bid schedules that are “too steep” do so by including large bid-ask spreads around their contract quantity. Consistently large bid-ask spreads are not consistent with collusion. Large bid-ask spreads result in *zero* sales whereas dynamic pricing would result in sales that are smaller than ex-post optimal levels but are still *positive*. In particular, it is difficult to rationalize the behavior of the firms in the bottom half of Table 3 by resorting to windfalls from collusion.

Moreover, the bids most consistent with dynamic pricing are those of small, fringe firms rather than the large players. Potential candidates for dynamic pricing are those that bid above unilateral ex post optimal prices for quantities above their contract position and below optimal prices for quantities below their contract position. In general, the firms that fit this description are smaller firms that submit very inelastic bid schedules. The larger firms bid closer to ex-post optimal. Although we have not performed a systematic analysis of potential coalitions, we believe it unlikely that a collusive coalition would form that included the small but not large players.³⁹

³⁹Another explanation offered by the literature to explain observed deviations from static profit maximization is the presence of regulatory threat, as suggested by Wolfram (1999). The Texas balancing market was in its infancy

5.4 Bounded Rationality?

Yet another possibility is that observed deviations from static profit maximization are caused by boundedly rational behavior. Such an explanation is not *a priori* implausible. Bidders have indicated to us that this market is subject to a lot of uncertainty, in which complex decisions have to be made in short amounts of time, and that bidding decisions are often made based on heuristics. This conjecture also is consistent with Peter Cramton’s observations from advising electricity bidders. Cramton (2004) notes that some electricity bidders explicitly compute residual demand functions while others use much more ad hoc, experimental approaches.

We now discuss two pieces of evidence that are suggestive of boundedly rational behavior in this market, particularly by the smaller bidders. The small firms do not appear to select bids based upon sophisticated analyses of their strategic environment and, as a result, sacrifice profits in the balancing market. In section 5.5, we claim that this apparent sacrifice of profits by small firms may be rational, given the large fixed costs of establishing a sophisticated bidding operation.

5.4.1 Coarse-grained Bidding strategies

The bid rules for the ERCOT balancing market allow each bidder 40 price-quantity points in each geographic zone (20 points for each zone’s inc and dec bid stacks). This affords generators a large degree of flexibility in bidding, especially those firms that own generation in more than one zone.

However, none of the bidders in our sample make full use of the 40 price-quantity points that they can use to trace out their optimal bidding functions. Table 5 shows the average number of bidpoints used by all firms that serve as their own bidder. The firm earning the greatest fraction of ex-post profits (Reliant) also uses the largest number of bidpoints, averaging 22.2 elbow points per bid schedule. None of the other firms use more than 13 points on average. Although profitability is not strictly rising in average bidpoints, the two measures of bidding sophistication are positively correlated. Apparently, traders choose not to formulate refined bid strategies with desired quantities for many potential realizations of the market clearing price. This is consistent with anecdotal evidence in which one market participant said that traders would try to bid above marginal cost, and the choice of the number of bidpoints depended upon “how detailed you wanted to be.”

5.4.2 “Sunk cost” heuristics?

There is some evidence that smaller firms base their bidding decisions on heuristics. One heuristic that would work well for the small firms in this market is to recognize that they have very little market power, and to bid a small markup above their marginal cost. However, our data suggests

at the same time the revelations about suspicious trading strategies by firms such as Enron put many energy trading firm in the cross hairs of regulators. Production strategies such as unilateral profit-maximization that are legal in most industries in the U.S. were questioned as illegal and unethical in the electricity industry. Energy firms were in limbo as to the rules of the game and “how far they could go” to maximize profits. However, it is noteworthy that the observed bidding behavior is not consistent with firms that fear “going too far”. To the contrary, bid functions were steeper than unilateral market power levels – prices were too high on the INC side and too low on the DEC side. If anything, firms were “too greedy” rather than “too competitive”.

that these small bidders are doing exactly the opposite by bidding too high. Our interviews of several small players with low performance metrics yielded several interesting explanations for this. For example, several bidders who own newer generating units that are financed under debt claimed that they add an additional markup to cover debt. Although they acknowledged this is essentially a fallacy of sunk costs, company officials claimed this nevertheless is done. One firm said that it sets high prices during uncongested intervals in order to “make up for losses” during periods of congestion when compensation under special market rules is allegedly low. We find that those high prices often priced the firm out of the market and lowered profits. Finally, some smaller players offered an explanation for not participating in the market even if the firm had additional capacity. Some players claim to keep extra capacity available so they are not caught short if another one of their units were to suffer an outage. Economic logic suggests it is optimal to bid extra capacity into the balancing market regardless of whether another unit unexpectedly goes offline. If another generating unit does not have an outage, the firm loses the opportunity of selling the extra capacity to the balancing market. If another unit does go offline, there is no difference between selling extra capacity to the market and buying it back at the market price, i.e. ‘selling the capacity to oneself’ at an opportunity cost of the market price.

5.5 Participation Costs and Scale Economies in Bidder Performance

We now consider the hypothesis that small firms might not have sufficiently large dollar stakes to justify the fixed cost of participating in the balancing market. To do this, we first have to clarify what we mean by “participation.” Throughout sections 4 and 5, we have documented that small firms submit bid schedules with large bid-ask spreads that effectively price them out of the market for plausible realizations of residual demand (in fact, as Table 3 shows, some firms at the bottom of this table are almost always priced out of the market). As noted in Section 5.1, these large bid-ask spreads can be rationalized by an “adjustment cost” that prevents these small firms from supplying power to the balancing market. However, as discussed in that section, we fail to find evidence for purely technological sources of such “adjustment costs.”

An alternative interpretation of these large bid-ask spreads is that it does not pay for the small firms to bid the optimal markup over their marginal cost, even if this optimal markup would allow them to profit from incrementing/decrementing their power generation, or “participating” in the balancing market much more often. This interpretation is plausible if it is costly for these firms to calculate the optimal markup conditional on participating. As a result, firms may find it profitable to devote substantial strategic resources to the bilateral but not to the balancing market.

Indeed, there is evidence that performing optimal markup calculations on a real-time basis is costly.⁴⁰ Through field trips and phone calls, we have observed that firms with large stakes in the balancing market have invested substantial amounts of money in energy trading hardware and

⁴⁰Note that although expected profit maximization predicts that small firms should bid close to marginal cost due to their lack of market power, an easy enough decision to make, these firms would not want to bid *too* close to marginal cost, either. Such a strategy would also yield low profits, especially when the firm’s marginal cost schedule is essentially flat, as is the case for small firms with a single generation plant.

software, and hire experienced energy traders. Those firms that do not invest the fixed cost to establish a trading operation may simply bring over personnel from other operations who do not have energy trading experience. The fixed and variable costs of running a trading operation are likely not to be trivial. One market participant suggested that even a simple bidding operation would require an upfront expenditure of \$3 million with annual operating costs of \$1 million, and that most sophisticated trading operations could be much more expensive. The magnitude of these costs is not trivial compared to the “money-left-on-table” figures reported in Table 3.⁴¹

Another component of this “fixed cost of participation” is institutional. ERCOT only allows certified companies, called “qualified scheduling entities” (QSEs) to submit bids in the balancing market.⁴² All other firms have to route their bids through a QSE, or contract with a QSE to conduct their bidding operations. This suggests that only firms with greater dollar stakes may find it optimal to incur the fixed costs of becoming a QSE.

Even conditional on paying the fixed cost of becoming a QSE, scale economies still appear to matter. This is clearly seen in Figure 8, which displays the relationship between bidding performance and size for generation firms who conduct their own bidding – i.e. the firms who act as their own QSE. Our measure of performance is the percent of ex-post optimal profits, calculated in the manner described in section 4.2. Our measure of stakes in the balancing market is the volume of sales under ex-post optimal bidding (using other size measures, such as actual volume of sales, or firms’ total capacity, yield similar patterns). There is a positive relationship between “Percent Achieved” and optimal sales volume. The figure includes the fitted linear relationship which is positive and marginally significant when all firms are included. When Bryan is excluded, the relationship is even stronger and highly significant.

One literature that has looked at departures from profit-maximization is the literature on corporate governance.⁴³ One might, for example, expect municipal utilities to behave very differently from investor-owned utilities, or merchant power generators. A quick look at Figure 8, where we have labeled whether a firm is a municipal or investor owned utility (only Calpine is a merchant in this figure) appears to reveal a weak contrast; though, it is difficult to disentangle this outside of a multiple regression framework – which is what we do next.

We now examine the extent to which stakes are correlated with performance for the broader sample of firms in a regression context, along with other firm-specific factors that we believe might affect performance. We regress each firm’s measure of “Percent Achieved” on a measure of stakes in the balancing market – the volume of sales under ex-post optimal bidding (SIZE). Also included are firm-level covariates on firm type (independent power producer, municipal utility, and investor-owned utility) and whether the firm acts as its own bidder (OWNBIDDER). Finally, we include dummy variables for whether the firm’s generation technology is at least 50% comprised of two

⁴¹An annual expenditure of \$1 million corresponds to \$114 per hour for one year of operations (365 days, 24 hours), without correcting for the fact that the hour we are analyzing is a peak hour.

⁴²The reasoning behind this restriction is to guarantee that a bidder will be able to back up its “bid” with physical delivery of power.

⁴³See, for example, Demsetz and Lehn [1985].

technologies that are less flexible to quick changes in output (COAL and COMBINED-CYCLE).

Results are reported in Table 7. The baseline regression in column 1 yields a result that is consistent with the “scale hypothesis:” a 1000 MWh increase in sales is associated with a 52 percentage point increase in “Percent Achieved”.

Column 2 suggests that a “corporate governance” based explanation is not borne in the data. Controlling for size and technology, the performance of municipal utilities in the balancing market appears, in fact, to be slightly (2.5 to 4.8%) better than that of investor-owned utilities, though the regression coefficient is not statistically significant at conventional levels. Moreover, merchant firms actually seem to be doing the worst.

We also find that the technology mix of a firm does not appear to affect its performance on the balancing market. In column 3 we add measures of technology type and, find that owning a large fraction of coal and combined-cycle generation units in fact does not negatively impact performance. This result reinforces our argument in Section 5.1.

We can best measure scale effects if we focus on those firms that choose to establish their own bidding operation rather than those that outsource. In column 4, we control for whether the firm performs its own bidding and allow the effect of SIZE to vary by OWNBIDDER status. Larger stakes are associated with higher performance for firms that serve as their own bidders. For firms that perform their own bidding, a 1000 MWh increase in optimal sales volume is associated with an 86 percentage point increase in Percent Achieved.

Moreover, if one were to view the choice to serve as their own bidder as a revealed preference, then the threshold volume of sales where it becomes profitable to construct an in-house bidding operation rather than to outsource is 71 MWh ($=-(-.071/.001)$). The results are similar when we control for firm type and technology (column 5). A 1000 MWh increase in optimal sales volume is associated with an 97 percentage point increase in Percent Achieved and the threshold size for ownbidding is 163 MWh.

Our interviews with traders support this positive correlation between stake-size and bidder performance conditional on participation. Larger firms hired many experienced energy traders and purchased software to retrieve and visualize the vast array of operational, bid, and weather data that is available in real-time. At least one firm developed in-house software to retrieve and analyze data from ERCOT and formulate bids in real-time. Many of the large firms also purchase the Genscape data on real-time production. In contrast, firms with smaller stakes typically do not have the same hardware and software assets, and rather than hiring experienced traders, they bring in personnel from their generation and transmission divisions. Our interviews also revealed that there is a fluid market for electricity traders – successful traders at small firms are hired away by larger firms. Finally, Bryan’s large “Percent Achieved” despite relatively small size deserves special note. Bryan claims to have invested heavily in a trading operation in order to become a bidder to which other small firms will outsource their bidding responsibilities.

5.6 Learning

One might also expect bidders to gradually learn the rules of competition in this market, and gradually improve their bidding behavior over time. Indeed, the downward trend in bid-ask spreads that we document in Section 5.1 can be interpreted as a sign of learning to participate in this market. To explore this hypothesis further, we now examine whether there are any time trends in bidder performance.

Table 9 reports the results of a regression of individual bidder’s daily performance (measured by the percentage of ex-post profit achieved for that day) on a time trend and controls for seasonality and the bidder’s ex-post optimal generation in that period. The specification includes bidder fixed-effects to account for firm-level sources of variation. Notice that, for the entire sample of firms, the estimated coefficient for the time trend is positive, and indicates a 3 percentage point improvement in performance for every 100 days (or roughly 10% over a year). It is interesting to note that this time trend is not significant (though of the same estimated magnitude) for the top six bidders. For the rest of the firms, the time trend in performance is strongly significant.

Thus, our data appears to support a learning hypothesis; though the rate of learning in this market strikes us as being rather slow (especially compared to rates of learning reported in laboratory experiments).⁴⁴ However, we should note that firms in this market face a considerable amount of uncertainty, which may slow bidders’ ability to infer optimal decision rules from their experiences. Moreover, several bidders have told us that they do not have the resources to perform detailed “ex-post” analyses that will enable them to assess how successful their bidding has been, and that they view our exercise as providing useful information in this regard.

6 Quantifying Efficiency Losses

Inefficiencies arise in electricity markets when the bids do not lead to least-cost production.⁴⁵ In a balancing market, firms bidding above marginal cost on the INC side may not be called upon to produce despite the fact they have low cost generators. Similarly, firms bidding below marginal cost on the DEC side may not be called to reduce production even if they have high cost plants operating.

We can measure the cost of production in the balancing market implied by actual bids and compare those costs to the production costs under various counterfactual bidding behaviors. Our benchmark for efficiency is competitive bidding, which is an equilibrium under an alternative market design – the multi-unit Vickrey auction. The multi-unit Vickrey auction can result in higher prices but pays firms information rents to reveal their true cost. We calculate the production costs if

⁴⁴It is difficult to make a direct comparison between “rounds” in the laboratory and “days” in electricity markets. However, if we are willing to equate “rounds” with “days,” bidding behavior appears to converge (in percentage profit terms) to theoretical predictions quicker in many laboratory experiments. See Kagel (1995), Section F for examples from auction experiments.

⁴⁵Because total demand is perfectly inelastic, prices higher or lower than competitive levels do not cause suboptimal levels of consumption. All inefficiencies are productive rather than allocative. For a more general discussion of the efficiency properties of uniform price multi-unit auctions, see Ausubel and Cramton (2002).

firms instead were to bid their marginal cost functions. These calculations suffer from one data limitation. The generation data are incomplete for a few of the smaller firms. This does not affect our ability to analyze bidding behavior for the remaining firms but does prevent us from calculating efficient dispatch for many days in our sample. Therefore, these measures of efficiency loss should be interpreted with caution for they only represent a fraction of our sample period.

Results are shown in Table 10. The average hourly dispatch costs in the balancing market are \$29,874 under actual bidding and \$23,571 under marginal cost bidding, implying that actual production costs are 27% higher than least cost production.

We can decompose the welfare losses into the two major sources of inefficiency. The first source of inefficiency is the strategic exercise of market power – large firms face steeper residual demand functions and thus have incentives to bid steeper than marginal cost. This may result in inefficient production when low cost units are withheld from the market while higher cost units are called to generate.⁴⁶

The second inefficiency is somewhat less standard. As we have noted, many of the smaller participants submit extremely steep bid functions – even though static profit maximization suggests that they should bid very close to marginal cost. As a result, they are often not called to produce despite the fact that it would be efficient to do so. This source of inefficient production can arise from a variety of sources as we discuss above (e.g. fixed cost of establishing a sophisticated trading operation).

To decompose the total amount of production inefficiency into these two components, we separate the firms into two groups: “strategic” bidders who exercise market power optimally, and “non-strategic” bidders who bid exceedingly steep schedules that effectively minimizes their participation in the market. We first compute counterfactual generation costs in this market under the assumption that “strategic” firms ignore their market power and bid competitively, i.e. we calculate what the total generation cost would have been if the “strategic” firms were bidding their marginal costs. In this counterfactual, we assume that the “non-strategic” firms continue to bid what they are observed to bid, i.e. that they do not respond to the (counterfactual) change in the behavior of their strategic competitors.⁴⁷

We assume that the three largest firms – Reliant, TXU, and Calpine – as well as the other 3 firms in the top 6 in Table 3 (Brownsville, Bryan and Tenaska) are “strategic.” We find counterfactual dispatch costs when these six firms bid marginal cost and all other firms submit their actual bid schedules. The difference in dispatch costs between this counterfactual bidding strategy and the actual bids can be interpreted as the “efficiency loss due to market power.” The remaining efficiency

⁴⁶Borenstein, Bushnell and Wolak (2002) find the actual production costs to be 14% higher than competitive levels in the California market in 2000 due to market power. Mansur (2004b) refines the methodology from Borenstein, Bushnell and Wolak to correct for production constraints such as start-up costs, and finds that actual production costs in the Pennsylvania, New Jersey, Maryland market exceed competitive costs by 3%. Note that our analysis only calculates productive inefficiencies in the *balancing* market for units that have already started and submitted day-ahead schedules.

⁴⁷Note that the “reverse” of this counterfactual, where we set “non-strategic” bids equal to marginal cost, and do not allow “strategic” bidders to respond, would be less realistic, since our results show that the “strategic” bidders do respond to changes in the residual demand they are facing.

loss is due to non-strategic firms that bid so as to not participate in the market.

Table 10 decomposes the efficiency losses. Again, the total efficiency loss due to misrepresentation of marginal costs is on average \$6,303 per hour, or 27% of the total cost of efficient generation in the balancing market. If the “strategic” firms were to bid their marginal costs, the total efficiency loss would have been only \$1203 less than the actual efficiency loss. This means that most (81%) of the observed efficiency loss is due to the steep bid schedules submitted by the “non-strategic” bidders.

We should note, once again, that the above calculation relies on relatively few intervals (62 out of a total of 220 periods – the latter periods are especially prone to the missing data problem). This calculation is largely based upon the first six months of the market’s operation. However, this calculation points out that the observed deviations from static profit maximization are not without economic consequence. In fact, in terms efficiency, they cause larger efficiency losses due to the “near-optimal” exercise of market power by the “strategic” firms. Submitting bid functions that are too steep not only sacrifices producer surplus, but also prevents technologically efficient firms from supplying energy to the balancing market.

One would expect that these efficiency losses due to the “non-strategic” firms’ behavior would dissipate over time. Unfortunately, as we noted above, we face data limitations that prevent us from conducting a long time series analysis, especially in the later part of the sample. However, market forces are at play to reduce the inefficiencies. To the extent there are scale effects, the small firms may outsource bidding decisions or consolidate bidding activities across plants. It is noteworthy that there are a variety of power marketers and large energy trading firms that offer energy asset management services to generators in ERCOT. It will be interesting to see if more firms outsource bidding in the future, and whether this increases the efficiency of production in this marketplace.

7 Conclusion

Our analysis of the ERCOT balancing market yields a comforting conclusion for economic theory. In a marketplace that is marked by considerable uncertainty and institutional complexity, two factors that may pose analytical challenges for both the firms competing in the market and the economists who are observing them, firms with large stakes in the market behave close to theoretical predictions of a strategic model of oligopolistic interaction. Indeed, as two empirical industrial organization researchers who have previously utilized such models to infer supply and demand parameters in other markets, we were swayed to interpret the behavioral pattern displayed by Reliant as good news for empirical analyses of oligopolistic markets. More specifically, confirming the basic predictions of the uniform price share auction model is important, as one could use this model to forecast bidder behavior in restructured electricity markets that are being put into operation in many different parts of the world.

However, our results also suggest some amount of caution when analyzing and predicting the

behavior of smaller players in newly restructured/deregulated markets. Smaller firms submit bids that differ substantially from the benchmarks we construct for optimal bidding. This finding is not inconsistent with rational economic behavior by these bidders, however. As argued in Section 5.5, “participation” in the balancing market may have nontrivial costs, and the behavioral pattern *across* firms appears to confirm this hypothesis.

The fact that participation costs may affect small firms’ behavior is not without consequence in this market. In Section 6, we calculated the small firms’ bidding patterns led to nontrivial losses in productive efficiency. This suggests interesting new avenues for market design that explicitly takes into account the strategic complexity, hence the participation costs imposed by proposed market mechanisms. Such a consideration would favor dominant strategy implementable mechanisms, such as the Vickrey-Clarke-Groves (VCG) mechanism, over others. However, as pointed out by Milgrom [2004] (chapter 2.5), the VCG mechanism suffers from serious pitfalls of its own. Nevertheless, we view theoretical research in this area to prove extremely fruitful for real-world applications.

APPENDIX A: CONSTRUCTION OF BID AND COST DATA

In this appendix, we describe the data on bids and total balancing demand, as well as how we construct the hourly marginal cost function of each firm. Bids into the balancing market are available for each hour of the sample. The bids are not unit-specific but rather are the aggregate bids for all units for which the bidder, or qualified scheduling entity (QSE), bids. The actual deployment is based on other factors as well. ERCOT takes the QSE portfolio deployments and allocates the incremental and decremental generation among the generating units based upon a specific algorithm which considers factors such as the rate at which units can ramp up generation and the location of units within a zone. In the event that units are allocated to solve intrazonal congestion, the generators also receive unit-specific premia payments. However, we restrict our sample to periods when these factors are unlikely to substantially alter the deployment that would be implied by the bids and least cost dispatch. (We analyze hours without interzonal congestion for a non-shoulder period). The bids implied by the bid stack alone predict the actual prices with only 5% error, so we believe the bid stack is a good model of actual deployment, or at least the bidder intent.

QSEs occasionally bid for more than one firm. For example, in the South zone in 2001, the QSE named Reliant bid for both Reliant and the City of San Antonio Public Service Company. We match the bid functions to all units for which the QSE bids. So for all units owned by both Reliant and the City of San Antonio in the South in 2001, we match the bid function to the generation data. This allows us to plot the bid function on top of the marginal cost of supplying into the balancing market. However, interpretation of the results becomes problematic when an observed bid function represents the bids by more than one firm. Because the results are some combination of two firms' behavior, we will not interpret results in such situations. We only interpret our results as firm-level behavior when at least 90% of all electricity generated by owners using that QSE can be attributed to a single owner. We make one exception to this 90% rule. TXU Generation comprises 87% of the generation for TXU the QSE in North 2002. Because TXU is such a major player, we nevertheless refer to this as TXU the generation company.

We measure the variable costs of output using data on each unit's fuel costs and the rate at which the unit converts the fuel to electricity. Only certain units are sufficiently flexible to supply power on short notice. We assume that nuclear, wind, and hydroelectric units cannot supply to the balancing market. In addition, we assume that only units that are currently operating can supply into the balancing market. Data on the availability of generating units are from ERCOT. For each 15-minute interval, we have data on whether a generating unit is operating, and its hourly available generating capacity. The ERCOT system is largely separated from other electricity grids in the country so there are virtually no imports into ERCOT.

We measure the marginal cost of units that burn natural gas and coal. For each unit, we have data on the fuel efficiency (i.e. average heat rate). Each unit is assumed to have constant marginal cost up to its hourly operating capacity. This assumption, which is common in the literature, assumes that average heat rates are good approximations of incremental heat rates

for the relatively small changes in output implied by balancing sales. This assumption is most problematic for combined-cycle gas turbines at lower levels of utilization.

Daily gas spot prices measure the opportunity cost of fuel for natural gas units. We use prices at the Agua Dulce, Katy, Waha, and Carthage hubs for units in the South, Houston, West, and North zones, respectively. We assume a gas distribution charge of \$0.10/mmBtu. Coal prices are monthly weighted average spot price of purchases of bituminous, sub-bituminous, and lignite in Texas, reported in Form FERC-423 by the Energy Information Administration.

Coal-fired plants in Texas are required to possess federal emission permits under the Clean Air Act Amendments for each ton of SO₂ emissions. Therefore, we require a measure of the emission permit cost for each MWh of electricity production. In order to measure average emission rates, we merge hourly net metered generation data from ERCOT with hourly emission data from EPA's Continuous Emission Monitoring System to calculate each unit's average pounds of SO₂ emissions per net MWh of electricity output. The emissions each hour are priced at the monthly average EPA permit price reported on the EPA website (this price is the average of Cantor-Fitzgerald and Fieldston prices).

In order to deal with complications posed by transmission congestion, we restrict our sample to daily intervals 6:00-6:15pm during which there is no interzonal transmission congestion during the 6-7pm bidding hour. We believe *intrazonal* (or local) congestion is also likely to be rare during the intervals.

We do not account for the sales of operating reserves. In electricity markets, the grid operator procures the right to purchase power from generating units that are to be operating with excess capacity available to provide power on short notice. We do not incorporate the possibility that some of the available capacity to INC in our data may be sold as reserves. However, the amount of operating reserves procured are small as a fraction of total demand.

We measure the marginal cost of INCing or DECing from the day-ahead schedule of output. We account for the fact that units cannot DEC down to zero output without incurring costs of startup and facing constraints on minimum downtime. It is unlikely that revenue from the balancing market would be sufficiently lucrative to compensate a unit for shutting down. Rather, firms are likely to supply balancing energy services along the operational levels of currently operating units. Therefore, we assume that each operating unit cannot DEC to a level below 20% of its maximum generating capacity. Therefore, each firm's marginal cost function is essentially the portfolio's marginal cost for the upper 80% of its units' maximum generating capacity.

APPENDIX B: Derivation of the first-order condition of optimality

Suppressing the i and t indices, the maximization problem becomes:

$$\max_{S(\cdot)} \int_{\underline{p}}^{\bar{p}} U(pS(p) - C(S(p)) - (p - PC)QC) dH(p, S(p); QC)$$

Integration by parts of the objective function yields, modulo a constant term:

$$- \int_{\underline{p}}^{\bar{p}} U'(pS(p) - C(S(p)) - (p - PC)QC) (pS'(p) + S(p) - C'(S(p))S'(p) - QC) H(p, S(p); QC) dp$$

Now label the integrand:

$$F(p, S, S') = -U'(pS(p) - C(S(p)) - (p - PC)QC) (pS'(p) + S(p) - C'(S(p))S'(p) - QC) H(p, S(p); QC)$$

The Euler-Lagrange necessary condition for the optimal $S(p)$ is given by:

$$\frac{d}{dp} F_{S'} = F_S$$

Evaluating the derivatives:

$$\begin{aligned} -F_S &= H_S U'(\cdot) [pS' + S - C'S' - QC] + H U''(\cdot) (p - C') [pS' + S - C'S' - QC] + H U'(\cdot) [1 - C''S'] \\ -F_{S'} &= H U'(\cdot) [p - C'] \end{aligned}$$

and taking the total derivative of $F_{S'}$ with respect to p :

$$-\frac{d}{dp} F_{S'} = H_p U'(\cdot) [p - C'] + H_S S' U'(\cdot) [p - C'] + H U''(\cdot) [pS' + S - C'S' - QC] (p - C') + H U'(1 - C''S')$$

Equating and cancelling terms, we get:

$$H_S U'(\cdot) (S - QC) = H_p U'(\cdot) (p - C')$$

or

$$p - C' = \frac{H_S}{H_p} (S - QC)$$

as claimed.

APPENDIX C: PROOF OF PROPOSITION 3

Given the additively separable form of the bidding strategies $S_i(p, QC_i) = \alpha_i(p) + \beta_i(QC_i)$, use the market clearing condition (1) above to represent the event $\{p_t^c \leq p | QC_i, S_i\}$, i.e. there is excess supply at p , conditional on firm i bidding S_i at this price:

$$\sum_{j \neq i} \beta_j(QC_j) - \varepsilon \geq D(p) - S_i - \sum_{j \neq i} \alpha_j(p)$$

The left hand side of this inequality can be labeled as a (bidder-specific) random variable, θ_i , that does not depend on p , and the right hand side is a deterministic function of price. Let $\Gamma_i(\cdot)$ denote the cdf of θ_i and $\gamma_i(\cdot)$ denote the pdf (both conditional on the bidder's contract quantity, QC_i). Given these:

$$\begin{aligned} H_p(p, S_i; QC_i) &= \frac{\partial}{\partial p} \Pr(p_t^c \leq p | QC_i, S_i) \\ &= \frac{\partial}{\partial p} \Pr(\theta_i \geq D(p) - S_i - \sum_{j \neq i} \alpha_j(p)) \\ &= \frac{\partial}{\partial p} [1 - \Gamma_i(D(p) - S_i - \sum_{j \neq i} \alpha_j(p))] \\ &= -\gamma_i(D(p) - S_i - \sum_{j \neq i} \alpha_j(p)) \frac{\partial}{\partial p} (D(p) - S_i - \sum_{j \neq i} \alpha_j(p)) \end{aligned}$$

and

$$\begin{aligned} H_S(p, S_i; QC_i) &= \frac{\partial}{\partial S_i} \Pr(p_t^c \leq p | QC_i, S_i) \\ &= -\gamma_i(D(p) - S_i - \sum_{j \neq i} \alpha_j(p)) \frac{\partial}{\partial S_i} (D(p) - S_i - \sum_{j \neq i} \alpha_j(p)) \end{aligned}$$

Evaluating the derivatives gives $\frac{H_p(p, S_i; QC_i)}{H_S(p, S_i; QC_i)} = -[D'(p) - \sum_{j \neq i} \alpha'_j(p)]$. Now, observe that with the above restrictions, the residual demand function faced by firm i (for a given realization of the random variables $\{\varepsilon, QC_{-i}, i = 1, \dots, N\}$), is given by:

$$RD_i(p, \varepsilon, QC_{-i}) = D(p) + \varepsilon - \sum_{j \neq i} \alpha_j(p) - \sum_{j \neq i} \beta_j(QC_j) \quad (8)$$

with derivative:

$$RD'_i(p) = D'(p) - \sum_{j \neq i} \alpha'_j(p)$$

which yields the result in the proposition.

Note that in one other case where we can collapse the multiple stochastic terms into a scalar random variable, we do not obtain the ex-post optimality result. This is the case when $S_i(p, QC_i) = \alpha_i(p) + \beta_i p QC_i$ and $\tilde{D}(p) = D(p) + p\varepsilon$, i.e. private information and aggregate uncertainty leads to pure *rotations* of the residual demand curve. In this case market clearing condition becomes

$$\theta_i = \sum_{j \neq i} \beta_j QC_j - \varepsilon \geq \frac{1}{p} \left[D(p) - S_i - \sum_{j \neq i} \alpha_j(p) \right], \text{ but } \frac{H_p(p, S_i)}{H_S(p, S_i)} \neq \frac{1}{RD'_i(p, \varepsilon, QC_{-i})}.$$

APPENDIX D: An Analytically Solvable Example for Ex-Post Optimal Equilibrium

Assume a symmetric setup with $N > 2$ firms in the market (the $N > 2$ restriction will become apparent in the solution).

Let the total cost function be quadratic of the form: $TC(Q) = k + aQ + \frac{1}{2}bQ^2$.

This gives the marginal cost function: $MC(Q) = a + bQ$, or the “true” supply function: $S^{true}(p) = \frac{1}{b}p - \frac{a}{b}$

Let the contract position of firm i be QC_i , where QC_i are iid random variables with distribution $F(QC)$.

We are going to look for symmetric Bayesian-Nash equilibria of the game in linear supply functions, of the form:

$$S_i(p) = S(p, QC_i) = \alpha + \beta p + \gamma QC_i$$

Given such supply functions, the probability that the market clearing price is below a given price p , conditional firm i 's strategy, $S_i(p)$, is defined by the condition:

$$S_i(p) + \sum_{j \neq i} S_j(p) \geq Q^{tot}$$

ie. there is “excess supply at p .” Equivalently:

$$Q^{tot} - \gamma \sum_{j \neq i} QC_j \leq (N-1)\alpha + (N-1)\beta p + S_i(p)$$

Notice that the left-hand side of this equation collects together all of the random elements of this problem, from the perspective of bidder i .

That is,

$$H(p, S_i(p)) = \Pr\{Q^{tot} - \gamma \sum_{j \neq i} QC_j \leq (N-1)\alpha + (N-1)\beta p + S_i(p)\}$$

where $H(p, S_i(p))$ depends on the distribution of the random variable $Z_i = Q^{tot} - \gamma \sum_{j \neq i} QC_j$.

Recall, however, the first-order condition for optimality derived in the paper:

$$p - MC(S_i(p)) = (S_i(p) - QC_i) \frac{\frac{\partial H(p, S_i(p))}{\partial S_i}}{\frac{\partial H(p, S_i(p))}{\partial p}}$$

A look at the expression for $H(p, S_i(p))$ reveals:

$$\frac{\frac{\partial H(p, S_i(p))}{\partial S_i}}{\frac{\partial H(p, S_i(p))}{\partial p}} = \frac{1}{(N-1)\beta}$$

This “simplification” does not depend on the distribution of QC_i 's. It holds for any (non-atomic) distribution with full support over the range of prices.

How does this help? Plugging in for $MC(Q) = a + bQ$:

$$\begin{aligned} p - (a + bS_i(p)) &= \frac{1}{(N-1)\beta}(S_i(p) - QC_i) \\ \beta(N-1)p - \beta(N-1)a + QC_i &= S_i(p)[1 + \beta(N-1)b] \end{aligned}$$

Now, use the expression: $S_i(p) = \alpha + \beta p + \gamma QC_i$

$$\beta(N-1)p - \beta(N-1)a + QC_i = (\alpha + \beta p + \gamma QC_i)[1 + \beta(N-1)b]$$

For this equality to hold, we need to have the conditions (equating coefficients on like terms):

$$\begin{aligned} \frac{\beta(N-1)}{1 + \beta(N-1)b} &= \beta \\ \frac{-\beta(N-1)a}{1 + \beta(N-1)b} &= \alpha \\ \frac{1}{1 + \beta(N-1)b} &= \gamma \end{aligned}$$

The first equation has two solutions, $\beta = 0$ and $\beta = \frac{N-2}{(N-1)b}$. We focus on the second solution, since $\beta = 0$ yields the bidding strategy, $S_i(p) = QC_i$, which can actually be equilibrium bidding strategies, since the residual demand curve seen by every bidder is vertical.

Using the second solution, we get:

$$\begin{aligned} \beta &= \frac{N-2}{(N-1)b} \\ \alpha &= -\frac{a(N-2)}{b(N-1)} \\ \gamma &= \frac{1}{N-1} \end{aligned}$$

Hence, the equilibrium strategies are: $S_i(p) = S_i(p, QC_i) = \frac{N-2}{(N-1)b}p - \frac{a(N-2)}{b(N-1)} + \frac{1}{N-1}QC_i$
 Contrasting this with the “true” supply functions given by the marginal cost functions, $S_i^{true}(p) =$

$\frac{1}{b}p - \frac{a}{b}$, yields that equilibrium strategies are "above" the true supply function (intercept term $-\frac{a}{b} \frac{N-2}{N-1} > -\frac{a}{b}$), and steeper ($\frac{N-2}{(N-1)b} < \frac{1}{b}$, but we are looking at slope when quantity is on the horizontal axis).

Notes:

1. When $N = 2$, the only equilibrium candidate that survives is $S_i(p) = QC_i$.
2. If the market demand function, $Q^{tot}(p)$ is not perfectly inelastic, then the "trivial" equilibrium $S_i(p) = QC_i$ does not survive, since the residual demand function faced by each firm is no longer vertical.
3. What if firms were asymmetric, i.e. $MC_i(Q) = a_i + b_iQ$? In this case, we look for asymmetric equilibria of the form: $S_i(p, QC_i) = \alpha_i + \beta_i p + \gamma_i QC_i$, and get a system of $3N$ equations to solve for $\{\alpha_i, \beta_i, \gamma_i\}$, since:

$$\frac{\frac{\partial H_i(p, S_i(p))}{\partial S_i}}{\frac{\partial H_i(p, S_i(p))}{\partial p}} = \frac{1}{\sum_{j \neq i} \beta_j}$$

yielding, for each $i = 1, \dots, N$

$$\begin{aligned} \frac{\sum_{j \neq i} \beta_j}{1 + b_i \sum_{j \neq i} \beta_j} &= \beta_i \\ \frac{-a_i \sum_{j \neq i} \beta_j}{1 + b_i \sum_{j \neq i} \beta_j} &= \alpha_i \\ \frac{1}{1 + b_i \sum_{j \neq i} \beta_j} &= \gamma_i \end{aligned}$$

Note that the first set of N equations, $\frac{\sum_{j \neq i} \beta_j}{1 + b_i \sum_{j \neq i} \beta_j} = \beta_i$ can be solved numerically for β_i as a function of b_1, \dots, b_N (we have not been able to find an analytic solution). Given β_i , one can then solve for α_i and γ_i .

REFERENCES

- L.M. Ausubel and P. Cramton. Demand Reduction and Inefficiency in Multi-Unit Auctions, working paper, University of Maryland , 2002.
- R. Baldick, R. Grant, and E. Kahn. Theory and Application of Linear Supply Function Equilibrium in Electricity Markets, *Journal of Regulatory Economics*, 2004.
- P. Bajari, A. Hortaçsu. Are Structural Estimates of Auction Models Reasonable? Evidence from Experimental Data, NBER working paper no. 9889, 2003.
- S. Borenstein, J.B. Bushnell, and F.A. Wolak. Measuring Market Inefficiencies in California's Restructured Wholesale Electricity Market. *American Economic Review*, 92(5), 1376-1405, 2002.
- J.B. Bushnell and C. Saravia. An Empirical Assessment of the Competitiveness of the New England Electricity Market. Center for the Study of Energy Markets working paper CSEMWP-101, 2002.
- P. Cramton. Competitive Bidding Behavior in Uniform-Price Auction Markets, Report to Federal Energy Regulatory Commission, Dockets EL00-95-075 and EL00-98-063.
- G. Crawford, J. Crespo, and H. Tauchen, Bidding Asymmetries in Multi-Unit Auctions: Implications of Bid Function Equilibria in the British Spot Market for Electricity, working paper, University of Arizona, 2003.
- H. Demsetz and K. Lehn, The Structure of Corporate Ownership: Causes and Consequences. *Journal of Political Economy*, 93, 1155-1177.
- L. Einav, A Note on The Analogies between Empirical Models of Auctions and of Differentiated Product Markets, working paper, Stanford University, 2004.
- D. Genesove, W. Mullin. Testing Static Oligopoly Models: Conduct and Cost in the Sugar Industry, 1890-1914. *RAND Journal of Economics*, 29(2), 1998.
- R. Green. Increasing Competition in the British Electricity Spot Market. *Journal of Industrial Economics*, 1992.
- R. Green and D. Newbery. Competition in the British Electricity Spot Market. *Journal of Political Economy*, 100(5), 929-953, 1990.
- L.P. Hansen. Large Sample Properties of Generalized Method of Moments Estimators. *Econometrica*, 50, 1029-54, 1982.
- A. Hortaçsu. Mechanism Choice and Strategic Bidding in Divisible Good Auctions: An Empirical Analysis of the Turkish Treasury Auction Market, working paper, University of Chicago, 2002a.
- A. Hortaçsu. Bidding Behavior in Divisible Good Auctions: Theory and Evidence from the Turkish Treasury Auction Market, working paper, University of Chicago, 2002b.
- P. Joskow. Deregulation and Regulatory Reform in the U.S. Electric Power Sector. In S. Peltzman and C. Winston, editors, *Deregulation of Network Industries: The Next Steps*. Brookings Press, 2000.

- P. Joskow and E. Kahn. A Quantitative Analysis of Pricing Behavior In California's Wholesale Electricity Market During Summer 2000. *The Energy Journal*, 23(4), 1-35, 2002.
- J. Kagel. Auctions: A Survey of Experimental Research. *Handbook of Experimental Economics* J. Kagel and A. Roth (eds.) Princeton University Press, 1995.
- P. Klemperer. Why Every Economist Should Learn Some Auction Theory. *Advances in Economics and Econometrics: Invited Lectures to 8th World Congress of the Econometric Society*. M. Dewatripont, L. Hansen and S. Turnovsky (eds.) Cambridge University Press, 2003.
- P. Klemperer and Meyer, M. Supply Function Equilibria in Oligopoly under Uncertainty, *Econometrica*, 57(6), 1243-1277, 1989
- K. Kuhn and M. Machado, Bilateral Market Power and Vertical Integration in the Spanish Electricity Spot Market, CEPR Working Paper No. 4950, 2004.
- E.T. Mansur. Upstream Competition and Vertical Integration in Electricity Markets, working paper, Yale University, 2004a.
- E.T. Mansur. Measuring Welfare in Restructured Electricity Markets, working paper, Yale University, 2004b.
- P. Milgrom. *Putting Auction Theory To Work: The Churchill Lectures in Economics*. Cambridge University Press, 2004.
- S.L. Puller. Pricing and Firm Conduct in California's Deregulated Electricity Market, working paper, Texas A&M University, 2004.
- A. Rudkevich. Supply Function Equilibrium in Power Markets: Learning All the Way. TCA Technical Report Number 1299-1702, Tabors Caramanis and Associates, 1999.
- A. Sweeting. Market Power in the England and Wales Wholesale Electricity Market: 1995-2000, working paper, Northwestern University, 2004.
- S. Viswanathan, J. Wang, and T. Witeliski, Optimal Bidding In Multi-Unit Discriminatory Auctions, working paper, Duke University, 2002.
- N.H. von der Fehr and D. Harbord. Spot Market Competition in the UK Electricity Industry. *Economic Journal*, 103, 531-546, 1993.
- J.J.D. Wang and J.F. Zender. Auctioning Divisible Goods. *Economic Theory*, 19(4):673-705, 2002.
- R. Wilson. Auctions of Shares. *Quarterly Journal of Economics*, 93, 675-689, 1979.
- R. Wilson. Architecture of Power Markets. *Econometrica*, 70(4), 1299-1340, 2002.
- F.A. Wolak. "Identification and Estimation of Cost Functions Using Observed Bid Data: An Application to Electricity Markets," in M. Dewatripont, L.P. Hansen, and S.J. Turnovsky, eds., *Advances in Economic and Econometrics: Theory and Applications, Eight World Congress, Volume II*. New York: Cambridge University Press, 2003a, pp. 133-169.
- F.A. Wolak. Measuring Unilateral Market Power in Wholesale Electricity Markets: The California Market 1998 to 2000. *American Economic Review*, 93(2), 425-430, 2003b.
- C.D. Wolfram. Measuring Duopoly Power in the British Electricity Spot Market. *American Economic Review*, 1999.

Table 1: Generation Ownership

| Owner | % of Installed Capacity |
|------------------------------------|------------------------------------|
| TXU | 24 |
| Reliant | 18 |
| City of San Antonio Public Service | 8 |
| Central Power & Light | 7 |
| City of Austin | 6 |
| Calpine | 5 |
| Lower Colorado River Authority | 4 |
| Lamar Power Partners | 4 |
| Guadalupe Power Partners | 2 |
| West Texas Utilities | 2 |
| Midlothian Energy | 2 |
| Dow Chemical | 1 |
| Brazos Electric Power Cooperative | 1 |
| Others | 16 |

Table 2: Volume of Sales into Balancing Market

| Firm | Average Sales (MWh) |
|------------------------------------|----------------------------|
| Reliant | 431 |
| TXU | 133 |
| Extex Laporte LP | 116 |
| Calpine Corp | 102 |
| Tenaska Gateway Partners | 72 |
| City of San Antonio Public Service | 49 |
| City of Austin | 48 |
| Central Power & Light | 35 |
| Cogen Lyondell Inc | 34 |
| City of Bryan | 30 |
| Lamar Power Partners | 30 |
| Lower Colorado River Authority | 28 |
| Ingleside Cogeneration | 25 |
| BP Energy | 17 |
| Brownsville Public Utility Board | 17 |
| Tractebel Power | 15 |
| Midlothian Energy | 15 |
| Gregory Power Partners | 14 |
| Guadalupe Power Partners | 12 |
| West Texas Utilities | 11 |

Note: Sales implied by bidstack and total balancing demand.

Table 3: Outcomes Under Actual, Ex-Post Optimal and Naive Best Response Bidding

| Firm | Percent Achieved | | Producer Surplus | | | Quantity Sales | | Upper |
|----------------------------|------------------|----------|------------------|----------|------------|----------------|--------|-------|
| | Relative To | | (\$/hour) | | | (MWh) | | Bound |
| | XP Optimal | Naive BR | Actual | Naive BR | XP Optimal | XP Optimal | Actual | Total |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | Prof |
| | | | | | | | | (8) |
| Reliant | 79% | 80% | 3,422 | 4,268 | 4,333 | 507 | 431 | 98% |
| Brownsville PUB | 50% | 50% | 173 | 343 | 343 | 42 | 17 | 88% |
| City of Bryan | 45% | 45% | 221 | 488 | 488 | 56 | 30 | 85% |
| Tenaska Gateway Partners | 41% | 41% | 456 | 1,111 | 1,115 | 182 | 72 | 88% |
| TXU | 39% | 41% | 1,243 | 3,056 | 3,159 | 441 | 133 | 97% |
| Calpine Corp | 37% | 38% | 820 | 2,168 | 2,214 | 408 | 102 | 91% |
| Denton Municipal Electric | 35% | 35% | 11 | 31 | 31 | 3 | 1 | 98% |
| Ingleside Cogeneration | 31% | 31% | 171 | 541 | 541 | 81 | 25 | 79% |
| City of Austin | 30% | 31% | 581 | 1,889 | 1,907 | 271 | 48 | 84% |
| Rio Nogales LP | 28% | 28% | 109 | 393 | 393 | 60 | 10 | 93% |
| Lower Colorado River Auth | 25% | 25% | 367 | 1,471 | 1,488 | 274 | 28 | 88% |
| City of San Antonio | 23% | 24% | 290 | 1,221 | 1,241 | 266 | 49 | 90% |
| Gregory Power Partners | 20% | 20% | 143 | 720 | 722 | 96 | 14 | 82% |
| Midlothian Energy | 17% | 17% | 171 | 1,016 | 1,024 | 175 | 15 | 86% |
| Cogen Lyondell Inc | 16% | 16% | 408 | 2,523 | 2,523 | 269 | 34 | 67% |
| Tractebel Power Inc | 16% | 16% | 127 | 795 | 795 | 97 | 15 | 79% |
| Brazos Electric Power Coop | 15% | 15% | 101 | 676 | 677 | 82 | 6 | 79% |
| Lamar Power Partners | 15% | 15% | 266 | 1,800 | 1,808 | 272 | 30 | 79% |
| Mirant Wichita Falls | 14% | 14% | 16 | 114 | 114 | 18 | 2 | 83% |
| BP Energy | 14% | 14% | 134 | 993 | 994 | 135 | 17 | 80% |
| City of Garland | 13% | 13% | 128 | 1,018 | 1,019 | 115 | 5 | 80% |
| Hays Energy | 8% | 8% | 64 | 775 | 777 | 111 | 8 | 82% |
| West Texas Utilities | 8% | 8% | 132 | 1,635 | 1,635 | 224 | 11 | 82% |
| Central Power & Light | 8% | 8% | 185 | 2,375 | 2,407 | 352 | 35 | 80% |
| Guadalupe Power Partners | 6% | 6% | 140 | 2,356 | 2,380 | 396 | 12 | 77% |
| Tenaska Frontier Partners | 5% | 5% | 52 | 1,044 | 1,051 | 144 | 7 | 80% |
| South Texas Electric Coop | 3% | 3% | 8 | 298 | 298 | 44 | 1 | 81% |
| Sweeney Cogeneration | 2% | 2% | 10 | 409 | 409 | 46 | 2 | 85% |
| Brazos Valley Energy LP | 0% | 0% | 1 | 134 | 134 | 12 | 0 | 68% |
| AES Deepwater | 0% | 0% | 1 | 969 | 969 | 92 | 0 | 60% |
| Frontera General LP | 0% | 0% | 0 | 984 | 1,004 | 197 | 0 | 62% |
| TGC | 0% | 0% | 0 | 405 | 405 | 70 | 0 | 81% |
| South Houston Green Power | 0% | 0% | 0 | 60 | 60 | 7 | 0 | 70% |
| Air Liquide America | -8% | -8% | -181 | 2,174 | 2,176 | 103 | 9 | 55% |
| Extex Laporte LP | -81% | -81% | -1,230 | 1,497 | 1,497 | 92 | 116 | 41% |

Producer surplus measures are all relative to profitability under not bidding into the balancing market and are measured in average dollars for each 6-7pm.

Percent Achieved is the ratio of producer surplus under actual bidding to producer surplus under the benchmark for optimal bidding (either ex-post optimal or Naive Best Response).

Actual is the actual bids submitted by the firms.

XP Optimal is ex-post optimal bidding constructed as we describe in section 3.

Naive BR is the naive best-response bid to rivals' lagged bids as we describe in section 4.3.2.

Quantity of sales are total volumes and, therefore, may sum to more than the average net demand for balancing energy.

In column 8, the Upper Bound for Total Profitability = (Percent Achieved XP Optimal*%Sales in Balancing)+(100% * %Sales in Bilateral) as defined in section 4.2.

Table 4: J-Tests of Expected Profit Maximization

| Firm | TJ_T | P-value | T | Reject $E[\pi]$ max? |
|------------------------|--------|---------|-----|----------------------|
| Reliant (pre-Nov 2002) | 13.1 | 0.71 | 116 | No |
| Reliant | 33.4 | 0.99 | 154 | Yes |
| Brownsville | 24.27 | 0.99 | 53 | Yes |
| Bryan | 91.3 | 1 | 220 | Yes |
| Tenaska | 75.61 | 1 | 82 | Yes |
| TXU | 146.1 | 1 | 220 | Yes |
| Calpine | 115.3 | 1 | 171 | Yes |

Notes: Section 4.3.2, Equation (6) was used form the moments. Specifically, the quantity derivative of the firm's profit function was taken at 11 equally spaced prices between \$16 and \$36. Hansen's J-statistic, J_T , is given by Equation (7). The p-value is based on a χ^2 distribution with 11 degrees of freedom. The 5% critical value is 19.7.

Table 5: Characteristics of Bid Functions By Percent Achieved

| Firm | Percent Achieved | Mean Bid Points | Mean Bid-Ask Spread (\$/MWh) |
|--------------------------------|------------------|-----------------|------------------------------|
| Reliant | 79% | 22.2 | \$2.06 |
| City of Bryan | 45% | 6.5 | \$22.58 |
| TXU | 39% | 12.6 | \$20.60 |
| Calpine Corp | 37% | 7.5 | \$12.55 |
| City of Austin | 30% | 10.0 | \$25.92 |
| Lower Colorado River Authority | 25% | 9.1 | \$25.98 |
| City of Garland | 13% | 6.1 | \$20.32 |
| South Texas Electric Coop | 3% | 3.7 | \$68.66 |

Notes: Percent Achieved is percent of possible producer surplus achieved under actual bidding relative to not bidding. Bid-Ask Spread is price difference between first INC and first DEC bid at balancing quantity=0.

Table 6: Relationship Between Profitability and Transmission Congestion

| Dependent Variable: Monthly <i>PercentAchieved</i> by Firm <i>i</i> | | |
|---|-----------------|-----------------|
| | All Firms | Own Bidders |
| Pct Intervals Congested | -0.16 (0.09) | -0.17 (0.14) |
| Volume Optimal Output (GWh) | 0.34 (0.21) | 0.09 (0.25) |
| Monthly Fraction DEC | -0.05 (0.06) | 0.03 (0.11) |
| R-squared | 0.48 | 0.47 |

Model includes firm fixed effects.

Robust standard errors in parentheses.

Note: *PercentAchieved* is the monthly fraction of potential profits achieved relative to not bidding, defined in section 4.2.

Table 7: Relationship Between Profitability and Firm Characteristics

| | (1) | (2) | (3) | (4) | (5) |
|----------------|-----------------------|-----------------------|----------------------|------------------------|-------------------------|
| Size (MWh) | 0.00052 (0.00031)* | 0.00047 (0.00026)* | 0.00035 (0.00028) | -0.00015 (0.00025) | -0.00035 (0.00024) |
| Size*OwnBidder | | | | 0.00101 (0.00044)** | 0.00132 (0.00042)*** |
| OwnBidder | | | | -0.071 (0.111) | -0.215 (0.109)* |
| Merchant Firm | | -0.098 (0.135) | -0.167 (0.146) | | -0.170 (0.098)* |
| Municipal | | 0.048 (0.135) | 0.025 (0.142) | | 0.047 (0.089) |
| Coal | | | -0.021 (0.113) | | 0.043 (0.120) |
| Combined-Cycle | | | 0.082 (0.058) | | 0.127 (0.071)* |
| Constant | 0.100 (0.054)* | 0.158 (0.142) | 0.200 (0.152) | 0.158 (0.058)** | 0.262 (0.098)** |
| Observations | 34 | 34 | 34 | 34 | 34 |
| R-squared | 0.15 | 0.28 | 0.31 | 0.34 | 0.50 |

Notes: Dependent variable is the Percent Achieved for firm i over the entire sample as defined in section 4. Parameters are estimated by OLS. White standard errors are in parentheses. The excluded category for firm type is Investor-Owned utility.

* significant at 10%

** significant at 5%

*** significant at 1%

Table 8: Evolution of Bid-Ask Spread Over Time

| | Top 6 Firms | Non Top 6 Firms |
|-------------------------|--------------------|-------------------|
| Days Since Market Began | -0.039 (0.015)* | -0.000 (0.010) |
| Off-Peak Season | 12.43 (6.98) | 6.00 (3.15) |
| Constant | 19.45 (4.72)** | 23.38 (4.78)** |
| Observations | 840 | 2596 |
| R-squared | 0.04 | 0.12 |

Notes: Model includes firm fixed effects. Dependent variable is the daily bid-ask spread, or difference in price between first INC and first DEC at $q=0$. Top 6 firms are the six firms with the largest *PercentAchieved* as reported in Table 3 – Reliant, Brownsville, City of Bryan, Tenaska Gateway Partners, TXU and Calpine. Off-peak season is October-April. Parameters are estimated by OLS. White standard errors are in parentheses.

* significant at 5%

** significant at 1%

Table 9: Evolution of Performance and Learning

| | All Firms | Top 6 Firms | Non Top 6 Firms |
|-----------------------------|------------------------|------------------------|------------------------|
| Days Since Market Began | 0.00031 (0.00010)** | 0.00033 (0.00029) | 0.00028 (0.00009)** |
| Volume Optimal Output (GWh) | 0.16 (0.06)** | 0.25 (0.10)* | 0.09 (0.06) |
| Off-Peak Season | 0.01361 (0.01892) | -0.04437 (0.03799) | 0.03164 (0.01885) |
| Constant | 0.02330 (0.03030) | 0.21053 (0.07503)** | -0.01441 (0.02886) |
| Observations | 9765 | 2103 | 7662 |
| R-squared | 0.25 | 0.18 | 0.11 |

Notes: Model includes firm fixed effects. The dependent variable is the *PercentAchieved* for firm i on day t . Parameters are estimated by OLS. White standard errors in parentheses.

* significant at 5%

** significant at 1%

Table 10: Decomposition of Efficiency Losses from Observed Bidding Behavior

| Bidding Counterfactual | Average Production Cost |
|---|--------------------------------|
| Actual Bids for all firms | \$29,874 |
| “Strategic” firms Bid MC, Others Bid Actual | \$28,671 |
| All Firms Bid MC (Vickrey auction) | \$23,571 |
| | |
| Total Efficiency Losses | \$6,303 |
| “Strategic Bidders” | \$1,203 |
| “Non-Strategic Bidders” | \$5,100 |

“Strategic” firms are Reliant, TXU, Calpine, Brownsville, Bryan and Tenaska Gateway Partners. The average production cost is the hourly average of the absolute value of the cost of dispatch (either costs in INC hours or cost savings in DEC hours) during uncongested weekdays for the 6-7pm bidding hour.

Figure 1: Empirical Strategy

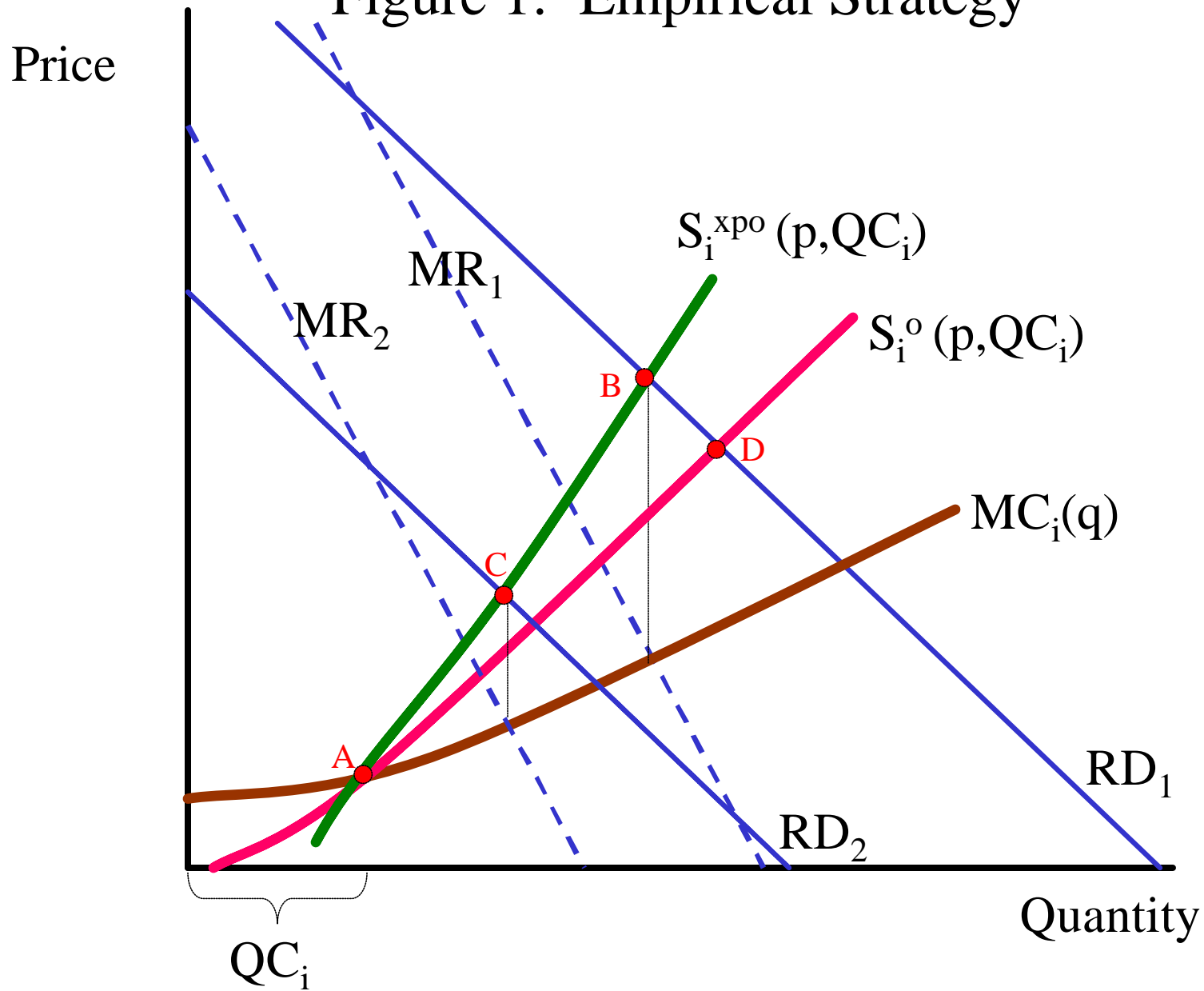


Figure 2

Example of Actual and Ex-Post Optimal Bidding

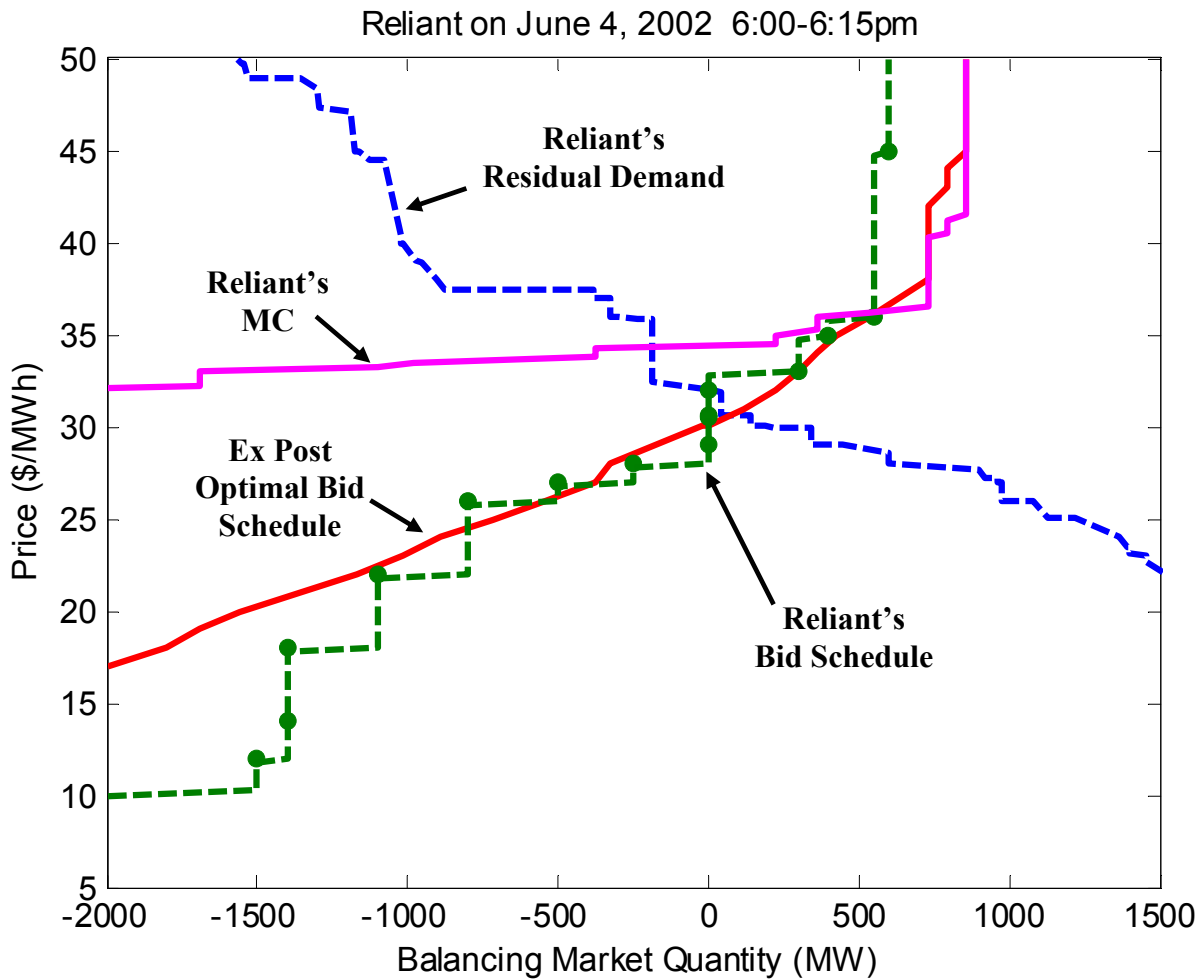


Figure 3 Sample Bidding Interface

ACS BTU ERCOT

Wednesday 7-Jan-2004 Local Time: 11:44

Date
Trade Date: 15 minutes for the day of Wednesday 7-Jan-2004

View
Scheduling for: BTU (QSE) Bryan Texas Utilities
Display

Filter
 Perspective: Balance
 Description: Balance Provision Equals ERCOT AS Deployment Balance Up Energy OR ERCOT

Grid: Averaged
Current Hour:

Filter Set: ERCOT
Filter on:

Energy balance
AS Balance
AS Bid / Award
AS Dispatch

Market Bid Strategy: 010704_DBES

Bid Type: ERCOT Bid

| Period Ending | Quantity | Bid Price |
|---------------|----------|-----------|
| 1:00 | 0 | 68.85 |
| 1:00 | 15 | 68.84 |
| 1:00 | 16 | 15 |
| 1:00 | 45 | 14 |
| 2:00 | 0 | 68.85 |
| 2:00 | 5 | 68.84 |
| 2:00 | 6 | 15 |
| 2:00 | 45 | 14 |
| 3:00 | 0 | 68.85 |
| 3:00 | 4 | 68.84 |
| 3:00 | 5 | 15 |
| 3:00 | 35 | 14 |

Delete Row
Clear

New
Save
Delete

OK
Cancel

Sum by: Schedule
To Date
Balance Of

Order: >Commodity type
 Provision Type
 Delivery Point
 Provision

| Bid BE | Premiums - Public | Dynamic Balance | Interchange | Tot | | | | |
|--------|-------------------|-----------------|-------------|-------|-------|-------|-------|------|
| 00 | 12:15 | 12:30 | 12:45 | 13:00 | 13:15 | 13:30 | 13:45 | Tot |
| 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 |
| 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 |
| 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| 9 | 8 | 8 | 8 | 8 | 8 | 8 | 8 | 8 |
| 5.97 | 5.84 | 5.84 | 5.84 | 5.84 | 6.59 | 6.59 | 6.59 | 6.59 |
| 9 | 8 | 8 | 8 | 8 | 8 | 8 | 8 | 8 |
| 1.49 | 1.46 | 1.46 | 1.46 | 1.46 | 1.65 | 1.65 | 1.65 | 1.65 |
| 7.80 | 7.80 | 7.80 | 7.80 | 5.00 | 5.00 | 5.00 | 5.00 | 4.99 |
| 4.99 | 4.99 | 4.99 | 4.99 | 4.99 | 4.99 | 4.99 | 4.99 | 4.99 |
| 5 | 5 | 5 | 5 | 15 | 15 | 15 | 15 | 14 |
| 1.95 | 1.95 | 1.95 | 1.95 | 1.25 | 1.25 | 1.25 | 1.25 | 1.25 |
| 2 | 2 | 2 | 2 | 3 | 3 | 3 | 3 | 3 |

| Schedule Description | 9:15 | 9:30 | 9:45 | 10:00 | 10:15 |
|---|-------|-------|-------|-------|-------|
| BTU-BTU D F Dansby Down Balance | 100 | 100 | 100 | 100 | 100 |
| | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| *: DownBalancingEnergy | 100 | 100 | 100 | 100 | 100 |
| | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| BTU-BTU E F ERCOT AS Deployment Down | | | | | 15 |
| | | | | | 48.39 |
| BTU-BTU E F ERCOT AS Deployment Up Er | | | | | |
| *: Energy | | | | | 15 |
| | | | | | 12.10 |
| BTU-BTU (ERCO) N F ERCOT AS Bid - Awar | 14 | 14 | 14 | 14 | 30 |
| | 11.69 | 11.69 | 11.69 | 11.69 | 7.11 |
| BTU-BTU N F Atkins7 Non-Spinning Reserv | 20 | 20 | 20 | 20 | 20 |
| | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| BTU-BTU N F Dansby Non-Spinning Reserv | 0 | 0 | 0 | 0 | 16 |
| | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| *: NonSpinningReserve | 34 | 34 | 34 | 34 | 66 |
| | 1.20 | 1.20 | 1.20 | 1.20 | 0.81 |
| BTU-BTU (ERCO) R F ERCOT AS Bid - Awar | 5 | 5 | 5 | 5 | 11 |
| | 6.25 | 6.25 | 6.25 | 6.25 | 4.88 |
| *: RegulationDown | 5 | 5 | 5 | 5 | 11 |
| | 1.56 | 1.56 | 1.56 | 1.56 | 1.22 |
| BTU-BTU (ERCO) R F ERCOT AS Bid - Awar | 5 | 5 | 5 | 5 | 15 |
| | 7.80 | 7.80 | 7.80 | 7.80 | 5.00 |
| *: RegulationUp | 5 | 5 | 5 | 5 | 15 |
| | 1.95 | 1.95 | 1.95 | 1.95 | 1.25 |
| BTU-BTU (ERCO) R F ERCOT AS Bid - Awar | 2 | 2 | 2 | 2 | 3 |
| | 4.00 | 4.00 | 4.00 | 4.00 | 2.60 |

Figure 4
Sample Bidder's Operations Interface

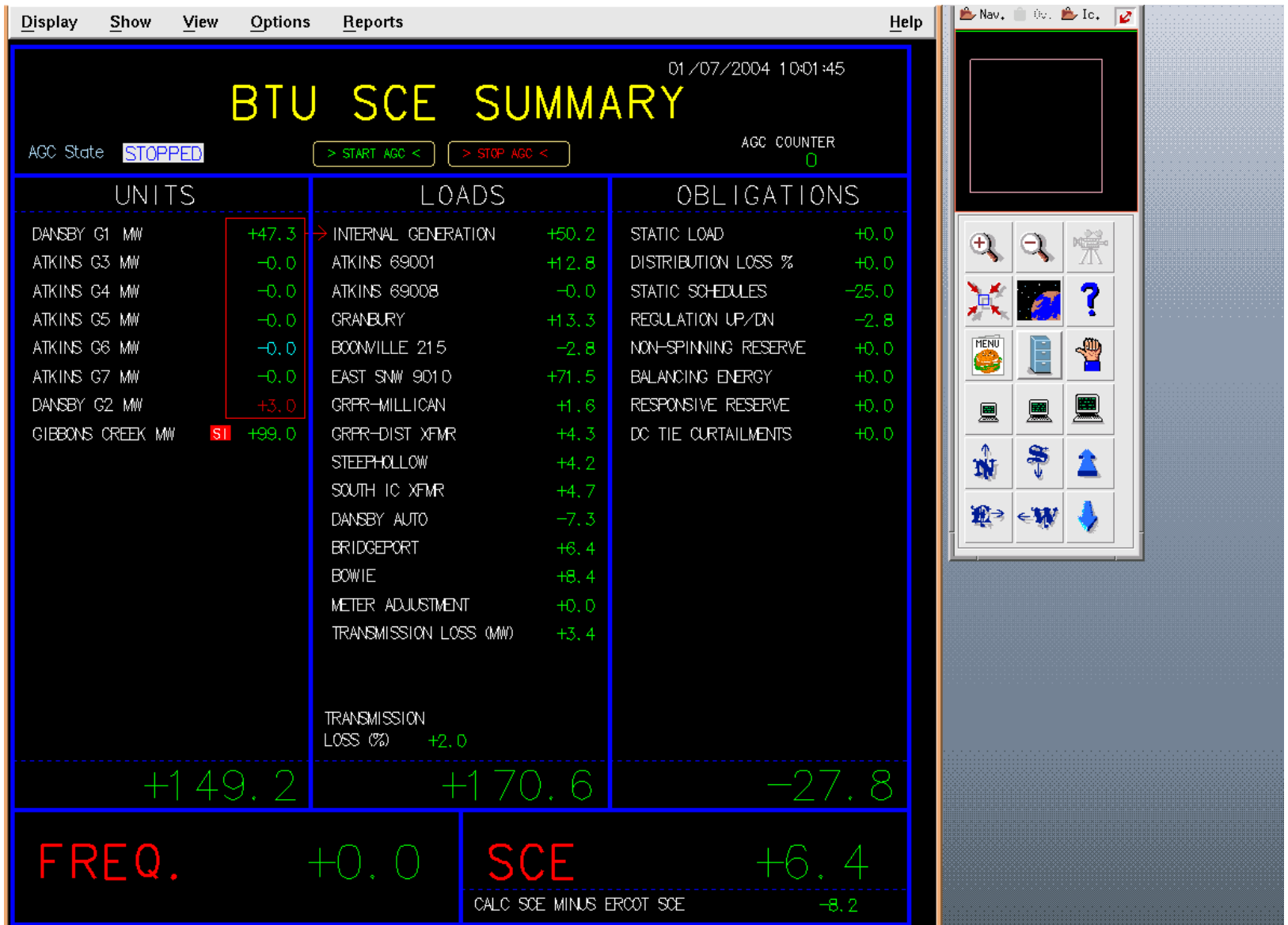


Figure 5
Sample Genscape Interface

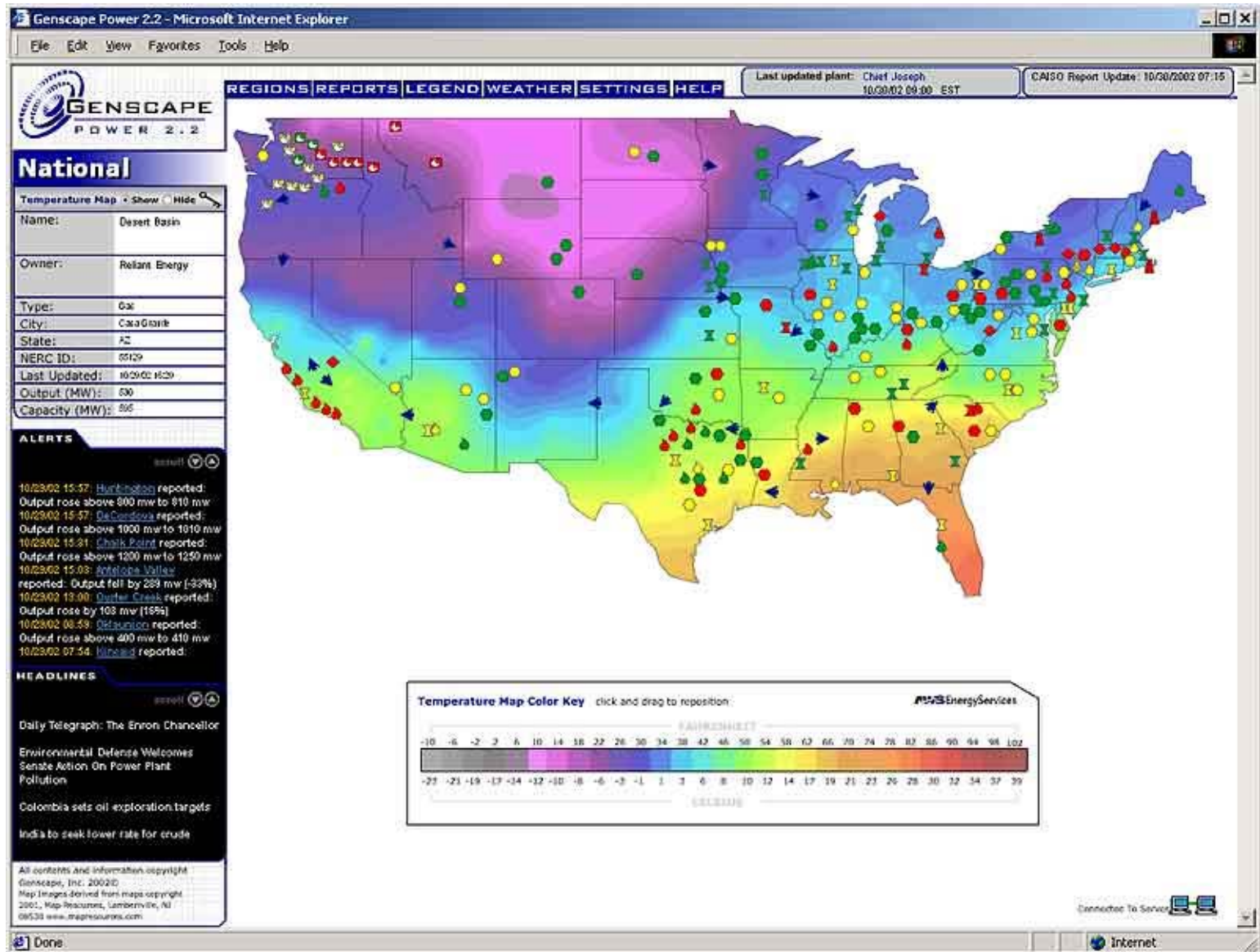
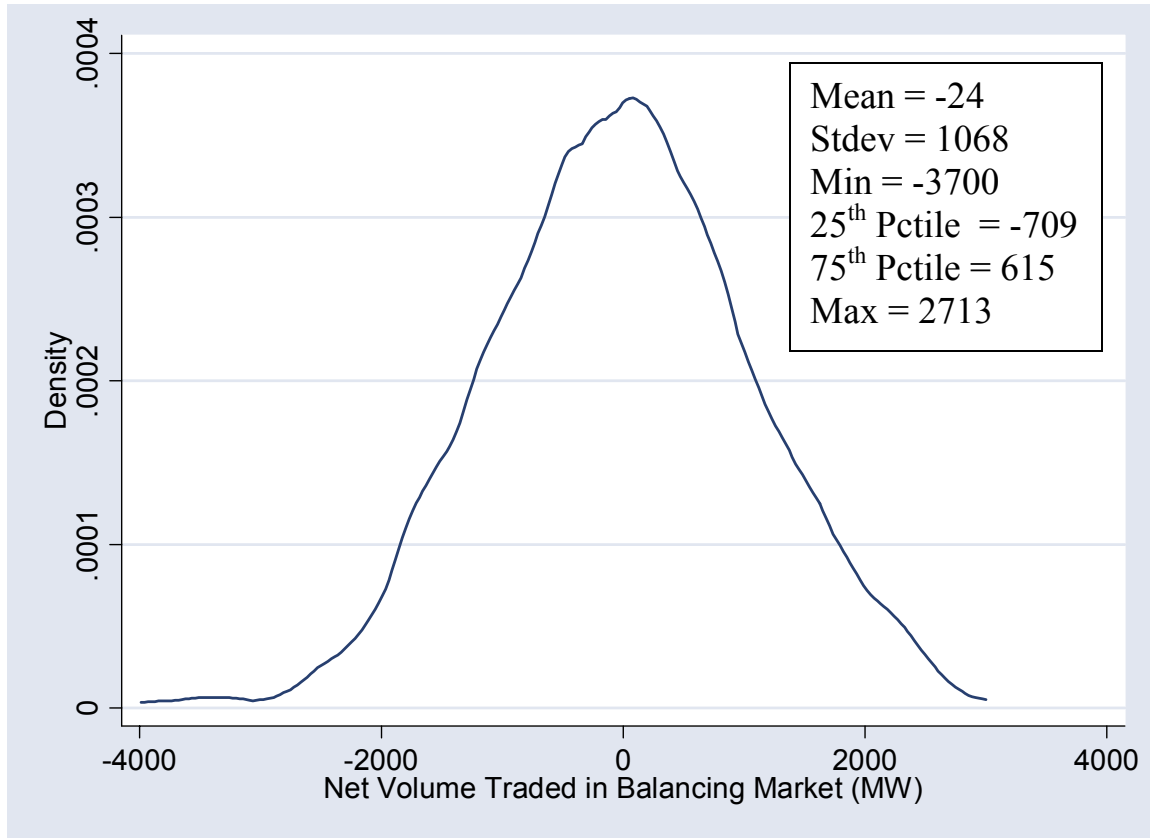


Figure 6

Density of Balancing Demand



Sample is uncongested weekdays 6:00-6:15pm from September 2001 - January 2003.

Figure 7: Examples of Actual and Optimal Bid Functions

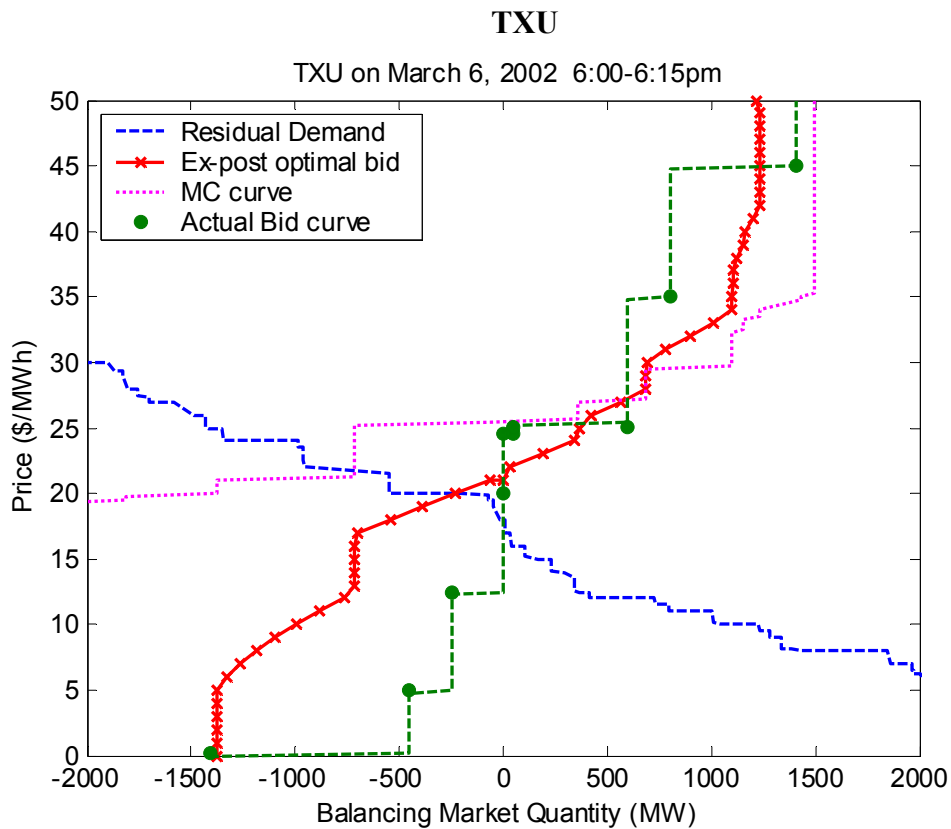
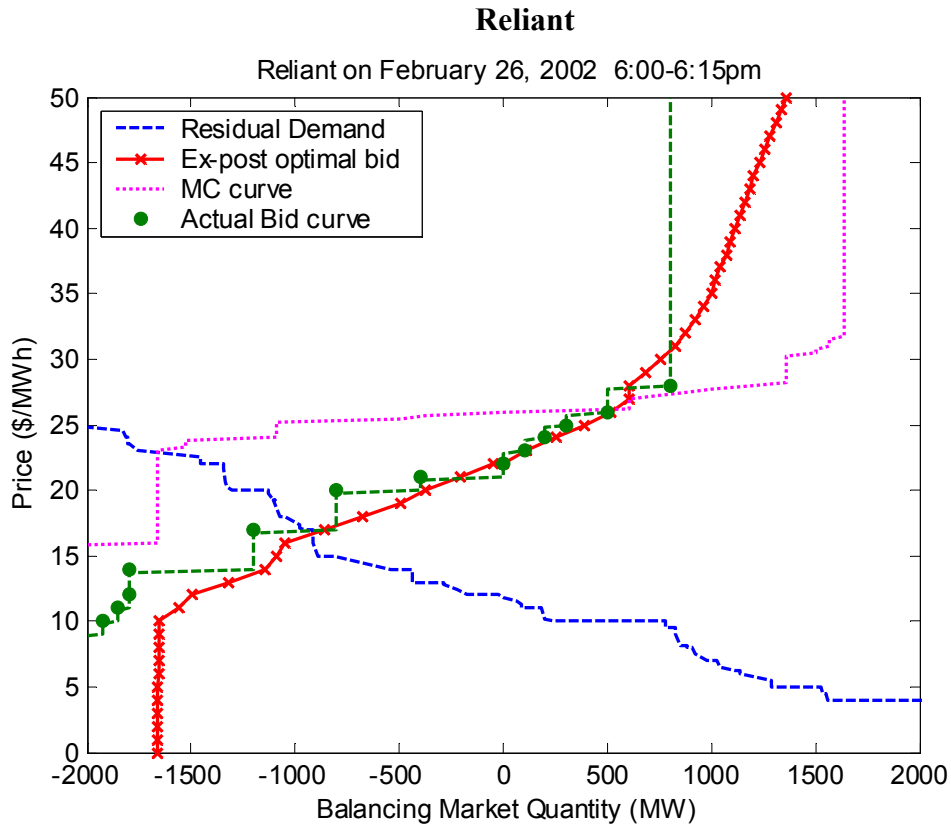
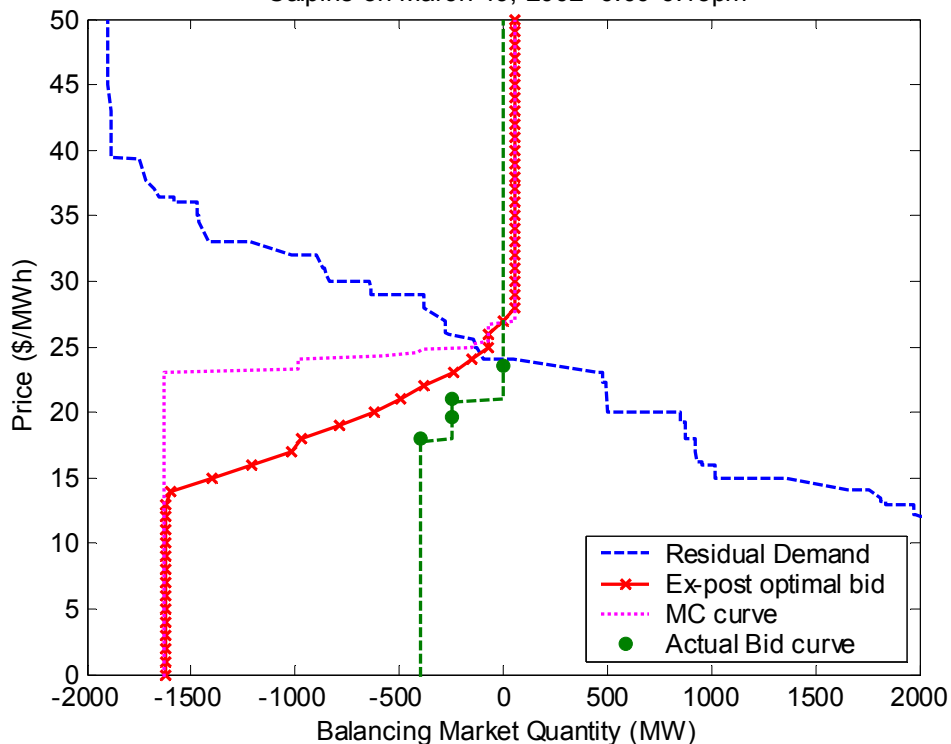


Figure 7 (contd)

Calpine

Calpine on March 19, 2002 6:00-6:15pm



Guadalupe Power Partners

Guadalupe on May 3, 2002 6:00-6:15pm

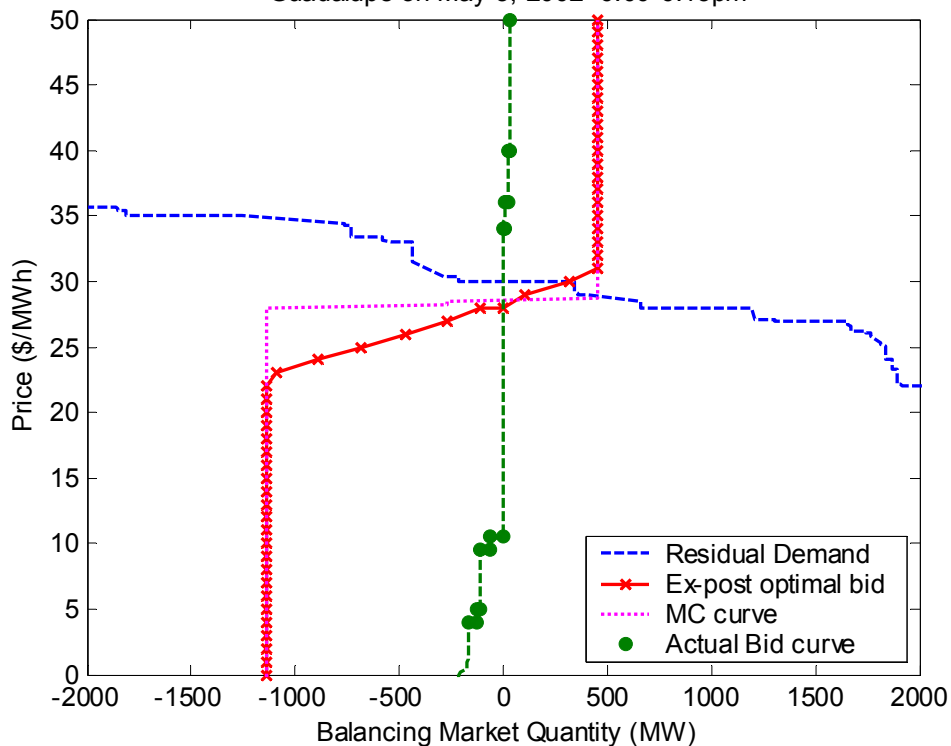
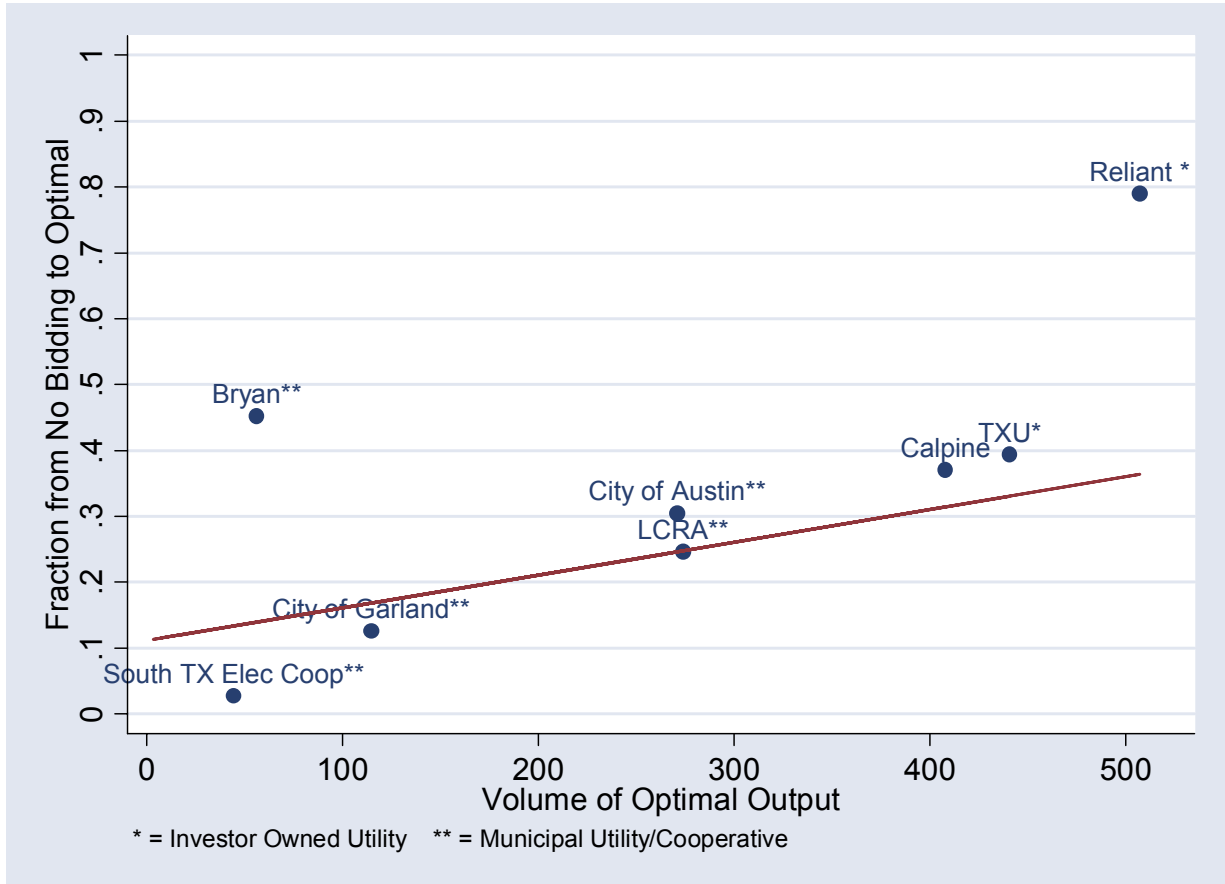


Figure 8

Performance vs. Stakes in Balancing Market For Firms Serving as Their Own Bidding Entities



Percent Achieved is the percent of possible producer surplus achieved relative to not bidding, as defined in section 4.2 and reported in column 1 of Table 3. Volume of optimal output is the mean of the absolute value of quantity sales under ex-post optimal bidding, as reported in column 6 of Table 3. We restrict this analysis to firms that serve as their own bidding entities because we can best measure “stakes” of a bidding entity by focusing on firms that serve as their own bidders.