Does Strategic Ability Affect Efficiency? Evidence from Electricity Markets*

Ali Hortaçsu Fernando Luco

Steven L. Puller Dongni Zhu

January 2017

Abstract

Standard 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 the presence of heterogeneity in strategic sophistication. This paper studies such a setting in the Texas electricity market, in which bidding behavior of some firms persistently and significantly deviates from Nash-equilibrium bidding. We leverage a unique dataset that contains information on bids and valuations to estimate the level of strategic sophistication of specific firms in the market. We do this by embedding a Cognitive Hierarchy model into a structural model of bidding into auctions. We show that larger firms have higher levels of strategic sophistication than smaller firms, and there is significant heterogeneity across firms. We then use the estimated distribution of types of strategic sophistication 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. We find that exogenously increasing sophistication of small firms increases productive efficiency. Furthermore, mergers that create no cost synergies and increase concentration may nevertheless increase efficiency by improving the strategic sophistication of smaller firms.

*Hortaçsu: University of Chicago and NBER, hortacsu@uchicago.edu. Luco: Texas A&M University, fluco@tamu.edu. Puller: Texas A&M University and NBER, puller@econmail.tamu.edu. Zhu: Texas A&M University, dongnizhu@email.tamu.edu. Luco and Puller gratefully acknowledge financial support from the NSF (SES 1628864). We thank Avi Goldfarb, Karen Palmer, Mo Xiao, and various seminar participants for their helpful comments. Any errors are our own.

1 Introduction

Models of strategic equilibrium form the foundation of many studies in industrial organization that investigate market efficiency in oligopoly settings. In these studies, firms are modeled as playing some form of Nash Equilibrium, and that model of supply-side behavior is used to estimate parameters with implications for policy. 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 market outcomes under alternative market structures. When studying auctions, researchers use a Bayesian Nash model of bidding to "invert" bids to estimate valuations and then conduct counterfactual experiments to predict market outcomes under alternative auction formats.

However, some research has suggested caution at applying such strategic equilibrium models in all settings. In a Nash equilibrium, each firm is best responding to its beliefs about each rival firm's behavior *and* all of those beliefs are mutually consistent. In some settings, the rationality assumption or mutual consistency assumption may break down, and firms may not be playing at the fixed point that equilibrium models characterize. Firms that compete in the same market can vary substantially along a number of dimensions including production capacity, personnel, and general core competency. As a result, firms in a market can vary in the level of strategic behavior they exhibit when competing against one another. Indeed, evidence exists that real-world firms may be boundedly rational, engaging in some level of strategic thinking, but the degree of strategic thinking may "fall short" of playing the Nash equilibrium strategy (e.g. Hortacsu and Puller [2008] and Goldfarb and Xiao [2011]).

Deviations from Nash equilibrium play can be economically significant and have implications for efficiency. Indeed, Hortacsu and Puller [2008] (hereafter HP) identify a set of firms that submit bids into electricity auctions that persistently deviate from Nash bidding and do so substantially. The consequence of these deviations is that low-cost powerplants are not called to produce, and this substantially raises total production costs. Overall, 81% of productive inefficiencies are caused by small firms departing from Nash bidding.

This suggests that models allowing for boundedly rational firm behavior can be valuable for explaining the outcomes of certain real-world markets. To this extent, theoretical research has developed models that help organize strategic behavior that deviates from the Nash equilibrium.[1] Examples of such models include level-k thinking and Cognitive Hierarchy in which players best-respond to (perhaps incorrect) beliefs about their rival behavior. The

[1] A rich literature in experimental economics has studied the behavior of laboratory participants in strategic games such a beauty contest games, documented deviation from Nash equilibrium play, and developed hierarchy models that can explain such behavior. For examples, see Nagel [1995], Stahl and Wilson [1995], Costa-Gomes et al. [2001], Crawford et al. [2008], and Arad and Rubinstein [2012].

Cognitive Hierarchy model (hereafter CH) allows for heterogeneity in the levels of strategic thinking by firms in a market. 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 submit bids that correspond to the best response to all other players behaving as such. 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 submit bids corresponding to the best response to those beliefs. The limiting case of this model corresponds to the Nash equilibrium.[2] In this setting, CH maintains the assumption that players best respond, but it allows for firms to have beliefs about their rival strategies that are not consistent with the rivals' actual behavior. This model has the appealing feature that it allows for a hierarchy of levels of sophistication by different players in a market.

Despite the availability of theoretical models of boundedly rational behavior, it is difficult to use data from field settings to apply such models, as there is a critical identification problem if the goal is to uniquely identify market fundamentals. To see this, consider studies that apply the standard "IO inversion" approach – use a model that maps marginal cost (or valuation) to prices (or bids), and then "invert" the model so that data on prices (or bids) can be used to estimate the underlying marginal cost (or valuation). This approach – used in many oligopoly and auction settings – hinges on the assumption of a "unique" model of firm strategic behavior. Otherwise, multiple combinations of behavior and costs or valuations are consistent with the observed prices or bids. Bounded rationality models, such as cognitive hierarchy, allow for multiple forms of strategic behavior, so that researchers, in general, cannot separately identify the cognitive hierarchy structure from costs or valuations.[3]

However, this empirical challenge can be overcome if researchers have data on both the prices (bids) *and* the marginal cost (valuation). In this paper, we exploit such a data-rich environment in the context of electricity auctions. In these auctions, firms owning powerplants bid hourly to supply power to the 'spot market' that balances real-time supply and demand of electricity in Texas. Firms submit offers to supply different quantities of power at different prices. The grid operator clears this market using a multi-unit, uniform-

[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 has been 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 could 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 [2013] 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.

price auction – essentially aggregating supply bids and finding the market-clearing price that equates aggregate supply and demand. A unique feature of this setting is that we have data on each firm's hourly marginal cost of supply and each firm's hourly supply bids.

In this setting, HP show evidence that firms deviate from Bayesian Nash equilibrium bidding, suggesting a fruitful environment to apply bounded rationality models. HP test whether each firm submits bids that correspond to the best response to rivals' actual bids (as required if firms play a Nash equilibrium) and find that a few firms – typically larger firms – submit bids close to best-response bidding. Small firms, however, 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.[4]

The puzzling behavior of firms in the Texas market generates important questions:

1. What type of strategic behavior are the small firms engaging in? And the large firms?

2. Can an existing model of bounded rationality organize the behavior and usefully predict bidding out-of-sample?

3. Could mergers that increase strategic sophistication (but do not create cost synergies) increase efficiency?

4. How much would an (exogenous) increase in strategic sophistication by a firm or group of firms affect the efficiency of the market?

In this paper, we address each of these questions. Specifically, we embed a Cognitive Hierarchy model into a structural model of bidding behavior to capture the heterogeneity in the observed deviations from Bayesian Nash equilibrium bidding. The Texas electricity market has firms of various sizes, organizational structures, and personnel backgrounds that bid into the spot market, and we use firm observables to parameterize the determinants of firm type.

After estimating the model, we turn to study how strategic sophistication affects productive efficiency. We do this by simulating a number of scenarios in which the level of strategic sophistication of low-type firms is increased either exogenously or through mergers with high-type firms. Importantly, the application of the CH model to multi-unit auctions has very valuable methodological benefit in this setting. As shown in Klemperer and Meyer [1989], in general there are multiple equilibria in multi-unit, uniform-price auctions that can range from competitive to Cournot-like behavior. The multiplicity of equilibria presents a challenge for conducting counterfactual calculations of market outcomes under changes in

[4] HP rule out a number of alternative explanations for such steep bids such as collusion, the presence of transmission constraints, and unmeasured adjustment costs.

cost or market structure. One way to address this problem has been to impose mathematical restrictions on permissible form of bids, such as restricting bid functions to be linear (Baldick et al. [2004]). The CH model provides a means to address the multiple equilibria problem without imposing such restrictions. The mutually consistent beliefs assumption – a source of the multiple equilibria problem – is not imposed in CH. Instead, given a firm's belief about its rivals' type distribution, one can calculate the (unique) best-response bid.[5] Therefore, the iterative nature of strategic thinking under CH allows us to calculate unique counterfactual market outcomes. We exploit this feature by computing market efficiency under possible mergers between firms with different levels of strategic sophistication, which would not be possible under a Nash equilibrium model. Thus, not only does CH allow for more realistic models of real-world bidding behavior, but it allows researchers to more precisely simulate outcomes under changes in market structure or changes in costs.

Thus, we are able to simulate unique predictions of market outcomes under various policy counterfactuals. For example, consider a merger between a large and small bidder in this electricity market. Such a merger is unlikely to lead to substantial cost synergies because the costs of generating electricity is almost entirely driven by the model and vintage of the electric generator. Thus, 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 caused the large firm to take over bidding operations, then the generation resources of the small firm would subsequently be controlled by a higher level strategic thinker. This could increase efficiency because the low-k 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 any firms in the Texas market. More generally, we can calculate market prices and efficiency under any counterfactual level of strategic sophistication by any firm.

Our results show that efficiency increases with strategic sophistication, 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% relative to the status quo. However, efficiency improvements are smaller when firms with median levels of sophistication are given higher sophistication levels. Finally, mergers can increase efficiency even with no

[5] Put differently, computing Bayesian Nash equilibria in multi-unit auctions is challenging because it requires solving for fixed points in function space. In contrast, the CH model only requires solving a sequence of best-response problems. In addition, CH serves as an equilibrium refinement technique. Camerer, Ho, and Chong [2004] note a related feature that the CH model can be viewed as a behavioral refinement that can eliminate the multiplicity of equilibria in coordination games.

cost synergies and increasing market concentration. Our results show that when a small, low-type firm merges with a large, high-type firm, that efficiency can 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 are Goldfarb and Xiao [2011] who apply a cognitive hierarchy model to field data and Doraszelski, Lewis, and Pakes [2014] who study the evolution of competitive behavior in an electricity market. 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 [2014] use models of learning to predict the evolution of pricing in a newly opened electricity ancillary service market. Our project 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]). Finally, 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. First, section 2 describes the Texas electricity market that is the focus of this paper. Then, 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 equilibrium bidding and discusses identification. Section 6 discusses estimation and results. Section 7 studies the impact of a number of policy and merger counterfactuals. Finally, section 8 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.[6]

In the bilateral market, generating firms contract with utilities that serve customers. Then, 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 and consumption. Importantly, the production levels that are scheduled one day-ahead can differ from the quantities that the generating firm has financially contracted in the bilateral market, so a generating 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. It is bidding behavior in this auction that we study. 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 (vertical).

The grid operator clears the market every 15-minute interval by finding the market-

[6] Hortacsu, Mandanizadeh, and Puller [2016] study competition in the Texas retail market.

clearing price where hourly aggregate supply (a monotonically increasing step function) equals each quarter hour's balancing demand (a vertical function). 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 sell into the wholesale market but 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.

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's operation. The market began in August 2001, and we study the period of August 2002-January 2003. Therefore we study behavior for a snapshot of six months after the market has had one year to adapt to a new regime for trading power.[7] Because our sample period begins in the second year of the market's operation, the firms have 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 have submitted bids into the balancing market for every hour of every day for one year.

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

[7]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.

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. When balancing demand is positive, the grid procures more power than the aggregate day-ahead schedule so that real-time supply satisfies real-time demand. When balancing demand is negative, the day-ahead schedule exceeds real-time demand, so the auction calls upon firms to reduce output from the day-ahead schedule. 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. As we discuss above, bid functions take the form of monotonic step functions with up to 40 elbow points, however most firms use substantially less that the maximum number of elbow points.

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.[8]

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.[9] 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 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 the fact that the firm 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

[8]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.

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

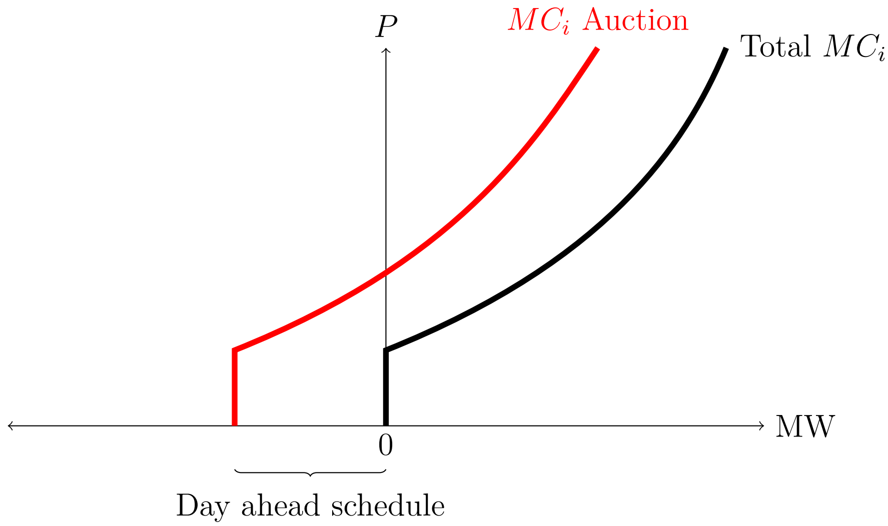


Figure 1: Stylized Marginal Cost of Supplying to Balancing Market

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, so we refer the reader to HP's Data Appendix for further details, and this is reproduced in Appendix A.

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 1 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 Genscape 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 in a given hour. However, in 26% of hours, transmission lines were congested which led to different market-clearing prices in different zones of the state. We exclude those auctions when there was transmission congestion; HP show that this does not affect 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.[10] Figure 2 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 an expectedly hot day and the firm faces RD_1 shown in Figure 2. 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 turn out to 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. 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 the red line S_i^{BR}.[11]

We can construct data analogs to these stylized pictures. Importantly, no estimation is

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

[11] In general, it is possible that the set of best-response points is not monotonic function, however we show in section 5.2 that in this setting the best-response points are monotonic.

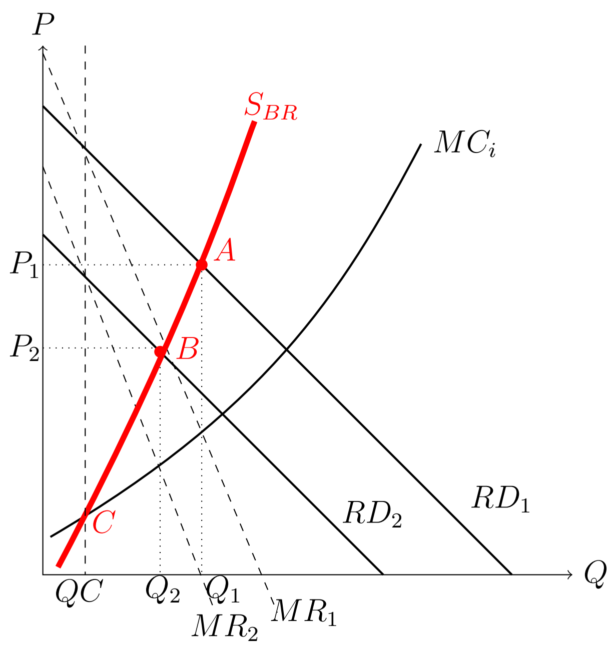


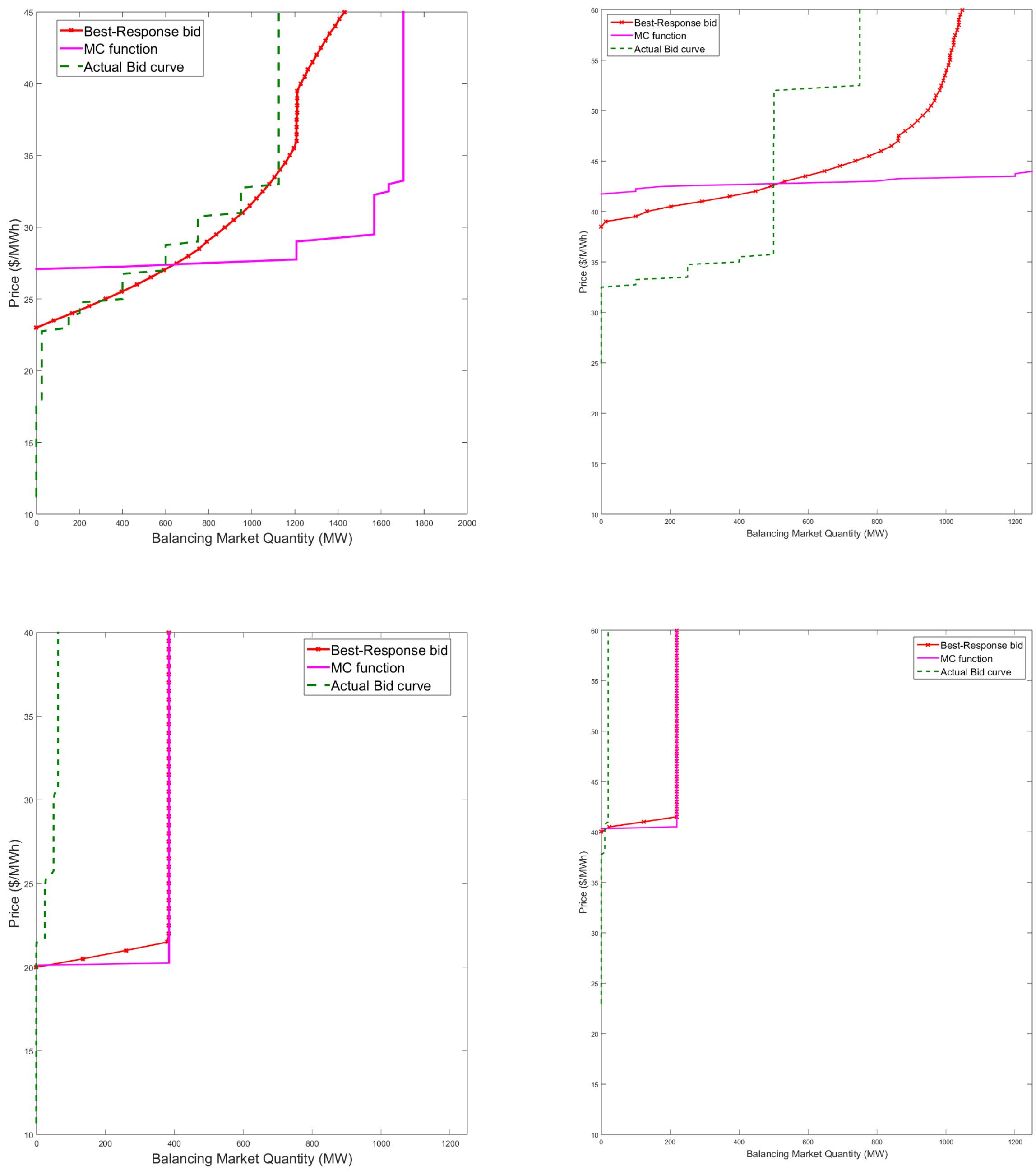
FIGURE 2: Best-Response Bidding in Spot Auction

required; the components of Figure 2 are available as *data* for each firm in each auction. We view this data-rich environment as a major strength of our approach. Data on marginal cost are critical to our identification strategy that allows us to identify strategic behavior.

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 3 displays representative bid functions for four different firms in this market. 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 are not depicted here.

The top-left panel of Figure 3 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 increase production relative to its day-ahead schedule by about 1800 MW. The firm has a contract position of about 600 MW 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.

Figure 3: Actual Bids vs. Best-Response Bids for Large, Medium, Small, and Very Small Firms



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 3 displays a representative 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 marginal cost for quantities greater than the contract position and below marginal cost for quantities 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 to increase production 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 function. The firm's actual bid function 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 3 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.

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 units must be called to produce instead. In the parlance of electricity markets, this bidding behavior inhibits least cost dispatch. As we document below, this productive inefficiency can be sizeable.

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' relative to the benchmark 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.

3.2 Ruling Out Alternative Explanations

In section 5, we use a model of boundedly rational bidding behavior to explain the heterogeneity in bidding. Before developing that model, we rule out other possible explanations that do not appear to explain the behavior. These explanations are discussed in additional detail in the NBER working paper version of Hortacsu and Puller [2008].

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 are 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 adjusts production when the balancing market clears. Note that the sample bid in the top-right panel of Figure 3 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 to see 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 – so 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 3 are very similar, yet the firms bid quite differently.

Second, it does not appear that there are simply too few dollars at stake in the balancing market to justify putting financial resources into a bidding operation. HP show that the profit gains over the course of one year are likely to surpass industry estimates of the costs of establishing a simple trading operation.

Third, one might be concerned that bidding rules prevent the 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 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 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 to test if firms systematically could raise profits by using only information available at the time that bids are submitted (and before any uncertainty is realized). It turns out that the Texas 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.[12] Then we use the naive best-response bids and clear the market with the actual (step function) residual demand for the auction in question. This simple trading rule using information that was available to each firm significantly outperforms the actual realized profits. The results of this test – reproduced from HP – are reported in Table 5 of the Appendix. The results show that actual bids generate profits substantially lower than profits achieved by best-responding to the most recent publicly available aggregate bid data. For example, TXU's actual bids yield 39.3% of the profits that would have been realized under our best-response benchmark. And it does not appear that this is driven by uncertainty/bidding rules because the "naive best-response" bids would yield nearly identical profits (96.7 %) as the best-response benchmark. Further details can be found in HP.

Fourth, the possibility of congested transmission lines does not affect our inference. Details can be found in HP.

Finally, we note that we can 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.

[12]The "naive best-response" bid functions respect bidding rules in the sense that they are monotonically increasing step functions.

All firms are observed to use at least 75% of stated capacity and on average to use 90% of stated capacity (see Table 4 in the Appendix). 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.

4 Theoretical Background on Cognitive Hierarchy

The theoretical literature has developed a rich set of models of boundedly rational strategic behavior that could explain deviations from Bayesian Nash equilibrium play. Generally speaking, bounded rationality models relax one of the two conditions of Nash Equilibrium; that (1) players maximize expected payoffs given beliefs about their rivals' actions and (2) that 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 behavioral models has been conducted by Costa-Gomes et al. [2001], Crawford and Iriberri [2007], and Camerer, Ho, and Chong [2004].

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

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)*. Comparing the two models, in one sense CH is a more flexible model. However, one could also view CH as a

[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 bidders in the electricity auctions estimate expected payoffs in an unbiased way, a key feature of the QRE model.

model that could be too flexible and explain "anything". In this paper, however, we sidestep this theoretical debate and use a model that is more general – cognitive hierarchy – and estimate the model with our data.

5 Empirical Strategy

5.1 Big Picture of Modeling and Estimation Strategy

The iterative nature of decision rules under CH facilitates a computationally tractable empirical strategy. For any firm i in auction t, we have data on the marginal cost of supplying power to the grid. We begin by defining the bidding behavior for a non-strategic 0-step player. This definition will be critical, so we spend time developing the rationale for that assumed behavior, but take that bidding behavior as given for the moment.

Consider firm i that is type k. The assumptions of the CH model imply that i believes its rivals are distributed between type-0 and type-(k-1), according to a normalized Poisson distribution with parameter τ. Firm i chooses its bids based upon these beliefs (that depend on its own type, contract position, and its rivals τ's) and marginal costs, in order to maximize its 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 level-0 players to 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 3). 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.

One of the advantages of this approach 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 also 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 small ones including the ones depicted in Figure 3. This, however, has the unintended cost of limiting the extent to which our counterfactual simulations can improve efficiency, as part of the inefficiencies that result from departure from Bayesian Nash bidding is generated by firms that we do not include in the cognitive hierarchy.

Once level-0 bidding is defined, we can use our data on each firm's marginal cost to calculate the predicted bidding behavior for a firm of any type $k > 0$. Specifically, given the assumption about level-0 players and a fixed vector $\tau = \{\tau_1, ..., \tau_N\}$ denoting N firms' levels of strategic sophistication, which depend on firm characteristics, X_i, we use an iterative process provided by the CH model to calculate each player's optimal theoretical bids under various sophistication levels. Then, based on information about players' type distribution $Poisson(\tau_i)$, we calculate players' theoretical optimal bids.

We then compare these bids to the firm's actual bidding behavior. The estimation process finds the parameters of τ – 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 realized costs, one would not be able to identify types from bid data without additional assumptions regarding the cost function.[14] In other words, instead of using data on observed bids and a Bayesian Nash equilibrium model of behavior to recover valuations, we use that we observe valuations and bids to recover the type that rationalizes observed behavior.

5.2 Modeling in Detail

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 have developed a formulation that incorporates modeling features of share auction models (Wilson [1979] and Hortacsu and Puller [2008]) and the Poisson

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

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}. As in HP, 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 $S^k_{it}(p, QC_{it})$ to produce different quantities at different prices. Let the bid function by rival j of type l be denoted $S^l_{jt}(\cdot)$.

All sellers are paid the market-clearing price, which is determined by:

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

From the perspective of firm i with private information on k_i, QC_i, and submitting bid $\hat{S}_{it}(p)$, the uncertainty can be characterized by defining the following function $H(\cdot)$ which defines the probability that the market-clearing price p^c_t is below any price level p:

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

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

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

$F(\boldsymbol{QC_{-it}}, \boldsymbol{l_{-i}}, \varepsilon_t | \hat{S}_{it}(p), k_i, QC_{it})$ 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 (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}). In addition, 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 considers how bidders type 0 bid and allows us to solve the problem recursively. The second assumption considers the distribution of types in the Cognitive Hierarchy model. Finally, the third assumption considers the distribution of the remaining sources of uncertainty.

Assumption 1. *Bidders type 0 submit perfectly inelastic bids that are determined by their contract positions. This is,*

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

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

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 a scalar random variable θ_{it} that is the sum of rival contract positions $\sum_{j \neq i} QC_{jt}$ and the aggregate demand shock $(-\varepsilon_t)$.

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 $H(\cdot)$ becomes:

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

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

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

Proposition 1. *1. If bidders type 0 submit perfectly inelastic bids that are determined by their contract positions in the CH model, their bids are additive separable,* $S^0_{it}(p, QC_{it}) = QC_{it}$.

2. If bidders type 0 submit perfectly inelastic bids that are determined by their contract positions in the CH model, bids of bidders type 1 are also additive separable, $S^1_{it}(p, QC_{it}) = \beta^1_{it}(QC_{it}) + \alpha^1_{it}(p)$. *Moreover,* $S^1_{it}(p, QC_{it}) = QC_{it} + \alpha^1_{it}(p)$.

Proof. 1. It is straight forward to see that bids of bidders type 0 are additive separable because $S^0_{it}(p, QC_{it}) = QC_{it} = QC_{it} + f(p)$, where $f(p) = 0$, $\forall p \in [\underline{p}, \bar{p}]$.

2. 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 $[(p - C'_{it}(S^1_{it}(p))] \frac{\int_{l_{-i}} \gamma\left(D_t(p) - \hat{S}^1_{it}(p)\right) \cdot \Delta(l_{-i})}{-\int_{l_{-i}} \gamma\left(D_t(p) - \hat{S}^1_{it}(p)\right) D'_t(p) \Delta(l_{-i})}$ is a function of price p. Therefore, bids of bidders type 1 are additive separable and can be represented by $S^1_{it}(p) = \alpha^1_{it}(p) + QC_{it}$, where $\alpha^1_{it}(p) = [(p - C'_{it}(S^1_{it}(p))] \frac{\int_{l_{-i}} \gamma\left(D_t(p) - \hat{S}^1_{it}(p)\right) \cdot \Delta(l_{-i})}{-\int_{l_{-i}} \gamma\left(D_t(p) - \hat{S}^1_{it}(p)\right) D'_t(p) \Delta(l_{-i})}$.

$\square$

For a bidder type 2 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 additive separable bids. This is, for a firm bidding $\hat{S}^2_{it}(p)$

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

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 bidders type 0 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]:

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

Finally, for Γ_i we assume it is a uniform distribution.

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

Allowing for other distributions, such as Normal, is possible, though it increases the computational burden as one needs to solve the first-order condition by successive approximations. Under the assumption that $\Gamma_i(\cdot)$ is uniform, the value of $\gamma(\cdot)$ in H_S and H_p is independent of rival type, so the first-order condition simplifies to (in the case of bidders type 1)

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

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 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 $k \in \{0, \dots, K\}$. 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 $k \neq 0$, firms have beliefs about its rivals' type. 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' type. Note, however, that in a specific auction, even if two bidders are of the same type, differences in (observed) 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 of a firm of being of 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

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

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 guarantee that we find the global minimum, we run estimation starting from 50 sets of random initial points.

6.2 Results: Estimated Parameters

Results are reported in Table 1. We estimate six 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 as it pertains to the firm's participation in the balancing market, we seek a metric of the firm's potential stakes in the balancing market that is plausibly exogenous to its realized bidding behavior. We compute the magnitude of MW 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.[15] 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 4 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 45% chance of being type-0, about a 35% chance of being type-1, about a 15% 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. The estimated type distribution is shown in Figure 6 of the Appendix, and it resembles Figure 4.

Next we explore whether 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 could impact the nature of the trading operations that a firm establishes. In columns (3) and (4) 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 the last three specifications, 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 used 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

[15] 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.

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).[16] 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. Then we estimate our benchmark specification using *Size* and adding dummy variables for University type or Degree type.

Our specifications with personnel are reported in columns (5)-(7) of Table 1. AAU-trained traders may employ more sophisticated bidding strategies. But, of course, the decision to hire AAU-trained traders is endogenously determined by the firm, and we do not model this choice. We find that when we include AAU in addition to firm size in column (5), 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. As we note above, 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.

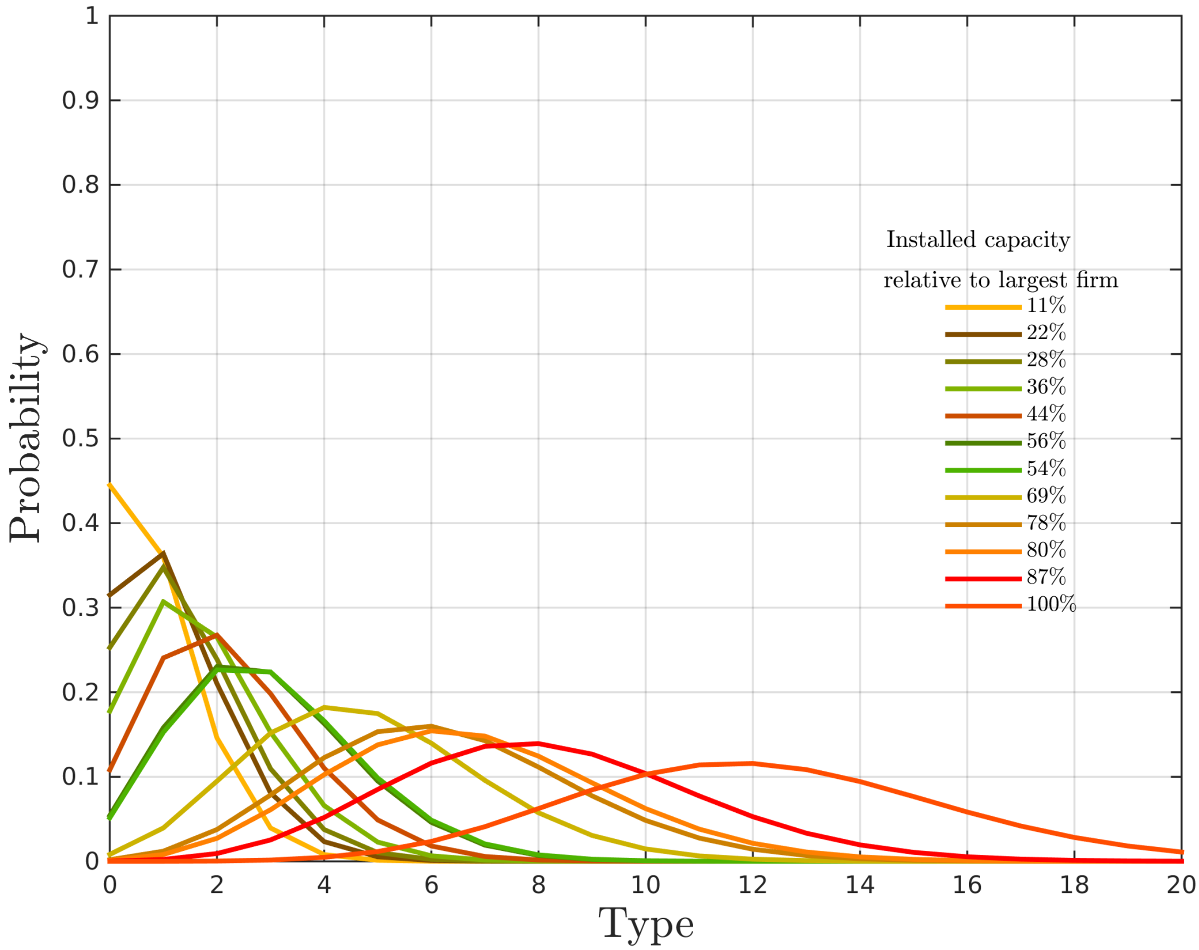
[16]The AAU includes 62 private and public research universities in the U.S. and Canada. A list of the AAU universities can be found at: http://www.aau.edu/about/default.aspx?id=16710.

Table 1: Structural Model: Estimated Parameters of Type Function

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Constant	-0.726	-0.196	-3.395	-0.292	-0.749	-3.493	-0.691
	(0.133)	(0.158)	(0.249)	(0.069)	(0.159)	(0.508)	(0.147)
Size	14.594	-1.163	25.789	14.856	13.619	3.090	11.933
	(1.369)	(2.892)	(3.575)	(2.326)	(1.865)	(0.671)	(1.409)
Size^2		86.191					
		(13.643)					
Merchant			-1.562	-0.771			
			(0.403)	(0.169)			
Merchant$\times$Size				-0.288			
				(0.186)			
AAU University					0.376		
					(0.069)		
Degree in Economics,						5.626	
Business or Finance						(1.935)	
Economics degree							1.633
							(0.340)
Number of auctions				99			

Note: Bootstrapped standard errors using 20 samples.

FIGURE 4: Estimated Distributions of Types for the 12 Firms (*Size* Specification)



7 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 three steps. We first ask how exogenous increases in strategic sophistication of specific firms affect market efficiency. We believe this is an important first step as market structure does not change with this intervention. Hence, though the actual intervention may appear as unreal (though consulting and hiring more qualified employees to operate the trading floor are good examples that fit in this description), it provides a way to isolate the impact of increasing sophistication in the absence of changes in market power. Then, we ask whether there are decreasing returns to increasing sophistication. Finally, we turn to studying how increases in strategic sophistication that result from low-type firms merging with higher-type firms, may affect market efficiency. In this case, bidding approaches that of Bayesian Nash but market concentration increases as well. Hence, the overall effect of the merger is ex-ante unknown.

To keep our results in perspective, it is important to note that there is an upper bound on the magnitude of the effects studied in our simulations. Indeed, while a social planner would minimize dispatch cost by inducing generation at marginal costs, we lack data on marginal costs for some bidders. This means that in our simulations the benchmark will not be the outcome of the social planner but that of a planner that forces firms in the CH to bid at their marginal costs but keeps bids of firms not included in the CH as they are in the data. 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 the first two specifications of Table 1.

7.1 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 rotate the belief about residual demand and induce bids with more slope. 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, if the increase in sophistication is publicly observable (e.g. rivals observe that the firm hires a bidding consultant), then there is a second effect. Rival firms recognize the increase in sophistication (even though their beliefs will continue to be wrong) and bid more competitively.

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 2 presents results using the $Size$ and $Size^2$ parameterization of τ_i. We report results separately for auctions with positive and negative balancing demand as this makes comparisons easier. First, we find that the direct effect ('Private') comprises the majority of the total effect ('Public'). Second, most of the gains in efficiency are driven by the change in behavior of small, low-type firms. Indeed, the three counterfactuals presented show that the most important gains in efficiency occur when the three smallest firms raise their type to the median, while increasing sophistication of "median" firms to the level of the firm with the maximum sophistication has a minor impact on efficiency. This also means that there are decreasing returns to increasing sophistication. Similarly, most of the gains obtained from increasing sophistication of the small firms to the median level correspond to the change in behavior of the three smallest firms, as the third row of both panels show.

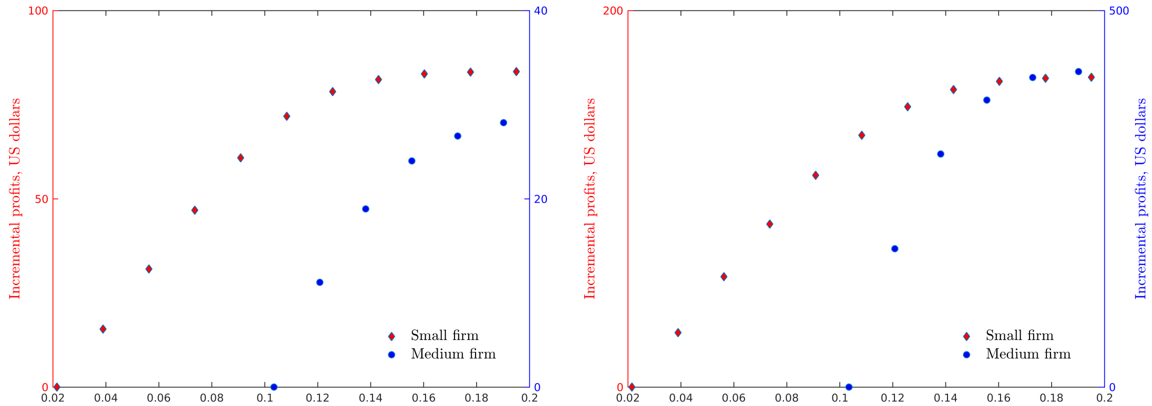
Table 2: Counterfactuals: Exogenous increase in sophistication

	INC side			DEC side		
Specification 1 (*Size*)	Public	Private	Δ PK	Public	Private	Δ PK
Small firms to median	9.43%	7.70%	1.73pp	16.49%	13.91%	2.58pp
Above median firms to highest	4.49%	3.72%	0.77pp	7.96%	6.50%	1.46pp
Three smallest to median	6.97%	6.04%	0.93pp	10.70%	9.94%	0.77pp
Specification 2 ($Size^2$)	Public	Private	Δ PK	Public	Private	Δ PK
Small firms to median	8.04%	6.90%	1.14pp	14.62%	11.94%	2.68pp
Above median firms to highest	4.35%	3.63%	0.73pp	7.91%	6.33%	1.58pp
Three smallest to median	5.30%	4.60%	0.71pp	8.53%	7.89%	0.64pp

Note: These numbers are computed using the estimates in table 1. Calculations are done separately for auctions with positive balancing demand (INC side, 60 auctions) and negative balancing demand (DEC side, 39 auctions). Δ PK: is the difference between efficiency gains under public and private knowledge of the change in sophistication.

7.2 Exogenous Increase: Marginal returns of sophistication

We now turn to studying whether the returns to increasing sophistication are decreasing or not. 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 5, 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 5: Marginal returns to increasing sophistication

7.3 Endogenous Increase: Mergers

We now turn to studying how mergers may affect efficiency. As mentioned above, we focus on potential mergers that do not generate cost synergies but do increase concentration. In this setting, mergers may increase efficiency both by relocating generation from high-cost, high-type firms to low-cost, low-type firms that have previously priced themselves out of the market, and by changing their bidding behavior. Mergers create a countervailing effect – increasing market concentration may create additional market power that leads to higher dispatch costs.

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.

In this setting, we explore three potential mergers. The first one corresponds to the merger between the smallest and the largest firms in the data. The second one considers the merger between the median and the largest firms. Finally, the last merger considers the case in which the two largest firms in the data merge. The results show that mergers may indeed increase efficiency, but only if the firms involved are not too large. Indeed, in the results we present below, efficiency increases when the smallest and largest firms merge, but the exploitation of more market power dominates the relocation of generation for the other two cases.

It is important to note that, in contrast with the case of exogenous increases in sophistication, in this case efficiency gains come from three sources. First, there is a direct effect of relocation of generation. Second, the newly formed firm also has more correct beliefs about its rivals and bids more competitively. Third, all rivals observe the increase in sophistication of the generating units that belonged to the smallest firm in the merger and bid more aggressively (conditional on their types). This is recognized by the newly created firm and induces this firm to bid more aggressively as well.

Table 3: Increasing Sophistication via Mergers: Impact on Generating Costs

	Specification 1		Specification 2		Specification 6	
	INC side	DEC side	INC side	DEC side	INC side	DEC side
Smallest and largest firms	-3.2%	-15.4%	-2.6%	-13.0%	-1.4%	+0.5%
Median and largest firms	+9.0%	+21.9%	7.7%	+18.0%	-4.2%	-5.9%
Two largest firms	+17.3%	+56.3%	14.1%	+57.0%	+11.6%	+13.6%

Note: These numbers are computed using the estimates reported in table 1. Calculations are done separately for auctions with positive balancing demand (INC side, 60 auctions) and negative balancing demand (DEC side, 39 auctions). All numbers correspond to the change in generating costs relative to those implied by the estimated model. We first compute the difference between the cost implied by the model and our efficient benchmark. We then repeat the calculation for the counterfactual. Finally, we compute the ratio between these two measures.

8 Conclusions

Models of strategic equilibrium form the foundation of many studies in industrial organization that investigate market efficiency in oligopoly settings. These models rely on the existence of a unique mapping from unobserved fundamentals, such as marginal costs or valuations, to observed prices or bids, to study questions about market efficiency and evaluate policy interventions, among others. 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. Furthermore, the literature has shown that these departures from (Bayesian) Nash 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 predicts, while bidding from 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 strategic sophistication of low-type firms, either exogenously or through mergers with higher-type firms, may affect efficiency. Our results show that not only exogenously increasing sophistication may increase efficiency significantly, but that also mergers that do not generate cost synergies but increase concentration may also increase efficiency as long as the higher sophistication of one of the merging parties is transferred to the rest of firms involved in the merger.

A Data Appendix 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.

Figure 6: Estimated Distributions of Types for the 12 Firms - $Size^2$ Specification

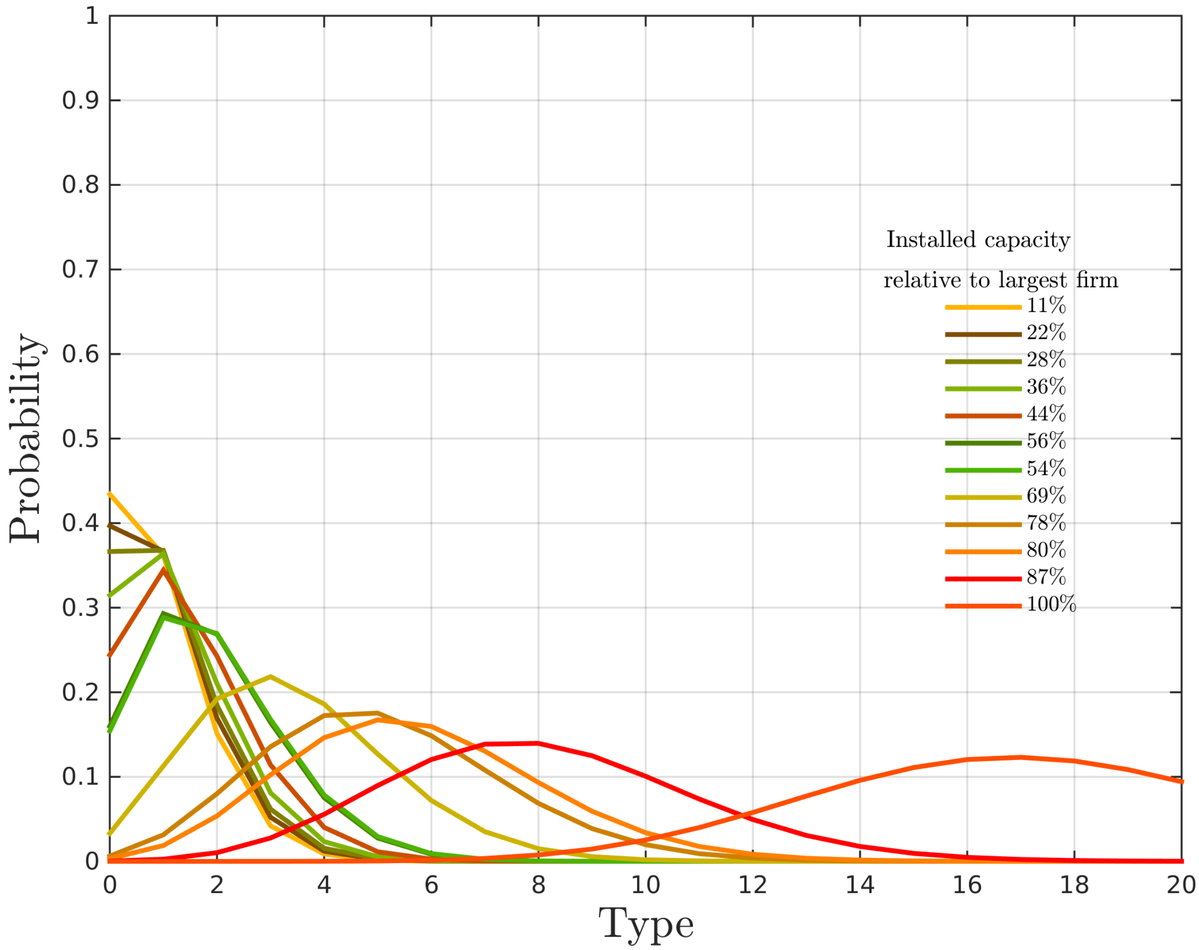


Table 4: APPENDIX: Capacity utilization relative to self-declared capacity

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.

Table 5: Evidence that Uncertainty and Bidding Rules
Do Not Bias the Best-Response Benchmark

	Percent of Best-Response Benchmark 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 report the percentage of potential profits achieved by the bids submitted by the firms (data) and by the bids that firms would submit if they were to best respond to lagged residual demand. The firms reported in the table correspond to the firms that we use in the Cognitive Hierarchy in estimation.

References

Yonghong An. Identification of first-price auctions with non-equilibrium beliefs: A measurement error approach. Texas A&M University working paper, 2013.

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.

Ross Baldick, Ryan Grant, and Edward Kahn. Theory and application of linear supply function equilibrium in electricity markets. *Journal of Regulatory Economics*, 25(2):143–167, March 2004.

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, 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 price competition in a new market. University of Pennsylvania Working Paper, 2014.

Natalia Fabra and Mar Reguant. Pass-through of emission costs in electricity markets. *American Economic Review*, 104(9):2872–2899, 2014.

David Gill and Victoria Prowse. Cognitive ability, character skills, and learning to play equilibrium: A level-k analysis. *Journal of Political Economy*, forthcoming.

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 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 multi-unit auctions: A case study of the Texas electricity spot market. *RAND Journal of Economics*, 39(1): 86–114, Spring 2008.

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

Ali Hortacsu, Seyed Ali Mandanizadeh, and Steven L Puller. Power to choose? an analysis of consumer inertia in the residential electricity market. working paper, June 2016.

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

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.

Paul D. Klemperer and Margaret A. Meyer. Supply function equilibria in oligopoly under uncertainty. *Econometrica*, 57(6):1243–1277, November 1989.

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.