American Economic Review 2019, 109(12): 4302–4342
https://doi.org/10.1257/aer.20172015

Does Strategic Ability Affect Efficiency? Evidence from Electricity Markets†

By Ali Hortaçsu, Fernando Luco, Steven L. Puller, and Dongni Zhu*

Oligopoly models of price competition predict that strategic firms exercise market power and generate inefficiencies. However, heterogeneity in firms' strategic ability also generates inefficiencies. We study the Texas electricity market where firms exhibit significant heterogeneity in how they deviate from Nash equilibrium bidding. These deviations, in turn, increase the cost of production. To explain this heterogeneity, we embed a cognitive hierarchy model into a structural model of bidding and estimate firms' strategic sophistication. We find that firm size and manager education affect sophistication. Using the model, we show that mergers which increase sophistication can increase efficiency despite increasing market concentration. (*JEL* D24, D43, G34, L13, L25, L94)

Firms that compete against one another can be quite different. Even if they compete in the same market, firms can vary across a number of dimensions: corporate structure, production capacity, market experience, and general core competency. Moreover, the managers who develop corporate strategy differ as well. For any particular market, a quick scan of resumes of firm managers often reveals differences in the type of academic training, the years of job tenure, and previous job experience. These differences could cause firms to adopt different types of strategic behavior when they compete against one another.

However, the potential existence of heterogeneity in strategic behavior does not play a role in typical models used to analyze strategic behavior in oligopoly. In many empirical studies, firms are modeled as being fully strategic and playing some form of a Nash equilibrium. In a strategic equilibrium, the assumption is that all firms are best responding to the beliefs of their rivals and that firms' beliefs are mutually consistent. Implicit in this modeling approach, the differences across firms, whether it be

* Hortaçsu: The Kenneth C. Griffin Department of Economics, University of Chicago, 1126 E. 59th Street, Chicago, IL 60637, and NBER (email: hortacsu@uchicago.edu); Luco: Department of Economics, Texas A&M University, 4228 TAMU, College Station, TX 77843 (email: fluco@tamu.edu); Puller: Department of Economics, Texas A&M University, 4228 TAMU, College Station, TX 77843, and NBER (email: puller@econmail.tamu.edu); Zhu: School of Economics, Shanghai University of Finance and Economics, 111 Wuchuan Road, Shanghai 200433, China (email: zhu.dongni@mail.shufe.edu.cn). Stefano DellaVigna was the coeditor for this article. Hortaçsu, Luco, and Puller gratefully acknowledge financial support from the NSF (SES 1426823 and SES 1628864). We thank Colin Camerer, Natalia Fabra, Dan Fragiadakis, Avi Goldfarb, Karen Palmer, Ariel Pakes, Mar Reguant, Mo Xiao, four anonymous referees, and many seminar participants for their helpful comments. We thank Hayden Parsley and Brittany Street for excellent research assistance. Any errors are our own. The authors declare that they have no relevant or material financial interests that relate to the research described in this paper.
† Go to https://doi.org/10.1257/aer.20172015 to visit the article page for additional materials and author disclosure statements.

in firm or manager characteristics, do not cause firms to adopt different competitive strategies. These models are used to study the nature of competition and the welfare consequences of alternative market structures. For example, when studying differentiated product industries, firms are modeled as engaging in Bertrand-Nash competition in order to predict prices if firms were to merge. When studying auctions, researchers use a Bayesian Nash model of bidding to "invert" bids in order to estimate valuations and to predict revenue and efficiency under alternative auction formats.

In this paper we ask: what if all firms engage in some level of strategic behavior, but some firms "fall short" of playing the Nash equilibrium? How does heterogeneity in strategic sophistication affect the efficiency of a market?

While models of strategic equilibrium are ubiquitous in empirical industrial organization, a large experimental literature on games demonstrates that subjects very often deviate from Nash equilibrium behavior (for example, see Kagel and Roth 1995; Camerer 2003; and Fudenberg, Rand, and Dreber 2012). To explain these deviations, a number of theoretical alternatives to Nash equilibrium play have been proposed. Among these, a popular framework is the cognitive hierarchy model of Camerer, Ho, and Chong (2004).[1] In a pioneering paper, Goldfarb and Xiao (2011) applies the cognitive hierarchy model to study the opening of the US local telephone market to competition, and finds that heterogeneity in strategic sophistication, driven by different manager characteristics, affects the amount of entry into different local markets, and that more sophisticated firms are more likely to survive.

In this paper, we argue that the cognitive hierarchy model can capture observed firm behavior in an oligopoly pricing setting very well. Our setting is the Texas electricity market, previously studied in earlier work (Hortaçsu and Puller 2008), where we document that some firms persistently deviate from Nash behavior. Our prior work did not provide a theory explaining these deviations. In this paper, we show that the cognitive hierarchy model captures the heterogeneity in deviations from Nash behavior quite well, in- and out-of-sample. Moreover, cognitive hierarchy also provides a computationally tractable framework to study firm behavior and market outcomes under counterfactual scenarios such as mergers.

In the cognitive hierarchy model, the least strategic players, level-0 players, are entirely non-strategic in their bidding. Level-1 players assume that all other players are level-0 players and choose actions that are the best-response to those beliefs. Level-2 players assume that all other players are some combination of level-0 or level-1 players and best respond to those beliefs. In general, level-k players assume that all other players are distributed between level-0 and level-$(k-1)$ and best respond to those beliefs.[2]

[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, Crawford, and Broseta (2001), Crawford, Gneezy, and Rottenstreich (2008), and Arad and Rubinstein (2012).

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

As noted above, Goldfarb and Xiao (2011) uses the cognitive hierarchy (CH) model in empirical work to study entry decisions. Utilizing the CH model to study oligopoly *pricing* decisions is not straightforward, however. This is due to a critical identification problem. Consider the large number of empirical studies that use a model of expected profit maximization that maps marginal cost to prices, and then "inverts" the model so that data on prices can be used to estimate the underlying marginal cost. This approach, used in many oligopoly and auction settings, hinges on the assumption of a particular form of strategic behavior. Otherwise, multiple combinations of behavior and costs may be consistent with the observed prices. Thus, in most empirical settings, the task of separately identifying parameters of the CH model from unobserved costs becomes a difficult exercise.[3] More generally, once the researcher allows for the possibility that firms deviate from Nash-type behavior, it is no longer possible to use observed prices to infer marginal cost.

This empirical challenge can be overcome if researchers have data on both the prices *and* the marginal cost. This paper exploits the data-rich environment of the Texas electricity spot market, in which many firms, which vary in size and other characteristics, compete. We have detailed data on individual firms' marginal cost of production and bids into power auctions. Having access to marginal cost data allows us to capture departures from Nash behavior and to identify the parameters of the CH model.

We find that the strongest determinant of a firm's level of strategic sophistication is size. Larger firms are higher type in the cognitive hierarchy, and thus are more strategically sophisticated. Manager characteristics, such as academic training, play a smaller but still significant role. Strikingly, there is substantial heterogeneity in the level of strategic sophistication across the firms in the Texas electricity market. This heterogeneity in sophistication impacts the efficiency of the market. Firms with lower levels of strategic sophistication submit bids that are so high that their plants are often priced out of the market despite the fact that their plants are often low cost. As a result, the power grid operator instead dispatches higher cost plants, resulting in inefficiently costly production. Finally, we do not find evidence of substantial learning in the early years of the market.

We also explore the extent to which our model predicts pricing behavior out-of-sample. We exploit a two-month outage at a nuclear plant in the middle of our sample period that significantly reduced nuclear power output. As a result, total demand for power intersected market marginal cost at a steeper point on the marginal cost function. We first test whether firms that are behaving strategically recognize that this publicly observable cost shock is likely to make their residual demand in the market less elastic. We find that firms estimated by our model to be a higher strategic type end up responding to the outage by recognizing that their residual demand is steeper, while lower-type firms do not. Second, we re-estimate

[3] One novel approach to address this problem in the auctions setting is proposed by Gillen (2010), which studies joint identification of types and valuations in the level-k setting. Gillen shows point identification of the joint distribution can be obtained by 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 (2017) also studies identification in the level-k model; he relaxes some of the 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.

our model of bidding behavior without using the data from the outage period, and use these estimates to predict profits during the outage period. We show that the CH model outperforms a model of unilateral best response when predicting realized profits.

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

We simulate the effect on market efficiency of increasing sophistication through actions such as hiring better managers. Our results show that increases in strategic sophistication improve efficiency, though at a decreasing rate. For example, exogenously increasing the sophistication of low-type firms to the level of median-type firms will increase market efficiency by 9–16 percent. However, efficiency improvements are smaller when firms with median levels of sophistication are given higher sophistication levels.

We also simulate market outcomes if heterogeneous firms were to merge. For example, consider a merger between a large and small firm that only affects the firms' bidding operations. Such a merger is unlikely to lead to cost synergies because the costs of generating electricity is almost entirely driven by technical characteristics of the power plants. One might expect the increase in market concentration to enhance market power and reduce economic efficiency. However, if the firms are boundedly rational, this merger could increase efficiency. Suppose that the large firm is a high-level strategic thinker and the small firm is a low-level strategic thinker. If the merger causes the large firm to take over bidding operations, then the power plants of the small firm will subsequently be controlled by a higher level strategic thinker. This can increase efficiency because the power plants of the low-type firm are more likely to be bid in at prices that cause low-cost plants to produce. We evaluate this conjecture by simulating mergers between various firms in the Texas market. We find that, in this setting, if a small, low-type firm were to merge with a large, high-type firm, then efficiency will improve despite the increase in concentration. However, when medium-sized firms merge with large firms, the market power effect dominates the sophistication effect, and efficiency decreases.

This paper makes four key contributions beyond the previous literature, including our earlier work on this market. First, this paper is constructive as we show that an important theoretical model of bounded rationality, cognitive hierarchy, can

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

explain deviations from Nash equilibrium in a field setting where firms engage in price competition. To our knowledge, this is the first paper that examines pricing/bidding decisions, using field data, through the lens of the cognitive hierarchy model. Second, we show that a behavioral game theory model can make better out-of-sample predictions of pricing behavior than a Nash-type model. Third, we show that behavioral game theory can be used to conduct computationally efficient counterfactuals for multi-unit auctions, which is not possible in standard Nash-type environments. And fourth, we quantify the insight that events that increase sophistication such as mergers can improve efficiency, which is a contribution to antitrust, both for academia and practitioners.

More broadly, this paper contributes to two emerging bodies of literature: behavioral industrial organization, and the study of firm sophistication and learning in new markets. The literature on behavioral industrial organization has largely focused on the case of boundedly rational behavior by *consumers*, with less emphasis on boundedly rational behavior by *firms*. However, evidence on boundedly rational firm behavior has been growing. (For a recent comprehensive survey of structural behavioral economics, see DellaVigna 2018.) On the pricing margin, Cho and Rust (2010) finds, through a field experiment, that rental car companies could increase profits by making rental rates a declining function of car mileage, and by holding on to their inventory of cars longer. Massey and Thaler (2013) finds evidence that NFL teams consistently over-value top draft picks. Ellison, Snyder, and Zhang (2018) estimates a model of price adjustment to capture price inertia and managerial inattention in an online market for computer components. DellaVigna and Gentzkow (2019) and Hitsch, Hortaçsu, and Lin (2019) report that many supermarket chains engage in near-uniform pricing and promotions across outlets located in zip codes with very different household income levels. Doraszelski, Lewis, and Pakes (2018) uses models of learning to predict the evolution of pricing in a newly opened electricity market. Finally, other work also has shown that larger firms perform better than smaller firms (Bloom and Van Reenen 2007 and DellaVigna and Gentzkow 2019).

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

The paper proceeds as follows. Section I describes the Texas electricity market. Section II introduces the data and provides descriptive evidence that motivates our modeling assumptions. Section III describes the cognitive hierarchy model and Section IV introduces our model of non-Nash bidding. Section V discusses identification, estimation, and results. Section VI discusses a reduced-form test of strategic versus non-strategic behavior using a nuclear plant shutdown, which provides additional support for our modeling assumptions. Section VII shows that our model predicts bidding out-of-sample. Section VIII measures the efficiency impact of increases in strategic behavior and of counterfactual mergers. Section IX concludes.

I. 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 power plants sell wholesale power to transmission and distribution utilities that serve customers. The wholesale market allowed power trading via both bilateral transactions and an organized spot auction. This paper focuses on competition in this wholesale market.

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

The second market for wholesale trading is an organized "day-of" spot market that is run by the grid operator to ensure that production and consumption exactly balance at every point in time. For example, suppose that a summer afternoon turns out to be hotter than 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 firm's day-ahead schedule. In contrast, if total demand is smaller than the aggregate day-ahead schedule, then winning bidders decrease, or "DEC," production relative to the firm's day-ahead schedule. During our sample period, approximately 2–5 percent of all power transactions occurred in the balancing auction. We study bidding behavior in this auction. Although the balancing auction is small relative to the bilateral market in percentage terms, we show below that substantial profits can be earned with "sophisticated" bidding into the auction.

The auction proceeds as follows. On the supply side, each firm bids a set of price-quantity pairs to create a bid function. The bid function is a step function, and the firm is allowed to submit up to 40 steps in its hourly bid function. For example, a firm may have scheduled 2,000 MW to be produced in a given hour. In the balancing auction, the firm is bidding to, say, increase production by 100, 200, or 300 MW (to 2,100, 2,200, or 2,300 MW of total production) or to decrease production from the 2,000 MW that was scheduled.[5] Bid functions are not tied to specific generating

[5] To be precise, 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).

units; rather a firm's bid function represents 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. However, the price that customers pay, the retail price, is not tied to the hourly wholesale price in the balancing auction. For this reason, the balancing demand function for each hour is perfectly inelastic.

The grid operator accepts supply bids from generating firms, observes the total balancing demand, and clears the market. The format of the auction is a multi-unit, uniform-price auction. Therefore, the grid operator finds the market-clearing price where aggregate supply bids (a monotonically increasing step function) intersects the perfectly inelastic balancing demand. Each firm is called to supply to the balancing market the quantity that was bid at the market-clearing price, and it is paid the market-clearing price for all power called to produce in the balancing auction. Thus, if a firm is called to increase production from its day-ahead schedule, it is paid the market-clearing price for its incremental production. If a firm is called to decrease production from its day-ahead schedule, it purchases power at the market-clearing price to meet any existing contract obligations.

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

II. Data

We study firm bidding behavior into the balancing auctions in an early year of the market's operation. Specifically, we study the first half of the second year of the market, as depicted in Figure 1.[6] Because our sample period begins in the second year of the market's operation, the firms had time to build up their trading operations and develop bidding strategies by the time that our sample begins. By the beginning of our sample, firms had submitted bids into the balancing market for every hour of every day for one year.

[6] 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. We end the sample period in January 2003 because a major ice storm hit Texas in February 2003 that caused large disruptions of the electric transmission grid and general market operations.

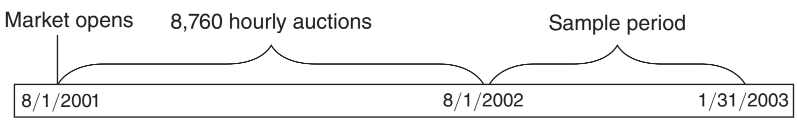

Figure 1. Market Time Line and Our Sample

Note: The figure presents the sample period considered in estimation.

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

Total balancing demand is perfectly inelastic because virtually no consumers face wholesale prices during the time of our study. Our balancing demand data are the hourly demand functions that were used by the grid operator to clear each auction. The bid data consist of each firm's bids to increase and decrease production relative to the firm's day-ahead schedule.

A key feature of our empirical strategy is that we measure each firm's marginal cost of supplying power to the balancing auction. We focus on the 6 PM hour of weekdays because the generators online during this interval are the most flexible type of generators that can respond to balancing calls without large adjustment costs.[7]

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.[8] 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 firms do not misstate their capacity.) In addition, we incorporate that firms cannot reduce generation below a minimum operating level.

The primary variable cost for electricity generation is fuel. For each natural gas and coal-fired unit, we have data on the "heat rate," the rate at which the generator converts the energy content of the fuel into electricity (Henwood Energy Services). Fuel costs for natural gas units are the daily natural gas spot prices at the nearest trading hub in Texas (Natural Gas Intelligence) plus a distribution charge. For coal units, we use the monthly average spot price for coal delivered to Texas (EIA). Variable costs also include a variable operating and maintenance cost per MWh

[7] 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. 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.

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

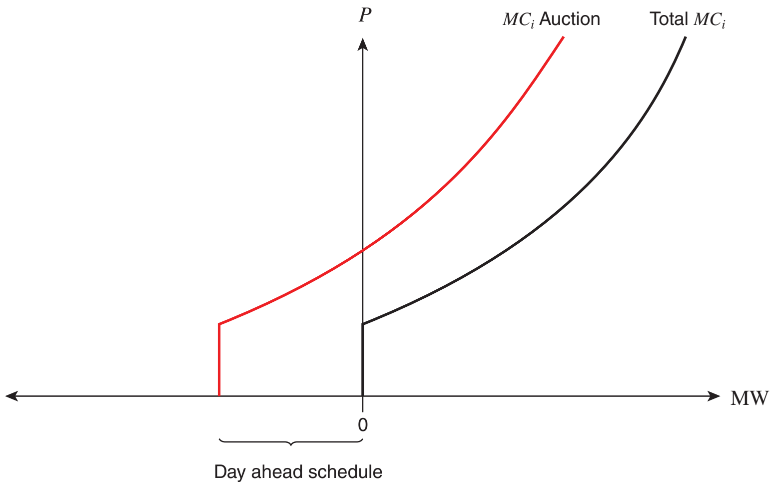

Figure 2. Stylized Marginal Cost of Supplying to Balancing Market

Note: The figure presents a stylized representation of how the marginal cost curve is shifted by the day-ahead schedule.

(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 refer the reader to Hortaçsu and Puller (2008, online Appendix B) for further details.

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

In our model below, marginal cost is public information. While this assumption may not hold in many industries, it holds in the electricity industry because the production technology is very similar across power plants 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 likely know when major generating units are on- or off-line; some firms purchase data from an energy information company that measures real-time output using remote sensors installed near power plants.

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

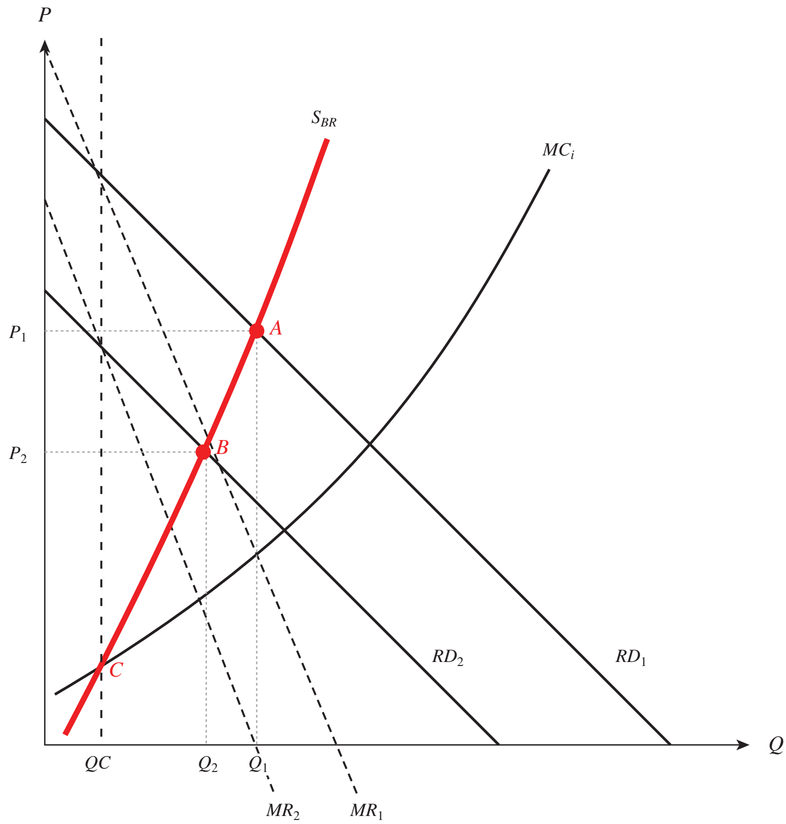

Figure 3. Best-Response Bidding in Spot Auction

Note: The figure presents a stylized representation of how best-response bids can be computed.

our sample to weekdays during the six-month sample period when there was no transmission congestion, we study 99 auctions.

A. *Descriptive Evidence*

As motivating evidence for our model, we show evidence that firms are not best-responding to the bidding behavior of their rivals. In fact, if each firm were to use publicly available data on recent bids, profits would significantly increase. To do so, first we explain how bids would be chosen if firms best respond to their rivals' actions.[9] Figure 3 explains the basic intuition of best-response bidding. Suppose that a firm has marginal cost of supplying to the balancing market given by MC_i. And the firm has signed forward contracts to supply an additional QC_i units of power beyond its day-ahead schedule. 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

[9] We characterize a formal model of bidding in Section IV.

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 markup will depend on the firm's residual demand elasticity. The residual demand function RD_i is equal to the total market demand minus the supply bids by all other bidders. Suppose that it is a hot day and the firm faces RD_1. A profit-maximizing firm will bid a quantity corresponding to the point where Marginal Revenue equals Marginal Cost ($MR_1 = MC_i$) and the price on the (inverse) Residual Demand function at that quantity. This is given by point A. Alternatively, it could be a cooler summer day so that total demand is lower and thus residual demand shifts 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 S^{BR}.[10]

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

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

Panel A of Figure 4 displays a 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 capacity to increase production relative to its day-ahead schedule by about 1,800 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.[11] But it will be a net buyer for quantities below 600 MW, so it has incentives to bid below marginal cost in order to drive down the market price. As seen by comparing the *Best-response bid* and *Actual bid curve*, the firm is bidding in a manner that very closely resembles best-response bidding.

However, other firms deviate from best-response bidding, and the magnitude of the deviation varies in the size of the firm. Panel B of Figure 4 displays a bid function for another firm with substantial generation capacity. This firm has a contract position of about 500 MW upon entering the balancing market, so best-response bids are above (below) marginal cost for quantities greater (less) than the contract position. The firm's actual bid function deviates from the best-response bid. For quantities below the contract position of about 500 MW, the firm submits bids at

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

[11] In Section IV, we show how we measure the contract position in our data.

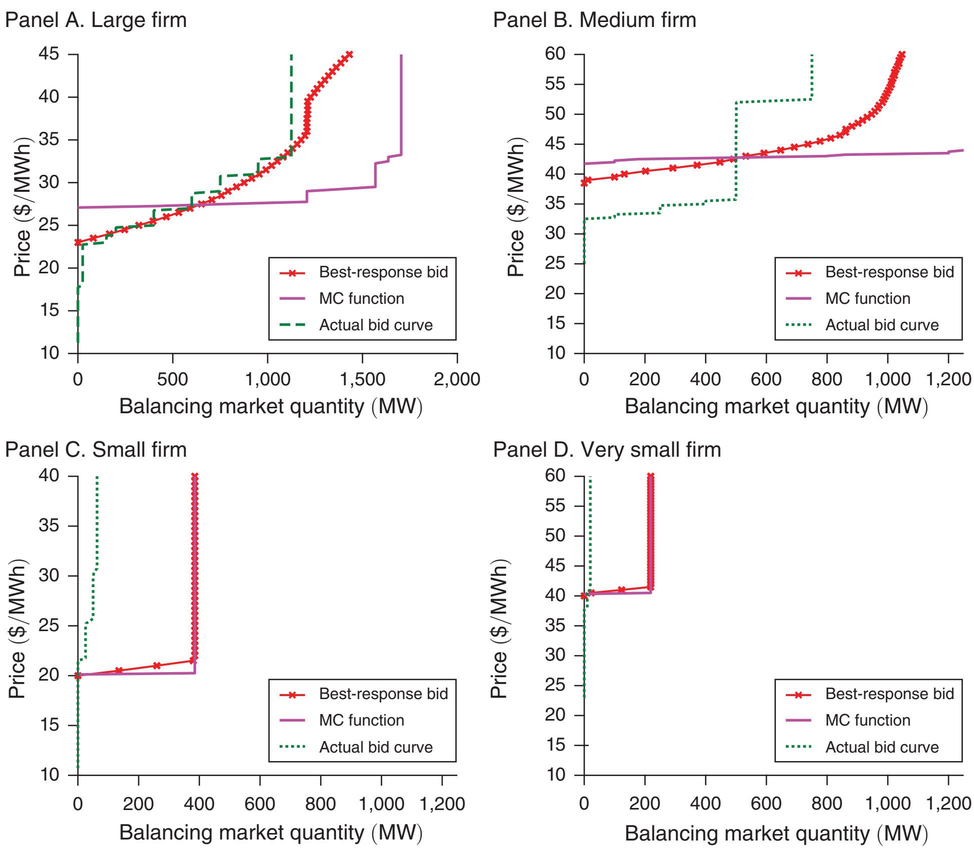

FIGURE 4. ACTUAL BIDS VERSUS BEST-RESPONSE BIDS FOR DIFFERENT-SIZED FIRMS

Notes: Each figure plots one example of the bidding behavior of a firm in an auction. *MC function* displays the firm's variable cost of supplying incremental power to the balancing market. The profit-maximizing *Best-response bid* is computed using the method described in Section IIA. *Actual bid curve* is the bid that the firm submitted into the auction.

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 500 MW, the firm submits bids at prices higher than the best-response bid prices. Loosely speaking, the firm submits a bid function that is "too steep" relative to the best-response bid. The firm's actual bid would 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.

Panels C and D of Figure 4 show 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 panel C 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 panel D bids in a similar fashion: only small quantities are offered into the market even though the firm has available capacity from power plants that are already operating.

Firms that submit bids that are "too steep" relative to best-response bids have two important consequences for the market. From the perspective of the firm, the bids effectively price the firm out of the market which reduces producer surplus relative to best-response bidding. But, importantly, the bidding behavior reduces the efficiency of the market. In some auctions, the firms exhibiting this type of bidding behavior have low-cost power plants 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 because higher-cost generators must be called to produce instead. As we document below, this productive inefficiency can be sizable.

Constructing similar figures for other firms in our sample generates systematic patterns. In particular, a few firms bid very close to "vertical bids" where very little generation capacity is offered into the market. Other firms offer substantial generation quantities into the market but offer that capacity at prices that are above the best-response prices. Also, for any given firm, the shapes of bids relative to best-response bids are very persistent across auctions; firms do not go back and forth between bidding "too steep" and bidding "too flat." Importantly, these deviations from best-response reduce profits, despite the fact that a simple trading rule based on publicly observable information could increase profits, as we show below.

These patterns in bid behavior create a puzzle: why do firms exhibit heterogeneity in bidding behavior relative to a benchmark of best-response bidding? Firms systematically submit bids that span a wide range, from close to the best-response benchmark to bids that are "too steep" to bids that are nearly vertical. These patterns motivate 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.

B. *Ruling Out Alternative Explanations*

Before describing our model, we explore possible alternative factors that could explain the observed behavior. After discussing these and arguing that none of them can rationalize our data, we turn to the cognitive hierarchy model.

First, we ask whether the potential profits in the balancing market are enough to justify setting up a bidding operation. We use observed bids to compute realized profits for each firm in each auction. Then, as in HP, we calculate profits under two scenarios: (i) best-response bidding and (ii) bidding vertically at the contract position (which is effectively not participating in the balancing market except to meet contract obligations). With these numbers in hand, we compute the fraction of potential profits that were achieved by actual bidding relative to non-participation. The results are summarized in online Appendix Table A.1 and show that, with the exception of the largest firm Reliant, none of the other firms achieved more than one-half of potential profits. One can calculate the profit increases of best-responding relative to bidding vertically at the contract position. If one conducts a back-of-the-envelope calculation to extrapolate these profits to every hour of a year, the increase in profits relative to bidding vertically ranges across the firms from \$4.3 million/year to \$37.4 million/year, which appears high enough to justify the labor and capital costs of a bidding operation.

Second, we show descriptive evidence that the phenomenon of offering small quantities into the market by submitting bids that are "too steep" is equally prevalent in the first and second year of the market. Specifically, for each firm-auction, we calculate the amount of generation capacity that the firm offers relative to the contract position at the market-clearing price, and we test whether firms offer additional generation into the market in the second year. The results, reported in online Appendix Table B.1, show that firms offer essentially the same capacity in the second year as they did in the first one.

Third, having shown that profits "left-on-the-table" are significant and that the behavior is persistent, one might expect that some firms that are forgoing profits will eventually exit the market or be acquired. While we cannot rule out such evolution of ownership in the long run, evidence suggests that any such dynamics are slow in this market. Among the firms that we analyze, only one firm was sold by 2005.[12] Thus, while we cannot rule out longer-term market responses to the foregone profits, it is clear that the firms that are deviating from best-response continue to be market players after four years of market operations.[13]

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

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

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

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

constraints were a driver of our main findings, we expect to find a strong positive correlation between ramping capability and size. We find no strong correlation; in fact, the correlation is -0.205, suggesting that if anything, larger firms have slightly *less* ramping capability.

Second, one might be concerned that bidding rules prevent firms from submitting bids that correspond to the best-response bids. The best-response benchmark in the descriptive evidence assumes that uncertainty shifts rather than pivots 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 (Hortaçsu and Puller 2005) includes moment-based tests for expected profit maximization. We find strong evidence that the form of uncertainty and bidding rules do not bias our best-response bids as a benchmark for expected profit-maximizing behavior.

The most straightforward evidence is that a simple trading rule would have systematically raised profits for all but the largest firm. This trading rule uses only information available at the time that bids are submitted and it respects all auction rules about the shapes of bid functions. The trading rule exploits an institutional feature of the Texas market: the grid operator publicly released the aggregate bid schedule with a two-day lag. Thus, firms can learn recent information about their rivals' aggregate bid behavior. Suppose a firm were to use the lagged aggregate bid data to create best-response bid functions to rivals' bids from three days prior to each auction, and submit these bids to the current auction. We compute these lagged best-response bids and use the bids to clear the market with the actual (step function) residual demand for the current auction. We find that this simple trading rule significantly outperforms the actual realized profits for all but the largest firm. The results of this test are reported in online Appendix Figure C.1.

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

Fourth, we show in online Appendix D that the observed deviations from best-response bidding are not driven by mismeasurement of firm capacity. Moreover, additional markets for ancillary services such as "regulation up" cannot explain all of the observed deviations from best-response because the total amount of ancillary services procured is significantly lower than the joint quantity of excess capacity that firms have available when submitting balancing market bids.[14]

Finally, it seems implausible that collusion explains the deviations from best-response. Many of the firms, especially the small firms, that deviate from

[14] Also, prices in the ancillary services markets were relatively low: prices were below \$20/MW in 95 percent of the intervals during the first 1.5 years of the market (Baldick and Niu 2005).

best-response submit bids that are vertical at the contract position. As a result, revenues are zero. Thus, a collusive regime would require side payments from the few large firms to numerous small firms.

III. Empirical Strategy to Estimate a Cognitive Hierarchy Model of Bidding

A. *Background on Cognitive Hierarchy*

The theoretical literature has developed a rich set of models of boundedly rational strategic behavior that can explain deviations from Nash Equilibrium play. Generally speaking, bounded rationality models relax one of the two conditions of Nash Equilibrium: (i) players maximize expected payoffs given beliefs about their rivals' actions and (ii) 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.[15] These models conceptualize players as having a hierarchical structure of strategic, or level-k, thinking. Seminal work on level-k models include Costa-Gomes, Crawford, and Broseta (2001); Camerer, Ho, and Chong (2004); and Crawford and Iriberri (2007).

Cognitive hierarchy (CH) developed by Camerer, Ho, and Chong (2004) conceptualizes players as engaging in different levels of strategic thinking ordered in a hierarchy. The least sophisticated players, 0-step players, engage in no strategic thinking, while higher types assume that all other players are distributed between 0-step and $(k-1)$-step players according to a Poisson distribution.[16] 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.[17]

B. *Big Picture of Modeling and Estimation Strategy*

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

[15] 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. Thus, it does not appear that our bidders estimate expected payoffs in an unbiased way, a key feature of the QRE model.

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

[17] 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 $(k-1)$.

One critical feature of estimating a CH model is how to define level-0 behavior. In the theoretical literature, a convenient assumption is that level-0 players uniformly randomize across all possible strategies, however Camerer et al. note that this is a placeholder assumption that can be modified to specific settings.[18] One conceptualization of a behavior that falls into level-0 is taking a very salient action that requires "low mental effort." In the context of the Texas electricity auctions, there is a natural assumption about non-strategic thinkers: bidding "vertically" at the contract positions for the range of plausible prices. (That is, the firms submit bids similar to panel D of Figure 4.) Our level-0 firms realize that they need to "true up" their contract position, so they bid so as to satisfy their contract position (a factor that they know). However, such firms do not view the balancing auction as a profit-making opportunity worthy of "strategic energy" to develop a sophisticated bidding strategy. A vertical bid at the contract position essentially indicates that even at very high prices, the firm does not want to sell power into the balancing market. Vertical bidding at the contract position is also empirically motivated in this setting because the firms that submit nearly vertical bids earn the lowest fraction of potential profits. Thus, "vertical bidding at the contract position" is a natural candidate for level-0 bidding behavior in this setting.

Our modeling choice about level-0 behavior is important because it anchors the bidding of higher types that respond to level-0 firms. While other types of bidding behavior are candidates for level-0 behavior, in Section VA we discuss the implications of such bidding patterns. Alternative level-0 assumptions are not consistent with the observed behavior in this market, as we show in Section VA. Moreover, in Section VI, we provide additional empirical evidence supporting the vertical bidding assumption.

One advantage of using this level-0 assumption is that we do not need to make strong assumptions about the form of the bid functions. Instead, as we show below, the assumption of level-0 firms 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 only a subset of firms enter into the cognitive hierarchy, while the rest form part of an unmodeled fringe. We do this because allowing for more firms makes the problem computationally challenging as each firm needs to compute its rivals bidding functions for all possible types, for all auctions. Furthermore, we do not have marginal cost data for all firms for all auctions, which imposes a constraint on the number of firms that we can include in the CH. For this reason, we select 12 firms to model with CH: these firms vary in firm characteristics and size, and jointly represent 66 percent of industry capacity. Nonetheless, in online Appendix E, we show that our baseline estimates are robust to this sample selection process. To do this, we sequentially add firms to the cognitive hierarchy and re-estimate our model. The outcome of this exercise is reported in online Appendix Figure E.1 and shows that our estimates are robust to the way in which we define the sample of firms to model as strategic players.

[18] For example, Goldfarb and Xiao's level-0 firms assume that no other firms enter the market.

Once level-0 bidding is defined, we can use our data on each firm's marginal cost to calculate the bidding behavior for a firm of any type $k > 0$. Firms observe characteristics of their rivals X_{-i} (e.g., size) and form beliefs about their rivals' type distributions. Specifically, the type distribution is given by $Poisson(\tau_i(X_i, \beta))$ where $\tau_i(\cdot)$ captures how firm characteristics map into type.[19] For any mapping from firm characteristics to type, we use an iterative procedure to calculate each player's optimal theoretical bids under various sophistication levels. Given level-0 players' bids, we calculate level-1 best-response bids for each firm. For each firm, level-1 bids are calculated as the best-response to level-0 bids by all the other CH firms and the bids by the unmodeled fringe. Then given our calculated level-0 and level-1 best-response bids, we calculate the level-2 best-response bids for each firm, and continue this recursive process up to the highest type K.

We then compare these calculated bids to the firm's actual bidding behavior. The estimation process finds the parameters of $\tau_i(X_i, \beta)$, how firm characteristics such as size affect strategic sophistication, that minimize the distance between actual bids and the bids predicted under CH. That is, in estimation, we use observed bids and realized marginal costs to recover the type of each firm. For this reason, it is critical that we observe marginal costs; in the absence of cost data, one would not be able to identify types from bid data without additional assumptions regarding the cost function.[20] In other words, instead of using data on observed bids and a Bayesian Nash equilibrium model of behavior to recover costs, we use data on costs and bids to recover the type that rationalizes observed behavior.

IV. Formal Cognitive Hierarchy Model of Bidding

This section formulates best-response bidding in a setting where firms have beliefs about rivals as characterized by the cognitive hierarchy model. We incorporate a model of bidding into share auctions (Wilson 1979 and Hortaçsu and Puller 2008) into the Poisson cognitive hierarchy model (Camerer, Ho, and Chong 2004).

Demand for power in each spot auction is given by $\tilde{D}_t(p_t) = D_t(p_t) + \varepsilon_t$ which is the sum of a deterministic and stochastic component.[21] The auctions occur in a private values setting where the private value is the firm's variable cost of providing power to the grid. Firm i has costs to supply power in period t given by $C_{it}(q)$. Prior to the auction, each firm has signed contracts to deliver certain quantities of power each hour QC_{it} at price PC_{it}, and we take these contracts to be predetermined. Note that $C_{it}(q)$ is modeled as public information because, as discussed above, in this industry power plant characteristics are public information and firms know when major generating units are online. In contrast, QC_{it} is private information because

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

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

[21] Although demand in the spot auction is inelastic (the retail price is not determined by the hourly wholesale price), we model demand as a function of the spot price because in estimation we model bidding behavior of a subset of firms that face a residual demand that is net of supply by an unmodeled fringe.

firms do not publicly disclose the terms of their forward contracts. As far as strategic ability, each firm is a k-step thinker. Firm i has private information on its own type k_i, but it only knows the distribution from which rival types are drawn. In each auction, firms simultaneously submit supply schedules $S_{it}^{k}(p, QC_{it})$ to produce different quantities at different prices. Let the bid function by rival j of type l be denoted $S_{jt}^{l}(\cdot)$.

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

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

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

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

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

The event that the market-clearing price p_t^c is less than any given price p is the event that there is excess supply at that p. Plugging the market-clearing condition (equation (1)) into (equation (2)), and letting $\mathbf{l}_{-i}$ represent the vector of types of i's rivals:

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

where $F\big(\mathbf{QC}_{-it}, \mathbf{l}_{-i}, \varepsilon_t \,|\, \hat{S}_{it}(p), k_i, QC_{it}\big)$ 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

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

This profit is 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 be risk averse or risk neutral. 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 solve

$$\max_{\hat{S}_{it}(p)} \int_{\underline{p}}^{\bar{p}} \Big(U\Big(p \cdot \hat{S}_{it}(p) - C_{it}\big(\hat{S}_{it}(p)\big) - (p - PC_{it})QC_{it}\Big)\Big)\, dH_{it}\big(p, \hat{S}_{it}(p); k_i, QC_{it}\big).$$

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

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

where H_s and H_p are given by derivatives of equation (3).

There is a simple intuition behind this condition. To see this, for the moment ignore the term H_S/H_p (it will be positive). The left 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 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 in auction t if the firm is type k: the best-response bid function for each firm given its type. 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 data and three identifying assumptions to "measure" each component of equation (5), which allows us to calculate the best-response function for each firm i in auction t for each type k. The first assumption defines the bidding behavior for type-0 bidders. The assumption has both the properties that it is natural in our setting and that it facilitates computation of CH outcomes by allowing us to solve the problem recursively. This recursive property yields an additive separability condition that makes it computationally straightforward to solve the firm's expected profit maximization problem, as we show below. The second and third assumptions define the distribution of types of the CH model and the distribution of the remaining sources of uncertainty.

Type-0 Bidding:

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

$$S^0_{it}(p, QC_{it}) = QC_{it}, \quad \forall p \in [\underline{p}, \bar{p}], \forall i \in \mathbf{I}_0,$$

where $\mathbf{I}_0$ *represents the set of bidders type* 0.

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

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

$$\begin{aligned} &H_{it}\big(p, \hat{S}_{it}^1(p); k_i = 1, QC_{it}\big) \\ &= \int_{\mathbf{QC}_{-it}\times \mathbf{l}_{-i}\times \varepsilon_t} \mathbf{1}\Big(\sum_{j\neq i} S_{jt}^0(p, QC_{jt}) + \hat{S}_{it}^1(p) \\ &\qquad\qquad \geq D_t(p) + \varepsilon_t\Big)\, dF\big(\mathbf{QC}_{-it}, \mathbf{l}_{-i}, \varepsilon_t \,\big|\, \hat{S}_{it}^1(p), k_i = 1, QC_{it}\big) \\ &= \int_{\mathbf{QC}_{-it}\times \mathbf{l}_{-i}\times \varepsilon_t} \mathbf{1}\Big(\sum_{j\neq i} QC_{jt} - \varepsilon_t \\ &\qquad\qquad \geq D_t(p) - \hat{S}_{it}^1(p)\Big)\, dF\big(\mathbf{QC}_{-it}, \mathbf{l}_{-i}, \varepsilon_t \,\big|\, \hat{S}_{it}^1(p), k_i = 1, QC_{it}\big) \\ &= \int_{\mathbf{QC}_{-it}\times \mathbf{l}_{-i}\times \varepsilon_t} \mathbf{1}\big(\theta_{it} \geq D_t(p) - \hat{S}_{it}^1(p)\big)\, dF\big(\mathbf{QC}_{-it}, \mathbf{l}_{-i}, \varepsilon_t \,\big|\, \hat{S}_{it}^1(p), k_i = 1, QC_{it}\big), \end{aligned}$$

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, because 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 is the distribution of θ_{it}.

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

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

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

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

The implication is that if type-0 bidders submit perfectly inelastic bids at their contract positions, then the bids of type-1 bidders are additively separable functions of price and private information on their contract positions. Bid functions will take the form $S^1_{it}(p, QC_{it}) = \alpha^1_{it}(p) + \beta^1_{it}(QC_{it})$, as shown in online Appendix F. This additive separability property is valuable because it implies that type-0 rivals' private information about their contract positions does not affect a firm's residual demand slope, as we show below. Unlike type-0 bidders who bid vertically, type-1 bidders submit bids that offer some quantity into the market.

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

(6)

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

where, as before, $\theta_{it} \equiv \sum_{j\neq i} QC_{jt} - \varepsilon_t$, but $l_j \in \{0,1\}$. In this way, we can write H_{it} just as before but taking into account that θ_{it} corresponds to the difference between the sum of contract position by rivals and ε_t.

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

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

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

Next, we make assumptions about $\Delta_i(\cdot)$ and $\Gamma_i(\cdot)$. For the distribution of types, $\Delta_i(\cdot)$, we follow the literature and assume that the distribution is Poisson. This distribution has a single parameter that characterizes firm types and has several intuitively appealing game theoretic properties, as described in Camerer, Ho, and Chong (2004). Specifically, we assume the following.

ASSUMPTION 2: $\Delta(\cdot)$ *is an independent multivariate Poisson distribution truncated at* $k-1$, *as given by the Poisson cognitive hierarchy model.*

Finally, we assume Γ_i to be a uniform distribution. We make this assumption primarily for computational convenience. Although other distributions could be used, this would increase computational cost significantly.[22] Thus, we assume the following.

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

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

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

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

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 calculate the $\hat{S}^k_{it}(p)$ that solves the equation above. For each firm-auction, we observe the bid function's price, marginal cost $\big(C'_{it}\big(\hat{S}^k_{it}(p)\big)\big)$, and the slope of residual demand $(-RD'_t(p))$. And we can measure the firm's contract position in the balancing market (QC_{it}) by calculating the quantity at which the bid function intersects the marginal cost function. The rationale for this method to measure the contract position is as follows. For type $k > 0$, the first-order condition above says that the firm bids a price above (below) marginal cost for quantities greater (less) than the contract position, and the markup (markdown) depends on beliefs about the shape of residual demand which is driven by beliefs about rival behavior.[23] For type-0 firms, we identify the contract position by assuming that type-0 firms use the balancing market to true-up any residual contract position not satisfied with the day-ahead schedule. Note that this assumption is consistent with non-strategic behavior; it says that the firm's action in the auction plays the role of fulfilling the contract position but does not depend on any beliefs about how rival firms behave.

[22] The researcher needs to solve the first-order condition looking for a fixed point in $S^{l_j}_{it}(p)$ for every bidder i of type l in auction t at point p. This is computationally very expensive in a setting with 12 firms. For this reason, instead of reducing the number of firms included in the CH and estimating the model using a different distribution Γ_i, we examine the extent to which our assumption may impact our results in a different way. Specifically, we use the estimated parameters that we report in Section VC and predict bidding behavior assuming that Γ_i is uniform and that Γ_i is Normal and compare the distribution of predicted bids. The results, presented in online Appendix G, show that the differences are minimal.

[23] This method of identifying the firm's contract position is shown formally in Proposition 1 of HP.

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.

In our implementation of the model, 12 firms are included in the cognitive hierarchy and the remaining firms are included as an unmodeled fringe, as we discuss above. Thus, the $D(p)$ in the theoretical model corresponds to balancing demand (which is perfectly inelastic) *net* of (elastic) supply by the unmodeled firms. In this context, our model would imply that there is no uncertainty about the *slope* of the supply of the unmodeled firms. While it may not be literally true that there is no uncertainty, to a first-order, there is not economically significant slope uncertainty, as we show in online Appendix H.

V. Model Estimation and Results

A. *Identification*

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

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

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

The identification of the parameters of firm type is intuitive. For example, consider the case with a single firm attribute X_i, so that the index of the $\tau_i(\cdot)$ function is $\exp(\gamma_0 + \gamma_1 X_i)$. First, consider the variation in the data that identifies γ_1 when

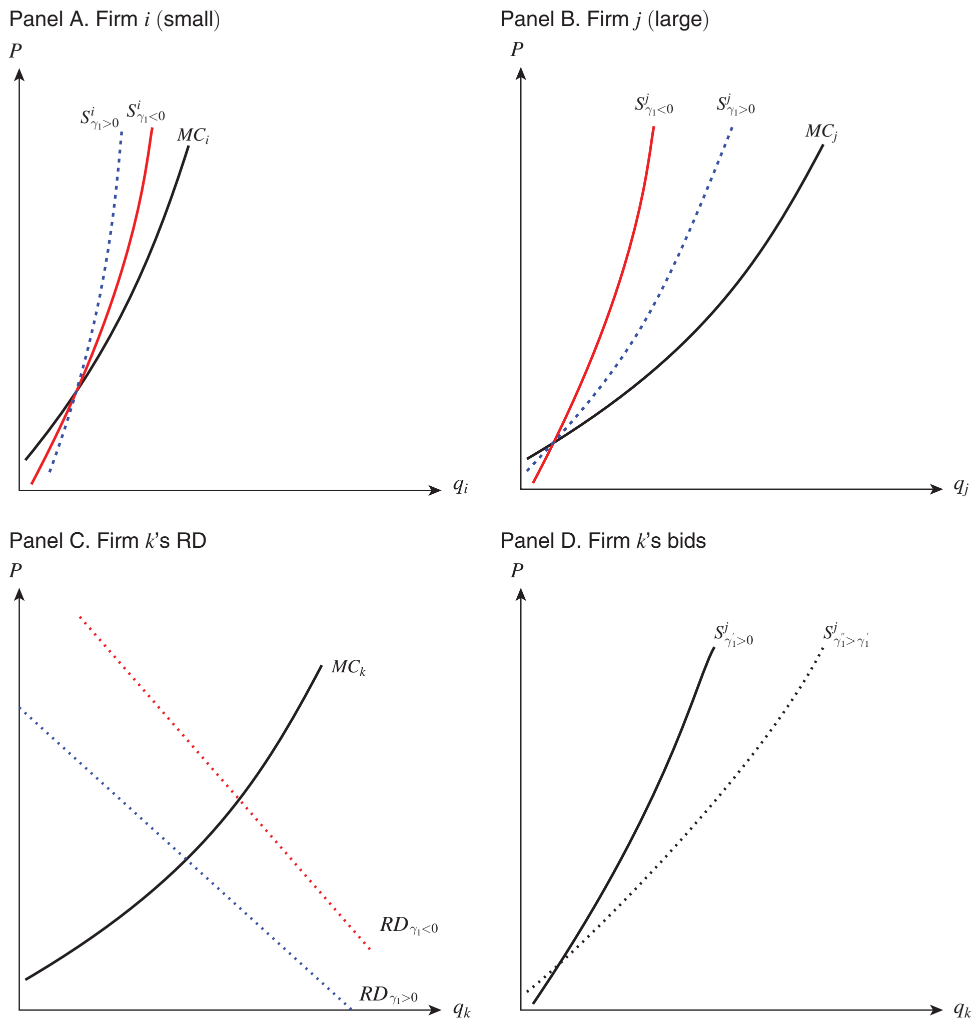

Figure 5. Example of the Identification Strategy

Notes: The figure presents the identification strategy using a stylized example with three firms. The figure explains how bids of firm k can be used to infer k's beliefs regarding the type of its rivals i and j.

we hold γ_0 fixed. To make the presentation clear, Figure 5 presents the identification strategy considering three firms, i, j, and k, and explains how k's bids can be used to identify γ_1, holding γ_0 fixed. Suppose that X_i is firm size so that $\gamma_1 > 0$ implies that larger firms are higher types and $\gamma_1 < 0$ implies that larger firms are lower types. Under CH, the observed bids are best-responses given beliefs about the residual demand function that firm k faces. For simplicity, suppose firm k faces a small (i) and a large (j) firm. If $\gamma_1 > 0$, firm i will offer less capacity into the market than firm j, both because it has less capacity to offer than j and because it is a lower type than j. These offers are presented in panels A and B of Figure 5. This will cause firm k to believe it faces a relatively elastic residual demand, presented in panel C of Figure 5 as $RD_{\gamma_1>0}$. In contrast, if $\gamma_1 < 0$, firm k believes that the large rival will not offer much capacity into the market, but rather that the small rival will do so. That is, k believes that firm i will submit $S^i_{\gamma_1<0}$, and that firm j will submit $S^j_{\gamma_1<0}$, presented in panel A and panel B of Figure 5, respectively. This causes the residual demand faced by firm k in this case ($RD_{\gamma_1<0}$ in panel C) to be more inelastic than the one it faces when $\gamma_1 > 0$, because the only rival firm with substantial capacity is a low type and low-type firms offer less capacity.

The arguments presented above imply that we can use k's bids to identify the relationship between size and firm type, as the identification question to be addressed is whether k's bids are more consistent with best responding to the relatively elastic ($RD_{\gamma_1>0}$) or relatively inelastic residual demand ($RD_{\gamma_1<0}$). Point identification of γ_1 is achieved by exploiting both the variation in firm characteristics X_i *and* that τ is a continuous function of γ_1. This is, if $\gamma_1' < \gamma_1''$, γ_1'' induces a flatter bid submitted by k than γ_1', holding X_i fixed for all i, because a higher τ_i shifts the type distribution to the right.

Consider now the identification of γ_0. This parameter is identified as long as not all firms bid as level-0 firms do. To see this, assume that firms can be classified into two groups. One group bids vertically (e.g., level-0), while another group bids in some capacity into the auction. Then, holding γ_1 constant, γ_0 allows the model to explain the fraction of firms that follow level-0 behavior. Therefore, if a firm that was classified in the vertical bidding group were to bid some capacity into the market (i.e., it is reclassified into the second group), this will be reflected by increasing γ_0 to rationalize the lower fraction of type-0 firms.

Implications of the Level-0 Assumption.—As discussed in Section IIIB, our assumption about the bidding behavior of level-0 players is supported by the behavior that is observed in the data. Nonetheless, it is important to discuss how alternative assumptions would impact our results. To organize our discussion, we divide alternative assumptions for level-0 behavior in two groups. Consider Figure 6, which depicts a simplified version of best-response bidding. One category of level-0 behavior yields low types bidding vertically and higher types submitting flatter bids, or loosely put, higher types approaching best-response bidding "from the left." Another category yields low types bidding relatively flat (e.g., marginal cost) and higher types submitting steeper bids, or higher types approaching best-response bidding "from the right." As we discuss in online Appendix I, our data refute marginal cost bidding, though we still discuss the implications of this assumption.

First, consider assumptions in which as firm type increases, bids get flatter and approach best-response bids from the left (area A in Figure 6). In this case, profits increase as firms bid flatter. Our assumption about level-0 behavior is one of these. An alternative, also non-strategic, is to assume that level-0 behavior is that of a monopoly. In this case, level-0 players best respond to the residual demand function that results from subtracting fringe supply from balancing demand. Higher type bidders would simply best-respond to level-0 players that assume they are monopolies. This assumption is similar to the one that we make, but it assumes that level-0 players bid with positive slope even for relatively low prices and small quantities, while our assumption is that they bid vertically on their contract positions. Because in our data level-0 bidders are effectively not providing power to the grid, and bid vertically for most prices, we believe our assumption reflects what we observe in a better way. This also implies that

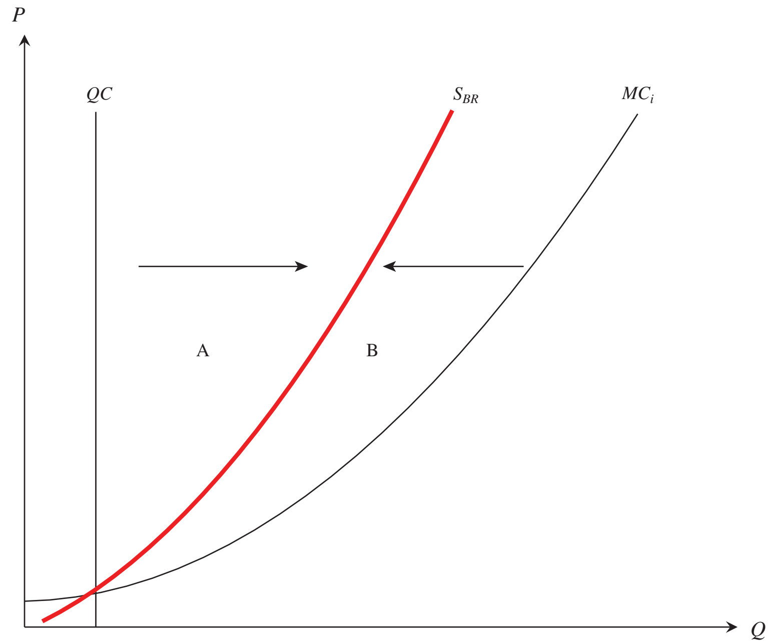

Figure 6. Bids Approximating Unilateral-Best Response as a Function of the Level-0 Assumption

Note: The figure presents the direction in which bids would approximate the unilateral best-response bid depending on the assumption made about level-0 bidding behavior.

if we were to assume that level-0 bidders behave as monopolists, the model would fit the data worse than under the assumption that these bidders bid vertically.

Next, consider alternatives in which as type increases, bids approach best-responses from the right (area B in Figure 6). Some of these alternatives include level-0 players that bid their marginal cost, or that bid a constant or proportional markup over marginal costs (i.e., a rule of thumb). These assumptions are also non-strategic but they result in higher-type rivals bidding steeper than level-0 players and both bids approaching best-responses from the right and profits increasing as firms bid steeper. Because these alternatives result in predicted bids far from those we observe in the data, any assumption that leads to bids approaching best-responses from the right will result in our model performing worse. Because bids would converge to best-responses from the right, the model would never predict bids that are in the neighborhood of those that we observe (i.e., bids would never be steeper than best responses).

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

B. *Details on Estimation*

Estimation follows a minimum-distance approach. Here, τ_i is a scalar that provides information about firm i's type. We assume that $\tau_i = \exp(X_i'\gamma)$ and, because X_i is public information, so is τ_i. Each firm i observes $\boldsymbol{\tau}_{-i}$, the vector of τs of its rivals. Also, each firm i has private information about its own type. Assume firm i is type $k \in \{0, \ldots, K\}$ and the type does not change over time. If $k = 0$, then, by Assumption 1, firm i would submit a vertical bid at its own contract position, regardless of its rivals. For all $k > 0$, firms have beliefs about its rivals' types. Specifically, by Assumption 2, these beliefs are assumed to follow a Poisson distribution truncated at k, meaning that firm i believes all its rivals to be type $k - 1$ or less. The probability associated with each type varies according to each rivals' τ.

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

Once firm i has computed what it expects its rivals to do for each possible type, it maximizes expected profits according to its beliefs about its rivals' types. This results in a bid function, conditional on i's type. Therefore, we compute i's bid for all possible types that firm i can be. Because i's actual type is unknown to the econometrician, we proceed in several steps.

First, we compute bid functions over a grid of price points that are centered at the realized market-clearing price. Second, for each auction t and bidder i, we compute the difference between the realized bid (data) and the bid predicted by the model, at each price point p, for each type k that firm i can be. However, because firms differ in capacity, we scale these differences by the quantity-difference between the predicted bid for each firm for types K and 0. Third, to compute the total difference between a predicted bid function and the bid data, we sum the quantity differences across all the price points in the grid, for bidder i in auction t, when i is type k. Because we are primarily concerned about differences that take place in the neighborhood of the market-clearing price, we weight price points by a triangular distribution centered at the market-clearing price.

Fourth, after we have computed the weighted quantity differences between the predicted bids for firm i (for each of its possible types k) and the bid data, we weight each of these differences by the probability that the firm is each type k. This probability is modeled as a Poisson distribution truncated at the number of possible types considered in estimation (level-0 and 20 levels of strategic sophistication). We use each firms' τ to compute this probability. Finally, we add over firms and auctions.

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

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

In estimation, we use a large grid of initial points to explore the reliability of our estimation routine in converging to the global minimum. Finally, we compute standard errors using 250 Bootstrap samples.

C. *Results: Estimated Parameters*

Results are reported in Table 1. We estimate seven specifications that differ in the observable characteristics (X_i) of the firms that affect firm τ_i. In our baseline specification column, type is determined by firm size. As a metric of size in the balancing market, we seek a metric of the firm's potential stakes in the balancing

Table 1—Structural Model: Estimated Parameters of Type Function

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Constant	−0.726	−0.196	−3.395	−0.749	−3.493	−0.691	−0.675
	(0.031)	(0.037)	(0.269)	(0.049)	(0.796)	(0.041)	(0.072)
Size	14.594	−1.163	25.789	13.619	3.090	11.933	13.776
	(0.301)	(0.464)	(3.385)	(0.603)	(0.847)	(0.546)	(0.629)
Size^2		86.191					
		(4.035)					
Merchant			−1.562				
			(0.295)				
AAU University				0.376			
				(0.059)			
Degree in economics, business or finance					5.626		
					(0.764)		
Economics degree						1.633	
						(0.115)	
Time trend							0.051
							(0.016)
Obj. fn./number of auctions	208.512	208.354	208.526	208.485	206.386	208.308	208.520

Notes: This table reports estimated parameters of our Cognitive Hierarchy model that are estimated using the minimum distance estimator described in Section VB. Each column reports estimates for different parameterizations of $\tau_i(X_i)$. Bootstrapped standard errors are calculated using 250 samples.

market that is exogenous to its realized bidding behavior. We compute the quantity of sales if the firm were to best-respond, averaged across all auctions. This is positively correlated with installed generation capacity.

As shown in column 1, we find that larger firms are higher types.[24] 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 7 plots the estimated type distribution for a set of firms ranging from the smallest to the largest. Consider the smallest firm with a size that is 11 percent of the size of the largest firm: the CDF farthest to the left in the figure. We estimate that the smallest firm has about a 50 percent chance of being type-0, about a 35 percent chance of being type-1, about a 10 percent chance of being type-2, and is higher than type-2 with very low probability. Each of the other CDFs in the figure show the estimated type distribution for other firms, with the larger firms having probability distributions further to the right. Overall, 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. Finally, it is important to note that our estimates are robust to the time period considered in estimation. Specifically, in online Appendix J we show that the relationship between size and type is robust to considering a different period of the day (7–8 pm).

[24] We expect the constant to be negative in order to rationalize level-0 players because 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.

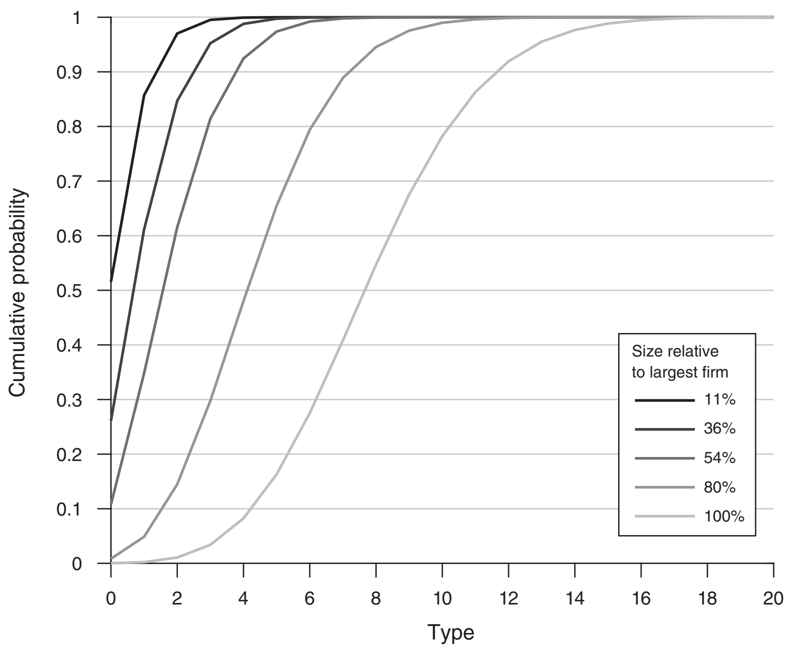

FIGURE 7. ESTIMATED CDFS OVER TYPES FOR FIRMS OF DIFFERENT SIZE (*Size* SPECIFICATION)

Note: The figure presents the cumulative distribution function of types, for four firms in the cognitive hierarchy.

We illustrate the model's fit for the same firms that we use to illustrate bidding in Section IIA. Figure 8 displays actual bid functions as well as bid functions for different type-k models, as predicted by the *Size* specification of our model. We plot the predicted CH bids for type-0 through type-5. The vertical CH bid is, by Assumption 1, submitting a perfectly inelastic bid at the contract position. The remaining bids depict predicted bids for type-1 through type-5, with higher types uniformly flatter. The largest differences between bid functions are for the lowest type; in this example, bids for type-4 and type-5 are very similar. Higher type bids converge rather quickly and we do not see any "cycles" in bids as types increase. Intuitively, because higher types believe their rivals will bid more capacity into the market, the best-response bid to those beliefs is to bid more capacity in the market.

We also explore other specifications of how firm characteristics relate to type. The second specification in Table 1 allows for nonlinearity in how size affects τ_i by adding size squared. The implied distribution of types is qualitatively very similar to our linear specification.

Next we explore if the organizational structure of the firm is associated with higher type. As discussed in Section I, 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 column 3 of Table 1, we test whether merchant firms tend to be higher types. However, we find that if anything, merchants are lower

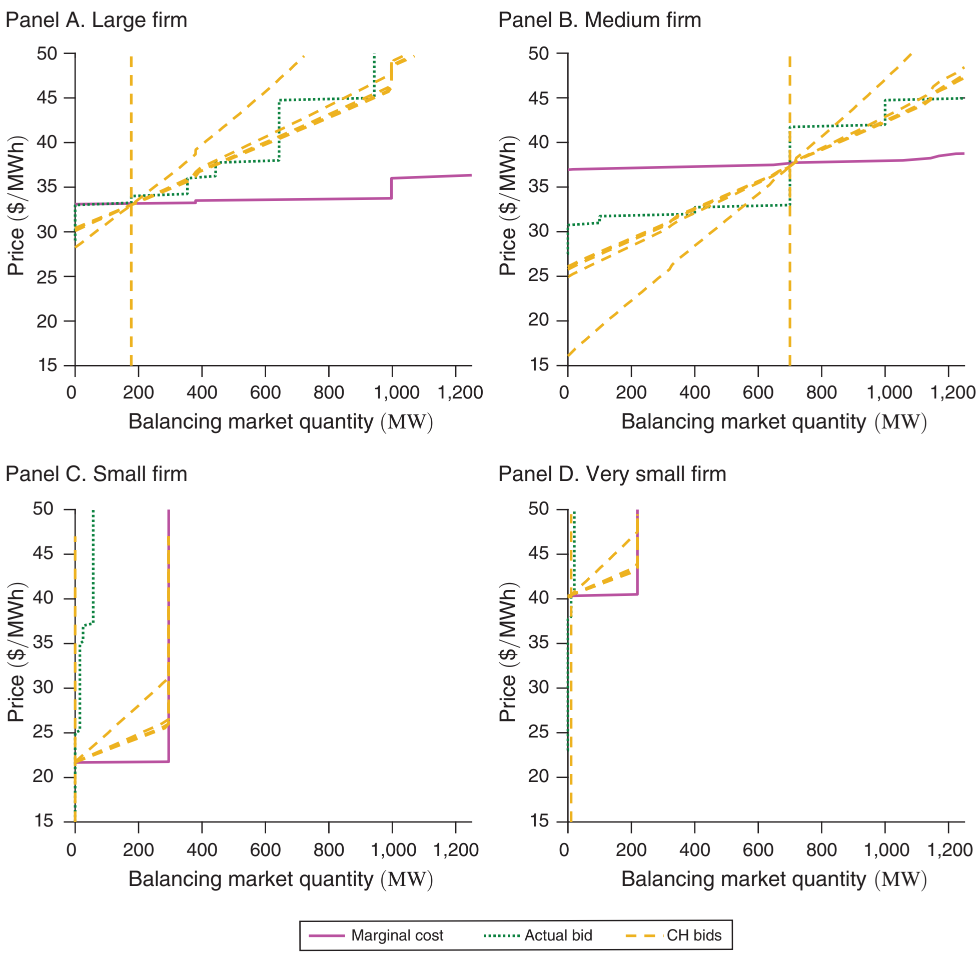

Figure 8. Actual Bids versus Cognitive Hierarchy Bids for Different-Sized Firms

Note: The figures replicate those in Figure 4, adding the bids predicted by the model for different levels k in the cognitive hierarchy.

types than former utilities and municipal utilities. However, the role of organizational structure is substantially smaller than the role of firm size.

In specifications 4–6, we investigate whether the personnel hired to run firm bidding operations is related to firm type. In order to assess the role of personnel, we use LinkedIn and other publicly available online data sources to make the best guess of the manager(s) who were responsible for each firm's power marketing operations at the time. In some cases, job titles were sufficiently clear to identify the power marketing manager, and in other cases we were only able to identify personnel who were involved in firm wholesale power operations. Therefore, the data used for this specification may not be as precise as the data on bids and costs. Nevertheless, this provides 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 a degree from a university that belongs to the American Association of Universities (AAU).[25] Five of the twelve firms have personnel who graduated from an AAU university. We also create a dummy variable for whether any personnel have a degree in either economics, business, or finance. Seven of the twelve firms have personnel with a degree in economics, finance, or an MBA while the most popular other type of degree is in engineering. We estimate our benchmark specification using *Size* and add dummy variables for university type or degree type.

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

Finally, we explore the possibility that firms may learn over time. To do so, we specify τ as a function of firm size and a linear time trend. The results are reported in the last column of Table 1 and show that the estimated coefficients for the constant and size are remarkably similar to those in column 1. Nonetheless, we find a positive and significant coefficient on the time trend, which suggests that firm types do change over time. However, the amount of learning is economically very small and the implied probability distributions over types, for the first and last week of our sample period overlap perfectly (see online Appendix Figure K.1). This suggests that learning is minimal and has no meaningful impact on the estimated probability distributions over types.

D. *Model Fit*

In this section we show that the CH model, with our parameterization of size determining beliefs, fits the actual bidding behavior quite well. To assess the model fit, we use the specification where *Size* determines firm type and subsequent bidding behavior (column 1 of Table 1). It is important to note that this specification is very parsimonious; we are using only one characteristic, size, to fully characterize a firm and to predict bids.

We compare the fit of the CH model to a model in which firms best-respond to their rivals' actual bid data; this best-response model is essentially a model in which firms *individually* best-respond and have consistent beliefs about their rivals' behavior. Table 2 reports results from a regression at the firm-auction level that predicts realized profits under actual bidding with profits under

[25] The AAU includes 62 private and public research universities in the United States and Canada.

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

	CH (1)	BR (2)	Both (3)
Profits under cognitive hierarchy	0.803		0.642
	(0.069)		(0.127)
Profits under best-response		0.428	0.137
		(0.044)	(0.062)
Constant	−328.17	−241.74	−374.17
	(141.98)	(120.72)	(125.79)
Observations	1,058	1,058	1,058
R^2	0.67	0.49	0.69

Notes: This table reports results from a regression of observed profits from actual bidding behavior on firm profits as predicted by the Cognitive Hierarchy model (column 1), firm profits that would be achieved from a model of unilateral best-response to actual rival bids (column 2), and both (column 3). An observation is a firm-auction. Standard errors clustered at the firm level are reported in parentheses.

either the CH model (column 1) or best-response model (column 2). The CH model fits the actual data better than the best-response model. The coefficient of modeled profits is much closer to unity for the CH model ($\hat{\beta}^{CH} = 0.803$ versus $\hat{\beta}^{BR} = 0.428$) and the CH model explains more of the variation in firm profits ($R^2_{CH} = 0.67$ versus $R^2_{BR} = 0.49$). When we include profits under both models of behavior in column 3, we find that once we include CH profits, the best-response profits do not help predict actual profits in an economically meaningful way. Moreover, we view the fit of the CH model as strong given that we are using only a single covariate, firm size, to explain the heterogeneity in behavior across firms.

E. *Interpretation of Model in Context of Management Practices*

In the context of the Texas electricity market, we interpret the CH framework as an "as-if" model capturing the real-world management practices of the market players. Our perspective into management practices of these firms is shaped by interviews we conducted with some of the firms in 2004. Unfortunately, we did not plan a systematic survey, so the information we gathered from these site visits is anecdotal in nature. However, one salient impression is that larger firms had many more resources devoted to trading operations, which included large trading floors, several employees with a PhD, and specially developed software applications to gather market information and compute bids. In contrast, trading operations at smaller firms appeared much more thinly staffed, with employees who appeared to specialize in operational/engineering details of the plant rather than in trading. In the smaller firms, the employees whom we spoke with had less time to devote to balancing market bidding and did not appear to conceptualize the market in terms of "residual demand."

The interviews suggest that the nature of scale/size effects is that larger firms have enough resources to hire technically sophisticated trading staff who have the latitude to reflect upon and optimize bidding operations to improve profits. As we discuss above, because most of the generation volume is sold in forward contracts, the balancing market volume comprises a small share of the quantity of electricity traded. While this does not mean that the balancing market is unimportant in terms of overall profits, the profit potential from this market may appear unimportant to smaller firms, who, in turn, may not devote many resources to "thinking strategically" about this market. This may rationalize strategies like vertical bidding, which essentially means that the firm is avoiding the balancing market altogether.

Although there is no direct mapping from type to business practice, one could conceptualize the correspondence in the following way: level-0 firms may misunderstand opportunity costs and fail to recognize the profit opportunities of selling excess capacity to the balancing market. Low-level types may understand the potential profitability but not invest great effort in back-end software and rather use rule-of-thumb heuristics. And high-level types may invest person-power to estimating residual demand, perhaps using the lagged bid data that the market operator posted. And, finally, because there was persistence in the personnel that ran bidding operations over our sample period, minimal evidence of learning may be present.

VI. A Reduced-Form Test of Strategic versus Non-Strategic Behavior: Evidence from a Nuclear Plant Outage

In this section, we show additional data-driven support for the CH model and our results. In general, in the CH model, level-0 behavior captures non-strategic agents. One type of non-strategic behavior is an agent who does not respond to changes in the (common knowledge) cost structure of its competitors, but may respond to changes in its own costs. Note that this definition of non-strategic behavior encompasses a large array of behavioral patterns that have been considered as level-0 behavior in the literature. Patterns such as vertical bidding, which is our definition of level-0 behavior, bidding marginal cost (or bidding truthfully, as in Crawford and Iriberri 2007, Gillen 2010, and An 2017), bidding a random number that is independent of competitors' costs (again, as in Crawford and Iriberri 2007, Gillen 2010, and An 2017), or behaving as if one is a monopolist regardless of one's competitors (as in Goldfarb and Xiao 2011) are consistent with this definition of non-strategic behavior.

We use a two-month outage at a nuclear plant to test whether firms change bids in response to competitor cost shocks. In the middle of our sample period, one large nuclear generator went off-line for about two months. This event suddenly reduced nuclear output by about 2,300 megawatts. As a result, total demand for power intersected aggregate system marginal cost at a steeper point on the marginal cost function. Firms that behave strategically will recognize that this publicly observable cost shock is likely to make their residual demand in the balancing market less elastic. Therefore, all else equal, we expect strategic firms to respond to the nuclear outage by submitting steeper bids.

We test this hypothesis and find that large firms respond to both competitor costs and own costs, but that small firms only respond to own costs. To do so, we analyze the slope of firms' bid functions in the months surrounding the nuclear outage and test whether bids are "steeper" during the outage. We create a panel of firm bids

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

	Largest three firms			Smallest three firms		
	(1)	(2)	(3)	(4)	(5)	(6)
Outage	−47.81	−31.13	−19.42	−0.02	0.17	0.36
	(7.50)	(8.30)	(5.19)	(0.18)	(0.19)	(0.18)
Own MC $\left(\frac{\partial q_{it}}{\partial MC}\right)$		0.15	0.26		0.04	0.09
		(0.05)	(0.04)		(0.01)	(0.01)
Constant	73.42	38.08	−8.61	1.32	0.65	0.63
	(7.18)	(12.37)	(8.84)	(0.12)	(0.18)	(0.19)
Bidder fixed effects	No	No	Yes	No	No	Yes
Observations	189	189	189	189	189	189
R^2	0.21	0.25	0.59	0.00	0.17	0.25

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

across each auction in our sample. The dependent variable is the (inverse) slope $(\partial S_{it}/\partial p)$ of firm i's bid in auction t.[26] The variable $Outage_t$ is an indicator equal to 1 if the auction occurred during the outage period of October 2, 2002 to November 27, 2002 and equal to 0 if the auction occurred prior to the outage. Firms strategically responding to the outage face less elastic residual demand and thus will submit bids with smaller inverse bid slope.

Results are shown in Table 3. We estimate the effect of the outage separately for the largest three and smallest three firms. In columns 1 and 4, we estimate the slope of bids only as a function of the outage. The largest three firms submit bid functions that are "steeper" (i.e., $\partial S_{it}/\partial p$ is smaller) during the outage, but the smallest three firms do not change the slope of their bids during the outage. In columns 2 and 5, we control for changes in own costs by including a measure of the firm's own marginal cost function: the (inverse) slope of the firm's MC $(\partial q_{it}/\partial MC)$. Both small and large firms submit steeper bids in response to changes in their own marginal costs; when marginal cost is flatter ($\partial q_{it}/\partial MC$ is larger), the firm submits a flatter bid function (i.e., $\partial S_{it}/\partial p$ is larger). However, only the large firms respond to the outage by submitting steeper bids. In columns 3 and 6, we add bidder fixed effects to allow for firms to face different shaped residual demand functions in general. We find robust evidence that all firms respond to their own cost shocks, but only large firms respond to the large rival cost shock induced by the nuclear outage. We find similar results when we analyze all 12 firms and define large (small) as all firms above (below) the median size firm.

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

VII. Out-of-Sample Prediction

We find that the CH model is better at predicting bidding behavior out-of-sample than a best-response model. To do so, we re-estimate the CH model using a subsample of auctions that excludes the outage period described in the previous section. Then, we use the estimated parameters from this restricted sample to predict equilibrium outcomes during the outage. Using the baseline CH model where type is parameterized to firm size, we predict each firm's profit in every auction for the auctions that occurred during the nuclear outage. Then we use a regression-based approach to measure whether profits achieved from the firms' actual bidding behavior is better explained by the CH model or a model of unilateral best-response.

Results are shown in Table 4. Column 1 shows that actual profits are only weakly correlated with profits predicted by a best-response model, while column 2 shows a much stronger correlation between actual profits and profits predicted by the CH model. Column 3 shows a "horse race" between the two models and suggests that our CH model outperforms a model of unilateral best-response.

VIII. Counterfactuals: Increasing Strategic Sophistication

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

First, we consider how increasing firms' sophistication, without changing market structure, affects market efficiency. Several events could cause firms to become more strategically sophisticated. For example, a firm could hire more qualified managers, perhaps with backgrounds associated with higher types (e.g., training in economics/business/finance or educated at an AAU institution). Alternatively, a firm could hire a consulting company to take over its bidding operations rather than use in-house personnel. Or perhaps the grid operator (which is a nonprofit entity with public interests) could distribute "teaching material" about how bidding vertically sacrifices profit-making opportunities. Under these counterfactuals, market efficiency will increase if more sophisticated bidding leads to more elastic bids and causes a firm's low-cost generation to be offered into the market.

Second, we estimate the impact on efficiency of increases in sophistication that are caused by large high-type firms merging with smaller lower-type firms. Suppose that when the two firms merge, the trading operation of the larger firm takes over all bidding operations, so the merged firm enjoys the sophistication level of the large firm. The effect of such mergers is ex ante ambiguous. On one hand, more sophisticated bidding of the small firm's generation assets is likely to cause lower-cost plants to be dispatched. On the other hand, mergers increase concentration and increase market power. We estimate which effect dominates for mergers between firms of different sizes.[27]

[27] The merger counterfactual also could represent a less extreme contractual relationship between two firms in which the low-type firm contracts out bidding operations to the high-type firm.

Table 4—Explaining Variation in Realized Profits during the Outage Period

	Realized profits		
	(1)	(2)	(3)
Profits under cognitive hierarchy		0.703	0.642
		(0.136)	(0.211)
Profits under best-response	0.263		0.061
	(0.052)		(0.091)
Constant	−64.484	−248.599	−264.619
	(156.308)	(101.941)	(97.348)
Observations	426	426	426
R^2	0.25	0.56	0.57

Notes: In all columns, the dependent variable is profits from the firms' actual bids. The covariates are firm profits as predicted by the Cognitive Hierarchy model (column 1), firm profits that would be achieved from a model of unilateral best-response to actual rival bids (column 2), and both (column 3). Predicted profits under CH are based on a sample of auctions that excludes the period of the nuclear outage. Standard errors clustered at the firm level are reported in parentheses.

Mechanically we compute counterfactuals by using the firm-specific estimated probability distribution over types (presented in Figure 7) to take 1,000 draws with replacement of the type of each firm in the CH. We then solve for the equilibrium for each auction and average generating costs over the 1,000 replications. It is important to note that only a subset of firms in the market are included in the cognitive hierarchy; the remaining firms are part of a unmodeled fringe. For this reason, we compute inefficiencies as the difference between the generating cost implied by the estimated model and our efficient benchmark in which all firms included in the CH bid marginal costs, while the rest of the firms bid according to their bids in the data. We present the counterfactuals for our benchmark specification where firm type is a function of size.

A. *Increasing Sophistication without Changing Market Structure*

Increases in sophistication that do not impact market structure, such as hiring a consulting firm to take over bidding operations or hiring better managers, will impact efficiency through two channels. First, if firms are induced to be higher-type thinkers, the bid functions will become "flatter" because beliefs that rivals are higher types imply that the firms believe their residual demand to be more elastic. As a result, more low-cost generation capacity will be offered into the balancing market and production costs will fall. This is the direct effect of increasing sophistication.

However, there is also a second indirect impact on efficiency. Suppose that the increase in sophistication is publicly observable (e.g., rivals observe that the firm hires a bidding consultant). Then rival firms recognize the increase in sophistication (even though their beliefs will continue to be wrong) and also submit more elastic bids.

We simulate the effects of increasing the strategic sophistication of firms of different sizes. It is a priori ambiguous which types of firms would most improve market efficiency by increasing sophistication. Small firms have smaller amounts of

TABLE 5—EXOGENOUS INCREASE IN SOPHISTICATION: CHANGE IN PRODUCTION COSTS

	INC side (%)		DEC side (%)	
Counterfactual	Public	Private	Public	Private
Small firms to median	−6.95	−6.22	−18.4	−17.6
Above median firms to highest	−2.71	−1.96	−13.42	−12.46
Three smallest to median	−4.67	−3.75	−14.24	−13.64

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

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 5 presents results using the *Size* parameterization of τ_i.[28] The first row reports estimated changes in total market production costs when the smallest six firms are given the sophistication level of the median-sized firm. When the increase in sophistication is publicly observed so that own-firm and rival firm bids adjust, production costs fall by 6.95 percent during periods of positive balancing demand. Most of the impact occurs through the channel of changing the firm's own bids: the production costs fall by 6.22 percent when rivals do not react to the change in sophistication. In the second row, we model the impact of increasing the sophistication of all firms that are larger than the median-sized firm to the sophistication level of the largest firm. Production costs fall by 2.71 percent, which is less than half of the efficiency improvement of increasing sophistication of the smaller firms. Finally, in the third row we focus on the three smallest firms and find that much of the room for efficiency improvements lies in increasing sophistication of very small firms. Although the small firms have less generation capacity to add to the market, the largest scope for efficiency improvement lies in the small firms that withhold so much capacity due to low sophistication. Finally, the two last columns of Table 5 show that during hours with negative balancing demand (DEC hours), the efficiency effects are even larger.

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

B. *Mergers*

We now turn to studying how mergers affect efficiency. As mentioned above, we focus on potential mergers that do not generate cost synergies but do increase

[28] Estimates using the *Size*^{2} parameterization are quite similar.

concentration. Specifically we imagine a merger that changes the bidding into the balancing market but does not change any other decisions of the merging firms, such as forward contracting decisions or day-ahead schedules. In reality, a complete merger of power marketing operations, including forward contracting, could change the positions at which bidders find themselves when balancing market bidding takes place. We are not in a position to simulate counterfactuals for full power marketing integration.[29] However, we can simulate the effect of a merger of balancing market bidding operations in order to illustrate the impact of increasing sophistication on the balancing market. In our merger simulations, the merging firm takes on the sophistication level of the most sophisticated firm (i.e., the larger firm) of the merging insiders.[30]

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

Our simulations show that mergers that increase sophistication can increase efficiency as long as the merging firms are not too large. The first row of Table 6 shows that a merger between the smallest and largest firms reduces production costs, despite the increase in concentration. However, for mergers involving medium-sized and larger firms, the market power effect dominates and mergers increase production costs.

IX. Conclusions

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

We study bidding in the Texas electricity market where bidding by some firms departs significantly from what Nash-type models predict while bidding by other firms closely resembles these predictions. We use this setting, as well as unique data

[29] Several factors prevent us from modeling the impact of mergers on both the forward contract market and the balancing market simultaneously. First, contracted quantities correspond to a mixture of long- and medium-term contracts that are signed in advance of the balancing auction and often refer to blocks of hours over periods involving days or weeks. Thus, contract positions are unlikely to be chosen in order to enter a specific balancing auction with a strategically chosen contract position. Second, from a practical perspective, forward bilateral contracts are proprietary data. However, firms often treat the two markets as separate with separate trading desks for long-term and balancing operations. Finally, considering the two markets as related necessarily means that the game to be modeled involves trading over long- and short-term contracts for power across multiple hours. Modeling this broader "super game" is beyond the scope of this paper.

[30] One could also assign the merging insiders the sophistication associated with the merged firm. However, the two assignment rules lead to very similar beliefs as long as one of the merging firms is small.

Table 6—Increasing Sophistication via Mergers: Change in Production Costs

Merging firms	INC side (%)	DEC side (%)
Smallest and largest firms	−2.62	−6.49
Median and largest firms	+10.29	+10.37
Two largest firms	+18.38	+48.72

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

on bids and marginal costs, to embed a cognitive hierarchy model into a structural model of bidding. We estimate heterogeneity in levels of strategic sophistication across firms. We find that while small firms behave as if they are boundedly rational in a cognitive hierarchy sense, large firms behave closely to what a Bayesian Nash model would predict. We then use the estimated levels of strategic sophistication to study how increasing the sophistication of low-type firms, either exogenously or through mergers with higher-type firms, may affect efficiency. Our results show that not only can exogenously increasing sophistication increase efficiency significantly, but that mergers that do not generate cost synergies but increase concentration may also increase efficiency.

REFERENCES

An, Yonghong. 2017. "Identification of First-Price Auctions with Non-Equilibrium Beliefs: A Measurement Error Approach." *Journal of Econometrics* 200 (2): 326–43.

Arad, Ayala, and Ariel Rubinstein. 2012. "The 11–20 Money Request Game: A Level-*k* Reasoning Study." *American Economic Review* 102 (7): 3561–73.

Baldick, Ross, and Hui Niu. 2005. "Lessons Learned: The Texas Experience." In *Electricity Deregulation: Choices and Challenges*, edited by James M. Griffin and Steven L. Puller, 182–224. Chicago: University of Chicago Press.

Bloom, Nicholas, and John Van Reenen. 2007. "Measuring and Explaining Management Practices across Firms and Countries." *Quarterly Journal of Economics* 122 (4): 1351–1408.

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

Bushnell, James B., Erin T. Mansur, and Celeste Saravia. 2008. "Vertical Arrangements, Market Structure, and Competition: An Analysis of Restructured US Electricity Markets." *American Economic Review* 98 (1): 237–66.

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

Camerer, Colin F., Teck-Hua Ho, and Juin-Kuan Chong. 2004. "A Cognitive Hierarchy Model of Games." *Quarterly Journal of Economics* 119 (3): 861–98.

Cho, Sungjin, and John Rust. 2010. "The Flat Rental Puzzle." *Review of Economic Studies* 77 (2): 560–94.

Costa-Gomes, Miguel, Vincent P. Crawford, and Bruno Broseta. 2001. "Cognition and Behavior in Normal-Form Games: An Experimental Study." *Econometrica* 69 (5): 1193–1235.

Crawford, Vincent P., Uri Gneezy, and Yuval Rottenstreich. 2008. "The Power of Focal Points Is Limited: Even Minute Payoff Asymmetry May Yield Large Coordination Failures." *American Economic Review* 98 (4): 1443–58.

Crawford, Vincent P., and Nagore Iriberri. 2007. "Fatal Attraction: Salience, Naïveté, and Sophistication in Experimental 'Hide-and-Seek' Games." *American Economic Review* 97 (5): 1731–50.

DellaVigna, Stefano. 2018. "Structural Behavioral Economics." In *Handbook of Behavioral Economics* Vol. 1, edited by B. Doug Bernheim, Stefano DellaVigna, and David Laibson, 613–723. Amsterdam: Elsevier.

DellaVigna, Stefano, and Matthew Gentzkow. 2019. "Uniform Pricing in U.S. Retail Chains." *Quarterly Journal of Economics*. 134: 2011–84

Doraszelski, Ulrich, Gregory Lewis, and Ariel Pakes. 2018. "Just Starting Out: Learning and Equilibrium in a New Market." *American Economic Review* 108 (3): 565–615.

Ellison, Sara Fisher, Christopher M. Snyder, and Hongkai Zhang. 2018. "Costs of Managerial Attention and Activity as a Source of Sticky Prices: Structural Estimates from an Online Market." Unpublished.

Fabra, Natalia, and Mar Reguant. 2014. "Pass-Through of Emissions Costs in Electricity Markets." *American Economic Review* 104 (9): 2872–99.

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

Gillen, Ben. 2010. "Identification and Estimation of Level-*k* Auctions." Unpublished.

Goldfarb, Avi, and Mo Xiao. 2011. "Who Thinks about the Competition? Managerial Ability and Strategic Entry in US Local Telephone Markets." *American Economic Review* 101 (7): 3130–61.

Hitsch, Guenter, Ali Hortaçsu, and Xiliang Lin. 2019. "Prices and Promotions is U.S. Retail Markets: Evidence from Big Data." Chicago Booth Research Paper 17–18.

Hortaçsu, Ali, Fernando Luco, Steven L. Puller, and Dongni Zhu. 2019. "Does Strategic Ability Affect Efficiency? Evidence from Electricity Markets: Dataset." *American Economic Review*. https://doi.org/10.1257/aer.20172015.

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

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

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

Kagel, John H., and Alvin E. Roth. 1995. *The Handbook of Experimental Economics.* Princeton, NJ: Princeton University Press.

Massey, Cade, and Richard H. Thaler. 2013. "The Loser's Curse: Decision Making and Market Efficiency in the National Football League Draft." *Management Science* 59 (7): 1479–95.

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

Nagel, Rosemarie. 1995. "Unraveling in Guessing Games: An Experimental Study." *American Economic Review* 85 (5): 1313–26.

Stahl, Dale O., and Paul W. Wilson. 1995. "On Players' Models of Other Players: Theory and Experimental Evidence." *Games and Economic Behavior* 10 (1): 218–54.

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

Wilson, Robert. 1979. "Auctions of Shares." *Quarterly Journal of Economics* 93 (4): 675–89.

Wolak, Frank A. 2003. "Measuring Unilateral Market Power in Wholesale Electricity Markets: The California Market, 1998–2000." *American Economic Review* 93 (2): 425–30.

Wolfram, Catherine D. 1998. "Strategic Bidding in a Multiunit Auction: An Empirical Analysis of Bids to Supply Electricity in England and Wales." *RAND Journal of Economics* 29 (4): 703–25.