NBER WORKING PAPER SERIES

DOES STRATEGIC ABILITY AFFECT EFFICIENCY? EVIDENCE FROM ELECTRICITY MARKETS

Ali Hortaçsu
Fernando Luco
Steven L. Puller
Dongni Zhu

Working Paper 23526
http://www.nber.org/papers/w23526

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
June 2017

Hortaçsu, Luco and Puller gratefully acknowledge financial support from the NSF (SES 1426823 and SES 1628864). We thank Natalia Fabra, Dan Fragiadakis, Avi Goldfarb, Karen Palmer, Ariel Pakes, Mar Reguant, Mo Xiao, and seminar participants at the AEA 2017, Berkeley Energy Camp 2016, Darthmouth, Hal White Antitrust Conference 2016, the FTC Ninth Annual Microeconomics Conference, HEC Montreal, IIOC 2016, Indiana, Mannheim Energy Conference 2016, NBER I.O. Winter Program Meeting 2017, Northwestern, QME 2016, Stanford, Stata Texas Empirical Micro Conference, Universidad de los Andes (Chile), U.C. Davis, U.T. Austin, University of Toronto, University of Virginia, and the Vanderbilt Conference in Honor of Rob Porter for their helpful comments. Any errors are our own. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

Does Strategic Ability Affect Efficiency? Evidence from Electricity Markets
Ali Hortaçsu, Fernando Luco, Steven L. Puller, and Dongni Zhu
NBER Working Paper No. 23526
June 2017
JEL No. D03,D22,L1

ABSTRACT

Oligopoly models of short-run price competition predict that large firms can exercise market power and generate inefficiencies. Inefficiency, however, can arise from other sources as well, such as from heterogeneity in strategic sophistication. We study such a setting in the Texas electricity market, in which bidding behavior of some firms persistently and significantly deviates from Nash-equilibrium bidding. We use information on bids and valuations to estimate the level of strategic sophistication of specific firms in the market. We do this embedding a Cognitive Hierarchy model into a structural model of bidding into auctions. We show that strategic sophistication increases with the size of the firm and it is also related to managers' educational backgrounds. We then use our model to perform counterfactual calculations about market efficiency under different scenarios that increase strategic sophistication of low-type firms either exogenously or through mergers with more sophisticated firms.

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

Fernando Luco
Department of Economics
Texas A&M University
3044 Allen
College Station, Texas 77845
fluco@tamu.edu

Steven L. Puller
Department of Economics
Texas A&M University
3046 Allen
College Station, TX 77843-4228
and NBER
puller@econmail.tamu.edu

Dongni Zhu
School of Finance
Shanghai Lixin University of Accounting
and Finance
995 Shangchuan Road
Shanghai, China, 201209
dongnizhu@lixin.edu.cn

1 Introduction

Models of strategic equilibrium are ubiquitous in empirical industrial organization, and those models can be powerful at studying the nature of competition and the welfare consequences of alternative market structures. For example, when studying differentiated product industries, firms are modeled as engaging in Bertrand-Nash competition in order to estimate marginal costs or to predict prices if firms were to merge. When studying auctions, researchers use a Bayesian Nash model of bidding to "invert" bids in order to estimate valuations and to predict revenue and efficiency under alternative auction formats.

However, a vast literature on experimental work on games shows that subjects very often deviate from Nash equilibrium behavior (for example, see Kagel and Roth [1995], Camerer [2003], Fudenberg et al. [2012], Agranov et al. [2016], and the references therein). Of course, one may think that experimental payoffs or the experience level of the subjects may not provide a good representation of behavior by professionals in the field. That said, there is a lot of heterogeneity across field participants as well: for any particular market, a quick scan of resumes of firm managers often reveals differences in the type of academic training, the years of job tenure, and previous job experience. One might expect that these differences could lead to different levels of strategic behavior when firms compete against one another.

In this paper we ask: what if all firms engage in *some* level of strategic behavior, but some firms "fall short" of playing the Nash Equilibrium? Does heterogeneity in strategic sophistication affect the efficiency of a market in an economically significant way?

To address this question in a quantitative manner, we build upon our earlier work in Hortacsu and Puller [2008] (hereafter HP) on the Texas electricity market, where we show evidence that firms systematically deviate from Bayesian Nash equilibrium bidding. Specifically, HP test whether each firm submits bids that correspond to the best response to rivals' actual bids (a requirement of Nash equilibrium). HP find that a few firms – typically larger firms – submit bids close to best-response bidding. Some medium-sized firms actively participate in the market, but submit bids that are far "steeper" than best response bids. And, small firms tend to submit bids with prices that are so far above their marginal costs that they often "bid themselves out of the market" and are not called to produce despite having low-cost generation available. These differences in bidding behavior across firms create a puzzle as to why firms bid differently and what factors are determinants of the heterogeneity.

We use a popular model of boundedly rational behavior in games to organize the heterogeneity in bidding behavior that we observe in the Texas market. In particular, we embed a Cognitive Hierarchy model of Camerer, Ho, and Chong [2004] into a structural model of bidding behavior to capture the heterogeneity in the observed deviations from Bayesian Nash equilibrium bidding.

The Cognitive Hierarchy model (hereafter CH) allows for heterogeneity in the levels of strategic thinking by firms in a market.[1] In the CH model, the least strategic players – level-0 players – are entirely non-strategic in their bidding. Level-1 players assume that all other players are level-0 players and choose actions that are the best-response to those beliefs. Level-2 players assume that all other players are some combination of level-0 or level-1 players and best respond to those beliefs. In general, level-k players assume that all other players are distributed between level-0 and level–*k-1* and best respond to those beliefs. The limiting case of this model corresponds to the Nash equilibrium.[2]

The use of the CH model in empirical work on competition was pioneered by Goldfarb and Xiao [2011] who study entry decisions into newly opened markets for local telephone competition. They find that heterogeneity in strategic sophistication – driven by different manager characteristics – affects the amount of entry into different local markets, and that more sophisticated firms are more likely to survive. Utilizing the CH model to study pricing decisions in empirical settings is not straightforward, however. This is due a critical identification problem. Consider the large number of empirical studies that use a model of expected profit maximization that maps marginal cost to prices, and then "invert" the model so that data on prices can be used to estimate the underlying marginal cost. This approach – used in many oligopoly and auction settings – hinges on the assumption of a particular form of firm conduct, or strategic behavior. Otherwise, multiple combinations of behavior and costs may be consistent with the observed prices. Thus, in most empirical settings, the task of separately identifying parameters of the CH model from unobserved costs becomes a difficult exercise.[3]

This empirical challenge can be overcome if researchers have data on both the prices *and* the marginal cost. In this paper, we exploit the data-rich environment of the Texas electricity spot market, in which many firms, that vary in size and other characteristics, compete. We have detailed data on individual firms' marginal cost of production and bids into power

[1]The CH model is part of a rich literature in experimental economics on hierarchy models to explain deviations from Nash behavior. For examples, see Nagel [1995], Stahl and Wilson [1995], Costa-Gomes, Crawford, and Broseta [2001], Crawford, Gneezy, and Rottenstreich [2008], and Arad and Rubinstein [2012].

[2]As noted in Camerer, Ho, and Chong [2004], the limiting case of the Poisson-CH model corresponds to the Nash equilibrium as long as the Nash equilibrium is reached by finitely-many iterations of weakly dominated strategies; other Nash equilibria may not correspond to this case.

[3]One novel approach to address this problem in the auctions setting is proposed by Gillen [2010], who studies joint identification of types and valuations in the level-k setting. Gillen shows point identification of the joint distribution can be obtained exploiting variation in the number of bidders and assuming constant valuations across auctions. However, in the absence of either of these, only set identification is possible. An [forthcoming] also studies identification in the level-k model; he relaxes some of these assumptions present in Gillen's work but imposes constraints on the structure of the data to identify both the number of types in the data and the type of each firm.

auctions. Having access to marginal cost data allows us to capture departures from Nash behavior, and identify the parameters of the CH model.

We find that the strongest determinant of a firm's level of strategic sophistication is size. Larger firms are higher type firms in a cognitive hierarchy, and thus are more strategically sophisticated. Manager characteristics, such as academic training, play a smaller but still significant role. Strikingly, there is substantial heterogeneity in the level of strategic sophistication across the firms in the Texas electricity market, and this heterogeneity leads the grid operator to dispatch higher cost power plants when lower cost plants are available. Moreover, we do not find evidence of substantial learning in the early years of the market.

To supplement our structural model estimates with a "reduced form" test of strategic vs. non-strategic behavior, in Section 7, we utilize a significant change in the cost structure of firms in the market. Our definition of non-strategic behavior is being unresponsive to (common knowledge) changes in competitors' costs, which encompasses several alternative definitions of level-0 behavior considered in prior literature. Our test of nonstrategic behavior is to see which firms changed their bidding behavior in response to the shutting down of a large nuclear reactor, which took a lot of supply off the market. We find, in concordance with the results from our model, that small firms did not react to the nuclear shutdown in detectable fashion, whereas larger firms did.

After estimating the model, we study how increases in strategic sophistication affect efficiency. We use the model parameters to calculate the outcomes under various scenarios in which the level of strategic sophistication of low-type firms is increased either exogenously or through mergers with high-type firms. An important benefit of using the CH model to study these multi-unit auctions is that, unlike with Nash equilibrium models, we are able to simulate unique predictions of market outcomes under various policy counterfactuals. For multi-unit auctions, solving for Nash equilibria is difficult because the researcher is searching for a fixed point in a multidimensional function space. Because CH specifies beliefs, solving for outcomes is computationally straightforward because it is a sequence of best responses, as we discuss below.[4] Thus, not only does CH allow for more realistic models of real-world bidding behavior, but it provides a convenient computational strategy for researchers who want to simulate outcomes under changes in strategic sophistication and market structure.

Our results show that increases in strategic sophistication improve efficiency, though at a decreasing rate. For example, exogenously increasing the sophistication of low-type firms to the level of median-type firms will increase market efficiency by 9-16 percent. However, efficiency improvements are smaller when firms with median levels of sophistication are given

[4] Camerer, Ho, and Chong [2004] note a related feature that the CH model can be viewed as a behavioral equilibrium refinement in certain classes of games.

higher sophistication levels.

We also simulate unique predictions of market outcomes under various policy counterfactuals. For example, consider a merger between a large and small firm that only affects the firms' bidding operations. One might expect the increase in concentration induced by the merger to enhance market power and reduce economic efficiency. However, in a merger between two boundedly rational firms, this merger could increase efficiency. Suppose that the large firm is a high-level strategic thinker and the small firm is a low-level strategic thinker. If the merger causes the large firm to take over bidding operations, then the power plants of the small firm will subsequently be controlled by a higher level strategic thinker. This can increase efficiency because the low-type firm would be less likely to bid prices so high that its efficient productive capacity is priced out of the market. We evaluate this conjecture by simulating mergers between various firms in the Texas market. Our results show that, *in this specific setting*, if a small, low-type firm were to merge with a large, high-type firm, then efficiency will improve despite the increase in concentration. However, when medium-sized firms merge with large firms, the market power effect dominates the sophistication effect, and efficiency decreases.

This paper contributes to an emerging body of literature that empirically models sophistication and learning in new markets. The two most closely related papers study firm behavior in newly opened markets. As mentioned above, the pioneering paper of Goldfarb and Xiao [2011] study the entry decisions into newly opened markets for local telephone competition. They apply the cognitive hierarchy model to an entry game and find that manager characteristics such as experience and education are determinants of strategic ability that predict firm performance. Doraszelski, Lewis, and Pakes [2017] use models of learning to predict the evolution of pricing in a newly opened electricity market. They find that monthly strategic interaction leads to a process of learning that converges to a rest point after three and a half to four years.

Our paper also contributes to the literature on how electricity generating firms formulate bids (e.g. Fabra and Reguant [2014]) and research that models oligopoly competition in the electricity sector (e.g. Wolfram [1998], Borenstein, Bushnell, and Wolak [2002], Wolak [2003] and Bushnell, Mansur, and Saravia [2008]). And more broadly, this work relates to the literature that studies differences in productivity across firms (e.g. Syverson [2004] and Hsieh and Klenow [2009]) and how managerial practices affect productivity (e.g. Bloom and Reenen [2007]).

The structure of the paper is the following. Section 2 describes the Texas electricity market that is the focus of this paper. Section 3 introduces the data and descriptive evidence that motivates our modeling assumptions. Section 4 provides background on the Cognitive

Hierarchy model and section 5 introduces our model of non-Nash bidding. Section 6 discusses identification, estimation, and results. Section 7 discusses our "reduced form" test of strategic vs. non-strategic behavior using the nuclear plant shutdown. Section 8 studies the impact of a number of policy and merger counterfactuals. Section 9 concludes.

2 Institutional Setting

We study an early year of the restructured electricity market in Texas. Prior to 2001, the Texas electricity industry consisted of vertically-integrated monopolies regulated by rate-of-return regulation. In 2001, the industry was restructured with former utilities divested into separate firms for power generation, transmission/distribution, and retailing. In August 2001, a wholesale market was opened through which generating firms that own powerplants sell wholesale power to transmission and distribution utilities that serve customers. The wholesale market allowed power trading via both bilateral transactions and an organized 'spot' auction. This paper focuses on competition in this wholesale market.

In the bilateral market, generating firms contract with utilities that serve customers. One day before production and consumption occur, each generating firm schedules a fixed quantity of production for each hour of the following day with the grid operator. This 'day-ahead schedule' serves the role of an initial plan for the next day's production. Importantly, the production levels that are scheduled one day-ahead can differ from the quantities that the firm has financially contracted in the bilateral market, so a firm can be net short or net long on its contract position with its day-ahead schedule.

The second market for wholesale trading is an organized 'day-of' spot market that is run by the grid operator to ensure that production and consumption exactly balance at every point in time. For example, suppose that a summer afternoon turns out to be hotter than previously anticipated so that realized demand for power exceeds the amount of generation that was scheduled one day-ahead. Then the spot market, or 'balancing market' in electricity parlance, is used to procure the additional supply needed to meet demand via an auction.

In the spot market, generating firms submit supply functions to increase or decrease production relative to their day-ahead schedule. If total electricity demand is higher than the aggregate day-ahead schedule, then the auction procures additional power and calls upon winning bidders to increase, or 'inc', production relative to the day-ahead schedule. In contrast, if total demand is smaller than the aggregate day-ahead schedule, then winning bidders decrease, or 'dec', production from the day-ahead schedule. During our sample period, approximately 2-5% of all power transactions occurred in the balancing auction. We study bidding behavior in this auction. Although the balancing auction is small relative to the bilateral market in percentage terms, we show below that substantial profits can be earned with 'sophisticated' bidding into the auction.

The auction format is a multi-unit, uniform-price auction. Each generating firm submits monotonically increasing step functions with up to 40 elbow points (up to 20 points to 'inc' production and 20 points to 'dec' production from the firm's day-ahead schedule). Bid functions are not tied to specific generating units; rather a firm's bid function represent offers to sell from the firm's portfolio of power plants. The firm submits a separate bid function for each hour of the day, and bids are finalized one hour before the operating hour. The demand side of the market is driven by customer usage. Because no customers during this time period responded to real-time wholesale prices, the balancing demand function for each hour is perfectly inelastic (i.e. vertical).

The grid operator clears the market every 15 minutes by finding the price where hourly aggregate supply (a monotonically increasing step function) equals each quarter hour's perfectly inelastic balancing demand. Each firm is called to supply to the balancing market the quantity that was bid at the market-clearing price, and it is paid the market-clearing price for all power called to produce in the balancing auction. Thus, if a firm is called to increase production from its day-ahead schedule, it is paid the market-clearing price for its incremental production. If a firm is called to decrease production from its day-ahead schedule, it purchases power at the market-clearing price to meet any existing contract obligations.

The generating firms that compete in the Texas market differ along a number of dimensions. Most importantly, firms vary in the size of their generating capacity. Two of the former investor-owned utilities – TXU and Reliant – are the two largest players, owning 24% and 18% of installed capacity, respectively. Other major investor-owned utilities include Central Power and Light (7% of installed capacity) and West Texas Utilities (2%). Private firms without any historical connection to utilities – so called 'merchant generators' – include firms such as Calpine (5% of installed capacity), Lamar Power Partners (4%), and Guadalupe Power Partners (2%). Small municipal utilities such as Garland Power & Light and Bryan Texas Utilities each comprise less than 1% of total capacity. The power plants are primarily fueled by natural gas and coal, although there are small amounts of hydroelectric, nuclear, and wind generation. Firms also vary in the education background and job experience of personnel in charge of power marketing operations, as we discuss below.

3 Data

We study firm bidding behavior into the balancing auctions in an early year of the market's operation. Specifically, we study the first half of the second year of the market, as depicted in Figure 1.[5] Because our sample period begins in the second year of the market's operation, the firms had time to build up their trading operations and develop bidding strategies by the time that our sample begins. By the beginning of our sample, firms had submitted bids into the balancing market for every hour of every day for one year.

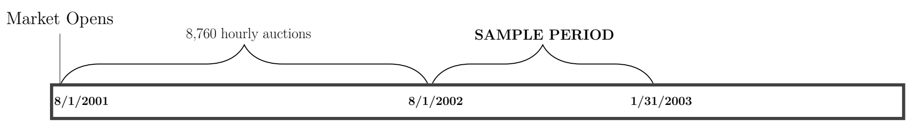


Figure 1: Market Timeline and Our Sample

One appealing feature of studying electricity markets is that detailed data are available on firm operations and costs. For each hourly auction, we have data on total demand for balancing power, each firm's bid functions, and each firm's marginal cost of providing power to the balancing market. These are the same data used in Hortacsu and Puller [2008].

Total balancing demand is perfectly inelastic because virtually no consumers face wholesale prices during the time of our study. Our balancing demand data are the hourly demand functions that were used by the grid operator to clear each auction.

The bid data consist of each firm's bids to increase and decrease production relative to the firm's day-ahead schedule. Bids are offers to supply power from the firm's portfolio of powerplants and are not tied to specific plants.

A key feature of our empirical strategy is that we can measure each firm's marginal cost of supplying power to the balancing auction. As in HP, we focus on weekdays between 6:00 and 6:15pm because the most flexible type of generators that can respond to balancing calls without large adjustment costs are online during this time interval. The technologies that are able to quickly adjust consumption in response to balancing calls are natural-gas fired units and to a lesser extent coal-fired units.[6]

We measure the marginal cost that each firm faced in each hour to change production from its day-ahead schedule. Our marginal cost function for a given firm consists of all the firm's generating units that are verified to be 'on-line' and operating during the hour of the auction.[7] Our data from ERCOT indicate which generating units are operating and the day-ahead scheduled quantity of each unit. Each unit is assumed to have constant marginal

[5]We obtained our data through a one-time arrangement with the Public Utility Commission of Texas, and unfortunately we are unable to extend our sample period to later years of the market.

[6]Nuclear and wind generated units are not marginal production units during these hours. Texas has very few hydroelectric units, and we do not study the behavior of the few firms that own hydro units.

[7]Because the units are already operating when the balancing auction clears, we do not need to include any startup costs.

cost up to capacity. For each generating unit, we observe the amount of capacity that the firm declares the unit can produce on a given day. (Below we provide evidence that the daily capacity declarations correspond to the generating units' rated capacity and that firms do not overstate their capacity). In addition, we incorporate that firms cannot reduce generation below a minimum operating level.

The primary variable cost for electricity generation is fuel. For each natural gas and coal-fired unit, we have data on the 'heat rate' – the rate at which the generator converts the energy content of the fuel into electricity (Henwood Energy Services). Fuel costs for natural gas units are the daily natural gas spot prices at the nearest trading hub in Texas (Natural Gas Intelligence) plus a distribution charge. For coal units, we use the monthly average spot price for coal delivered to Texas (Energy Information Administration). Variables costs also include a variable operating and maintenance cost per MWh (Henwood Energy Services). Finally, units that emit SO_2 incur permit costs (EPA). This approach to measuring variable costs is standard in the literature on electricity markets. We use the same data that are used in HP, and we refer the reader to Appendix A for further details.

Using these data, we calculate each firm's marginal cost of production in a given hour. Because each firm is bidding to change production relative to its day-ahead schedule, we subtract the day-ahead scheduled quantity from its total marginal cost to measure the marginal cost of supplying power to the balancing market. A stylized representation of this function is shown in Figure 2 with $MC_i^{Auction}$. This function's values in the first quadrant represents the firm's marginal cost of increasing production beyond its day-ahead schedule, i.e. supplying positive power to the balancing market. And the function's values in the second quadrant represents the firm's marginal savings of reducing production from its day-ahead schedule, i.e. supplying negative production to the balancing market.

In our model below, marginal cost is public information. While this assumption may not hold in many industries, it is likely to hold in the electricity industry because the production technology was very similar across powerplants in Texas and fuel costs are publicly available. This was confirmed in conversations with several market participants suggesting that traders have good information about their rivals' marginal cost. Moreover, firms are likely to know whether major generating units are on- or off-line at any time; some firms purchase data from an energy information company that measures real-time output using remote sensors installed near transmission lines.

In the majority of hours during our sample, the Texas market was fully integrated so that all powerplants face the same selling price. However, in 26% of hours, transmission lines were congested which led to different prices in different zones of the state. We exclude those auctions when there was transmission congestion; HP show that this does not affect

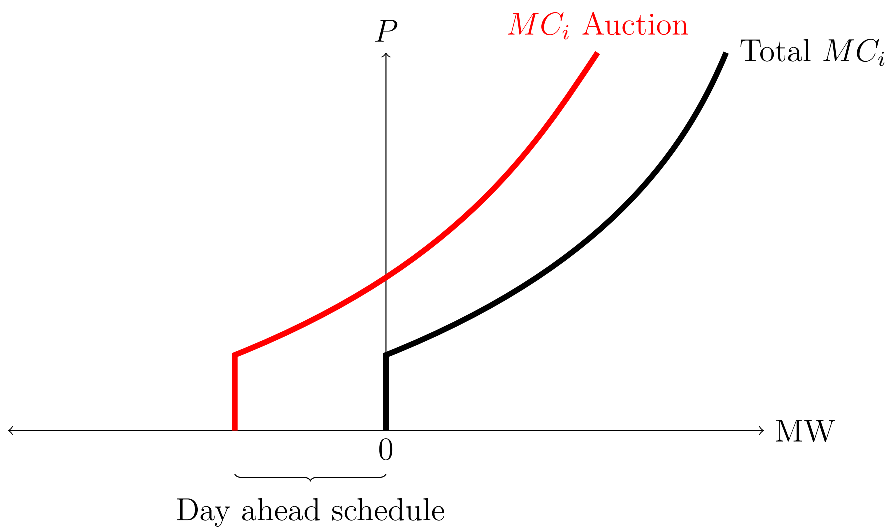


FIGURE 2: Stylized Marginal Cost of Supplying to Balancing Market

our inference about bidding behavior. After restricting our sample to weekdays during the six month sample period when there was no transmission congestion, we study 99 auctions.

3.1 Descriptive Evidence

We start this section explaining how bids would be chosen if firms best respond to their rivals actions.[8] Figure 3 explains the basic intuition of best-response bidding in this market. Suppose that a firm has marginal cost of supplying to the balancing market given by $MC_i(q)$. In addition, assume that the firm has forward contracts to supply QC_i units of power. Because the firm is a net seller after it has covered its contract position, the firm has an incentive to bid prices above marginal cost for quantities greater than QC_i. Likewise, the firm is a net buyer for quantities less than QC_i, so it has an incentive to bid prices below MC for quantities less than the contract position in order to drive down the market price. The size of the mark-up will depend on the firm's residual demand elasticity. The residual demand function RD_i is equal to the total market demand minus the supply bids by all other bidders. Suppose that it is a hot day and the firm faces RD_1. Then the firm has the incentive to bid a quantity corresponding to the point where Marginal Revenue equals Marginal Cost ($MR_1 = MC_i$) and a price corresponding to the (inverse) Residual Demand function at that quantity. This point is given by point A in the figure. Alternatively, it could be a cooler summer day which means that total demand is lower and thus residual demand is shifted in, as given by RD_2. In that case, the same logic implies that the best response is point B.

[8]We characterize a formal model of bidding in section 5.

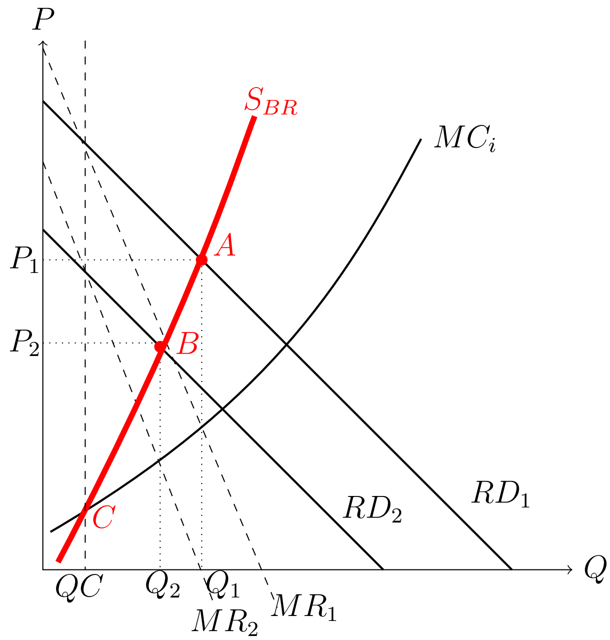


FIGURE 3: Best-Response Bidding in Spot Auction

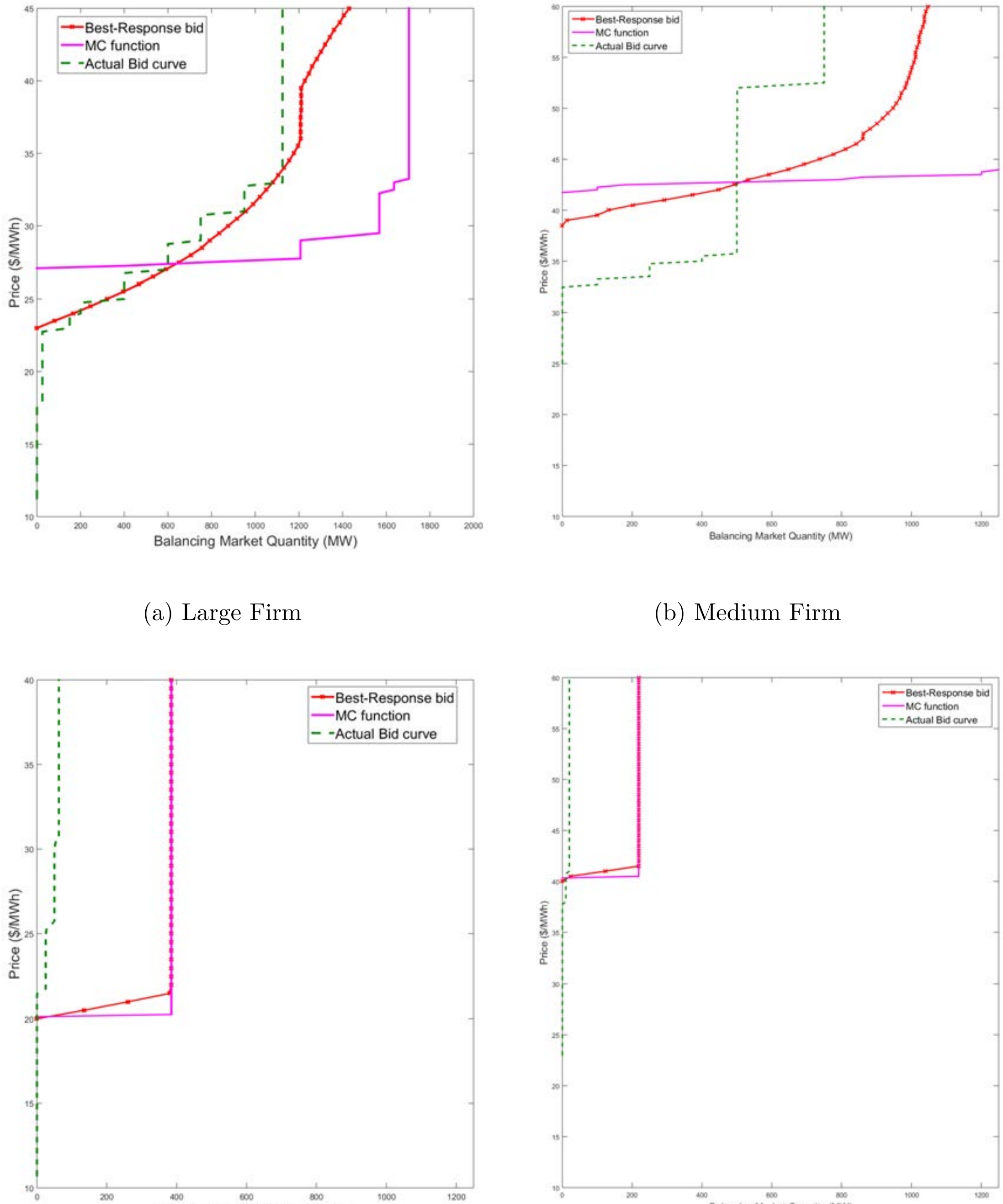
Because the firm can submit a large number of (price, quantity) points, it can consider a continuum of different residual demand functions. Thus, the firm can 'trace out' the set of best-response bids, and submit a best-response bid function given by S_i^{BR}.[9]

We can construct data analogs to these stylized pictures. Importantly, no estimation is required; the components of Figure 3 are available as *data* for each firm in each auction. We view this data-rich environment as a major strength of our approach.

We now present descriptive evidence that some firms deviate from Bayesian Nash equilibrium bidding, and we use this evidence to motivate our modeling assumptions. Figure 4 displays representative bid functions for four different firms. The top-left and top-right firms both have large quantities of generation capacity. The bottom-left firm has smaller generation capacity and the bottom-right firm is very small. Each of these figures shows the bids on the 'inc' side of the market (i.e. the horizontal axis includes only positive balancing market quantities). The firms also compete on the 'dec' side of the market (negative balancing market quantities) which we include in our analysis but for exposition are not depicted here.

The top-left panel of Figure 4 displays a representative bid function for a large firm that submits bids that correspond closely to the best-response to actual rival bids. As shown by the marginal cost function, the firm has the ability to increase production relative to its day-ahead schedule by about 1800 MW. The firm has a contract position of about 600 MW

[9]In general, it is possible that the set of best-response points is not a monotonic function, however we show in section 5 that in this setting the best-response points are monotonic.



(a) Large Firm (b) Medium Firm

(c) Small Firm (d) Very Small Firm

FIGURE 4: Actual Bids vs. Best-Response Bids for Large, Medium, Small, and Very Small Firms

upon entering the balancing market, so it has incentives to bid prices above marginal cost for quantities above 600 MW. But it will be a net buyer for quantities below 600 MW, so it has incentives to bid below marginal cost in order to drive down the market price. As seen by comparing the 'Best-Response bid' and 'Actual Bid Curve', the firm is bidding in a manner that very closely resembles best-response bidding.

However, other firms deviate from best-response bidding, and the magnitude of the deviation varies in the size of the firm. The top-right panel of Figure 4 displays a bid function for another firm with substantial generation capacity. This firm has a contract position of about 500 MW upon entering the balancing market, so best-response bids are above (below) marginal cost for quantities greater (less) than the contract position. The firm's actual bid function deviates from the best-response bid. For quantities below the contract position of about 500MW, the firm submits bids at prices of approximately \$35 which is below the marginal cost of \$43. However, the best-response to actual rivals bids is around \$40. For quantities above the contract position of 500MW, the firm submits bids at prices higher than the best-response bid prices. Loosely speaking the firm submits a bid function that is 'too steep' relative to the best-response bid. The firm's actual bid could correspond to best-responding only if the firm faced a residual demand function that is less elastic than the realized residual demand function. Thus, the bid function is consistent with the firm believing that its residual demand is less elastic that it actually is.

The bottom two panels of Figure 4 show representative bids for small and very small firms. For each firm, the contract position is zero. As shown by the best-response bids, each firm has some market power despite being small, so it is optimal to bid prices several dollars above marginal cost. However, the firm in the bottom-left panel submits high-priced bids and only offers a small quantity into the market – the firm has nearly 400 MW of available capacity yet it only offers 35 MW at relatively high prices. The firm in the bottom-right panel bids in a similar fashion – only small quantities are offered into the market even though the firm has available capacity from powerplants that are already operating.

Firms that submit bids that are 'too steep' relative to best-response bids have two important consequences for the market. From the perspective of the firm, the bids effectively price the firm out of the market which reduces producer surplus relative to best-response bidding. But, importantly, the bidding behavior reduces the efficiency of the market. In some auctions, the firms exhibiting this type of bidding behavior have low cost powerplants available to supply additional power to the balancing market, yet the generators are not called to produce because bid prices are higher that market-clearing prices. This creates productive inefficiency as higher cost generators must be called to produce instead. As we document below, this productive inefficiency can be sizable.

Constructing similar figures for other firms in our sample generates systematic patterns. In particular, a few firms bid very close to 'vertical bids' where very little generation capacity is offered into the market. Other firms offer substantial generation quantities into the market but offer that capacity at prices that are above the best-response prices. Also, for any given firm, the shapes of bids relative to best-response bids are very persistent across auctions; firms do not go back-and-forth between bidding 'too steep' and bidding 'too flat'.

These patterns in bid behavior create a puzzle – why do firms exhibit heterogeneity in bidding behavior relative to a benchmark of best-response bidding? We observe firms that systematically submit bids over a six month period that fit in a wide range – from close to the best-response benchmark to bids that are 'too steep' to bids that are nearly vertical. These patterns serve as motivation for our model of boundedly rational bidding within a cognitive hierarchy structure. In our model below, we allow firms with different characteristics to differ in level of strategic sophistication.

However, before describing our model in detail, we explore possible alternative mechanisms that could explain the observed behavior. After discussing these in detail and arguing that none of them can rationalize our data, we turn to discuss our model and estimation.

We start the discussion of possible alternative explanations by asking whether the potential profits in the balancing market are enough to justify setting up a bidding operation. We make the following calculation. First, we use observed bids to compute realized profits for each firm in each auction. Then, as in Hortacsu and Puller [2008], we calculate profits under two scenarios: (1) best-response bidding and (2) bidding vertically at the contract position (which is effectively not participating in the balancing market except to meet contract obligations). With these numbers in hand, we compute the fraction of potential profits that were achieved by actual bidding relative to non-participation. The results are summarized in Table 6 in Appendix B and show that, with the exception of the largest firm Reliant, none of the other firms achieved more than one half of potential profits. The firm-level profits 'left-on-the-table' average between \$1000 and \$4000 every hour.

Second, we show descriptive evidence that the phenomenon of offering small quantities into the market by submitting bids that are 'too steep' is equally prevalent in the first and second year of the market. We test whether firms offer more generation capacity into the market in the second versus the first year. Specifically, for each firm-auction, we calculate the amount of generation capacity that the firm offers relative to the contract position at the market-clearing price, and we test whether firms offer additional generation into the market in the second year. The results, reported in Table 7 in Appendix C, show that firms offer essentially the same capacity in the second year as they did in the first one.

Third, having shown that profits 'left-on-the-table' are significant and that bidding behavior is persistent over time, one might expect that some of the firms that are foregoing profits will eventually exit the market or be acquired by a competitor. While we cannot rule out such evolution of ownership in the long-run, evidence suggests that any such dynamics are slow in this market. Indeed, among the 12 firms that we analyze, only one firm was sold by 2005.[10] Thus, while we cannot rule out longer-term market responses to the foregone profits, it is clear that the firms that are deviating from best-response continue to be market players after four years of market operations.[11]

Next we explore possible explanations based upon technical features of electricity market operations. First, we do not believe that there are unmeasured variable costs that we fail to incorporate. Recall that our marginal cost function incorporates generating units that are 'on and operating' and that the measure of capacity is declared by the firm each day. We incorporate fuel, operating and maintenance, and emission costs which comprise all of the major sources of variable costs. It is worth noting that even if one of these variable costs is biased up to a *level* shift, this would not affect our finding that firms deviate in the *slope* of their bid functions. One might be concerned that there are unobserved costs to adjust production in the balancing market. Based on our discussions with industry officials, there are no meaningful costs to increasing or decreasing production on short notice. Firms have invested in hardware and software that automatically adjust production when the balancing market clears. For example, the sample bid in the top-right panel of Figure 4 is not consistent with there being unmeasured adjustment costs. The firm is bidding prices too *low* for positive balancing quantities below its contract position. If there were unmeasured adjustment costs to changing production from the day-ahead position, one would expect prices *above* the best-response benchmark for quantities just above zero.

The primary means through which adjustment impacts output are constraints on the rate at which generating units can increase production – called 'ramprates'. In the vast majority of intervals in our sample, the marginal units to adjust output are natural gas-fired units which generally are flexible and have 'fast' ramprates. Ramprates are unlikely to drive the cross-firm heterogeneity in bidding behavior because many of the firms have generating units with similar ramprates. For example, the ramprates of the generating units of the top-left and bottom-left firms in Figure 4 are very similar, yet the firms bid quite differently.

Second, one might be concerned that bidding rules prevent firms from submitting bids that correspond to the best-response bids. The best-response benchmark that we show with the descriptive evidence assumes that all uncertainty results in shifts rather than pivots in residual demand. As a result, the set of best-response bids will be monotonic which is a

[10]Central Power & Light was acquired by Sempra Energy and a private equity group in March 2004.

[11]We are unable to acquire additional years of data to analyze bidding behavior in later years. Therefore, we cannot assess whether the under-performing firms made changes in internal bidding operations.

requirement of the bidding rules. In general, it is possible that uncertainty results in pivots in residual demand. The published version of HP includes tests for this possibility, and the NBER working paper version (Hortacsu and Puller [2005]) includes moment-based tests for expected profit maximization. We find strong evidence that the form of uncertainty and bidding rules do not bias our best-response bids as a benchmark for expected profit-maximizing behavior.

The most straightforward evidence is that a simple trading rule would have systematically raised profits for all but the largest firm. This trading rule uses only information available at the time that bids are submitted and it respects all auction rules about the shapes of bid functions. The trading rule exploits an institutional feature of the Texas market – the grid operator publicly released the aggregate bid schedule with a 2 day lag. Thus, firms can learn recent information about their rivals' aggregate bid behavior. Suppose a firm were to use the lagged aggregate bid data to create best-response bid functions to rivals' bids from 3 days prior to each auction, and submit these bids to the current auction. We compute these lagged best-response bids and use the bids to clear the market with the actual (step function) residual demand for the current auction. We find that this simple trading rule significantly outperforms the actual realized profits for all but the largest firm. The results of this test are reported in Table 8 of Appendix D.

Third, the possibility of congested transmission lines does not explain why actual bids differ from best-response. Recall that we exclude auctions where transmission lines were congested between zones in Texas, but the possibility of congestion could impact bids even if congestion was not realized. However, this does not appear to explain the deviations from best-response bids that we observe. First, expected congestion might explain bids that are 'too steep' in import-constrained zones, but some of the firms are in export-constrained zones and nevertheless submit bids that are 'too steep'. Second and more formally, Hortacsu and Puller [2005] finds that firm profitability is not strongly related to the frequency of transmission congestion.

Fourth, we rule out the possibility that our measure of capacity – the firm's self-declared capacity for each day – overstates the actual capacity. We compare each firm's stated capacity to the highest amount of production that we observe during our sample. All firms are observed to use at least 75% of stated capacity and on average to use 90% of stated capacity (see Table 9 in Appendix E). This suggests that our finding that firms do not bid significant capacity into the balancing market is not driven by overstating capacity. Moreover, the concern of overstated capacity does not apply to periods when firms *decrease* production, or 'dec', and we observe deviations from Nash equilibrium in 'dec' intervals as well.[12]

[12] Additional markets for ancillary services such 'regulation up' cannot explain all of the observed deviations

Finally, it seems implausible that collusion explains the deviations from best-response. Many of the firms - especially the small firms – that deviate from best-response submit bids that are vertical at the contract position. As a result, revenues are zero. Thus, a collusive regime would require side payments from the few large firms to numerous small firms.

4 Empirical Strategy to Estimate a Cognitive Hierarchy Model of Bidding

4.1 Background on Cognitive Hierarchy

The theoretical literature has developed a rich set of models of boundedly rational strategic behavior that can explain deviations from Bayesian Nash equilibrium play. Generally speaking, bounded rationality models relax one of the two conditions of Nash Equilibrium: (1) players maximize expected payoffs given beliefs about their rivals' actions and (2) player beliefs about rivals' actions are consistent. Hierarchy models (such as Cognitive Hierarchy and level-k) maintain the assumption of best-response but relax the assumption of consistent beliefs.[13] These models conceptualize players as having a hierarchical structure of strategic, or level-k, thinking. Seminal work on level-k models include Costa-Gomes, Crawford, and Broseta [2001], Camerer, Ho, and Chong [2004], and Crawford and Iriberri [2007].

Cognitive Hierarchy (CH) developed by Camerer, Ho, and Chong [2004] conceptualizes players as engaging in different levels of strategic thinking ordered in a hierarchy. The least sophisticated players – 0-step players – engage in no strategic thinking, while higher types assume that all other players are distributed between 0-step and k-1-step players according to a Poisson distribution.[14] Importantly, a player's belief about rivals need not be correct; hence, the beliefs are not mutually consistent. However, each player rationally best-responds

from best-response because the total amount of ancillary services procured is significantly lower than the joint quantity of excess capacity that firms have available when submitting balancing market bids.

[13] Another model used in the bounded rationality literature – Quantal Response Equilibrium (McKelvey and Palfrey [1995]) – does not appear to be suitable in our particular setting. QRE has the property that players play more profitable strategies with higher probability. However, small players in our setting systematically play low-profit strategies as shown in the sample bid functions above. In other words, it does not appear that our bidders estimate expected payoffs in an unbiased way, a key feature of the QRE model.

[14] The model does not require the distribution be Poisson. However, Camerer et al. [2004] note that the Poisson has the property that as k rises, fewer players perform the next step of thinking, which is consistent with increasing working memory being required for an additional step of iterative calculation. This cognitive hierarchy structure is conceptually appealing because it captures behavior in which firms have limits to the level of strategic thinking and/or firms are overconfident about their own abilities. Note that it might seem peculiar that a firm believes that all rivals are strictly lower types rather than equal types. However, if a firm believed that its rivals were playing the same level strategy, that firm would best respond which would make it a higher than k-level player.

given its (perhaps incorrect) beliefs, meaning that CH maintains the rationality assumption of Nash Equilibrium but relaxes the assumption of mutually consistent beliefs.[15]

4.2 Big Picture of Modeling and Estimation Strategy

The recursive nature of decision rules under CH facilitates a computationally tractable empirical strategy. Consider firm i that is type k. Under the CH model, firm i believes its rivals are distributed between type-0 and type-(k-1), according to a normalized Poisson distribution with parameter τ. Given its marginal cost and beliefs about rivals' types, firm i chooses bids to maximize expected profits.

One critical feature of estimating a CH model is how to define level-0 behavior (or in the language of Camerer et al. "0-step players"). In the theoretical literature, a common assumption is that level-0 players (uniformly) randomize across all possible strategies, although that assumption can be relaxed to match a particular setting (i.e., Goldfarb and Xiao [2011] assume that level-0 players believe they are monopolists in an entry game). In the context of the Texas electricity auctions, there is a natural assumption about non-strategic thinkers: we observe some firms bidding "vertically" at their contract positions for the range of plausible prices. (That is, the firms submit bids similar to the bottom-right panel of Figure 4). In other words, these firms are indicating that at even very high prices, they do not want to sell power into the balancing market. This clearly violates any standard model of (expected) profit-maximization; the firms have low cost generation to offer into the market, but they choose not to do so. Thus, "vertical bidding at the contract position" is a natural candidate for level-0 bidding behavior.

Our modeling choice about level-0 behavior is important because it anchors the bidding of higher types that respond to level-0 firms. Vertical bidding at the contract position (i.e. "non-participation in the balancing market") is a natural level-0 assumption because the firms that submit nearly vertical bids earn the lowest fraction of potential profits. While other types of bidding behavior may also be considered level-0 behavior, in section 6.1.1 we discuss the implications of such bidding patterns and find that they are inconsistent with observed behavior. Moreover, in section 7, we provide additional empirical evidence supporting the vertical bidding assumption. In all applications of the cognitive hierarchy model we are aware of, level-0 agents are "non-strategic," in the sense that their actions do not respond to changes in the (common knowledge) payoff/cost structure of their competitors. Section 7 reports evidence that firms submitting vertical bids indeed do not respond to significant

[15] The level-k model is a specific form of the CH model where a level-k player assumes that *all* other players are level-*(k-1)*. In other words, rather than rivals coming from a *distribution* of types *(k-1)* and below, in the level-k model, rival firms are type *(k-1)*.

changes in competitors' cost structure.

Of course, in general non-strategic behavior could be consistent with bidding marginal cost. However, we do not observe any of the firms systematically bidding marginal cost. As shown in the representative bid functions in Figure 4, firms that deviate from best-response submit bids that are "too steep" rather than bids that are "too flat" as one would see if the firms bid their marginal cost. In Appendix F we provide further evidence that firms do not bid marginal cost. If firms submit bids that deviate from best response in the direction of marginal cost bidding, then the actual quantity sold into the auction should be *greater* than predicted sales under best-response bidding. However, this is not the case in our setting – Appendix Table 10 shows that firms systematically sell less output that the best-response level of output, which supports our assumption that firms deviate by submitting bids that are steeper than best-response.

One advantage of using this level-0 assumption is that we do not need to make strong assumptions about the form of the bid functions. Instead, as we show below, the assumption of level-0 bidders bidding vertically at their contract positions together with the recursive solution method of the CH model allow us to completely characterize the bidding functions without further assumptions about how private information enters the bidding decision.

Finally, we assume that not all firms engage in strategic thinking or even enter the Cognitive Hierarchy. Indeed, we allow only a subset of firms to enter the hierarchy, while the rest form part of an unmodeled fringe. We do this because allowing for more firms makes the problem computationally challenging as each firm needs to compute its rivals bidding functions for all possible types, for all auctions. Furthermore, we do not have marginal cost data for all firms for all auctions, which imposes a constraint on the number of firms that we can include in the Cognitive Hierarchy. Accordingly, we model all "big" firms entering the Cognitive Hierarchy plus a number of medium-sized and small ones including the ones depicted in Figure 4.

Once level-0 bidding is defined, we can use our data on each firm's marginal cost to calculate the bidding behavior for a firm of any type $k > 0$. Firms observe characteristics of their rivals X_{-i} (e.g. size) and form beliefs about their rivals' type distributions. Specifically, the type distribution is given by $Poisson(\tau_i(X_i, \beta))$ where $\tau_i(\cdot)$ captures how firm characteristics map into type.[16] For any mapping from firm characteristics to type, we use an iterative procedure to calculate each player's optimal theoretical bids under various sophistication levels. Given level-0 players' bids, we calculate level-1 best-response bids for each firm. Then given our calculated level-0 and level-1 best-response bids, we calculate the level-2 best-response

[16] This implies that firms with the same characteristics can be a different type; $\tau_i(\cdot)$ determines the *distribution* from which type is drawn. However, rivals with the same observable characteristics generate the same (distribution of) beliefs.

bids for each firm, and continue this recursive process up to the highest type K.

We then compare these calculated bids to the firm's actual bidding behavior. The estimation process finds the parameters of $\tau_i(X_i, \beta)$ – how firm characteristics such as size affect strategic sophistication – that minimize the distance between actual bids and the bids predicted under CH. That is, in estimation, we use observed bids and realized marginal costs to recover the type of each firm. For this reason, it is critical that we observe marginal costs; in the absence of cost data, one would not be able to identify types from bid data without additional assumptions regarding the cost function.[17] In other words, instead of using data on observed bids and a Bayesian Nash equilibrium model of behavior to recover costs, we use data on costs and bids to recover the type that rationalizes observed behavior.

5 Formal Cognitive Hierarchy Model of Bidding

A formal model of bidding into the Texas electricity auctions needs to formulate best-response bidding in a setting where firms have beliefs about rivals as characterized by the Cognitive Hierarchy model. We incorporate a model of bidding into share auctions (Wilson [1979] and Hortacsu and Puller [2008]) into the Poisson Cognitive Hierarchy model (Camerer, Ho, and Chong [2004]).

Demand for power in each spot auction is given by $\tilde{D}_t(p_t) = D_t(p_t) + \varepsilon_t$ which is the sum of a deterministic and stochastic component. The auctions occur in a private values setting where the private value is the firm's variable cost of providing power to the grid. Firm i has costs to supply power in period t given by $C_{it}(q)$. Prior to the auction, each firm has signed contracts to deliver certain quantities of power each hour QC_{it} at price PC_{it}, and we take these contracts to be pre-determined. $C_{it}(q)$ is public information and QC_{it} is private information. Each firm is a k-step thinker. Firm i has private information on its own type k_i, but it only knows the distribution from which rival types are drawn. In each auction, firms simultaneously submit supply schedules $S_{it}^k(p, QC_{it})$ to produce different quantities at different prices. Let the bid function by rival j of type l be denoted $S_{jt}^l(\cdot)$.

All N sellers are paid the market-clearing price (p_t^c), which is determined by:

$$\sum_{i=1}^{N} S_{it}(p_t^c, QC_{it}) = D_t(p_t^c) + \varepsilon_t \tag{1}$$

From the perspective of firm i with private information on k_i, QC_{it}, and submitting bid $\hat{S}_{it}(p)$, the uncertainty can be characterized by defining the following function $H(\cdot)$ which

[17] Specifically, without any assumption on the form of the cost function, it is always possible to recover a cost function that rationalizes observed bids.

defines the probability that the market-clearing price p_t^c is below any price level p:

$$H_{it}(p, \hat{S}_{it}(p); k_i, QC_{it}) \equiv Pr(p_t^c \leq p | \hat{S}_{it}(p), k_i, QC_{it}) \tag{2}$$

There are three sources of uncertainty – (1) the shock to demand (ε_t), (2) each rival's type of k-step thinking, and (3) each rival's contract position QC_{jt} which affects the rival's bids.

The event that the market-clearing price p_t^c is less than any given price p is the event that there is excess supply at that p. Plugging the market-clearing condition (Equation 1) into (Equation 2):

$$\begin{aligned} H_{it}(p, \hat{S}_{it}(p); k_i, QC_{it}) &= Pr(\sum_{j \neq i} S_{jt}^l(p, QC_{jt}; k_i) + \hat{S}_{it}(p) \geq D_t(p) + \varepsilon_t | \hat{S}_{it}(p), k_i, QC_{it}) \\ &= \int_{\boldsymbol{QC_{-it}} \times \boldsymbol{l_{-i}} \times \varepsilon_t} 1(\sum_{j \neq i} S_{jt}^l(p, QC_{jt}; k_i) + \\ &\quad \hat{S}_{it}(p) \geq D_t(p) + \varepsilon_t) dF(\boldsymbol{QC_{-it}}, \boldsymbol{l_{-i}}, \varepsilon_t | \hat{S}_{it}(p), k_i, QC_{it}) \end{aligned} \tag{3}$$

where $F(\boldsymbol{QC_{-it}}, \boldsymbol{l_{-i}}, \varepsilon_t | \hat{S}_{it}(p), k_i, QC_{it})$ is the joint density of each source of uncertainty from the perspective of firm i.

A firm's realized profit in this setting (after the realization of uncertainty) is given by:

$$p \cdot \hat{S}_{it}(p) - C_{it}\left(\hat{S}_{it}(p)\right) - (p - PC_{it})QC_{it} \tag{4}$$

This profit represents spot market revenues minus costs plus the payoff from its contract position.

We model the bidder's expected utility maximization problem, where we allow for bidders to potentially be risk averse. We denote the utility enjoyed by the bidder earning π dollars of profit as $U(\pi)$. Under the CH model, best-response $k-$step thinking bidders will solve:

$$\underset{\hat{S}_{it}(p)}{\text{Max}} \int_{\underline{p}}^{\overline{p}} \left(U\left(p \cdot \hat{S}_{it}(p) - C_{it}\left(\hat{S}_{it}(p)\right) - (p - PC_{it})QC_{it}\right)\right) dH_{it}\left(p, \hat{S}_{it}(p); k_i, QC_{it}\right)$$

One can show that the Euler-Lagrange necessary condition for the (pointwise) optimality of the supply schedule is given by:

$$p - C'_{it}\left(S_{it}^*(p)\right) = \left(S_{it}^*(p) - QC_{it}\right) \frac{H_s\left(p, S_{it}^*(p); k_i, QC_{it}\right)}{H_p\left(p, S_{it}^*(p); k_i, QC_{it}\right)} \tag{5}$$

where H_s and H_p are given by derivatives of Equation 3.

There is a simple intuition behind this condition. To see this, for the moment ignore the term $\frac{H_S}{H_p}$ (it will be positive). The left hand side is the difference between bid prices and marginal cost. Suppose that the firm is a net seller into the market because it is supplying more than its contract position (i.e. $S(\cdot) > QC_{it}$). Then the firm will have an incentive to bid above marginal cost, i.e. $p > C'_{it}$, in order to "exercise market power". The amount of market power is determined by the term $\frac{H_S}{H_p}$. The denominator of this term is simply the density of the market clearing price. The numerator is the "market power term" – how much the firm can change the (distribution of) the market price by changing its supply bid.

The goal is to find $S^*_{it}(p)$ for firm i if the firm is type k – the best-response bid function for each firm i in auction t if the firm is type k. And in our empirical exercise, we will compare the firm's actual bid to each of these best-response functions to make inferences about what type of k-step thinker the bidder is.

We use detailed data and several identifying assumptions to "measure" each component of Equation 5, which allows us to calculate the best-response function for each type. In our data, we observe the marginal cost function C'_{it}, and we follow the strategy developed in HP to measure QC_{it}.

Ideally, one would like to (non-parametrically) estimate $\frac{H_S}{H_p}$ as is common in the T-bill literature (e.g. Hortacsu and McAdams [2010], Hortaçsu et al. [2012], Kang and Puller [2008]). However, in this institutional setting it is not credible to pool across auctions or to assume that some subsets of bidders in a given auction are ex ante symmetric. Therefore, HP follow the approach of assuming that bid strategies are additively separable in private information (QC_{it}). HP also show that expected profit-maximizing bids are ex-post optimal. The intuition is that in the absence of uncertainty about rivals types, all other sources of uncertainty affect the intercept but not the slope of residual demand. As a consequence, the single observed realization of uncertainty is sufficient to calculate $RD'(p)$ under all possible realizations of uncertainty.

This approach will not work in the Cognitive Hierarchy model. In CH, there is an additional source of uncertainty – firms have private information on their own type and uncertainty about their rivals' types (though the uncertainty is fully characterized for a firm of given type k). Intuitively, higher type rivals are likely to submit bid functions that are "closer" to best response, which in our setting means "flatter". As a result, uncertainty affects the slope of residual demand, so the expected profit-maximizing bid function does not reduce to the simple formula developed by HP.

For this reason, we now make three identifying assumptions so that we can "measure" $H_{it}\left(p, \hat{S}_{it}(p); k_i, QC_{it}\right)$ and thus its derivatives H_s and H_p. The first assumption defines the bidding behavior for type-0 bidders. The assumption has both the properties that it is natural in our setting and that it facilitates computation of CH outcomes by allowing us to solve the problem recursively. This recursive property yields an additive separability condition that makes it computationally straightforward to solve the firm's expected profit maximization problem, as we show below. The second and third assumptions define the distribution of types of the CH model and the distribution of the remaining sources of uncertainty.

Type-0 Bidding:

Assumption 1. *Type-0 bidders submit perfectly inelastic bids at their contract positions. That is,*

$$S^0_{it}(p, QC_{it}) = QC_{it} \quad \forall p \in [\underline{p}, \overline{p}], \ \forall i \in \boldsymbol{l_0},$$

where $\boldsymbol{l_0}$ represents the set of bidders type 0.

This formalizes our observation in Section 3.1 that the least sophisticated bidders use the balancing market to meet any remaining contract obligations but otherwise do not participate in the market – they bid vertically at their contract positions.

Type-1 Bidding: Given the assumption about type-0 bidding, we can characterize bids for type-1 firms. For a bidder type-1, all rivals are type-0 under the CH model. Thus, we can write $H(\cdot)$ (Equation 3) for a type-1 firm submitting bid $\hat{S}^1_{it}(p)$:

$$\begin{aligned} H_{it}(p, \hat{S}^1_{it}(p); k_i = 1, QC_{it}) &= \int_{\boldsymbol{QC_{-it}} \times \boldsymbol{l_{-i}} \times \varepsilon_t} 1(\sum_{j \neq i} S^0_{jt}(p, QC_{jt}) + \hat{S}^1_{it}(p) \geq \\ &\quad D_t(p) + \varepsilon_t) dF(\boldsymbol{QC_{-it}}, \boldsymbol{l_{-i}}, \varepsilon_t | \hat{S}^1_{it}(p), k_i = 1, QC_{it}) \\ &= \int_{\boldsymbol{QC_{-it}} \times \boldsymbol{l_{-i}} \times \varepsilon_t} 1(\sum_{j \neq i} QC_{jt} - \varepsilon_t \geq \\ &\quad D_t(p) - \hat{S}^1_{it}(p)) dF(\boldsymbol{QC_{-it}}, \boldsymbol{l_{-i}}, \varepsilon_t | \hat{S}^1_{it}(p), k_i = 1, QC_{it}) \\ &= \int_{\boldsymbol{QC_{-it}} \times \boldsymbol{l_{-i}} \times \varepsilon_t} 1(\theta_{it} \geq D_t(p) - \hat{S}^1_{it}(p)) dF(\boldsymbol{QC_{-it}}, \boldsymbol{l_{-i}}, \varepsilon_t | \hat{S}^1_{it}(p), k_i = 1, QC_{it}) \end{aligned}$$

where the second equality follows from Assumption 1 and the third equality from defining $\theta_{it} \equiv \sum_{j \neq i} QC_{jt} - \varepsilon_t$.

This tells us that, as a bidder type-1 believes all its rivals are type-0, she expects all her rivals to submit perfectly inelastic bids determined by her rivals contract positions (which are private information). Furthermore, conditional on rivals' types, uncertainty in rivals' QC_{jt} and the aggregate demand shock act as shifters in residual demand (but not pivots). Thus, all that matters with respect to uncertainty about $(\boldsymbol{QC_{-it}} \times \varepsilon_t)$ is the distribution of the random variable θ_{it}.

Let $\Gamma(\cdot)$ denote the conditional distribution of θ_{it} (conditional on the realization of all $N-1$ draws from the joint distribution of rival types) and let $\Delta(l_{-i})$ denote the marginal distribution of the vector of rival firm types. Then $H(\cdot)$ for a type-1 bidder becomes:

$$H_{it}(p, \hat{S}^1_{it}(p); k_i = 1, QC_{it}) = \int_{l_{-i}} \left[1 - \Gamma\left(D_t(p) - \hat{S}^1_{it}(p)\right)\right] \cdot \Delta(l_{-i})$$

Taking derivatives of $H(\cdot)$ to find H_S and H_p and plugging into to solve for $\frac{H_S}{H_p}$:

$$\frac{H_s\left(p, S^*_{it}(p); k_i, QC_{it}\right)}{H_p\left(p, S^*_{it}(p); k_i, QC_{it}\right)} = \frac{\int_{l_{-i}} \gamma\left(D_t(p) - \hat{S}^1_{it}(p)\right) \cdot \Delta(l_{-i})}{-\int_{l_{-i}} \gamma\left(D_t(p) - \hat{S}^1_{it}(p)\right) D'_t(p)\Delta(l_{-i})}.$$

The implication is that if type-0 bidders submit perfectly inelastic bids at their contract positions, then the bids of type-1 bidders are additively separable functions of price and private information on their contract positions. Bid functions will take the form: $S^1_{it}(p, QC_{it}) = \alpha^1_{it}(p) + \beta^1_{it}(QC_{it})$. Specifically, $S^1_{it}(p, QC_{it}) = \alpha^1_{it}(p) + QC_{it}$.

To see this, note that bids of bidders type-1, $S^1_{it}(p)$, can be calculated from equation (5), which can be rewritten as

$$\begin{aligned} S^1_{it}(p) &= \left[(p - C'_{it}\left(S^1_{it}(p)\right)\right] \frac{H_p\left(p, S^1_{it}(p); k_i, QC_{it}\right)}{H_s\left(p, S^1_{it}(p); k_i, QC_{it}\right)} + QC_{it} \\ &= \left[(p - C'_{it}\left(S^1_{it}(p)\right)\right] \frac{\int_{l_{-i}} \gamma\left(D_t(p) - \hat{S}^1_{it}(p)\right) \cdot \Delta(l_{-i})}{-\int_{l_{-i}} \gamma\left(D_t(p) - \hat{S}^1_{it}(p)\right) D'_t(p)\Delta(l_{-i})} + QC_{it} \\ &= \alpha^1_{it}(p) + QC_{it} \end{aligned}$$

because the argument $[(p - C'_{it}\left(S^1_{it}(p)\right)] \frac{\int_{l_{-i}} \gamma\left(D_t(p) - \hat{S}^1_{it}(p)\right) \cdot \Delta(l_{-i})}{-\int_{l_{-i}} \gamma\left(D_t(p) - \hat{S}^1_{it}(p)\right) D'_t(p)\Delta(l_{-i})}$ is a function of price p. Therefore, bids of type-1 bidders are additively separable and can be represented by $S^1_{it}(p) = \alpha^1_{it}(p) + QC_{it}$, where $\alpha^1_{it}(p) = [(p - C'_{it}\left(S^1_{it}(p)\right)] \frac{\int_{l_{-i}} \gamma\left(D_t(p) - \hat{S}^1_{it}(p)\right) \cdot \Delta(l_{-i})}{-\int_{l_{-i}} \gamma\left(D_t(p) - \hat{S}^1_{it}(p)\right) D'_t(p)\Delta(l_{-i})}$.
This additive separability property is valuable because it implies that type-0 rivals' private information about their contract positions does not affect a firm's residual demand slope, as we show below.

Type-k Bidding for $k > 1$:

For a type-2 bidder, the procedure to derive optimal bids is exactly the same, with one difference. Rival firms j are now either type-0 *or* type-1 with additively separable bids. That is, for a firm bidding $\hat{S}^2_{it}(p)$

$$
\begin{aligned}
H_{it}(p, \hat{S}^2_{it}(p); k_i = 2, QC_{it}) = & \int_{\boldsymbol{QC_{-it}} \times \boldsymbol{l_{-i}} \times \varepsilon_t} 1(\sum_{j \neq i} S^{l_j}_{jt}(p, QC_{jt}) + \hat{S}^2_{it}(p) \geq \\
& D_t(p) + \varepsilon_t) dF(\boldsymbol{QC_{-it}}, \boldsymbol{l_{-i}}, \varepsilon_t | \hat{S}^2_{it}(p), k_i = 2, QC_{it}) \\
= & \int_{\boldsymbol{QC_{-it}} \times \boldsymbol{l_{-i}} \times \varepsilon_t} 1(\sum_{j \neq i} QC_{jt} + \sum_{j \neq i} \alpha^{l_j}_{jt}(p) + \hat{S}^2_{it}(p) \geq \\
& D_t(p) + \varepsilon_t) dF(\boldsymbol{QC_{-it}}, \boldsymbol{l_{-i}}, \varepsilon_t | \hat{S}^2_{it}(p), k_i = 2, QC_{it}) \\
= & \int_{\boldsymbol{QC_{-it}} \times \boldsymbol{l_{-i}} \times \varepsilon_t} 1(\theta_{it} \geq D_t(p) - \sum_{j \neq i} \alpha^{l_j}_{jt}(p) - \hat{S}^2_{it}(p)) \qquad (6) \\
& dF(\boldsymbol{QC_{-it}}, \boldsymbol{l_{-i}}, \varepsilon_t | \hat{S}^2_{it}(p), k_i = 2, QC_{it})
\end{aligned}
$$

where, as before, $\theta_{it} \equiv \sum_{j \neq i} QC_{jt} - \varepsilon_t$, but $l_j \in \{0, 1\}$.

In this way, we can write H_{it} just as before but taking into account that θ_{it} corresponds to the difference between the sum of contract position by rivals and ε_t.
Taking derivatives of $H(\cdot)$ to find H_S and H_p and plugging into to solve for $\frac{H_S}{H_p}$:

$$
\frac{H_s\left(p, S^*_{it}(p); k_i, QC_{it}\right)}{H_p\left(p, S^*_{it}(p); k_i, QC_{it}\right)} = \frac{\int_{l_{-i}} \gamma\left(D_t(p) - \sum_{j \neq i} \alpha^{l_j}_{jt}(p) - \hat{S}^2_{it}(p)\right) \cdot \Delta(l_{-i})}{-\int_{l_{-i}} \gamma\left(D_t(p) - \sum_{j \neq i} \alpha^{l_j}_{jt}(p) - \hat{S}^2_{it}(p)\right) D'_t(p) \Delta(l_{-i})}.
$$

Therefore, when solving for any type-k bidder for $k > 0$, we use this iterative procedure that relies on the assumption that type-0 bidders submit perfectly inelastic bid functions.

Next, we make two assumptions about and $\Delta_i(\cdot)$ and $\Gamma_i(\cdot)$. For $\Delta_i(\cdot)$, we adopt the Poisson assumption from Camerer, Ho, and Chong [2004]. And we assume that Γ_i is a uniform distribution (we can use an alternative distribution, such as Normal, though it increases the computational burden as one needs to solve the first-order condition by successive approximations.)

Assumption 2. $\Delta(\cdot)$ *is an independent multivariate Poisson distribution truncated at* $k-1$*, as given by Poisson Cognitive Hierarchy model.*

Assumption 3. $\Gamma_i(\cdot)$ *is a uniform distribution.*

We are now prepared to characterize bid functions for each firm type in a manner so that we can use realized data from each auction to characterize ex ante bids submitted by each firm-type under the CH model.

For a type-1 bidder, under these assumptions, the first order-condition can be written as:

$$p - C'_{it}\left(\hat{S}^k_{it}(p)\right) = \frac{1}{-D'_t(p)} * \left[\hat{S}^k_{it}(p) - QC_{it}\right] = \frac{1}{-RD'_t(p)} * \left[\hat{S}^k_{it}(p) - QC_{it}\right],$$

where the second equality follows from the fact that for $RD(p) = D(p) + \varepsilon - \sum_{j\neq i} S_{jt}(p) = D(p) + \varepsilon - \sum_{j\neq i} QC_{jt}$. Hence, $RD'(p) = D'(p)$ for all p.

It is computationally straightforward to solve for the $\hat{S}^k_{it}(p)$ that solves the above equation. This yields a straightforward method to calculate firm i's best-response bid function for any type k. To see this, note that the equation above is just the familiar "inverse elasticity pricing rule". Firm markups of bid over marginal cost are inversely proportional to their residual demand elasticity. Each component of the residual demand function can be iteratively solved for, using our data and Assumptions 1-3.

6 Estimation and Results

6.1 Identification

Firms' beliefs about rival types determine bidding strategies. For this reason, it is important to discuss the process of how beliefs are formed before we turn to estimation. At the same time, because discussing how beliefs are formed is equivalent to discussing identification of the parameters of interest, this section will informally discuss how our data identify the relationship between firm characteristics and strategic type.

Under CH, all firms best-respond given their beliefs about their rivals' bidding behavior. If a firm is deviating from (realized) best-response in an auction, the model provides beliefs about rival behavior that rationalize the observed bid as a best-response to those beliefs. Rivals' characteristics are informative about the distribution of rivals' types, so beliefs about rivals' types are modeled as a function of rivals' observable characteristics. The CH model proposes that this relationship is given by a truncated Poisson distribution with a parameter τ, which is a parameterized function of firms' characteristics (i.e., $\tau_i = \exp(X_i'\boldsymbol{\gamma})$).

Consider a given parameterized relationship between type and a firm characteristic such as size. A firm observes each rival's size and forms beliefs about the rival's type distribution based on that relationship. The firm can calculate each rival's bid function under those beliefs. For example, a firm that is type-5 can use the relationship between size and type to compute how rivals will bid; this yields a residual demand function to which the type-5 firm best-responds. As analysts, we can perform the same calculation and compute the best-response, not only for a type-5 firm but for any type from $k = 0...K$. Then we can compare the firm's actual bid to each of the type-k bids. This calculation is based on a particular parameterized relationship between size and type. Our estimation strategy is to search for the parameter relating size to type that minimizes a metric of distance between observed and CH-computed bids, as we describe in detail in subsection 6.2.

We illustrate the intuition of identification with the following example. Suppose that X_i is firm size so that $\gamma > 0$ implies that larger firms are higher types (stochastically) and $\gamma < 0$ implies that larger firms are lower types. Bids by firm i help identify γ in the following manner. Under CH the observed bids are best-responses given beliefs about the residual demand function that firm i faces. For simplicity, suppose firm i faces a large and a small firm. If $\gamma > 0$, the large rival firm offers substantial capacity into the market – both because $\gamma > 0$ implies the firm is a higher type and thus offers more of its capacity into the market and because the large firm has more capacity to begin with. This will cause firm i to believe it faces a relatively elastic residual demand. In contrast, if $\gamma < 0$, firm i believes that the large rival will not offer much capacity into the market, but rather that the small rival will do so. Hence, the residual demand faced by firm i in this case is steeper–less elastic–than the one it faces when $\gamma > 0$, because the only rival firm with substantial capacity is a low type and low type firms offer less capacity. Our bid data identify γ by estimating whether firm i's bid is more consistent with best responding to the relatively elastic ($\gamma > 0$) or relatively inelastic residual demand ($\gamma < 0$). Because we have bid data for 12 firms across 99 auctions, our data provide rich variation to identify the relationship between firm size and the level of strategic thinking.

6.1.1 Implications of the Level-0 Assumption

It is important to understand the role that the level-0 assumption plays in our estimates of strategic behavior and market efficiency. Given our level-0 assumption, lower type firms bid "closer to vertical" and limit the amount of generation capacity bid into the market. To see this, consider how beliefs about rivals affect bids. A level-1 firm believes that all CH rivals submit vertical bids so that its residual demand is relatively steep (recall, there is an unmodeled fringe that generates some slope to residual demand). Thus, a level-1 firm submits a steep, but not vertical, bid function. A level-2 firm believes that some of its rivals are level-0 and some are level-1, and because level-1 firms submit bids with some slope, the level-2 firm believes its residual demand is flatter than does a level-1 firm, so the best response bid is flatter. In general, the higher a firm's type, the more that it believes its rivals are higher type and thus its residual demand is flatter. Therefore, higher type firms bid more capacity into the market and thus bid "more competitively". As a result, increases in type will tend to increase efficiency.

As discussed in Section 4.2, our assumption about the bidding behavior of level-0 players is supported by the behavior that is observed in the data. Nonetheless, it is important to discuss how alternative assumptions would impact our estimation. To organize our discussion, we divide alternative assumptions for level-0 behavior in two groups. Consider Figure 3 which depicts best response bidding for auctions when balancing demand is positive. One category of level-0 behavior yields low types bidding vertically and higher types submitting flatter bids, or loosely put, higher types approaching best-response bidding "from the left". Another category yields low types bidding relatively flat (e.g. marginal cost) and higher types submitting steeper bids, or higher types approaching best-response bidding "from the right". As we discussed in Appendix F, our data seems to refute marginal cost bidding, though we will still consider the implications of this assumption.

Consider, first, assumptions in which as firm type increases, bids get flatter and approximate best-responses from the left. Our assumption about level-0 behavior is one of these. An alternative, also non-strategic, is to assume that level-0 behavior is that of a monopoly. In this case, level-0 players would best respond to the residual demand function that results from subtracting supply from the fringe from balancing demand. Higher type bidders would simply best-respond to level-0 players that assume they are monopolies. This assumption is similar to the one that we make, but it assumes that level-0 players bid with positive slope even for relatively low prices and small quantities, while our assumption is that they bid vertically on their contract positions. Because in our data level-0 bidders are effectively not providing power to the grid, and bid vertically for most prices, we believe our assumption reflects what we observe in a better way.

Consider now alternatives in which as type increases, bids approach best-responses from the right. Some of these alternatives include level-0 players that bid their marginal cost, or that bid a constant or proportional markup over marginal costs (i.e., a rule of thumb). These assumptions are also non-strategic but they result in higher-type rivals bidding steeper than level-0 players and bids approaching best-responses from the right. Because these alternatives result in predicted bids far from those we observe in the data, any assumption that leads to bids approaching best-responses from the right will result in our model performing worse. Because bids would converge to best-responses from the right, the model would never predict bids that are in the neighborhood of those that we observe.

In summary, although one could make alternative assumptions about level-0 behavior, our assumption is driven by what we observe in the data and our knowledge of how these firms determine their bidding strategies.

6.2 Details on Estimation

Estimation follows a minimum-distance approach. Critical to this approach is τ_i, a scalar that provides information about firm i's type. We assume that $\tau_i = \exp(X_i'\boldsymbol{\gamma})$ and, because X_i is public information, so is τ_i. Each firm i observes $\boldsymbol{\tau}_{-i}$, the vector of τ's of its rivals. Also, each firm i has private information about its own type. Assume firm i is type $k \in \{0, \ldots, K\}$. If $k = 0$, then, by Assumption 1 above, firm i would submit a vertical bid on its own contract position, regardless of its rivals. For all $k > 0$, firms have beliefs about its rivals' types. Specifically, by Assumption 2, these beliefs are assumed to follow a Poisson distribution truncated at k, meaning that firm i believes all its rivals to be type $k-1$ or less. The specific probability associated with each type varies according to each rivals' τ.

Then, we can use the model to compute, for each firm i and auction t, the optimal bid function given i's type and its beliefs over its rivals' types. Note, however, that in a specific auction, even if two bidders are of the same type, differences in marginal costs will generate differences in predicted bids.

Once firm i has computed what it expects its rivals to do for each possible type, it maximizes expected profits according to its beliefs about its rivals' types. This results in a bid function, conditional on i's type. However, types are unknown to the econometrician. For this reason, we proceed as follows. First, we compute bid functions over a grid of price points. Denote a price point by p. Second, we compute the square of the scaled difference between the bid data for bidder i in auction t at price point p and the bid predicted by the model for i when we assume i to be of type k. Scaling is done using the quantity-difference between the predicted bid for each firm for types K and 0. Third, we sum these differences across price points for bidder i in auction t, weighting price points by a triangular distribution centered at the market clearing price. Fourth, as all of this is done conditional on bidder i being type k, we weight each of these sums by the probability that the firm is each type. This probability is modeled as following a Poisson distribution truncated at the number of possible types considered in estimation (level-0 and 20 levels of strategic sophistication) and not truncated at each firms' beliefs. We use each firms' τ to compute this probability. Finally, we add over firms and auctions.

In this context, our estimate $\hat{\boldsymbol{\gamma}}$ is

$$\hat{\boldsymbol{\gamma}} = \underset{\gamma}{\operatorname{argmin}} \sum_i \sum_t \left[\sum_k \Big[\sum_p \left(\frac{b_{it}^{\text{data}}(p) - b_{it}^{\text{model}}(p|k)}{b_{it}^{\text{model}}(p|K) - b_{it}^{\text{model}}(p|0)} \right)^2 \times \mathbb{P}(p) \Big] \mathbb{P}_i(k|\ |K|, \hat{\boldsymbol{\gamma}}) \right],$$

where $\mathbb{P}(p)$ corresponds to the probability of observing a price point p as given by the triangular distribution and $\mathbb{P}_i(k|\ |K|, \hat{\boldsymbol{\gamma}})$ corresponds to the probability of bidder i being type k, conditional on there being $|K|$ possible types and $\hat{\gamma}$ being the estimated parameters. To ensure that we find the global minimum, we run estimation starting from 50 sets of random initial points.

6.3 Results: Estimated Parameters

Results are reported in Table 1. We estimate seven specifications that differ in the observable characteristics (X_i) of the firms that affect firm τ_i. In our baseline specification column, type is determined by firm size. Therefore in column (1) of Table 1, X_i includes a constant and firm size. As a metric of size in the balancing market, we seek a metric of the firm's potential stakes in the balancing market that is exogenous to its realized bidding behavior. We compute the quantity of sales if the firm were to best-respond, averaged across all auctions. This is positively correlated with installed generation capacity.

As shown in column (1), we find that larger firms are higher types.[18] In order to interpret the positive coefficient on $Size$, we calculate the implied distribution of firm type for each of the 12 firms that we include in the Cognitive Hierarchy model. Figure 5 plots the estimated type distribution for each firm. Consider the smallest firm with a size that is 11% of the size of the largest firm – the pdf farthest to the left in the figure. We estimate that the smallest firm has about a 50% chance of being type-0, about a 35% chance of being type-1, about a 10% chance of being type-2, and is higher than type-2 with very low probability. Each of the other pdfs in the figure show the estimated type distribution for the other firms, with the larger firms having probability distributions further to the right. Thus, we find that larger firms are likely to be higher type, and importantly, there is substantial heterogeneity across these firms in the estimated types. This means that only the largest firms actually engage in behavior that is similar to what a Bayesian Nash model would predict.

Our second specification allows for non-linearity in how size affects τ_i by adding size squared. The implied distribution of types is qualitatively very similar to our linear specification.

Next we explore if the organizational structure of the firm is associated with higher type. As discussed in section 2, some firms are merchant firms that have never been part of a regulated utility while other firms are either municipal utilities or generation firms that were formerly integrated into an investor-owned utility. It is possible that organizational structure

[18] We expect the constant to be negative in order to rationalize level-0 players, as a positive constant would decrease the probability of observing a level-0 player significantly. Note, however, that this is not required by the CH model as one need not observe level-0 behavior in the data. However, as we have specified level-0 behavior according to what we observe in our data, a negative constant shows that the type of level-0 behavior that we have assumed is not uncommon.

could impact the nature of the trading operations that a firm establishes. In column (3) of Table 1, we test whether merchant firms tend to be higher types. However, we find that if anything, merchants are lower types than former utilities and municipal utilities. However, the role of organizational structure is substantially smaller than the role of firm size.

In specifications (4)–(6), we investigate whether the personnel hired to run firm bidding operations is related to firm type. In order to assess the role of personnel, we use LinkedIn and other publicly available online data sources to make the best guess of the manager(s) who were responsible for each firm's power marketing operations for Texas in 2003. In some cases, job titles were sufficiently clear to identify the power marketing manager, and in other cases we were only able to identify personnel who were involved in firm wholesale power operations. Therefore, the data used for this specification may not be as precise as the data on bids and costs. Nevertheless, this should provide suggestive evidence on the role of power trading personnel. For each firm manager whom we identify, we collect information on job title and education. For each firm, we create an $AAU\,University$ dummy variable to indicate whether any of the firm's power marketing personnel have an academic degree from a university that belongs to the American Association of Universities (AAU).[19] Five of the twelve firms have personnel who graduated from an AAU university. We also create a dummy variable for whether any personnel have a degree in either Economics, Business, or Finance. Seven of the twelve firms have personnel with a degree in economics, finance, or an MBA while the most popular other type of degree is in engineering. We estimate our benchmark specification using $Size$ and add dummy variables for University type or Degree type.

Our specifications with personnel are reported in columns (4)-(6) of Table 1. We find that when we include AAU in addition to firm size in column (4), the AAU coefficient is positive and the coefficient of size is slightly lower. We find similar patterns when we include a dummy for degree in Economics, Finance or Business – the coefficient is positive and the coefficient of size falls significantly. These results suggest one mechanism through which size may affect the level of strategic sophistication. Discussions with industry personnel suggest that the dollar stakes of each firm are likely sufficient to cover the costs of establishing a basic trading operation. But only larger firms may have sufficient dollar stakes to hire high quality and well-trained traders and to build sophisticated trading operations. This is consistent with our finding that once we control for whether personnel are trained in economics, finance or have an MBA, that the relationship between size and type is weaker.

Finally, we explore the possibility that firms may learn over time. To capture this, we

[19] The AAU includes 62 private and public research universities in the U.S. and Canada. A list of the AAU universities can be found at: http://www.aau.edu/about/default.aspx?id=16710.

specify τ as a function of firm size and a linear time trend. The results are reported in the last column of Table 1 and show that the estimated coefficients for the constant and size are remarkably similar to those in column (1). Nonetheless, we find a positive and significant coefficient on the time trend, which suggests that firm types do change over time. However, the amount of learning is economically very small. To see this, Figure 6 plots the estimated probability distributions over types for both the smallest and largest firm in the market. The distribution plotted in solid yellow depicts the type distribution in the first week of our sample, i.e. the first week of the second year of the market. The distribution plotted with green markers depicts the distribution six months later in the last week of the sample. For each firm, the distributions are nearly identical which suggests that in practice learning is minimal and has no impact on the estimated probability distributions over types.

FIGURE 5: Estimated Distributions of Types for the 12 Firms (*Size* Specification)

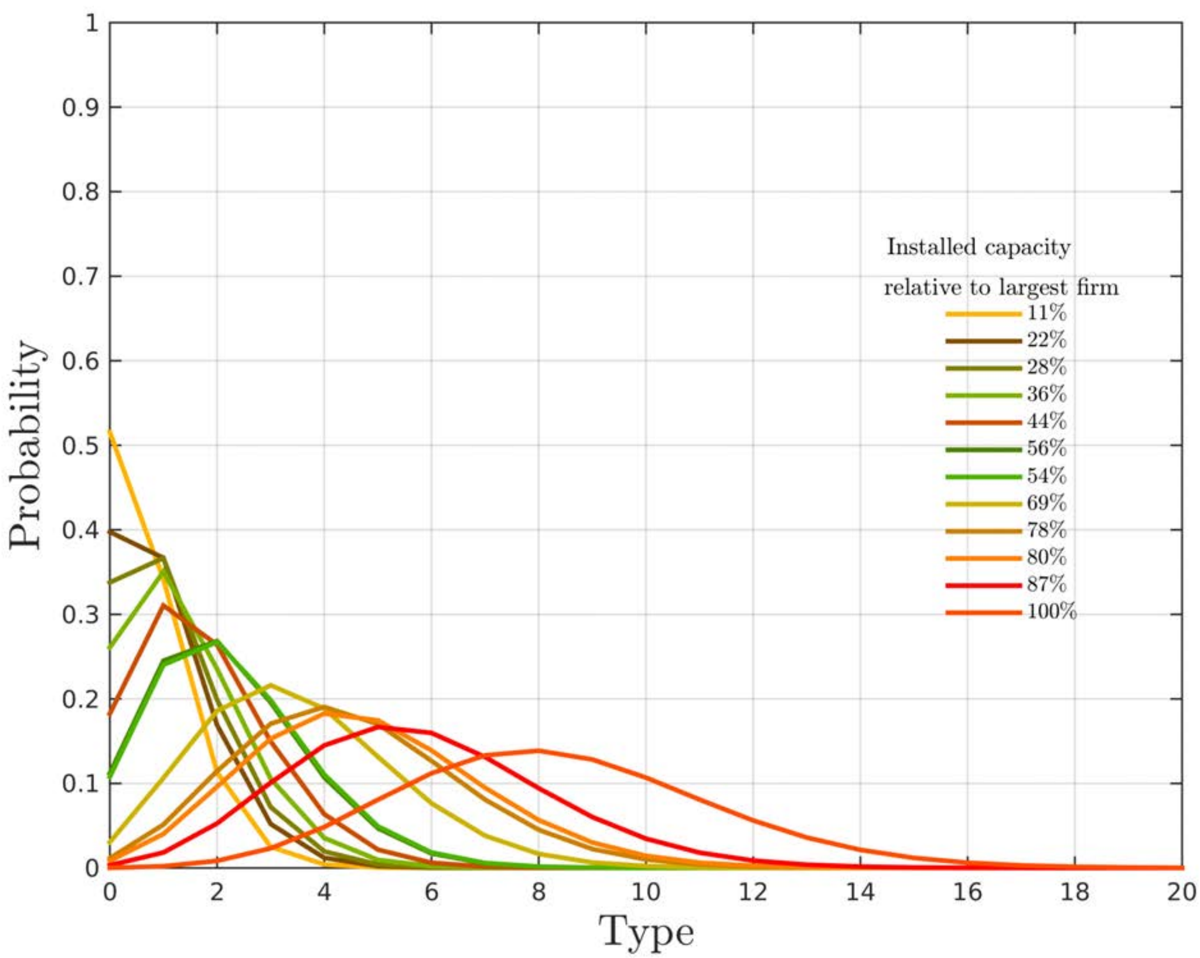
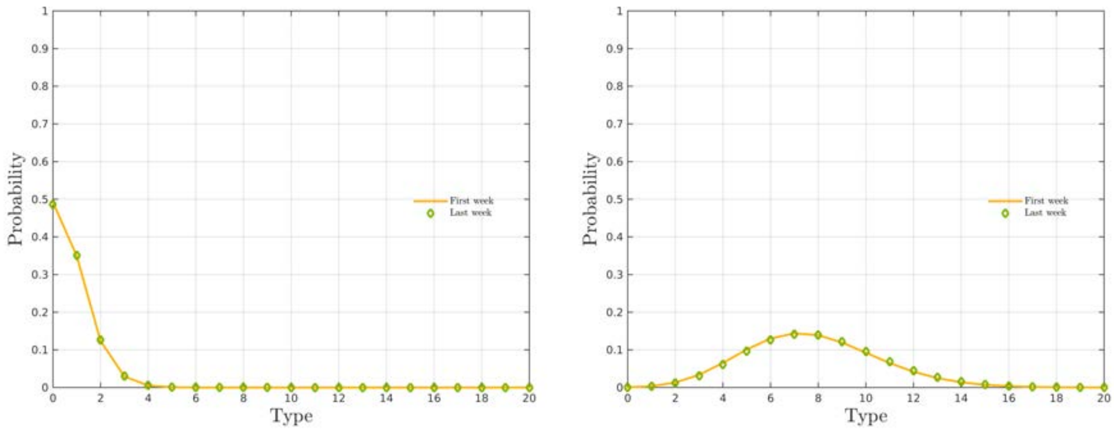


Table 1: Structural Model: Estimated Parameters of Type Function

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Constant	-0.726	-0.196	-3.395	-0.749	-3.493	-0.691	-0.675
	(0.133)	(0.158)	(0.249)	(0.159)	(0.508)	(0.147)	(0.085)
Size	14.594	-1.163	25.789	13.619	3.090	11.933	13.776
	(1.369)	(2.892)	(3.575)	(1.865)	(0.671)	(1.409)	(0.826)
Size^2		86.191					
		(13.643)					
Merchant			-1.562				
			(0.403)				
AAU University				0.376			
				(0.069)			
Degree in Economics,					5.626		
Business or Finance					(1.935)		
Economics degree						1.633	
						(0.340)	
Time Trend							0.051
							(0.014)
Obj. Fn. / #Auctions	208.512	208.354	208.526	208.485	206.386	208.308	208.520

Note: Bootstrapped standard errors using 20 samples.



(a) Smallest Firm

(b) Largest Firm

FIGURE 6: Estimated Distributions of Types and Learning (*Size* and time trend specification)

6.4 Model Fit

In this section we show that the cognitive hierarchy model – with our parameterization of size determining beliefs – fits the actual bidding behavior quite well. To assess the model fit, we use the specification where *Size* determines firm type and subsequent bidding behavior (column (1) of Table 1). We compare the fit of the CH model to a model in which firms best-respond to their rivals' actual bid data; this best-response model is essentially a model in which firms individually best-respond and have consistent beliefs about their rivals' behavior. Table 2 reports results from a regression at the firm-auction level that predicts realized profits under actual bidding with profits under either the CH model (column 1) or best-response model (column 2). The CH model fits the actual data better than the best-response model. The coefficient of modeled profits is much closer to unity for the CH model ($\hat{\beta}^{CH} = .803$ vs. $\hat{\beta}^{BR} = .428$) and the CH model explains more of the variation in firm profits ($R^2_{CH} = 0.67$ vs. $R^2_{BR} = 0.49$). Moreover, we view the fit of the CH model as strong given that we are using only a single covariate – firm size – to explain the heterogeneity in behavior across firms.

Table 2: Model Fit: Comparison of Cognitive Hierarchy Model
to Unilateral Best-Response Model

	(1) CH Model	(2) Best-Response
Dependent Variable: Profits from Actual Bids		
Profits under Cognitive Hierarchy	0.803	–
	(0.069)	–
Profits under Best-Response	–	0.428
	–	(0.044)
Constant	-328.17	-241.74
	(141.976)	(120.722)
Observations	1058	1058
R^2	0.67	0.49

Note: This table reports results from a regression of observed profits from actual bidding behavior on either firm profits as predicted by the Cognitive Hierarchy model (column 1) or firm profits that would be achieved from a model of unilateral best-response to rival bids (column 2). An observation is a firm-auction. Standard errors clustered at the firm-level are reported in parentheses.

7 A "Reduced Form" Test of Strategic vs. Non-strategic behavior: Evidence from a Nuclear Plant Outage

In this section, we show additional data-driven support for the CH model and our results. In general, in the CH model, level-0 behavior captures non-strategic agents. One type of non-strategic behavior is an agent who does not respond to changes in the (common knowledge) cost structure of its competitors, but may respond to changes in its own costs. Note that this definition of non-strategic behavior encompasses a large array of behavioral patterns that have been considered as level-0 behavior in the literature. Including vertical bidding, which is our definition of level-0 behavior, bidding marginal cost (or bidding truthfully, as in Crawford and Iriberri [2007], Gillen [2010], An [forthcoming]), bidding a random number that is independent of competitors' costs (again, as in Crawford and Iriberri [2007], Gillen [2010], An [forthcoming]), or behaving as if one is a monopolist regardless of one's competitors (as in Goldfarb and Xiao [2011]) is consistent with this definition of non-strategic behavior.

We exploit a large discrete change in competitor costs – an outage at a large nuclear plant – and show that large firms changed their bids in response to this shock while small firms did not. Given that the small firms in our sample submitted nearly vertical bids (see Figure 4), this evidence suggests that vertical bidding is an appropriate level-0 assumption in this setting.

We use a two month outage at a nuclear plant to test whether firms change bids in response to competitor cost shocks. In the middle of our sample period, one large nuclear generator in Texas went off-line for approximately two months. This event reduced nuclear output by about 2300 megawatts as shown in Figure 7. As a result, total demand for power intersected aggregate system marginal cost at a steeper point on the marginal cost function. Firms that are behaving strategically will recognize that this publicly observable cost shock is likely to make their residual demand in the balancing market less elastic. Therefore, all else equal, we expect to see strategic firms responding to the nuclear outage by submitting steeper bids.

We test this hypothesis and find that large firms respond to both competitor costs and own costs, but that small firms only respond to own costs. To do so, we analyze the slope of firms' bid functions in the months surrounding the nuclear outage and test whether bids were steeper during the outage. That is, we create a panel of firm bids across each auction in our sample. The dependent variable is the slope ($\frac{\partial S_{it}}{\partial p}$) of firm i's bid in auction t.[20] $Outage_t$ is an indicator equal to one if the auction occurred during the outage period of October 2,

[20]We measure slope as the bid linearized plus and minus \$10 around the auction's market-clearing price (\$10 is the standard deviation of the market-clearing price in our sample).

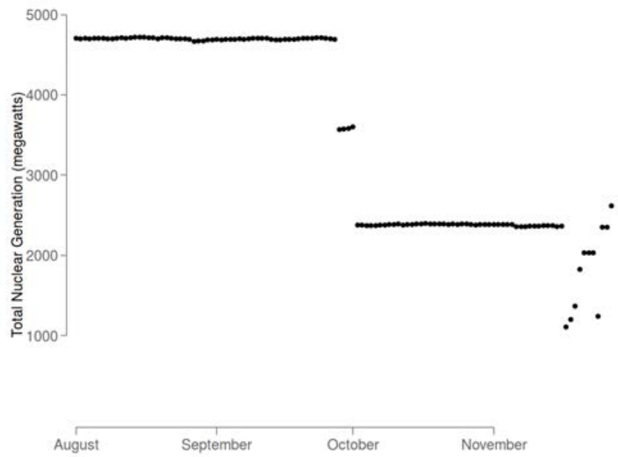


Figure 7: Nuclear Production During Fall 2002 Outage

2002 to November 27, 2002 and equal to zero if the auction occurred prior to the outage.

Results are shown in Table 3. The top panel estimates the effect of the outage separately for the largest three and smallest three firms. In the first two columns, we estimate the slope of bids only as a function of the outage. The largest three firms submit steeper bid functions during the outage but the smallest three firms do not change the slope of their bids during the outage. In the next two columns, we control for changes in own costs by including the slope of the firm's own marginal cost. We find that both small and large firms submit steeper bids when their own marginal costs are steeper. However, only the large firms respond to the nuclear outage by submitting steeper bids. And in the final two columns, we add bidder fixed effects to allow for firms to face different shaped residual demand functions in general. We find robust evidence that while all firms respond to their own cost shocks, only large firms respond to the large rival cost shock induced by the nuclear outage. In the bottom panel of Table 3, we estimate the same models using all 12 firms in the cognitive hierarchy – defining Large (Small) as all firms above (below) the median size firm – and we continue to find that small firms do not respond to the nuclear outage.

Table 3: Evidence of Non-Strategic Bidding: Bidding Response to Nuclear Outage

	Largest Three	**Smallest Three**	**Largest Three**	**Smallest Three**	**Largest Three**	**Smallest Three**
Outage	-47.81*	-0.02	-31.13*	0.17	-19.42*	0.36
	(7.50)	(0.18)	(8.30)	(0.19)	(5.19)	(0.18)
Own MC			0.15*	0.04*	0.26*	0.09*
			(0.05)	(0.01)	(0.04)	(0.01)
Constant	73.42*	1.32*	38.08*	0.65*	-8.61	0.63*
	(7.18)	(0.12)	(12.37)	(0.18)	(8.84)	(0.19)
Bidder Fixed Effects	No	No	No	No	Yes	Yes
N	189	189	189	189	189	189
R^2	0.21	0.00	0.25	0.17	0.59	0.25

	Largest Six	**Smallest Six**	**Largest Six**	**Smallest Six**	**Largest Six**	**Smallest Six**
Outage	-26.27*	-0.64	-9.80*	0.4	-8.40*	-0.03
	(4.69)	(0.42)	(2.92)	(0.38)	(2.05)	(0.25)
Own MC			0.27*	0.18*	0.30*	0.11*
			(0.03)	(0.02)	(0.03)	(0.02)
Constant	40.28*	3.75*	2.82	0.19	-21.13*	0.76*
	(4.49)	(0.32)	(2.41)	(0.37)	(6.55)	(0.21)
Bidder Fixed Effects	No	No	No	No	Yes	Yes
N	378	378	378	378	378	378
R^2	0.09	0.01	0.40	0.31	0.67	0.68

Note: Each column reports estimates from a separate regression of the slope of a firm's bid function on an indicator variable that the auction occurred during the fall 2002 nuclear outage. An observation is a firm-auction. The dependent variable is the slope ($\frac{\partial S_{it}}{\partial p}$) of firm i's bid in auction t where the slope is linearized plus and minus \$10 around the market-clearing price. $OwnMC$ is the slope of the firm's own marginal cost function linearized plus and minus \$10 around the market-clearing price. White standard errors are reported in parentheses. + p<0.05, * p<0.01

8 Counterfactual Analysis: Increasing Strategic Sophistication

Having estimated our model of bidding behavior that allows for heterogeneity in strategic sophistication, we now turn to a key question of this paper: How does the lack of strategic sophistication affect market efficiency? We address this question in two steps.

First, we estimate how exogenous increases in strategic sophistication of specific firms affect market efficiency. We conceptualize an exogenous increase in sophistication as an event such as the firm outsourcing bidding to a consulting firm or hiring more qualified employees to operate the trading floor. This counterfactual provides us a way to isolate the impact of increasing sophistication without changes in market structure that could increase market power. Market efficiency will improve if more sophisticated bidding leads to more elastic bids and causes a firm's low-cost generation to be offered into the market.

Second, we estimate the effect on market efficiency of increases in sophistication that are caused by large high-type firms merging with small low-type firms. Suppose that when the two firms merge, the trading operation of the larger firm takes over all bidding operations, so the merged firm enjoys the sophistication level of the large firm. The effect of such mergers is ex-ante ambiguous. On one hand, more sophisticated bidding of the small firm's generation assets is likely to cause lower-cost plants to be dispatched. On the other hand, mergers increase concentration and increase market power. We estimate which effect dominates for mergers between firms of different sizes.

Is is important to note that only a subset of firms in the market are included in the cognitive hierarchy; the remaining firms are part of a unmodeled fringe. For this reason, we compute inefficiencies as the difference between the generating cost implied by the estimated model and our efficient benchmark in which all firms included in the CH bid marginal costs, while the rest of the firms bid according to their bids in the data. We present the counterfactuals for our benchmark specifications where firm type is a function of size.

8.1 Exogenous Increase

An exogenous increase in sophistication will impact efficiency through two channels. First, if firms are induced to be higher-type thinkers, the bid functions will become "flatter" because beliefs that rivals are higher types imply that the firms believe their residual demand to be more elastic. As a result, more low-cost generation capacity will be offered into the balancing market and dispatch costs will fall. This is the direct effect of increasing sophistication. However, there is also a second indirect impact on efficiency. Suppose that the increase in sophistication is publicly observable (e.g. rivals observe that the firm hires a bidding consultant). Then rival firms recognize the increase in sophistication (even though their beliefs will continue to be wrong) and also submit more elastic bids.

We simulate the effects of increasing the strategic sophistication of firms of different sizes. It is a priori ambiguous which types of firms would most improve market efficiency by increasing sophistication. Small firms have smaller amounts of generation capacity to offer into the market, but it is the small firms that we find are bidding with the least sophistication (i.e. "too steep").

Table 4 presents results using the *Size* parameterization of τ_i.[21] The first row reports estimated changes in total market production costs when the smallest six firms are given the sophistication level of the median-sized firm. We find that when the increase in sophistication is publicly observed so that own-firm and rival firm bids adjust, then production costs fall by 9.43 percent during periods of positive balancing demand. Most of the impact occurs through the channel of changing the firm's own bids – the production costs fall by 7.70 percent when rivals do not react to the change in sophistication. In the second row, we model the impact of increasing the sophistication of all firms that are larger than the median-sized firm to the sophistication level of the largest firm. Production costs fall by 4.49 percent which is only about half of the efficiency improvement of increasing sophistication of the smaller firms. Finally, in the third row we focus only on the three smallest firms and find that much of the room for efficiency improvements lies in increasing sophistication of very small firms. Although the small firms have less generation capacity to add to the market, the largest scope for efficiency improvement lies in the small firms that withhold so much capacity due to low sophistication.

Finally, we estimate that there are diminishing marginal private returns to increases in sophistication. As shown in Appendix G, when firms exogenously increase sophistication to the level of larger firms (but maintain their existing production capacity), the average increases in profits diminish with size.

8.2 Endogenous Increase: Mergers

We now turn to studying how mergers affect efficiency. As mentioned above, we focus on potential mergers that do not generate cost synergies but do increase concentration. Specifically we imagine a merger than changes the bidding into the balancing market but does not change any other decisions of the merging firms such as forward contracting decisions. In reality, a complete merger of power marketing operations – including forward contracting –

[21] Estimates using the $Size^2$ parameterization are quite similar.

Table 4: Exogenous increase in sophistication: Change in Production Costs

	INC side		DEC side	
Counterfactual	Public	Private	Public	Private
Small firms to median	-9.43%	-7.70%	-16.49%	-13.91%
Above median firms to highest	-4.49%	-3.72%	-7.96%	-6.50%
Three smallest to median	-6.97%	-6.04%	-10.70%	-9.94%

This table reports the changes in total production costs when different subsets of firms are modeled to increase sophistication (τ_i). These counterfactual calculations use parameter estimates from the first specification in Table 1. *Small firms to median* simulates production when the smallest six firms are given the sophistication level of the median firm. *Above median firms to highest* simulates production when all firms above the median-sized firm are given the sophistication of the largest firm. *Three smallest to median* simulates production when the three smallest firms are given the sophistication level of the median firm. Counterfactuals are calculated separately for auctions with positive balancing demand (INC side) and negative balancing demand (DEC side). *Public* indicates that the change in sophistication is observed by rival firms so that rival bids change due to increase in sophistication. *Private* indicates that the change in sophistication is not observed by rivals so that only the bids of the modeled firm change.

could change the positions at which bidders find themselves when balancing market bidding takes place. We are not in a position to simulate counterfactuals for full power marketing integration. However, we can simulate the effect of a merger of balancing market bidding in order to illustrate the countervailing effects of increasing sophistication in the balancing market.[22]

In this context, mergers have countervailing effects on efficiency. Mergers that increase sophistication of one of the merging firms increase efficiency through the same mechanisms as the exogenous mergers studied above. This reallocates production from high-cost, high-type firms to low-cost, low-type firms that have previously priced themselves out of the market. However, mergers create a countervailing effect – increasing market concentration will create additional market power that leads to higher production costs.

Our simulations show that mergers that increase sophistication can increase efficiency as long as the merging firms are not too large. The first row of Table 5 shows that a merger between the smallest and largest firms reduces production costs, despite the increase on concentration. However, for mergers involving medium-sized and larger firms, the market power effect dominates and mergers increase production costs.

[22]In order to model a merger between two firms, we incorporate the two firms' individual marginal costs of production and day-ahead schedules. We horizontally add the marginal cost functions and the day-ahead schedules to compute the marginal cost of supplying power to the grid for the merged firm, relative to the aggregate day-ahead schedule.

Table 5: Increasing Sophistication via Mergers: Change in Production Costs

Merging Firms	INC side	DEC side
Smallest and largest firms	-3.2%	-15.4%
Median and largest firms	+9.0%	+21.9%
Two largest firms	+17.3%	+56.3%

This table reports the changes in total production costs when different pairs of firms merge. These counterfactual calculations use parameter estimates from the first specification in Table 1. Counterfactuals are calculated separately for auctions with positive balancing demand (INC side) and negative balancing demand (DEC side).

9 Conclusions

Models of strategic equilibrium form the foundation of many studies in Industrial Organization that investigate market efficiency in oligopoly settings. However, there is some evidence suggesting that the application of such strategic equilibrium models to all settings has to be done with caution, as in some settings observed behavior may depart significantly and persistently from what equilibrium models predict. These departures from Nash equilibrium behavior may have significant implications for efficiency.

In this paper we study bidding in the Texas electricity market, a market in which bidding by some firms departs significantly from what Bayesian Nash models predict, while bidding by other firms closely resembles these predictions. We use this setting, as well as a unique dataset containing information on bids and marginal costs, to embed a Cognitive Hierarchy model into a structural model of bidding behavior. Our unique dataset, in addition to our model, allows us to identify and estimate heterogeneity in levels of strategic sophistication across electricity generators. Our results show that while small firms appear to behave as if they are boundedly rational in a Cognitive Hierarchy sense, large firms behave closely to what a Bayesian Nash model would predict. We then use the estimated levels of strategic sophistication to study how increasing the sophistication of low-type firms, either exogenously or through mergers with higher-type firms, may affect efficiency. Our results show that not only can exogenously increasing sophistication increase efficiency significantly, but that mergers that do not generate cost synergies but increase concentration may also increase efficiency.

References

Marina Agranov, Guillaume Fréchette, Thomas Palfrey, and Emanuel Vespa. Static and Dynamic Underinvestment: An Experimental Investigation. *Journal of Public Economics*, 143:125–141, 2016. ISSN 0047-2727.

Yonghong An. Identification of First-Price Auctions with Non-Equilibrium Beliefs: A Measurement Error Approach. *Journal of Econometrics*, forthcoming.

Ayala Arad and Ariel Rubinstein. The 11-20 money request game: Evaluating the upper bound of k-level reasoning. *American Economic Review*, 102(7):3561–3573, 2012.

Nicholas Bloom and John Van Reenen. Measuring and Explaining Management Practices Across Firms and Countries. *Quarterly Journal of Economics*, CXXII(4):1351–1408, November 2007.

Severin Borenstein, James B. Bushnell, and Frank A. Wolak. Measuring Market Inefficiencies in California's Restructured Wholesale Electricity Market. *American Economic Review*, 92(5):1376–1405, December 2002.

James Bushnell, Erin Mansur, and Celeste Saravia. Vertical Arrangements, Market Structure, and Competition: An Analysis of Restructured U.S. Electricity Markets. *American Economic Review*, 98(1):237–266, 2008.

Colin F. Camerer. *Behavioral Game Theory: Experiments in Strategic Interaction*. Princeton University Press, 2003.

Colin F. Camerer, Teck-Hua Ho, and Juin-Kuan Chong. A Cognitive Hierarchy Model of Games. *Quarterly Journal of Economics*, 119(3):861–898, 2004.

Miguel A. Costa-Gomes, Vincent P. Crawford, and Bruno Broseta. Cognition and Behavior in Normal-form Fames: an Experimental Study. *Econometrica*, 69(5):1193–1235, 2001.

Vincent P. Crawford and Nagore Iriberri. Fatal attraction: Salience, Naivete, and Sophistication in Experimental "Hide-and-Seek" Games. *American Economic Review*, 97(5): 1731–1750, December 2007.

Vincent P. Crawford, Uri Gneezy, and Yuval Rottenstreich. The Power of Focal Points Is Limited: Even Minute Payoff Asymmetry May Yield Large Coordination Failures. *American Economic Review*, 98(4):1443–1458, September 2008.

Ulrich Doraszelski, Gregory Lewis, and Ariel Pakes. Just Starting Out: Learning and Equilibrium in a New Market. University of Pennsylvania Working Paper, 2017.

Natalia Fabra and Mar Reguant. Pass-Through of Emission Costs in Electricity Markets. *American Economic Review*, 104(9):2872–2899, 2014.

Drew Fudenberg, David G. Rand, and Anna Dreber. Slow to Anger and Fast to Forgive: Cooperation in an Uncertain World. *American Economic Review*, 102(2):720–49, April 2012.

Ben Gillen. Identification and Estimation of Level-k Auctions. California Institute of Technology working paper, August 2010.

Avi Goldfarb and Mo Xiao. Who Thinks about the Competition: Managerial Ability and Strategic Entry in US Local Telephone Markets. *American Economic Review*, 101(7): 3130–3161, 2011.

Ali Hortaçsu, Jakub Kastl, et al. Valuing dealers' informational advantage: a study of Canadian Treasury auctions. *Econometrica*, 80(6):2511, 2012.

Ali Hortacsu and David McAdams. Mechanism choice and strategic bidding in divisible good auctions: An empirical analysis of the Turkish treasury auction market. *Journal of Political Economy*, 118(5):833–865, 2010.

Ali Hortacsu and Steven L Puller. Understanding Strategic Bidding in Restructured Electricity Markets: a Case Study of ERCOT. NBER Working Paper 11123, February 2005.

Ali Hortacsu and Steven L. Puller. Understanding Strategic Bidding in Multi-Unit Auctions: A Case Study of the Texas Electricity Spot Market. *RAND Journal of Economics*, 39(1): 86–114, Spring 2008.

Chang-Tai Hsieh and Peter J. Klenow. Misallocation and Manufacturing TFP in China and India. *Quarterly Journal of Economics*, 124(4):1403–48, 2009.

John H Kagel and Alvin E Roth. *The Handbook of Experimental Economics*. Princeton university press Princeton, NJ, 1995.

Boo-Sung Kang and Steven L Puller. The effect of auction format on efficiency and revenue in divisible goods auctions: A test using Korean Treasury Auctions. *The Journal of Industrial Economics*, 56(2):290–332, 2008.

Richard D. McKelvey and Thomas R. Palfrey. Quantal Response Equilibria for Normal Form Games. *Games and Economic Behavior*, 10(1):6–38, July 1995.

Rosemary Nagel. Unraveling in Guessing Games: An Experimental Study. *American Economic Review*, 85(5):1313–1326, December 1995.

Dale O Stahl and Paul W Wilson. On players' models of other players: Theory and experimental evidence. *Games and Economic Behavior*, 10(1):218–254, 1995.

Chad Syverson. Product Substitutability and Productivity Dispersion. *Review of Economics and Statistics*, 86(2):534–50, 2004.

Robert Wilson. Auctions of Shares. *Quarterly Journal of Economics*, 93(4):675–689, 1979.

Frank A. Wolak. Measuring Unilateral Market Power in Wholesale Electricity Markets: The California Market, 1998-2000. *American Economic Review Papers and Proceedings*, 93(2): 425–430, May 2003.

Catherine D. Wolfram. Strategic Bidding in a Multiunit Auction: An Empirical Analysis of Bids to Supply Electricity in England and Wales. *RAND Journal of Economics*, 29(4): 703–725, 1998.

A Appendix: Data Details (reproduced from Hortaçsu and Puller [2008])

Hourly bid schedules by each bidder, or qualified scheduling entity (QSE), are from ERCOT. 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. 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. 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 which comprises 87% of the generation for TXU the QSE in North 2002.

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. For each 15-minute interval, we have data from ERCOT on whether a generating unit is operating, its day-ahead scheduled generation, and its hourly available generating capacity. 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, an assumption that is common in the literature. The ERCOT system is largely separated from other electricity grids in the country so there are virtually no imports.

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. Coal-fired plants in Texas are required to possess federal emission permits for each ton of SO2 emissions. In order to measure average emission rates, we merge hourly net metered generation data from ERCOT with hourly emission data from EPA's CEMS 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.

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 *intra*zonal (or local) congestion is also likely to be rare during these intervals.

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. Therefore, we assume that each operating unit cannot DEC to a level below 20% of its maximum generating capacity.

B Appendix: Profitability of Actual Bids

In this appendix, we report metrics of 'money left-on-the-table' across firms during the first 1.5 years of the market, as in Hortacsu and Puller [2008]. Given our data on costs and bids, for each firm in each auction we calculate producer surplus under two scenarios: (1) best-response bidding and (2) bidding vertically at the contract position which is essentially not participating in the balancing market except to meet contract obligations. Then we calculate realized producer surplus under actual bidding, and we compute the fraction of potential profits relative to non-participation that were achieved by actual bidding. Table 6 reports the average percent of potential profits realized by different firms across the first 1.5 years of the market.[23] A large firm Reliant – which is the large firm depicted in panel (a) of Figure 4 – realized 79% of realized profits. However, all of the other firms realized less than one half of potential profits. The firm-level profits 'on-the-table' average between $1000 and $4000 each hour.

[23] In this table we report profitability only for the 12 firms that we will model in section 6.

Table 6: Firm-Level Profitability in First 1.5 Years of Market

Firm	Percent of Potential Profits Achieved
Reliant	79%
City of Bryan	45%
Tenaska Gateway Partners	41%
TXU	39%
Calpine Corp	37%
Cogen Lyondell Inc	16%
Lamar Power Partners	15%
City of Garland	13%
West Texas Utilities	8%
Central Power and Light	8%
Guadalupe Power Partners	6%
Tenaska Frontier Partners	5%

This reports the percent of potential profits that are achieved with actual bids relative to a benchmark where firms do not participate in the balancing market, i.e. bid vertically at the contract position. The figures represent the firm-level profitability averaged over auctions in the first 1.5 years of the market. Source: Hortacsu and Puller [2008].

C Appendix: Additional Evidence on Learning

In this appendix, we show descriptive evidence that firm patterns of offering small quantities into the balancing market via bids that are 'too steep' is a phenomenon equally prevalent in the first and second year of the market. We test whether firms offer more generation capacity into the market in the second versus the first year of the market, or whether bidding that is 'too steep' is equally present in both years. Specifically, for each firm-auction, we calculate the amount of generation capacity that the firm offers relative to the contract position at the market-clearing price. We call this variable *Participation Quantity*. We define *Participation Quantity*$_{it}$=$|(S_{it}(p_t^{mcp}) - QC_{it})|$, using absolute value to capture bidding behavior for quantities above and below the contract position. A firm bidding vertically at the contract position is measured as *Participation Quantity*=0, but firms bidding with more elasticity have positive measures of *Participation Quantity*. We test whether firms offer additional generation into the market in the second year. Results are shown in Table 7. In all specifications we include firm fixed effects so that we can test if firms participate more in the second year of the market relative to their participation in the first year. Column (1) shows that firms offer *less* generation in the second year, however the point estimate of *Year 2* (-35 MW) is neither economically nor statistically significant. Column (2) conditions on whether balancing demand is positive, and the point estimate is even smaller and not statistically different from zero. Column (3) conditions on the day of week and yields nearly an identical estimate. Finally, column (4) estimates the relationship for only the small firms and finds that these firms offer a very small amount of additional capacity in the second year – 1.52 MW – and this is not statistically different from zero. This persistence of small quantities offered into the market suggests that learning is slow in this market. Formal tests of learning are reported in section 6.3.

Table 7: Offered Quantities into Market in Year 2 vs Year 1

	All Firms	All Firms	All Firms	Small Firms
	(1)	(2)	(3)	(4)
Year 2	-34.76	-15.85	-16.15	1.52
	(42.42)	(34.24)	(34.70)	(2.90)
Firm Fixed Effects	Yes	Yes	Yes	Yes
INC Fixed Effects	No	Yes	Yes	Yes
Day of Week Fixed Effects	No	No	Yes	Yes
Observations	2264	2264	2264	1029
R^2	0.01	0.03	0.04	0.09

$^+$p<0.05; *p<0.01 The dependent variable *Participation Quantity*$_{it}$ is the megawatt quantity of output bid at the market-clearing price relative to the firm's contract position in auction t, i.e. $|S_{it}(p^{mcp}) - QC_{it}|$. The sample period is the first 1.5 years of the market and *Year 2* is a dummy variable for the second year. Standard errors clustered at the firm-level are reported in parentheses.

D Appendix: Evidence that Bidding Rules Do Not Bias Best-Response Bids as a Benchmark for Expected Profit Maximization

In this appendix, we test if firms would increase profits by following a simple trading rule that follows all bidding rules and places no restrictions on how uncertainty affects residual demand. The trading rule takes advantage of an institutional feature of the Texas market – the grid operator publicly released the aggregate bid schedule with a 2 day lag; therefore firms can learn their rivals' aggregate bid function with a 2 day lag. Suppose firms were to use the lagged bid data to create best-response bid functions to rivals' bids from 3 days prior to each auction, and submit these bids to the current auction. We compute lagged best-response bids and call these "naive best response" bids. Then we use the naive best-response bids and clear the market with the actual (step function) residual demand for the current auction. We find that this simple trading rule firm significantly outperforms the actual realized profits for all but the largest firm.

The results of this test are reported in Table 8. For example, TXU's actual bids yield 39.3% of the profits that would have been realized under our best-response benchmark. However, if TXU had used the simple trading rule, it would have earned (96.7%) of best-response profits, which indicates that there is strong persistence in the shape of residual demand across auctions. Similarly, all firms except Reliant would have significantly increased profits by following this simple trading rule.

Table 8: Evidence that Uncertainty and Bidding Rules
Do Not Bias the Best-Response Benchmark

	Percent of Best-Response Profits achieved by:	
Firm	**Actual bids**	**Best-responding to lagged bids**
Reliant	79.0	98.5
City of Bryan	45.3	100.0
Tenaska Gateway Partners	40.9	99.6
TXU	39.3	96.7
Calpine Corp	37.0	97.9
Cogen Lyondell Inc	16.2	100.0
Lamar Power Partners	14.7	99.6
City of Garland	12.6	99.9
West Texas Utilities	8.1	100.0
Central Power and Light	7.7	98.7
Guadalupe Power Partners	5.9	99.0
Tenaska Frontier Partners	4.9	99.3

Note: The table reports the percentage of best-response profits that are achieved by: (1) the firm's actual bidding data and (2) the simple trading rule where firms best-respond to the most recent publicly available data (the 2 day lagged residual demand). The firms reported in the table correspond to the firms that we use in the Cognitive Hierarchy in estimation. Source: Hortacsu and Puller [2008].

E Appendix: Evidence that Firms Do Not Misrepresent Capacity

Table 9: Capacity utilization relative to self-declared capacity

Firm	Maximum capacity utilization (%)
Reliant	81.72
City of Bryan	76.59
Tenaska Gateway Partners	125.88
TXU	97.13
Calpine Corp	83.84
Cogen Lyondell Inc	81.12
Lamar Power Partners	76.19
City of Garland	93.57
West Texas Utilities	92.92
Central Power and Light	98.82
Guadalupe Power Partners	74.69
Tenaska Frontier Partners	93.40

Note: The table reports maximum capacity utilization relative to self-declared capacity for each day, for the firms that we consider in the Cognitive Hierarchy.

F Appendix: Evidence that Firms Do Not Bid Marginal Cost

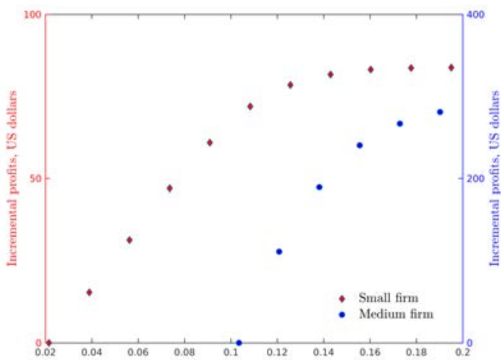
Table 10: Comparison of Actual Output to Best-Response Output

Firm	Actual Output	Best-Response Output
Reliant	431	507
TXU	133	441
Calpine	102	408
Guadalupe	12	396
Central Power & Light	35	352
Lamar Power Partners	30	272
Cogen Lyondell	34	269
West Texas Utilities	11	224
Tenaska Gateway	72	182
Tenaska Frontier	7	144
Garland	5	115
Bryan Texas Utilities	30	56

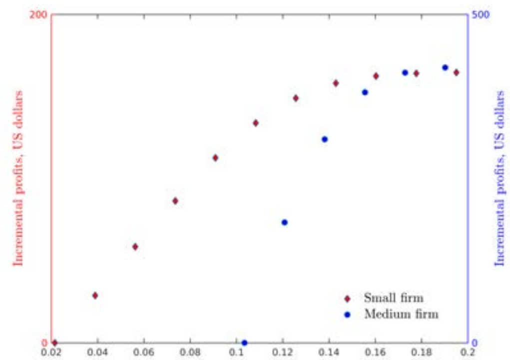
Note: This table reports average output under actual bids and best-response bids from the opening of the market until January 31, 2003. The figures are the average of the absolute value of sales so that output during both INC and DEC intervals has the property that marginal cost bidding will yield actual output greater than best-response output and that bidding steeper than best-response will yield actual output less than best-response output. Source: Hortacsu and Puller (2008).

G Appendix: Diminishing Returns to Increasing Sophistication

In this appendix, we investigate whether the private returns to increasing sophistication are decreasing or increasing. We do this for two firms, a small one (the one with the highest probability of being type 0) and a medium-size one. Because types are parameterized by size, we sequentially increase firm size until the firm reaches the same capacity as the largest firm in the market. As before, we do this for auctions that clear on the INC and DEC side separately. The results are reported in Figure 8, that reports incremental returns relative to the status quo of each firm. The figure confirms that there are decreasing marginal returns to increasing sophistication.



(a) INC side



(b) DEC side

FIGURE 8: Marginal returns to increasing sophistication