STRUCTURAL MODELS OF MERGERS

A Dissertation

by

DONGNI ZHU

Submitted to the Office of Graduate and Professional Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Chair of Committee,	Steven L. Puller
Co-Chair of Committee,	Qi Li
Committee Members,	Fernando A. Luco Echeverría
	Ximing Wu
Head of Department,	Timothy J. Gronberg

August 2016

Major Subject: Economics

ABSTRACT

Governments and researchers are frequently forced to predict the impact of perspective mergers on markets. This dissertation provides structural methods to empirically evaluate mergers.

We first build a static model in which players are boundedly rational to evaluate the welfare consequence of mergers in that environment. Then we use that model studying bidding in the Texas electricity market, a market in which bidding by some firms departs significantly from what Bayesian Nash models predict, while bidding from other firms closely resembles these predictions. Our results show that exogenously increasing sophistication may significantly increase efficiency and additionally, mergers may increase efficiency even without cost synergies.

The next chapter provides a structural method to empirically evaluate mergers in a dynamic setting. We build an infinite five-step repeated game. Then, we propose a three-step estimation method to estimate the game in which Markov perfect Nash equilibrium is played. Our three-step estimation method is flexible and can be easily modified to estimate various market structures.

These dissertation studies mergers in more realistic settings. We first show that mergers that do not generate cost synergies may also increase efficiency when some of the firms in the market are boundedly rational. Then we build a dynamic endogenous merger game and provide a new method to estimate it. Our simulation result shows that our estimation method is computational feasible and can be applied to real data.

DEDICATION

This dissertation is dedicated to my family.

ACKNOWLEDGEMENTS

My deepest gratitude is to my advisor, Dr. Steven L. Puller. I have been fortunate to meet an advisor whose patience and support helped me overcome many difficulties and finish this dissertation. My co-advisor Dr. Qi Li has been always been there to help me. I am deeply grateful to him for insightful discussions and practical advice. I learned a lot from Dr. Fernando A. Luco Echeverría who is my coauthor of the second chapter. I am indebted to him for helping me to understand, commenting on my views and enrich my ideas. I am also grateful to Dr. Ximing Wu for his help in the past five years. Finally, I would like to express my appreciation to my coauthor Dr. Ali Hortaçsu without whom I could never finish the second chapter of my dissertation.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

1. INTRODUCTION

Static oligopoly models are often used by governments to simulate welfare consequences of prospective mergers and challenge those mergers predicted to decrease welfare. However, firm behaviors might deviate from assumptions in these models. This dissertation provides methods to evaluate mergers in more realistic settings.

In the next chapter, we build and estimate a model in which firms are boundedly rational and simulate the effect of mergers in such settings. In most oligopoly models, firms are modeled as playing some form of Nash equilibrium. This model of supply-side behavior is used to estimate parameters with implications for policy. However, evidence exists that real-world firms may be boundedly rational, engaging in some level of strategic thinking, but the degree of strategic thinking may "fall short" of playing the Nash equilibrium strategy (e.g. Hortaçsu and Puller [2008] and Goldfarb and Xiao [2011]). Moreover, deviations from Nash equilibrium play can be economically significant and have implications for efficiency. Hortaçsu and Puller [2008] identify a set of firms that submit bids into Texas electricity auctions that persistently and substantially deviate from Nash bidding. The consequence of these deviations is that low-cost power plants are not called to produce, and this substantially raises total production costs. We embed a Cognitive Hierarchy model into a structural model of bidding behavior to capture the heterogeneity in the observed deviations from Bayesian Nash equilibrium bidding. We use this model to study the firms in the Texas electricity spot market. Our results show that efficiency increases with strategic sophistication and mergers may also increase efficiency even with no cost synergies and increasing market concentration.

In the third section, we study mergers in a dynamic setting where firms maximize their long-term profits. In standard merger simulation exercises, only changes in the period right after the mergers are considered in a merger review. There are many reasons to believe long-term welfare changes after a merger should be reviewed as well as a merger could have considerable impact on the dynamic evolution of the market. The literature has moved very slowly in developing such methods because of the complication of modeling dynamic mergers and computational burden involved in estimating the model. In recent years, researchers have proposed several new methods to ease the computational burden of estimating dynamic oligopoly model. In this chapter, we extend the classic Ericson and Pakes [1995] oligopoly dynamic model and include merger as a dynamic strategy in the game. In particular, we build an infinite five-step repeated game. Then, we propose a three-step estimation method to estimate the game in which Markov perfect Nash equilibrium is played. The first two steps follow Bajari et al. [2007] to ease the computational burden of estimating dynamic oligopoly models. The third step applies the moment inequality condition estimation method proposed in Tamer [2003], which solves the inference of multiple equilibria in discrete choice games. Our three-step estimation method is flexible and can be easily modified to estimate various market structures.

2. STRATEGIC ABILITY AND PRODUCTIVE EFFICIENCY IN ELECTRICITY MARKETS

2.1 Introduction

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

However, some research has suggested caution at applying such strategic equilibrium models in all settings, because while in a Nash equilibrium each firm is best responding to its beliefs about each rival firm's behavior *and* all of those beliefs are mutually consistent, in real-world settings, the rationality assumption or mutual consistency assumption may break down, and firms may not be playing at the fixed point that equilibrium models characterize. Indeed, evidence exists that real-world firms may be boundedly rational, engaging in some level of strategic thinking, but the degree of strategic thinking may "fall short" of playing the Nash equilibrium strategy (e.g. Hortaçsu and Puller [2008] and Goldfarb and Xiao [2011]).

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

This suggests that models allowing for boundedly rational firm behavior can be valuable for explaining the outcomes of certain real-world markets. To this extent, theoretical research has developed models that help organize strategic behavior that deviates from the Nash equilibrium.[1] Examples of such models include level-k thinking and Cognitive Hierarchy in which players best-respond to (perhaps incorrect) beliefs about their rival behavior. The Cognitive Hierarchy model (hereafter CH) allows for heterogeneity in the levels of strategic thinking by firms in a market. In the CH model, the least strategic players – Level-0 players – are entirely non-strategic in their bidding. Level-1 players assume that all other players are level-0 players and submit bids that correspond to the best response to all other players behaving as such. Level-2 players assume that all other players are some combination of level-0 or level-1 players and best respond to those beliefs. In general, level-k players assume that all other players are distributed between level-0 and level–k-*1* and submit bids corresponding to the best response to those beliefs. The limiting case of this model corresponds to the Nash equilibrium.[2] In this setting, CH maintains the assumption

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

[2] As noted in Camerer et al. [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.

that players best respond, but it allows for firms to have beliefs about their rival s-trategies that are not consistent with the rivals' actual behavior. This model has the appealing feature that it allows for a hierarchy of levels of sophistication by different players in a market.

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

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

[3]One novel approach to address this problem has been proposed by Gillen [2010] who studies joint identification of types and valuations in the level-k setting. Gillen shows point identification of the joint distribution could be obtained exploiting variation in the number of bidders and assuming constant valuations across auctions. However, in the absence of either of these, only set identification is possible. An [2013] also studies identification in the level-k model but he relaxes some of these assumptions present in Gillen's work but imposes constraints on the structure of the data to identify both the number of types in the data and the type of each firm.

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

In this setting, HP show evidence that firms deviate from Bayesian Nash equilibrium bidding, suggesting a fruitful environment to apply bounded rationality models. HP test whether each firm submits bids that correspond to the best response to rivals' actual bids (as required if firms play a Nash equilibrium) and find that a few firms – typically larger firms – submit bids close to best-response bidding. However, many small firms tend to bid to supply power at prices so far above their marginal costs that they "bid themselves out of the market" and are not called to produce despite having low-cost generation available.[4]

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

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

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

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

In this chapter, we address each of these questions. Specifically, we embed a Cognitive Hierarchy model into a structural model of bidding behavior to capture the heterogeneity in the observed deviations from Bayesian Nash equilibrium bidding. The Texas electricity market has firms of various sizes and organizational structures

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

that bid into the spot market, and we use firm observables to parameterize the determinants of firm type. We estimate the model using a minimum-distance approach.

We then turn to study how strategic sophistication affects productive efficiency. We do this by simulating a number of scenarios in which the level of strategic sophistication of low-type firms is increased either exogenously or through mergers with high-type firms. Importantly, the application of the CH model to multi-unit auctions has very valuable methodological benefit in this setting. As shown in Klemperer and Meyer [1989], in general there are multiple equilibria in multi-unit, uniform-price auctions that can range from competitive to Cournot-like behavior. The multiplicity of equilibria presents a challenge for conducting counterfactual calculations of market outcomes under changes in cost or market structure. One way to address this problem has been to impose mathematical restrictions on permissible form of bids, such as restricting bid functions to be linear (Baldick et al. [2004]). The CH model provides a means to address the multiple equilibria problem without imposing such restrictions. The mutually consistent beliefs assumption – a source of the multiple equilibria problem – is not imposed in CH. Instead, given a firm's belief about its rivals' type distribution, one can calculate the (unique) best-response bid.[5] Therefore, the iterative nature of strategic thinking under CH allows to calculate unique counterfactual market outcomes. We exploit this feature by computing market efficiency under possible mergers between firms with different levels of strategic sophistication, which would not be possible under a Nash equilibrium model. Thus, not only does CH allow for more realistic models of real-world bidding behavior, but it allows researchers to more precisely simulate outcomes under changes in market structure or changes in costs.

[5] Camerer et al. [2004] note a related feature that the CH model can be viewed as a behavioral refinement that can eliminate the multiplicity of equilibria in coordination games.

Thus, we are able to simulate unique predictions of market outcomes under various policy counterfactuals. For example, consider a merger between a large and small bidder in this electricity market. Such a merger is unlikely to lead to substantial cost synergies because the costs of generating electricity is almost entirely driven by the model and vintage of the electric generator. Thus, one might expect the increase in concentration induced by the merger to enhance market power and reduce economic efficiency. However, in a merger between two boundedly rational firms, this merger could increase efficiency. Indeed, suppose that the large firm is a high-level strategic thinker and the small firm is a low-level strategic thinker. If the merger caused the large firm to take over bidding operations, then the generation resources of the small firm would subsequently be controlled by a higher level strategic thinker. This could increase efficiency because the low-k firm would be less likely to bid prices so high that its efficient productive capacity is priced out of the market. We can evaluate this conjecture by simulating mergers between any firms in the Texas market. More generally, we can calculate market prices and efficiency under any counterfactual level of strategic sophistication by any firm.

Our results show that efficiency increases with strategic sophistication, though at a decreasing rate. Indeed, exogenously increasing sophistication of low-type firms results in reducing inefficiencies by up to 24% relative to the status quo. However, the number of firms whose sophistication matters and increasing sophistication of any single small firm has little impact on efficiency. On the other hand, we also find that it is not necessary to increase sophistication to the maximum estimated level to achieve significant efficiency gains. Indeed, increasing sophistication to that of the median firm is enough to generate essentially same efficiency gains as increasing sophistication to the maximum observed level. Finally, mergers may also increase efficiency even with no cost synergies and increasing market concentration. Our results show that when, for example, a small, low-type firm, merges with a large, high-type firm, as long as the resulting firm is of the same type as the largest involved in the merger, efficiency may increase by 68%.

This chapter contributes to an emerging body of literature that empirically models sophistication and learning in new markets. The most prominent paper that has applied a cognitive hierarchy model to field data is Goldfarb and Xiao [2011] who study the entry decisions into newly opened markets for local telephone competition. They apply the cognitive hierarchy model to an entry game and find that manager characteristics such as experience and education are determinants of strategic ability that predict firm performance. Related work by Doraszelski et al. [2014] use models of learning to predict the evolution of pricing in a newly opened electricity ancillary service market. This chapter also contributes to the literature on how electricity generating firms formulate bids (e.g. Fabra and Reguant [2014]) and research that models oligopoly competition in the electricity sector (e.g. Wolfram [1998] and Bushnell et al. [2008]).

2.2 Institutional Setting

Hortaçsu and Puller [2008] describe the Texas electricity market in detail, so here we focus only on key aspects of it. The Texas electricity market was restructured in August 2001. Since then, firms are no longer part of a natural monopoly and most of the electricity is now traded through bilateral forward contracts between generators and electricity consumers. However, aggregate demand may fall beyond or below contracted quantities at the last-minute. To meet the shortage or surplus, the generation firms submit bids to adjust their production relative to contracted quantities. They do this by participating in an hourly "balancing market" administered by ERCOT (Electric Reliability Council of Texas). This market trades between 2 and 5% of all power traded in Texas. The hourly market is cleared by a multi-unit, uniform-price auction in which firms bid supply functions and winning sellers earn the price at which aggregate supply bids equal demand. Firms do this by submitting schedules of quantities of electricity to inject and withdraw at specific locations on the transmission grid to ERCOT, the day before electricity transmission occurs. These are called day-ahead schedules and may differ from the firms' forward contract. These supply and demand schedules may also differ from actual production and consumption in real time because of a variety of reasons such as extreme hot or cold weather. Depending upon whether there is shortage or surplus of power relative to the day-ahead schedule, balancing demand can be positive or negative. Because demand does not respond to prices in real time, balancing demand is a perfectly inelastic.

In this setting, firms submit bids that increase or decrease the amount of power to supply relative to their day-ahead schedule. In addition, bids may be changed up until one hour prior to the operating hour. The market is then cleared every 15-minute using a uniform-price, multi-unit auction. ERCOT determines the market clearing price aggregating supply bids and intersecting the aggregate supply curve with total demand. A generator called to increase production is paid the market clearing price, while a generator called to decrease production purchases power from ERCOT at the same price to meet existing contract obligations.

In the Texas electricity market, the generation technology is mainly fueled by natural gas and coal with small amounts of hydroelectric, nuclear, and wind generation. TXU and Reliant are the two largest players and they are the two largest former incumbent utilities, owning 24% and 18% of installed capacity, respectively. Other major investor-owned utilities include Central Power and Light (7% of installed capacity) and West Texas Utilities (2%). There are also some municipal utilities, among which the largest are City of San Antonio Public Service (8% of installed capacity) and City of Austin (6%). Finally, there is a large number of merchant generation firms of various sizes, such as Calpine (5%).

2.3 Data

We study bidding behavior in the Texas electricity market starting after it was restructured in September 2001 until January 2003. As HP, we focus on non-congested weekdays between 6:00 and 6:15pm because the most flexible type of generators that can respond to balancing calls without large adjustment costs are likely to be online at this period. Non-congested periods account for 74% of all periods between September 2001 and January 2003. Finally, in our sample, we consider auctions for which we have bid and marginal cost data for all firms included in the Cognitive Hierarchy.

We argue that each firm's marginal cost data is public information.[6] The production technologies used by power plants in Texas are very similar to each other and data on the fuel cost of each generating units are publicly available. Moreover, firms know whether major generating units are on- or off-line at any time. Also, some firms purchase large plants' generation data from Genscape, an energy information company that measures real time output by remote sensors installed near the transmission lines.

Overall, our data consists on information that allow us to calculate each firm's best-response bids in each auction, for all possible levels of strategic sophistication. In particular, for each auction, we have data on total balancing demand, firm-level marginal cost functions, and firm-level bids, as well as information on firm characteristics.

[6] HP claimed that conversations with several market participants suggest that traders have good information on their rivals' marginal costs.

2.3.1 Descriptive Evidence

We study bidding into the Texas electricity spot market in the first two years after the market was created in 2001. The spot market – called the balancing market in electricity parlance – is an hourly auction to supply power to the grid operator in order to ensure that supply equals demand at every point in time. Electricity generating firms will have scheduled a certain amount of production the day-ahead, and the balancing market is a mechanism to adjust production up or down from that day-ahead production plan. Each firm submits an hourly bid function for its entire portfolio of plants to increase or decrease production relative to the day-ahead schedule. Market demand is determined by unexpected changes in the amount of power needed. For example, if the weather is unexpectedly hot on a summer afternoon, then balancing demand is positive. The grid operator uses a uniform-price, multi-unit auction; so the market is cleared each period by intersecting the hourly aggregate supply bids with the total balancing demand.

We start this section explaining how bids would be chosen in which firms best respond to their rivals actions. Figure 2.1 explains the basic intuition of best-response bidding in this market. Suppose that a firm has marginal cost of supplying to the balancing market given by $MC_i(q)$. In addition, assume that the firm has forward contracts to supply QC_i units of power. Because the firm is a net seller after it has covered its contract position, the firm has incentives to bid prices above marginal cost for quantities greater than QC_i.[7] The size of the mark-up will depend on the firm's residual demand elasticity. The residual demand function RD_i is equal to the total market demand minus the supply bids by all other bidders. Suppose that it is an expectedly hot day and the firm faces RD_1 shown in figure 2.1. Then the firm has

[7] Likewise, the firm is a net buyer for quantities less than QC_i, so it has incentives to bid prices below MC for quantities less than the contract position.

the incentive to bid a quantity corresponding to the point where Marginal Revenue equals Marginal Cost ($MR_1 = MC_i$) and a price corresponding to the (inverse) Residual Demand function at that quantity. This point is given by point A in the figure. However, suppose instead that it is a cool day and the firm faces a different residual demand function of RD_2. In that case, the same logic implies that the best response is point B. Because the firm can submit a large number of (price, quantity) points, it can consider a continuum of different residual demand functions. Thus, the firm can "trace out" the set of best-response bids, and submit a best-response bid function given by the red line S_i^{BR}.[8] Importantly, no estimation is required; the components of figure 2.1 are available as *data* for each firm in each auction. We view this data-rich environment as a major strength of our approach. Data on marginal cost are critical to our identification strategy that allows us to identify strategic behavior.

We now present descriptive evidence that some firms deviate from Bayesian Nash equilibrium bidding, and we use this evidence to motivate our modeling assumptions. Figure 2.2 displays representative bid functions for a large firm that submits bids close to best-response bids. For two different auctions, the figure shows the firm's marginal cost of supplying more power to the grid. In the left figure, the firm had a contract position of nearly 600 MW, so it bid above (below) marginal cost for quantities above (below) that contract position. This firm submitted an actual bid function that was very close to our calculation of the best-response bid function. The right graph shows a similar outcome.

However, many of the small firms in this market do not submit bids close to the best-response bid function. Figure 2.3 shows representative bids for one of these

[8] In general, it is possible that the set of best-response points is not monotonic function, however we show in section 2.5.2 that in this setting the best-response points are monotonic.

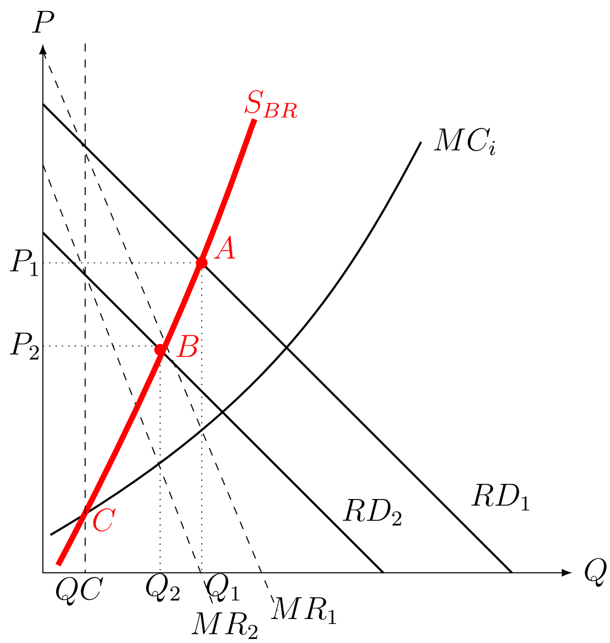


FIGURE 2.1: Best-response bidding in spot auction

small firms. This firm has a contract position of zero. As shown by the best-response bid that we calculate, this firm has some market power despite being small, so it is optimal to bid prices several dollars above marginal cost. However, this firm submitted bids at extremely high prices for small quantities, making it unlikely that the firm will be called to produce. The consequence of these bids is two-fold: (1) the firm reduces profits, and (2) the market is not efficient because the firm's relatively low cost production is not utilized.

The patterns displayed above are persistent throughout the early years of the market. We estimate that the firms in this market left between \$3-\$18 million per year "on the table", depending on firm size. This type of bidding behavior serves as motivation for our model of boundedly rational bidding within a cognitive hierarchy structure.

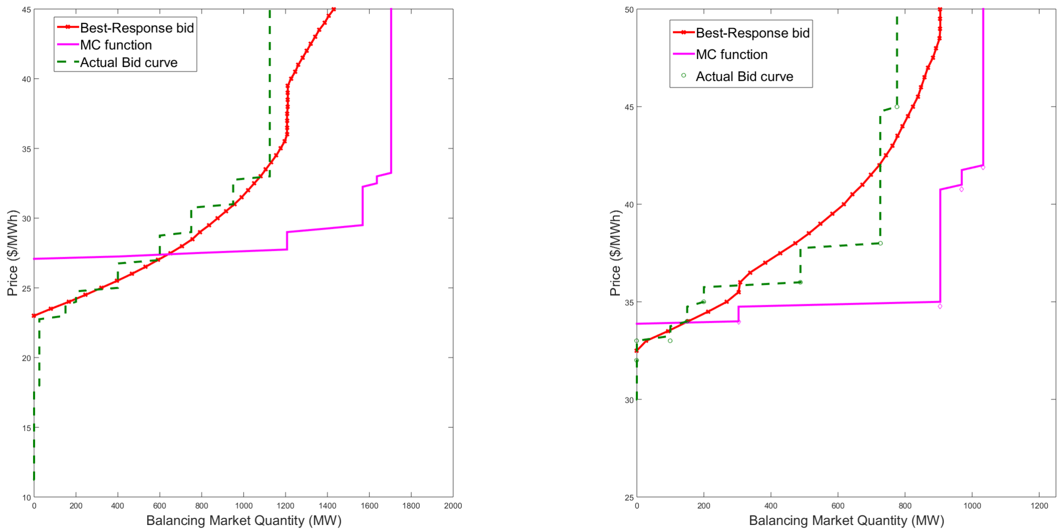


FIGURE 2.2: Large firm: actual bids vs. best-response bids

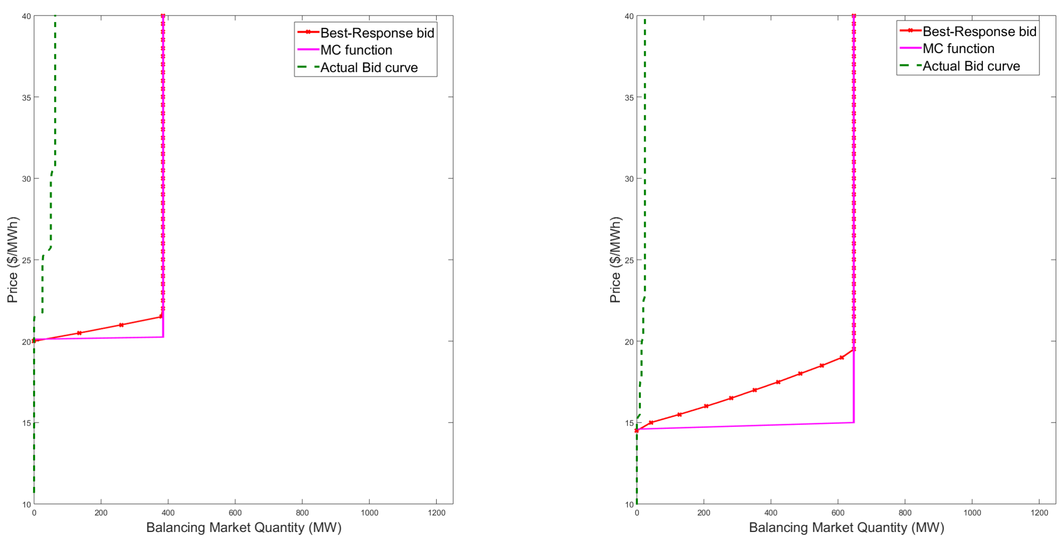


FIGURE 2.3: Small firm: actual bids vs. best-response bids

2.4 Theoretical Background on Cognitive Hierarchy

The theoretical literature has developed a rich set of models of boundedly rational strategic behavior that could explain deviations from Bayesian Nash equilibrium play. Generally speaking, bounded rationality models relax one of the two conditions of Nash Equilibrium; that (1) players maximize expected payoffs given beliefs about their rivals' actions and (2) that player beliefs about rivals' actions are consistent. Hierarchy models (such as Cognitive Hierarchy and level-k) maintain the assumption of best-response but relax the assumption of consistent beliefs.[9] These models conceptualize players as having a hierarchical structure of strategic, or level-k thinking. Seminal work on level-k behavioral models has been conducted by Costa-Gomes et al. [2001], Crawford and Iriberri [2007], and Camerer et al. [2004].

Cognitive Hierarchy (CH) developed by Camerer et al. [2004] conceptualizes players as engaging in different levels of strategic thinking ordered in a hierarchy. As explained above, the least sophisticated players – 0-step players – engage in no strategic thinking, while higher types (say, k) assume that all other players are distributed between 0-step and k-1-step players according to a Poisson distribution. [10] 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.

[9] 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; for example, see figure 2.3. In other words, it does not appear that bidders in the electricity auctions estimate expected payoffs in an unbiased way, a key feature of the QRE model.

[10] The model does not require the distribution be Poisson. However, Camerer et al. [2004] note that the Poisson has the property that as k rises, fewer players perform the next step of thinking, which is consistent with increasing working memory being required for an additional step of iterative calculation.

This cognitive hierarchy structure is conceptually appealing because it captures behavior in which firms have limits to the level of strategic thinking and/or firms are overconfident about their own abilities.

The level-k model is a specific form of the CH model where a level-k player assumes that *all* other players are level-*(k-1)*. In other words, rather than rivals coming from a *distribution* of types *(k-1)* and below, in the level-k model, rival firms are type *(k-1)*. Comparing the two models, in one sense CH is a more flexible model. However, one could also view CH as a model that could be too flexible and explain "anything". In this paper, however, we sidestep this theoretical debate and rather write down an empirical model that is most general – cognitive hierarchy – and estimate the model with our data.

2.4.1 Big Picture of Modeling and Estimation Strategy

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

Consider firm i that is type k. The assumptions of the CH model imply that i believes its rivals are distributed between type-0 and type-(k-1), according to a normalized Poisson distribution with parameter τ. Firm i decides its bidding strategy according to these beliefs (that depend on its own type, contract position, and its rivals τ's) and valuations (marginal costs), to maximize its expected profits.

One critical feature of estimating a CH model is how to define level-0 behavior (or in the language of Camerer et al. "0-step players"). In the theoretical literature, a common assumption is that level-0 players (uniformly) randomize across all possible strategies, although that assumption can be relaxed to match a particular setting (i.e., Goldfarb and Xiao [2011] assume level-0 players to believe they are monopolists in an entry game). In the context of the Texas electricity auctions, there is a natural assumption about non-strategic thinkers: we observe some firms bidding "vertically" at their contract positions for the range of plausible prices. (That is, the firms submit bids similar to the second panel of figure 2.3. In other words, these firms are indicating that at even very high prices, they do not want to sell power into the balancing market. This clearly violates any standard model of (expected) profit-maximization; the firms have low cost generation to offer into the market, but they choose not to do so. Thus, "vertical bidding at the contract position" is a natural candidate for level-0 bidding behavior.

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

Finally, we assume that not all firms engage in strategic thinking or even enter the Cognitive Hierarchy. Indeed, we allow only a subset of firms to enter the hierarchy, while the rest form part of an unmodeled fringe. We do this because allowing for more firms makes the problem computationally challenging as each firm needs to compute its rivals bidding functions for all possible types, for all auctions. Furthermore, we do not have marginal cost data for all firms for all auctions, which also imposes a constraint on the number of firms that we can include in the Cognitive Hierarchy. Accordingly, we model all "big" firms entering the Cognitive Hierarchy plus a number of small ones including the one depicted in figure 2.3. This, however, has the unintended cost of limiting the extent to which our counterfactual simulations can improve efficiency, as part of the inefficiencies that result from departure from Bayesian Nash bidding is generated by firms that we do not include in the cognitive hierarchy.

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

We then compare these bids to the firm's actual bidding behavior. The estimation process finds the parameters of τ – how firm characteristics such as size affect strategic sophistication – that minimize the distance between actual bids and the bids predicted under CH. That is, in estimation, we use observed bids and realized marginal costs to recover the type of each firm. For this reason, it is critical that we observe marginal costs, as in the absence of realized costs, one would not be able to identify types from bid data without additional assumptions regarding the cost function.[11] In other words, instead of using data on observed bids to recover valuations, we use that we observe valuations and bids to recover the type that rationalizes observed behavior.

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

2.4.2 Modeling in Detail

A formal model of bidding into the Texas electricity auctions needs to formulate best-response bidding in a setting where firms have beliefs about rivals as characterized by the Cognitive Hierarchy model. We have developed a formulation that incorporates modeling features of share auction models (Wilson [1979] and Hortaçsu and Puller [2008]) and the Poisson Cognitive Hierarchy model (Camerer et al. [2004]).

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

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

$$\sum_{i=1}^{N} S_{it}(p_t^c, QC_{it}) = D_t(p_t^c) + \varepsilon_t \tag{2.1}$$

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

$$H_{it}(p, \hat{S}_{it}(p); k_i, QC_{it}) \equiv Pr(p_t^c \leq p | \hat{S}_{it}(p), k_i, QC_{it}) \tag{2.2}$$

There are three sources of uncertainty – (1) the shock to demand (ε_t), (2) each rival's type of k-step thinking, and (3) each rival's contract position QC_{jt} which affects the rival's bids. The event that the market-clearing price p_t^c is less than any given price p is the event that there is excess supply at that p. Plugging the market-clearing condition (2.1) into (2.2):

$$\begin{aligned} H_{it}(p, \hat{S}_{it}(p); k_i, QC_{it}) &= Pr(\sum_{j \neq i} S_{jt}^l(p, QC_{jt}; k_i) + \hat{S}_{it}(p) \geq D_t(p) + \varepsilon_t | \hat{S}_{it}(p), k_i, QC_{it}) \\ &= \int_{\boldsymbol{QC_{-it}} \times \boldsymbol{l_{-i}} \times \varepsilon_t} 1(\sum_{j \neq i} S_{jt}^l(p, QC_{jt}; k_i) + \\ &\quad \hat{S}_{it}(p) \geq D_t(p) + \varepsilon_t) dF(\boldsymbol{QC_{-it}}, \boldsymbol{l_{-i}}, \varepsilon_t | \hat{S}_{it}(p), k_i, QC_{it}) \end{aligned} \tag{2.3}$$

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

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

$$p \cdot \hat{S}_{it}(p) - C_{it}\left(\hat{S}_{it}(p)\right) - (p - PC_{it})QC_{it} \tag{2.4}$$

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

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

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

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

$$p - C'_{it} \left(S^*_{it}(p) \right) = \left(S^*_{it}(p) - QC_{it} \right) \frac{H_s \left(p, S^*_{it}(p); k_i, QC_{it} \right)}{H_p \left(p, S^*_{it}(p); k_i, QC_{it} \right)} \tag{2.5}$$

where H_s and H_p are given by derivatives of (2.3).

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

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

We use detailed data and several identifying assumptions to "measure" each component of equation (2.5), which allows us to calculate the best-response function for each type. In our data, we observe the marginal cost function C'_{it}, and we follow the strategy developed in HP to measure QC_{it}.

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

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

For this reason, we now make three identifying assumptions so that we can "measure" $H_{it}\left(p, \hat{S}_{it}(p); k_i, QC_{it}\right)$ and thus its derivatives H_s and H_p. The first assumption considers how bidders type 0 bid and allows us to solve the problem recursively. The second assumption considers the distribution of types in the Cognitive Hierarchy model. Finally, the third assumption considers the distribution of the remaining sources of uncertainty.

Assumption 2.1 *Bidders type 0 submit perfectly inelastic bids that are determined by their contract positions. This is,*

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

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

For a bidder type 1, all rivals are type-0 under the CH model. Thus, we can write $H(\cdot)$ (equation 2.3) for a type-1 firm submitting bid $\hat{S}^1_{it}(p)$:

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

where the second equality follows from Assumption 2.1 and the third equality from defining $\theta_{it} \equiv \sum_{j \neq i} QC_{jt} - \varepsilon_t$.

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

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

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

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

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

Proposition 2.1 1. *If bidders type 0 submit perfectly inelastic bids that are determined by their contract positions in CH model, their bids are additive separable,* $S^0_{it}(p, QC_{it}) = QC_{it}$.
2. *If bidders type 0 submit perfectly inelastic bids that are determined by their contract positions in CH model, bids of bidders type 1 is also additive separable,* $S^1_{it}(p, QC_{it}) = \beta^1_{it}(QC_{it}) + \alpha^1_{it}(p)$. *Moreover,* $S^1_{it}(p, QC_{it}) = QC_{it} + \alpha^1_{it}(p)$.

Proof. 1. It is straight forward to see that bids of bidders type 0 are additive separable because $S^0_{it}(p, QC_{it}) = QC_{it} = QC_{it} + f(p)$, where $f(p) = 0$, $\forall p \in [\underline{p}, \bar{p}]$.
2. Bids of bidders type 1 $S^1_{it}(p)$ can be calculated from equation (2.5), which can be rewritten as

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

because the argument $[(p - C'_{it}\left(S^1_{it}(p)\right)] \frac{\int_{l_{-i}} \gamma\left(D_t(p) - \hat{S}^1_{it}(p)\right) \cdot \Delta(l_{-i})}{-\int_{l_{-i}} \gamma\left(D_t(p) - \hat{S}^1_{it}(p)\right) D'_t(p)\Delta(l_{-i})}$ is a function of price p.

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

■

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

$$\begin{aligned} H_{it}(p, \hat{S}^2_{it}(p); k_i = 2, QC_{it}) &= \int_{\boldsymbol{QC_{-it}} \times \boldsymbol{l_{-i}} \times \varepsilon_t} 1(\sum_{j \neq i} S^{l_j}_{jt}(p, QC_{jt}) + \hat{S}^2_{it}(p) \geq \\ &\quad D_t(p) + \varepsilon_t) dF(\boldsymbol{QC_{-it}}, \boldsymbol{l_{-i}}, \varepsilon_t | \hat{S}^2_{it}(p), k_i = 2, QC_{it}) \\ &= \int_{\boldsymbol{QC_{-it}} \times \boldsymbol{l_{-i}} \times \varepsilon_t} 1(\sum_{j \neq i} QC_{jt} + \sum_{j \neq i} \alpha^{l_j}_{jt}(p) + \hat{S}^2_{it}(p) \geq \\ &\quad D_t(p) + \varepsilon_t) dF(\boldsymbol{QC_{-it}}, \boldsymbol{l_{-i}}, \varepsilon_t | \hat{S}^2_{it}(p), k_i = 2, QC_{it}) \\ &= \int_{\boldsymbol{QC_{-it}} \times \boldsymbol{l_{-i}} \times \varepsilon_t} 1(\theta_{it} \geq D_t(p) - \sum_{j \neq i} \alpha^{l_j}_{jt}(p) - \hat{S}^2_{it}(p)) \\ &\quad dF(\boldsymbol{QC_{-it}}, \boldsymbol{l_{-i}}, \varepsilon_t | \hat{S}^2_{it}(p), k_i = 2, QC_{it}) \end{aligned} \tag{2.6}$$

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

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

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

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

Next, we make two assumptions about and $\Delta_i(\cdot)$ and $\Gamma_i(\cdot)$. For $\Delta_i(\cdot)$, we adopt the Poisson assumption from Camerer et al. [2004]:

Assumption 2.2 $\Delta(\cdot)$ *is an independent multivariate Poisson distribution truncated at* $k-1$*, as given by Poisson Cognitive Hierarchy model.*

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

Assumption 2.3 $\Gamma_i(\cdot)$ *is a uniform distribution.*

Allowing for other distributions, such as Normal, is possible, though it increases the computational burden as one needs to solve the first-order condition by successive approximations.

Proposition 2.2 *Under the assumption that* $\Gamma_i(\cdot)$ *is uniform, the value of* $\gamma(\cdot)$ in H_S *and* H_p *is independent of rival type, so the first-order condition simplifies to*

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

It is computationally straightforward to solve for the $\hat{S}^k_{it}(p)$ that solves the above equation. This yields a straightforward method to calculate firm i's best-response bid function for any type k. To see this, note that the equation above is just the familiar "inverse elasticity pricing rule". Firm markups of bid over marginal cost are inversely proportional to their residual demand elasticity. Each component of the residual demand function can be iteratively solved for, using our data and Assumptions 2.1-2.3.

2.5 Estimation and Results

2.5.1 Details on Estimation

Estimation follows a minimum-distance approach. Key to this approach is τ_i, a scalar that provides information about firm i's type. We assume that $\tau_i = \exp(X_i'\boldsymbol{\gamma})$ and, because X_i is public information, so it is τ_i.

Each firm i observes $\boldsymbol{\tau}_{-i}$, the vector of τ's of its rivals. Also, each firm i has private information about its own type. Assume firm i is type $k \in \{0, \ldots, K\}$. If $k = 0$, then, by Assumption 2.1 above, firm i would submit a vertical bid on its own contract position, regardless of its rivals. For all $k \neq 0$, firms have beliefs about its rivals' type. Specifically, by Assumption 2.2, these believes are assumed to follow a Poisson distribution truncated at k, meaning that firm i believes all its rivals to be type $k-1$ or less. The specific probability associated with each type varies according to each rivals' τ.

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

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

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

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

where $\mathbb{P}(p)$ corresponds to the probability of observing a price point p as given by the triangular distribution and $\mathbb{P}_i(k|\ |K|, \hat{\boldsymbol{\gamma}})$ corresponds to the probability of bidder i being type k, conditional on there being $|K|$ possible types and $\hat{\gamma}$ being the estimated parameters.

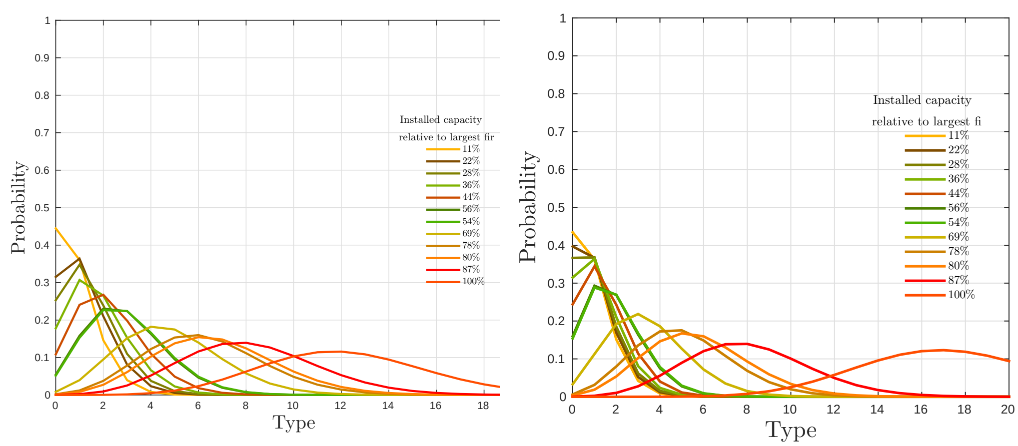
2.5.2 Results: Estimated Parameters

To guarantee that we find the global minimum, we run estimation starting from 50 sets of random initial points. We then re-estimate via Bootstrap starting each Bootstrap estimation from the set of points that minimized the objective function in the first stage. We do this for two specifications that differ in how τ_i is computed. The first specification includes only size in the parametrization of τ, while the second one uses size and size squared. The estimated parameters are reported in Table 2.1 and the implied probability distributions over type for the first two specifications are presented in figure 2.4.

Table 2.1: Structural model: estimated parameters

	(1)	(2)
Constant	-0.547	-0.126
	(0.164)	(0.084)
Size	15.625	-1.604
	(1.766)	(0.357)
Size^2		88.257
		(6.558)
Number of auctions	99	99

Note: Bootstrapped standard errors using 20 samples.



(a) $\tau_i = \exp(\gamma_0 + \gamma_1 size_i)$

(b) $\tau_i = \exp(\gamma_0 + \gamma_1 size_i + \gamma_2 size_i^2)$

Figure 2.4: Estimated distributions of types allowing for 21 types (level-0 and 20 levels of strategic sophistication)

It is important to make a number of comments regarding the estimates and figures. First, we expect the constant to be negative in order to rationalize level-0 players, as a positive constant would decrease the probability of observing a level-0 player significantly. Note, however, that this is not required by the CH model as one need not observe level-0 behavior in the data. However, as we have specified level-0 behavior according to what we observe in our data, a negative constant shows that the type of level-0 behavior that we have assumed is not uncommon.

Second, as described above, larger firms appear to be higher-level thinkers, though there is significant heterogeneity across firms, which shows up in that the first specification has a positive coefficient on size but the second specification has a negative one and a positive coefficient on sized squared. This means that only the largest firms actually engage in behavior that is similar to what a Bayesian Nash model would predict.

Third, the number of types allowed in estimation plays an important role. Specifically, if one allows for many types wanting the data to "talk" about which types are actually relevant, then for a large number of types, all predicted bids are identical as bids converge as type increases. As a consequence, starting estimation from multiple initial values is critical as one may reach local minima. Having said this, we do not find two set of different estimated parameters that result in the minimum objective function.

2.6 Counterfactual Analysis: Increasing Strategic Sophistication

Having estimated our model of bidding behavior that allows for heterogeneity in strategic sophistication, we now turn to a key question of this paper: How does the lack of strategic sophistication affect market efficiency? As described above, HP show that most of all productive inefficiencies can be explained by bidding departing from Bayesian Nash bidding, while the rest is explained by the exploitation of market power in an oligopoly setting. For this reason, we now turn to studying how increasing strategic sophistication of low-type firms may affect efficiency. We do this in two steps. We first ask how exogenous increases in strategic sophistication of specific firms affect market efficiency. We believe this is an important first step as market structure does not change with this intervention. Hence, though the actual intervention may appear as unreal (though consulting and hiring more qualified employees to operate the trading floor are probably good examples that could fit in this description), it provides a way to isolate the impact of increasing sophistication in the absence of changes in market power. We then turn to studying how increases in strategic sophistication that result from low-type firms merging with high-type firms, may affect market efficiency. In this case, bidding approaches that of Bayesian Nash but market concentration increases as well. Hence, the overall effect of the merger is ex-ante unknown.

To keep our results in perspective, it is important to note that there is an upper bound on the magnitude of the effects studied in our simulations. Indeed, while a social planner would minimize dispatch cost by inducing generation at marginal costs, we lack data on marginal costs for all bidders. These means that in our simulations the benchmark will not be the outcome of the social planner but that of a planner that forces firms in the CH to bid at their marginal costs but keeps bids of firms not included in the CH as they are in the data. For this reason, we measure all resulting inefficiencies with respect to this benchmark. Nonetheless, because the unmodeled fringe includes firms for which we do have marginal cost data for some auctions, we define as "Social Planner" the outcome of the simulation in which all firms for which we have marginal cost data bid their marginal costs and those firms for which we do not have marginal cost data bid according to their realized bids.

2.6.1 *Exogenous Increase*

Table 2.2 present results for the two specifications presented in table 2.1. We report results separately for auctions with positive and balancing demand as this makes comparisons easier.

The results show that, regardless of the specification, exogenously increasing sophistication of a large fraction of low-type firms has significant impact on efficiency. Indeed, when using the first specification of table 2.1 and considering auctions with positive balancing demand (INC side), we find that the estimated inefficiency decreases by 17.4% relative to the baseline (the model at the estimated parameters), with the remaining inefficiencies being caused by the exploitation of market power. In the case of auctions with negative balancing demand (DEC side, column 2), we find that increasing sophistication results in reducing inefficiencies by 8.3%. Similar results are obtained when considering the second specification of table 2.1.

The results also show that there are decreasing returns to increasing sophistication. Indeed, the last row of table 2.2 shows that increasing sophistication to the median type results in essentially identical efficiency gains to those that follow from increasing sophistication to the highest estimated type.

Finally, there is one important result that is not presented in the table and that has to be discussed. This is that increasing sophistication of just one firm has little effect on efficiencies. It is necessary for a group of low-type firms to increase sophistication for efficiency to increase significantly.

2.6.2 *Endogenous Increase: Mergers*

We now turn to studying how mergers may affect efficiency. As mentioned above, we focus on potential mergers that do not generate cost synergies but do increase concentration. In this setting, mergers can only increase efficiency by relocating

Table 2.2: Counterfactuals: exogenous increase in sophistication

	Average cost of generation (US dollars)			
	$\hat{\tau}_i = \exp(\hat{\gamma}_0 + \hat{\gamma}_1 \text{ size}_i)$		$\hat{\tau}_i = \exp(\hat{\gamma}_0 + \hat{\gamma}_1 \text{ size}_i + \hat{\gamma}_2 \text{ size}_i^2)$	
Scenario	INC side	DEC side	INC side	DEC side
Social Planner	8721.28	-46674.69	8721.28	-46674.69
CH firms bidding MC	16352.42	-39536.33	16352.42	-39536.33
Baseline	23663.82	-31850.44	24425.79	-31221.04
Low-type firms to high	22390.97	-32505.22	22480.30	-32418.45
Low-type to median	22420.11	-32488.84	24214.91	-32402.31

Note: These numbers are computed using the estimates in table 2.1. Calculations are done separately for auctions with positive balancing demand (INC side, 60 auctions) and negative balancing demand (DEC side, 39 auctions).

generation from high-cost, high-type firms to low-cost, low-type firms that have priced themselves out of the market. For this reason, to model the merger we take into account two sources of data. First, the marginal cost of each of the firms involved in the merger. Second, the day-ahead schedule of each of the merging parties. Then, we horizontally add the marginal cost functions and the day-ahead schedules to compute the marginal cost of supplying power to the grid for the merged firm, relative to the aggregate day-ahead schedule. This is shown in figure 2.5 for one of the auctions in the data.

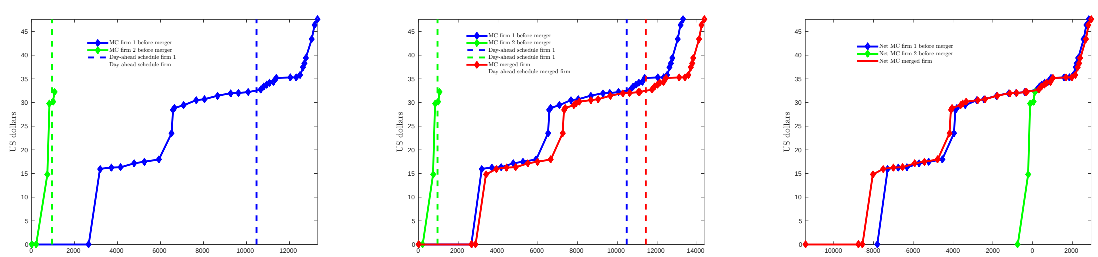


Figure 2.5: Marginal costs, day-ahead schedule, and net marginal costs

In this setting, we explore two potential mergers. The first one corresponds to a merger between a small firm and the largest one. The second one considers the merger of the two largest firms. These results, that only consider the second specification of table 2.1, show that mergers between the two largest firms result in the smallest increase in efficiency because there is little generation to rellocate and the increase in market power limits the gains from this rellocation. Furthermore, the increase in sophistication of the second largest firm has no impact on the bids of the smaller firms under the assumptions of the CH model as the second largest firm was of higher type than all small firms before the merger. Hence, the gains in efficiency from this merger are small and only due to rellocation of generation. Second, the biggest gain in efficiency is obtained from the merger between the largest and the smallest firm (19.6% reduction in inefficiency for auctions with positive balancing demand and 68% for auctions with negative balancing demand). These gains, however, come from three sides. First, there is a direct effect of rellocation of generation. Second, the newly formed firm also has correct beliefs about its rivals and bids more competitively. Third, all rivals observe the increase in sophistication of the generating units that belonged to the smallest firm in the merger and bid more aggresively (conditional on their types). This is recognized by the newly created firm and induces this firm to bid more aggresively too.

2.7 Summary

Models of strategic equilibrium form the foundation of many studies in industrial organization that investigate market efficiency in oligopoly settings. These models rely on the existence of a unique mapping from unobserved fundamentals, such as marginal costs or valuations, to observed prices or bids, to study questions about market efficiency and evaluate policy interventions, among others. However, there is

Table 2.3: Counterfactuals: increasing sophistication via mergers

	Average cost of generation (US dollars) $\hat{\tau}_i = \exp(\hat{\gamma}_0 + \hat{\gamma}_1 \text{ size}_i + \hat{\gamma}_2 \text{ size}_i^2)$	
Scenario	INC side	DEC side
Social Planner	8721.28	-46674.69
CH firms bidding MC	16352.42	-39536.33
Baseline	24425.79	-31221.04
Merger 1: Small and big firm	22840.84	-36850.03
Merger 2: Two largest firms	24070.73	-32727.13

Note: These numbers are computed using the estimates in the second specification of table 2.1. Calculations are done separately for auctions with positive balancing demand (INC side, 60 auctions) and negative balancing demand (DEC side, 39 auctions).

some evidence suggesting that the application of such strategic equilibrium models to all settings has to be done with caution, as in some settings observed behavior may depart significantly and persistently from what equilibrium models predict. Furthermore, the literature has shown that these departures from (Bayesian) Nash behavior may have significant implications for efficiency.

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

3. A DYNAMIC MODEL OF ENDOGENOUS MERGERS

3.1 Introduction

Governments are frequently forced to predict the impacts of mergers on the evolution of markets because breaking up consummated mergers could be extremely costly. The technique known as merger simulation is usually used by governments and researchers to forecast price and welfare changes caused by prospective mergers and to challenge mergers that are predicted to increase market price and decrease welfare. In a standard merger simulation exercise, the demand system is recovered using pre-merger data, marginal costs are estimated using firms' first order condition and price and welfare changes in the next period are simulated under the assumption of a static oligopoly game. As is mentioned in Weinberg and Hosken [2013], there might be bias in such simulation. The bias might come from the fact that only changes in the period right after the merger are considered in a merger review. There are many reasons to believe long-term welfare changes after a merger should be reviewed as well because a merger could have a considerable impact on the dynamic evolution of the market. Structural changes of the market after the merger would lead to post-merger changes in firms' behavior. Both merging and non-merging firms would adjust their entry, exit, investment or price strategies according to the new market structure. In addition, a current merger might trigger future mergers, therefore, mergers happening after a successful one should also be considered when reviewing a prospective merger.

Take the U.S. airline industry as an example. The Delta-Northwest merger was proposed in April 2008 and was cleared by the Department of Justice on October 29, 2008. A standard merger simulation exercise would suggest that this merger would reduce the competition, allow airlines to raise their ticket prices and provide poorer service. However, raw data exhibited exactly the opposite. Memphis, for example, was considered to be one of the worst affected cities by the Delta-Northwest merger in terms of market concentration. According to figure 3.1, the average ticket prices of Delta and Northwest airlines in Memphis decreased after the merger while the market price in Memphis stayed the same. Furthermore, figure 3.2 shows that the average flight frequency of Delta and Northwest increased after Northwest began to use Delta as its title in 2010. It seems that traditional merger simulations fail to forecast price and service quality changes after that merger. Figure 3.3 provides one of the many reasons as to why this happens. According to figure 3.3, the number of markets Delta used to serve in U.S. domestic market decreased from about 500-650 to 350-500. In fact, Delta began to cut off the routes served in Memphis after the merger and later the Memphis hub was officially closed in September, 2013. Since their market power after merger in Memphis was reduced because of exit behavior, average itinerary fare of Delta in Memphis decreased instead of increasing. Furthermore, Delta could concentrate on those remaining markets to provide higher flight frequency since they reduced the number of markets they served. Clearly that exit behavior and many other firm behaviors after a prospective merger could have significant effect on the evolution of a market, but traditional static merger simulation exercises might miss these changes and therefore provide poor forecast.

All of these suggest a need to build dynamic models of mergers. However, the merger literature has moved very slowly in doing so. It is not that people are not aware of the dynamic features mergers exhibit, but the complication that come with modeling dynamic mergers and the computational burden involved in estimating those models have kept researchers from developing a more flexible model. A small number of papers analyze mergers in dynamic settings. Gowrisankaran [1999] suc-

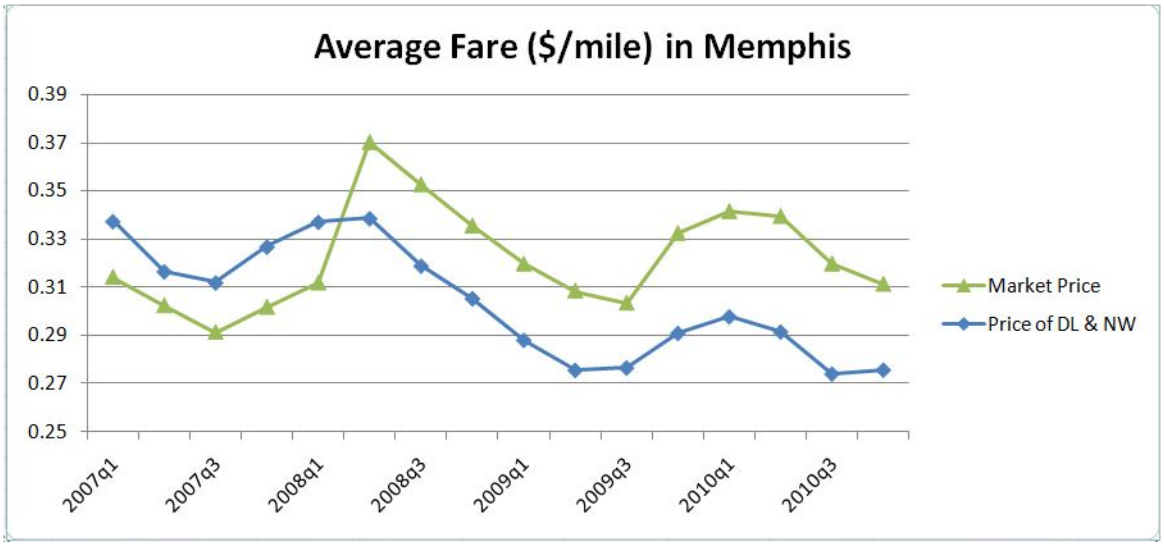


FIGURE 3.1: Average ticket price of Delta and Northwest and market average price in Memphis - The data comes from the Airline Origin and Destination Survey (DB1B) provided by the U.S. Department of Transportation (DOT).

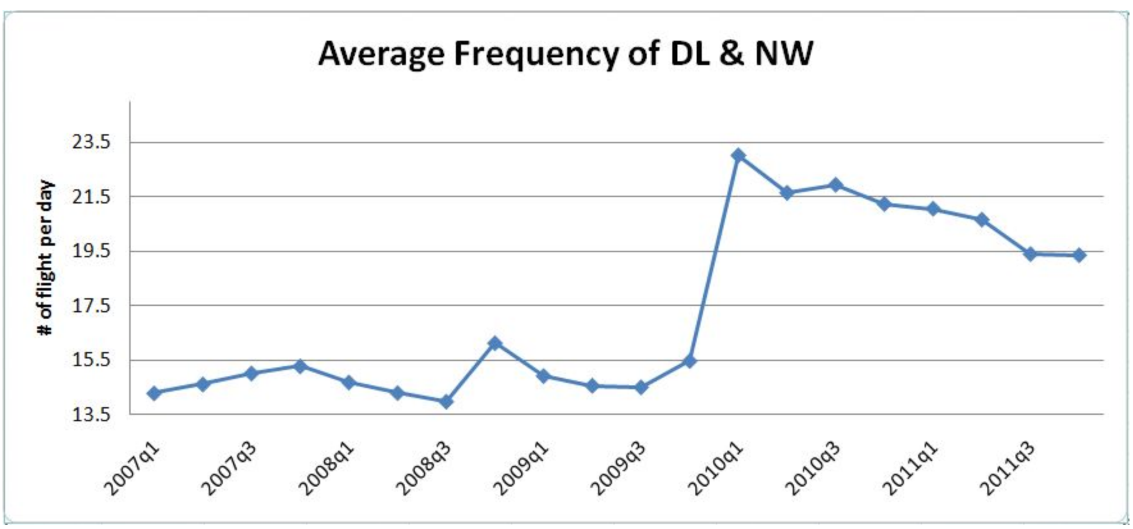


FIGURE 3.2: Average flight frequency of Delta and Northwest - The data comes from the Airline Origin and Destination Survey (DB1B) provided by the U.S. Department of Transportation (DOT).

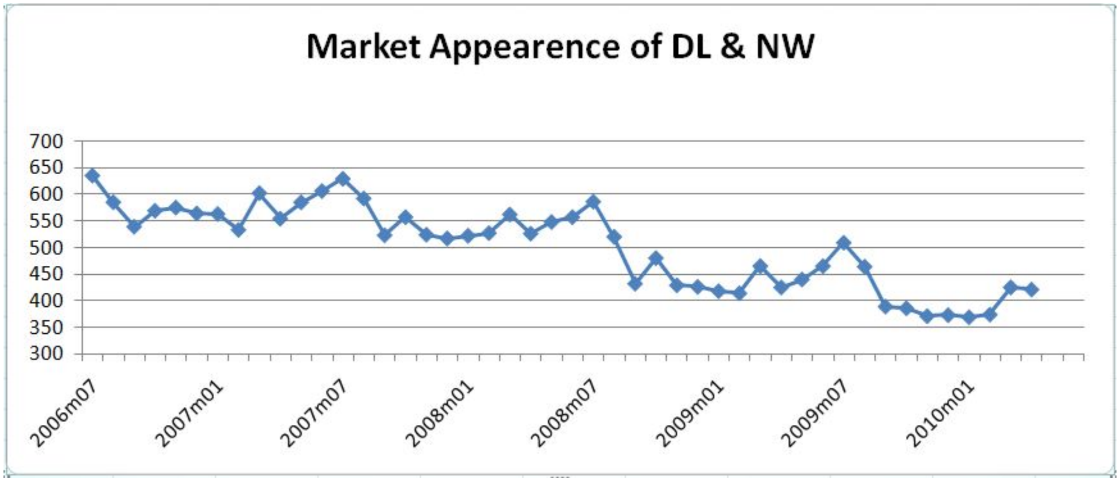


FIGURE 3.3: Number of markets that Delta and Northwest served in the U.S. domestic market - The data comes from the Airline Origin and Destination Survey (DB1B) provided by the U.S. Department of Transportation (DOT).

cessfully models a dynamic endogenous merger process. However, his paper imposes equilibrium selection rule on merger game's multiple equilibria and can only simulate the effect of artificial mergers. In his sequential game, larger firms have the priority to merge with smaller firms. Smaller firms can merge with other small firms only when all larger firms have failed in their merger attempts. Additionally, his model fails to include the situation where a small firm acquires a large firm. In Chen [2009], a dynamic oligopoly model under capacity constraint is used, but the merger is just a one-time exogenous shock in his model. These two methods are complicated and are hard to apply when considering prospective mergers.

Benkard et al. [2010] examines the effect of several mergers within the U.S. airline industry in a dynamic setting. The straightforward simulation method proposed in their paper provides a computationally easy way to implement model by avoiding repeatedly computing the Markov perfect Nash equilibrium in a dynamic oligopoly model. However, like Chen [2009], their method also assumes that a merger is an one-time exogenous shock. They can only exam the medium and long term effect of a merger without considering future mergers in markets. The dynamic endogenous merger process is not modeled in their method.

The theoretical equilibrium concept of dynamic oligopoly game, referred to as Markov perfect Nash equilibrium, is proposed in Maskin and Tirole [1988a, 1988b]. Ericson and Pakes [1995] add to this a framework by which we can empirically analyse dynamic oligopoly game. Their method has proven versatile in analysing dynamic evolution of markets. The early estimation methods for dynamic oligopoly game turn out to be quantitatively burdensome and difficult to apply to industries with a more complicated structure.

In recent years, researchers have proposed several new methods to ease the computational burden of estimating dynamic oligopoly model. These include Bajari et al. [2007] (hereafter BBL), Aguirregabiria and Mira [2007], and Pakes et al. [2007], etc. With newly available estimation methods, the Ericson and Pakes [1995] model has been used to tackle many empirical problems. Recent empirical applications of dynamic oligopoly models using new computational methods include advertising, capacity accumulation, and learning by doing. In this paper, we extend the classic Ericson and Pakes [1995] model and include merger as a dynamic strategy. To be more specific, we built a dynamic game with five steps in each period. These five steps are merger, exit, production, entry and investment respectively.

Our infinite repeated five-step game follows Gowrisankaran [1999], but our assumption of merger game diverges from his framework. As is mentioned in Gowrisankaran [1999], merger games involving more than two firms possess the problem of multiple equilibria, he has to impose equilibrium selection rule on his merger game because of the limitation of econometric method in solving a multi-equilibria game.

In recent years, much work has been done to solve the inference of multiple equilibria. One of the methods that has been fruitful in both applications and econometric theory is the moment inequality condition method.[1] With the help of newly available econometric estimation technique, we are able to build a more intuitive dynamic merger process and identify that merger game. In our game, every firm proposes a simultaneous bid for every other firm and it gets a asking price from other firms. As long as the post merger expected discounted value (value function) of the buyer is larger than the sum of the before merger expected discounted value (value function) of the two participants, the merger might happen. The particular merger which occurs depends on factors found unnecessary in our model. For instance, the final buyer may make her move earliest or act more aggressively. This more intuitive game can be identified with the necessary conditions in our game. The moment inequality conditions method proposed in Tamer [2003] identifies the binary choice in a two agent game with multiple equilibria and is applied in our multiple choice multiple agents game as well.

We propose an infinite repeated dynamic game with five steps in each period and a three-step estimation method to identify our model. We estimate the model from the last step to the first step using backwards induction. The last four steps of the game (exit, production, entry and investment) are simultaneously estimated following the spirit of the two-step estimation proposed by BBL [2007]. This method is an appealing solution to deal with complicated industry structures. It can capture the industry structures easily and the compute value function relatively quick by avoiding computing Markov perfect Nash equilibrium.

In the first step, we estimate policy functions and transition functions of the

[1] Ciliberto and Tamer [2009] uses moment inequality condition studying firm entry behavior in the U.S. airline industry.

dynamic game. In the second step, value functions of the game under various initial states are simulated. Then, in the third step, the moment inequality condition is implemented to identify the merger game using the value function estimators from the second step. With a fully identified structural model, we can forecast medium and long term effects of mergers. Additionally, the model forecasts the effect of a merger more accurately when there exists strategic behavior after a merger. The three-step estimation method presented here proves relatively simple to compute and easily applied to review prospective mergers.

3.2 Markov Perfect Nash Equilibrium

In a typical dynamic oligopoly model, time is infinite and firms repeatedly choose their strategies at each period. For example, in the classic Ericson and Pakes [1995] model, firms make entry and exit decisions and choose their investment and production level according to current state at each period. Markov perfect Nash equilibrium (hereafter MPNE), which is a generalization of Nash equilibrium, is the equilibrium concept used in dynamic oligopoly games. In the games, no firm wants to unilaterally deviate from a strategy set in MPNE. Every agent chooses strategies (policy functions) to maximize her expected discounted profit (value function) each period. The firm value function is represented as the conditional expectation of discounted long-term profits:

$$E[\Sigma_t^{\infty}\beta^{\tau-t}\pi_i(\boldsymbol{\sigma}_\tau, \mathbf{s}_\tau, v_{i\tau})|\mathbf{s}_\tau]$$

Here, π_i is profit function, $\boldsymbol{\sigma}_\tau$ is a vector that contains policy functions of each firm, $\mathbf{s}_\tau$ represents state such as firm capacity or product quality at time τ, $v_{i\tau}$ is firm specific random shock, and β is discount factor, which captures the fact that firms value current profit more than future profits.

The distribution of state in the next period only depends on all firms' policy function made in this period and current state. Therefore, state in the next period s_{t+1} is distributed as a Markov process:

$$P(\mathbf{s}_{t+1}|\boldsymbol{\sigma}_t, \mathbf{s}_t)$$

Since the behavior is given by Markov stratgy, we can rewrite value function as a Bellman equation:

$$V_i(\mathbf{s};\boldsymbol{\sigma}) = E_v\left[\pi_i(\boldsymbol{\sigma}(\mathbf{s},\mathbf{v}),\mathbf{s},v_i) + \beta\int V_i(\mathbf{s}',\boldsymbol{\sigma})dP(\mathbf{s}'|\boldsymbol{\sigma}(\mathbf{s},\mathbf{v}),\mathbf{s})|\mathbf{s}\right]$$

We claim $\boldsymbol{\sigma}$ is Markov Perfect Nash Equilibrium if for each firm, σ_i is her best response given that other firms do not change their policy function:

$$V_i(\mathbf{s};\boldsymbol{\sigma}) \geq V_i(\mathbf{s};\sigma_i',\boldsymbol{\sigma}_{-i})$$

Here,

$$\begin{aligned} V_i(\mathbf{s};\sigma_i',\boldsymbol{\sigma}_{-i}) = &E_v\Big[\pi_i(\mathbf{s};\sigma_i'(\mathbf{s},v_i),\boldsymbol{\sigma}_{-i}(\mathbf{s},\mathbf{v}_{-i}),v_i) \\ &+\beta\int V_i(\mathbf{s}';\sigma_i'(\mathbf{s},v_i),\boldsymbol{\sigma}_{-i}(\mathbf{s},\mathbf{v}_{-i}),\mathbf{s})dP(\mathbf{s}'|\mathbf{s},\sigma_i'(\mathbf{s},v_i),\boldsymbol{\sigma}_{-i}(\mathbf{s},\mathbf{v}_{-i}))|\mathbf{s}\Big] \end{aligned}$$

3.3 Model: An Infinite Repeated Five-Step Game

In a static merger game, a merger is usually followed by a one-period production which is assumed to be some production games as shown in Cournot or Bertrand. Our endogenous dynamic merger game is an infinitely repeated version of the static one. To be more specific, a merger process is followed by an exit, a Bertrand production game among the firms that are still in the market, and finally investment and entry decisions. These five steps are then repeated over and over. Our dynamic model is also an extension of the Ericson and Pakes [1995] model. We add merger process into their game, which involves exit, production, entry and investment. Our model captures firms' dynamic behaviors after a merger and that merger's impact on the evolution of markets.

3.3.1 Merger

Two obstacles impede the development of modeling dynamic merger procedure: nonexistence of equilibrium and multiple equilibria. We detail the reasons below and explain our methods to solve each setback.

Existence of equilibrium of dynamic model is generally shown by Brouwer's fixed-point theorem which requires the continuity of the operator, but here the after merger value function of nonparticipants are discontinuous. We follow Gowrisankaran [1999]'s method and add a source of randomness to the merger process, but our assumption on the source of randomness differs in that we assume that a firm's after merger value function is not only determined by the value function predicted under the post-merger industry structure, but also other factors like bargaining power. For instance, even if a potential buyer knows the after merger value of the firm that she is interested in, the potential seller might want to bargain a higher price. Therefore, the merger would proceed under a higher price. Here, we assume this randomness from bargaining is independently and identically drawn from a mean-zero uniform distribution $U[-\varsigma, \varsigma]$.

To the best of our knowledge, Gowrisankaran [1999] is the only paper which successfully models an endogenous merger process. He solves multiple equilibria by imposing a sequential merger process assumption. In his game, larger firms have the priority to merge with smaller firms. Smaller firms can merge with others only after all larger firms have failed in their attempts in mergers. However, his model fails to include the situation where a small firm acquires a large firm. Moreover, he claims that even under this assumption, the multiple equilibrium problem is not solved. A sequential merger process can only avoid some of the major sources of multiple equilibria. Instead of imposing restrictive assumption on merger game, we will give a more intuitive merger game and use moment inequality method proposed in Tamer [2003] to solve the game with multiple equilibria directly. We will explain how this method works in more detail in Section 3.4.

Let us first describe the merger game process. In our game, every firm proposes a simultaneous bid for every other firm and it gets a asking price from other firms. As long as the post merger expected future value of the buyer is bigger than the sum of the before merger expected future value of the two participants, the merger might happen. The particular merger which would happen depends on factors not specified in the game which will not affect our estimation. For instance, the final buyer might make her move earliest. In this case, let $\boldsymbol{m}_l = \{m_1, m_2, ..., m_n\}$ denote one possible merger outcome in a market where $m_i \in \{-1, 0, 1\}$. Here, -1 represents a scenario in which firm i is merged by another firm, 0 means that firm i does not participate in the merger process, and 1 means that firm i is the buyer.

Let M denote a set that contains all merger outcomes. Define M^M as $M/\{0, ..., 0\}$. Here, M^M contains all merger outcomes except the one that no merger happens ($\{0, ..., 0\}$). Let $V^{bf}_{buyer}(\mathbf{s}^{bf}, \boldsymbol{m}_l)$ denote the value function of buyer before merger, $V^{af}_{buyer}(\mathbf{s}^{bf}, \boldsymbol{m}_l)$ represent the value function of buyer after merger, u_{lk} denote the random shock we have mentioned earlier, and $V^{bf}_{seller}(\mathbf{s}^{bf}, \boldsymbol{m}_l)$ denote the value function of buyer before merger. For a specific merger outcome $\boldsymbol{m}_l$, it is one potential outcome in the market under before merger state $\mathbf{s}^{bf}$ if:

$$V^{af}_{buyer}(\mathbf{s}^{bf}, \boldsymbol{m}_l) + u_{lk} \geqslant V^{bf}_{buyer}(\mathbf{s}^{bf}, \boldsymbol{m}_l) + V^{bf}_{seller}(\mathbf{s}^{bf}, \boldsymbol{m}_l) \tag{3.1}$$

Let $M^M_{(1)}$ denote the set of merger outcomes in M^M that satisfy condition (3.1).

(a) If $|M^M_{(1)}| > 1$, there are more than one $\boldsymbol{m}_l$ which satisfy (3.1) and any one of them can be the merger outcome, i.e., $\forall j, m_j \in M^M_{(1)}$ merger outcomes exist. As mentioned previously, outside factors not considered here may determine the chosen merger.[2]

(b) If $|M^M_{(1)}| = 1$, there is only one $\boldsymbol{m}_l$ which satisfies (3.1). In this case, $\boldsymbol{m}_l$ is the merger outcome.

(c) If $|M^M_{(1)}| = 0$, there is no $\boldsymbol{m}_l$ which satisfies (3.1). Then no merger will happen making $\{0, 0, ..., 0\}$ the only outcome.

3.3.2 Exit

After the merger, every incumbent firm decides simultaneously whether or not to exit the market with a scrap value ϕ. Suppose there are n incumbent firms in the market and we define $\chi(\boldsymbol{s}) = \{\chi_1(\boldsymbol{s}), ..., \chi_i(\boldsymbol{s}), ..., \chi_n(\boldsymbol{s})\}$ as a vector of exit policy functions. An incumbent firm i will exit the market when its scrap value ϕ is larger than its value function.

$$\chi_i(\boldsymbol{s}) = \begin{cases} 1 & \text{if } s_i > 0 \;\&\; \phi > \pi_i(\bar{s}_i) - x_i(\bar{s}_i) + \beta \int V_i(\bar{\mathbf{s}}; \sigma(\mathbf{s})) dP(\bar{\mathbf{s}}|\sigma(\mathbf{s}), \mathbf{s}) \\ 0 & \text{otherwise} \end{cases}$$

where $\bar{\boldsymbol{s}} = (\chi_1, ..., \chi_{i-1}, 1, \chi_{i+1}, ..., \chi_n) \cdot \mathbf{s}$ represents after merger state in the next period, and $\sigma(\mathbf{s})$ is policy functions at state $\mathbf{s}$. $\chi_i(\boldsymbol{s}) = 1$ means the firm will stay in

[2] Here, $|\cdot|$ denotes the number of value in the set.

the market and $\chi_i(\boldsymbol{s}) = 0$ means the firm will exit the market.

3.3.3 Production

Following the typical assumption of the Ericson and Pakes model, we model production process as a static Bertrand game. We assume a logit demand system. Consumer r gets utility U_{ri} from consuming good i,

$$U_{ri} = \gamma_0 ln(s_i) + \gamma_1 ln(y_r - p_i) + \varepsilon_{ri}$$

where s_i represents the quality of good i, y_r is person r's income, p_i is the price of good i, and ε_{ri} is a independently and identically distributed logit error term. For simplicity, incomes for all consumers are assumed to be a constant, $y_r = y$. We also assume that each firm has a constant marginal cost:

$$mc(q_i, \mu) = \mu.$$

3.3.4 Entry and Investment

The last two steps of the game are entry and investment. It is common practice to assume that they happen simultaneously at each period because both of them contribute to the state evolution. A firms state s_{it+1} at period $t+1$ is affected by its investment in period t in the following way:

$$s_{it+1} = s_{it} + \nu_{it} - \bar{\nu}_t$$

Here, ν_{it} represents firm i's random return of investment x_{it} realized in the next period and α denotes the parameter.

$$\nu_{it} = \begin{cases} 1 & \text{with prob. } \alpha x_{it}/(1+\alpha x_{it}) \\ 0 & \text{with prob. } 1/(1+\alpha x_{it}) \end{cases}$$

$\bar{\nu}_t$ is an industry wide random depreciation at period t and market state decreases by 1 with probability δ.

$$\bar{\nu}_t = \begin{cases} 1 & \text{with prob. } \delta \\ 0 & \text{with prob. } 1-\delta \end{cases}$$

For a situation in which entry happened simultaneously with investment, each potential entrant receives a random entry cost from a uniform distribution $U(-e, e)$. The firm knows its entry cost before it makes entry decision. If it chooses to enter the market, potential entrants state at next period is $s^E - \bar{\nu}_t$.

3.4 Three-step Estimation Method

We estimate the model from the last step to the first step using backwards induction. The last four steps of the game (exit, production, entry and investment) are simultaneously estimated following the spirit of the two-step estimation proposed by BBL [2007]. The BBL [2007] method is chosen here because it is an appealing method to deal with complicated industry structures. It can capture the industry structures easily and compute the value function relatively quickly without computing MPNE even once. In the first step, we estimate policy functions and transition functions of the dynamic game. Second, value functions of the game under various initial states are simulated. Then, in the third step, we apply the moment inequality conditions proposed in Tamer [2003] to identify the merger game using the value function estimators from the second step.

To reduce computation, we assume firms' profit functions are linear in the unknown parameters like BBL [2007]. Therefore, their value functions are also linear in parameters in that:

$$V_i(\mathbf{s}) = W^1(\mathbf{s}) + W^2(\mathbf{s}) \cdot \xi + W^3(\mathbf{s}) \cdot \phi \tag{3.2}$$

Since $W^1(\mathbf{s})$, $W^2(\mathbf{s})$, $W^3(\mathbf{s})$ can be estimated from simulation in the 2nd step, estimation of $V_i(\mathbf{s})$ can be represented as a function with two parameters:

$$\hat{V}_i(\mathbf{s}) = \hat{W}^1(\mathbf{s}) + \hat{W}^2(\mathbf{s}) \cdot \xi + \hat{W}^3(\mathbf{s}) \cdot \phi \tag{3.3}$$

$$\equiv f(\xi, \phi, \mathbf{s}) \tag{3.4}$$

Now again, let us define V_i^{bf} as before merger incumbent value function and V_i^{af} as after merger incumbent value function for firm i. Define $\mathbf{s}^{bf} \in \mathbf{S}$ as before merger state and $\mathbf{s}_l^{af} \in \mathbf{S}$ as its after merger state when merger outcome is $\boldsymbol{m}_l$. We can estimate before merger value function from $\hat{V}_i^{bf}(\mathbf{s}^{bf}, \boldsymbol{m}_l) = f(\xi, \phi, \mathbf{s}^{bf})$ and after merger value function from $\hat{V}_i^{af}(\mathbf{s}^{bf}, \boldsymbol{m}_l) = f(\xi, \phi, \mathbf{s}_l^{af})$. Moving forward, we can identify the merger game with moment inequality conditions using these before and after merger value functions.

3.4.1 A Simple 2×2 *Entry Game*

In this section, a 2×2 game is used to show the identification problem of multiple equilibria and illustrate the main idea of moment inequality conditions which are proposed by Tamer [2003] to solve the problem of multi-equilibria. Then, we describe the moment inequality conditions which are used to identify our merger game.

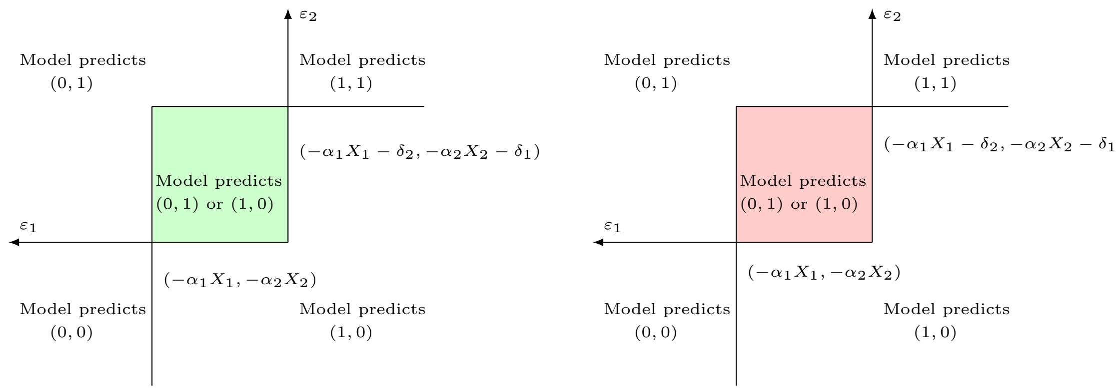


FIGURE 3.4: Entry game with multiple equilibria

Let us consider a binary simultaneous equation system:

$$y_1 = 1[\alpha_1' X_1 + \delta_2 y_2 + \varepsilon_1 \geq 0],$$
$$y_2 = 1[\alpha_2' X_2 + \delta_1 y_1 + \varepsilon_2 \geq 0],$$

where X_1 and X_2 are vectors of observed firm-specific exogenous regressors. This game can be viewed as a two agent entry game. A firm's entry decision depends on its own characteristics and whether its competitor enter the market. As is shown in Tamer [2003], this game has multiple equilibria if support for ε is large enough. In figure 3.4, multiple equilibria appear in the shade region.

This multiple equilibrium problem complicates the estimation process because it proves impossible to find out the theoretical probability of the realization of a market outcome when it is difficult to determine which equilibrium is realized in the overlapping multiple equilibrium region. Hence, an equilibrium selection rule must be specified in a traditional method. This rule is used to "pick up" an equilibrium from the multiple equilibria.

Usually, there is little guidance on how to choose the rule. To avoid an unrealistic assumption about the equilibrium selection, Tamer [2003] proposed the following method, which gets around imposing any assumption about the multiple equilibrium selection process.

Let us consider the case where $(y_1, y_2) = (1, 0)$ is the market outcome.

$$\begin{aligned} Pr(1,0|X) =& Pr((\varepsilon_1, \varepsilon_2) \in R_1(X, \theta)) + \\ & \int Pr(1,0|\varepsilon_1, \varepsilon_2, X) 1[(\varepsilon_1, \varepsilon_2) \in R_2(X, \theta)] dF_{\varepsilon_1, \varepsilon_2} \end{aligned}$$

where

$$\begin{aligned} R_1(\theta, X) =& \{(\varepsilon_1, \varepsilon_2) : (\varepsilon_1 \geq -\alpha_1' X_1; \varepsilon_2 \leq -\alpha_2' X_2\} \\ & \cup (\varepsilon_1 \geq -\alpha_1' X_1 - \delta_2; -\alpha_2' X_2 \leq \varepsilon_2 \leq -\alpha_2' X_2 - \delta_1)\}, \end{aligned}$$

$$\begin{aligned} R_2(\theta, X) =& \{(\varepsilon_1, \varepsilon_2) : (-\alpha_1' X_1 \leq \varepsilon_1 \leq -\alpha_1' X_1 - \delta_2; \\ & -\alpha_2' X_2 \leq \varepsilon_2 \leq -\alpha_2' X_2 - \delta_1)\}, \end{aligned}$$

$Pr(1, 0|\varepsilon_1, \varepsilon_2, X)$ is the selection mechanism for multi-equilibria.

Furthermore, the above equation implies the following inequality condition:

$$\Pr((\varepsilon_1, \varepsilon_2) \in R_1) \leq \Pr((1, 0)) \leq \Pr((\varepsilon_1, \varepsilon_2) \in R_1) + \Pr((\varepsilon_1, \varepsilon_2) \in R_2) \qquad (3.5)$$

In figure 3.5, the shaded area in the graph on the right hand side represents the region that would predict $(1, 0)$ uniquely, which is R_1. The shaded area in the graph on the left hand side represents the region that $(1, 0)$ is a possible outcome, which represents $R_1 + R_2$.

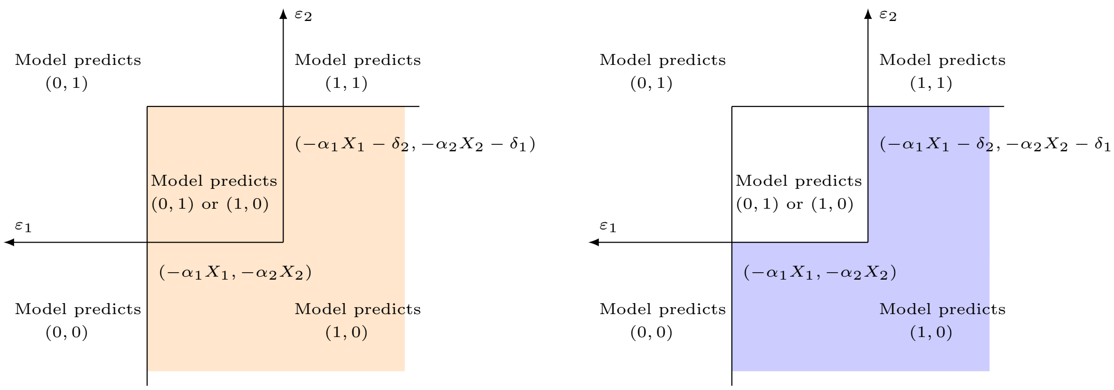


FIGURE 3.5: Upper and lower bounds for entry game

For the simple 2×2 game, we can represent all the inequality conditions in the following way:

$$\boldsymbol{H_1(\theta, X)} \equiv \begin{pmatrix} H_1^1(\boldsymbol{\theta}, X) \\ H_1^2(\boldsymbol{\theta}, X) \\ H_1^3(\boldsymbol{\theta}, X) \\ H_1^4(\boldsymbol{\theta}, X) \end{pmatrix} \leq \begin{pmatrix} Pr((0,0)|X) \\ Pr((1,0)|X) \\ Pr((0,1)|X) \\ Pr((1,1)|X) \end{pmatrix} \leq \begin{pmatrix} H_2^1(\boldsymbol{\theta}, X) \\ H_2^2(\boldsymbol{\theta}, X) \\ H_2^3(\boldsymbol{\theta}, X) \\ H_2^4(\boldsymbol{\theta}, X) \end{pmatrix} \equiv \boldsymbol{H_2(\theta, X)}$$

where $\boldsymbol{H_1(\theta, X)}$ is the lower bound function which represents the probability that a particular market structure is the unique equilibrium. $\boldsymbol{H_2(\theta, X)}$ is the upper bound function which also counts the probability that multiple equilibria happen in the market. $\boldsymbol{H_1(\theta, X)}$ and $\boldsymbol{H_2(\theta, X)}$ can be analytically solved when the distribution of $\boldsymbol{\varepsilon} = \{\varepsilon_1, \varepsilon_2\}$ is assumed to be known. For example, H_1^1 and H_2^1 can be solved from equation (3.5). The above inequality condition is then used to identify the 2×2 simple game.

3.4.2 Moment Inequality Conditions in the Merger Game

According to the merger game we propose in section 3.3, we can use the following moment inequality conditions to estimate the dynamic parameters. Suppose that $Pr(\boldsymbol{m}_l|\mathbf{s}^{bf})$ is the probability of merger outcome $\boldsymbol{m}_l$ happens when state is $\mathbf{s}^{bf}$. We can derive the inequality condition for the game as follows:

$$L(\boldsymbol{m}_l|\mathbf{s}^{bf}) \leq Pr(\boldsymbol{m}_l|\mathbf{s}^{bf}) \leq U(\boldsymbol{m}_l|\mathbf{s}^{bf})$$

Here, $L(\boldsymbol{m}_l|\mathbf{s}^{bf})$ is the lower bound for $\boldsymbol{m}_l$ to happen under state $\mathbf{s}^{bf}$. It is the probability that $\boldsymbol{m}_l$ is the only merger outcome that satisfy equation (3.1) at state $\mathbf{s}^{bf}$. $U(\boldsymbol{m}_l|\mathbf{s}^{bf})$ is the upper bound for $\boldsymbol{m}_l$ to happen under state $\mathbf{s}^{bf}$. It is the probability that $\boldsymbol{m}_l$ is the the merger outcome that satisfies equation (3.1) at state $\mathbf{s}^{bf}$. All the moment inequality conditions of the merger game are represented as follows:

$$\boldsymbol{L}(\boldsymbol{\vartheta}, \mathbf{s}^{\boldsymbol{bf}}) \equiv \begin{pmatrix} L^1(\boldsymbol{\vartheta}, \mathbf{s}^{bf}) \\ . \\ . \\ L^L(\boldsymbol{\vartheta}, \mathbf{s}^{bf}) \end{pmatrix} \leq \begin{pmatrix} Pr(\boldsymbol{m}_1|\mathbf{s}^{bf}) \\ . \\ . \\ Pr(\boldsymbol{m}_L|\mathbf{s}^{bf}) \end{pmatrix} \leq \begin{pmatrix} U^1(\boldsymbol{\vartheta}, \mathbf{s}^{bf}) \\ . \\ . \\ U^L(\boldsymbol{\vartheta}, \mathbf{s}^{bf}) \end{pmatrix} \equiv \boldsymbol{U}(\boldsymbol{\vartheta}, \mathbf{s}^{\boldsymbol{bf}})$$

where $\boldsymbol{\vartheta}$ is a vector that contains dynamic parameters ξ, ϕ and ς.

3.4.3 Simulate Upper and Lower Bounds

In the simple 2×2 game, analytical solutions of upper bound and lower bound can be easily derived. In our merger game, it is not easy to find the analytical solutions of the corresponding upper and lower bounds, but we follow Ciliberto and Tamer [2007]'s simulation method to find them. Let us rewrite equation (3.1) in another form:

$$\Pi(\mathbf{s}^{bf}, \boldsymbol{m}_l) \equiv V^{af}_{buyer}(\mathbf{s}^{bf}, \boldsymbol{m}_l) - V^{bf}_{buyer}(\mathbf{s}^{bf}, \boldsymbol{m}_l) - V^{bf}_{seller}(\mathbf{s}^{bf}, \boldsymbol{m}_l) + u_l \tag{3.6}$$

If $\Pi(\mathbf{s}^{bf}, \boldsymbol{m}_l) \geqslant 0$, $\boldsymbol{m}_l$ is one possible merger outcome when state is $\mathbf{s}^{bf}$.

We first draw R simulations of unobservable for each state $\mathbf{s}^{bf}$. Then we obtain the $\Pi(\mathbf{s}^{bf}, \boldsymbol{m}_l)$ for every possible $\boldsymbol{m}_l \in M$ as a function of states, observables, and parameters. If $\Pi(\mathbf{s}^{bf}, \boldsymbol{m}_l) \geqslant 0$ for some $l \in \mathbb{L}$, $\boldsymbol{m}_l$ is an potential outcome of that game (one of the equilibria). If this equilibrium is unique, 1 is added to the lower bound probability ($\hat{L}(\boldsymbol{m}_l|\mathbf{s}^{bf})$) for outcome $p(\boldsymbol{m}_l|\mathbf{s}^{bf})$ and 1 is added to the upper bound probability ($\hat{U}(\boldsymbol{m}_l|\mathbf{s}^{bf})$). If the equilibrium is not unique, then we only add a 1 to the upper bound. For example, the lower bound for merger outcome $\boldsymbol{m}_l$ under state $\mathbf{s}^{bf}$ is:

$$\hat{L}(\boldsymbol{m}_l|\mathbf{s}^{bf}) = \frac{1}{R}\sum_{j=1}^{R} 1\left[\Pi(\mathbf{s}^{bf}, \boldsymbol{m}_l) \geqslant 0, \Pi(\mathbf{s}^{bf}, \boldsymbol{m}_{-l}) < 0\right]$$

which simulate the probability that $\boldsymbol{m}_l$ is the only merger outcome at state $\mathbf{s}^{bf}$.

3.4.4 Identification

To recover the primitives of the model, we can minimize the following objective function,

$$Q_n(\boldsymbol{\theta}) = \frac{1}{n}\sum_{i=1}^{n}\left[||p(\boldsymbol{m}_l|\mathbf{s}^{bf}) - \hat{L}(\boldsymbol{m}_l|\mathbf{s}^{bf}_k)||_- + ||p(\boldsymbol{m}_l|\mathbf{s}^{bf}) - \hat{U}(\boldsymbol{m}_l|\mathbf{s}^{bf})||_+\right]$$

where $(A)_- = [a_1 1[a_1 \leqslant 0], ..., a_{n(n-1)+1} 1[a_{n(n-1)+1} \leqslant 0]]$ and similarly $(A)_+ = [a_1 1[a_1 \geqslant 0], ..., a_{n(n-1)+1} 1[a_{n(n-1)+1} \geqslant 0]]$ for a $n(n-1)+1$ vector A, and where $||\cdot||$ is the Euclidian norm.

3.5 Simulation Result

Table 3.1 shows Monte Carlo estimation result of the static primitives including demand system, marginal cost and investment evolution from our 3-step estimation method. In general, estimators are very close to true parameters. We follow Benkard et al. [2010] and Chernozhukov et al. [2007] to infer our dynamic primitives from the game. Chernozhukov et al. [2007], which is a iterative process involving subsampling from the second step, is used to do inference of the merger game. We only show the results from the first step in table 3.2 for now because it extremely time consuming to do subsampling. According to table 3.2, true value of investment cost falls in the estimated bounds. Moreover, lower and upper bounds are close to true value. However, our bounds of scrape value do not contain true value. Although true value of random shock fall in the estimated interval, the estimated interval are too wide to contain any information. We expect more accurate bound estimation after we finish iterative process of Chernozhukov et al. [2007].

Table 3.1: Endogenous merger game Monte Carlo results (static primitives)

Parameter	True Value	Mean	SE
Demand			
γ_0	0.1	0.100	0.000
γ_1	1.5	1.492	0.000
Marginal cost μ	3	2.991	0.000
Investment evolution			
δ	0.7	0.755	0.244
ρ	1.25	1.553	0.243

Table 3.2: Endogenous merger game Monte Carlo results (dynamic primitives)

Parameter	True Value	Bounds
Investment cost ξ	-1	[-2.152, 0.578]
Scrape value ϕ	6	[11.914, 30.000]
Random shock ς	2	[-28.578, 25.357]

4. CONCLUSION

This dissertation studies merger simulation exercises used by governments and researchers to evaluate prospective mergers. Specifically, we extend the traditional static merger models.

In the second chapter, we build and estimate a model in which firms are boundedly rational and simulate the effect of mergers in such settings. Models of strategic equilibrium form the foundation of many studies in industrial organization that investigate market efficiency in oligopoly settings. However, there is some evidence suggesting that the application of such strategic equilibrium models to all settings has to be done with caution, as in some settings observed behavior may depart significantly and persistently from what (Bayesian) Nash behavior models predict. In this chapter,we study bidding in the Texas electricity market, a market in which bidding by some firms departs significantly from what Bayesian Nash models predicts, while bidding from other firms closely resembles these predictions. Our unique dataset, in addition to our model, allows us to identify and estimate heterogeneity in levels of strategic sophistication across electricity generators. Our results show that not only exogenously increasing sophistication may increase efficiency significantly, but that also mergers that do not generate cost synergies but increase concentration may also increase efficiency as long as the higher sophistication of one of the merging parties is transfered to the rest of firms involved in the merger.

In the next chapter, we study mergers in a dynamic setting. In a standard merger simulation exercise, only the price and welfare changes in the next period are simulated under the assumption of a static oligopoly game. As is mentioned in Weinberg and Hosken [2013], there might be bias in such simulation. The bias might come from the fact that only changes in the period right after the merger are considered in a merger review. There are many reasons to believe long-term welfare changes after a merger should be reviewed as well because a merger could have a considerable impact on the dynamic evolution of the market. In this chapter, we build an infinite five-step repeated game under the framework of the Ericson and Pakes model [1995]. Then, we propose a three-step estimation method to estimate the game in which Markov perfect Nash equilibrium is played. Our three-step estimation method is flexible and can be easily modified to estimate various market structures.

REFERENCES

Victor Aguirregabiria and Pedro Mira. Sequential estimation of dynamic discrete games. *Econometrica*, 75(1):1-53, 2007.

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

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

Patrick Bajari, Lanier C. Benkard and Jonathan Levin. Estimating dynamic models of imperfect competition. *Econometrica*, 75(5):1331-1370, 2007.

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

Lanier C. Benkard, Aaron Bodoh-Creed, and John Lazarev. Simulating the dynamic effects of horizontal mergers: U.S. Airlines. Standford University working paper, 2010.

James Bushnell, Erin Mansur, and Celeste Saravia. Vertical arrangements, market structure, and competition: An analysis of restructured U.S. electricity markets. *American Economic Review*, 98(1):237-266, 2008.

Colin F. Camerer, Teck-Hua Ho, and Juin-Kuan Chong. A cognitive hierarchy model of games. *Quarterly Journal of Economics*, 119(3):861-898, 2004.

Jiawei Chen. The effects of mergers with dynamic capacity accumulation. *International Journal of Industry Organization*, 27(1):92-109, 2009.

Victor Chernozhukov, Han Hong and Elie Tamer. Estimation and confidence regions for parameter sets in econometric models. *Econometrica*, 75(5):1243-1284, 2007.

Federico Ciliberto and Elie Tamer. Market structure and multiple equilibria in airline markets. *Econometrica*, 77(6):1791-1828, 2009.

Miguel A. Costa-Gomes, Vincent P. Crawford, and Bruno Broseta. Cognition and behavior in normal-form fames: an experimental study. *Econometrica*, 69(5):1193-1235, 2001.

Vincent P. Crawford and Nagore Iriberri. Fatal attraction: Salience, naivete, and sophistication in experimental hide-and-seek games. *American Economic Review*, 97(5):1731-1750, 2007.

Vincent P. Crawford, Uri Gneezy, and Yuval Rottenstreich. The power of focal points is limited: Even minute payoff asymmetry may yield large coordination failures. *American Economic Review*, 98(4):1443-1458, 2008.

Ulrich Doraszelski, Gregory Lewis, and Ariel Pakes. Just starting out: Learning and price competition in a new market. University of Pennsylvania Working Paper, 2014.

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

Richard Ericson and Ariel Pakes. Markov-perfect industry dynamics: A framework for empirical work. *Review of Economic Studies*, 62(1):53-82, 1995.

Ben Gillen. Identification and estimation of level-k auctions. California Institute of Technology working paper, 2010.

Avi Goldfarb and Mo Xiao. Who thinks about the competition: Managerial ability and strategic entry in us local telephone markets. *American Economic Review*, 101(7):3130-3161, 2011.

Gautam Gowrisankaran. A dynamic model of endogenous horizontal mergers.

RAND Journal of Economics, 30(1):56-83, 1999.

Ali Hortaçsu and Jakub Kastl. Valuing dealers informational advantage: a study of canadian treasury auctions. *Econometrica*, 80(6):2511-2542, 2012.

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

Ali Hortaçsu and Steven L. Puller. Understanding strategic bidding in multi-unit auctions: A case study of the Texas electricity spot market. *RAND Journal of Economics*, 39(1):86-114, 2008.

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

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

Eric Maskin and Jean Tirole. A theory of dynamic oligopoly, I: Overview and quantity competition with large fixed costs. *Econometrica*, 56(3):549-569, 1988.

Eric Maskin and Jean Tirole. A theory of dynamic oligopoly, II: Price competition, kinked Demand curves, and edgeworth cycles. *Econometrica*, 56(3):571-599, 1988.

Richard D. McKelvey and Thomas R. Palfrey. Quantal response equilibria for normal form games. *Games and Economic Behavior*, 10(1):6-38, 1995.

Rosemary Nagel. Unraveling in guessing games: An experimental study. *American Economic Review*, 85(5):1313-1326, 1995.

Ariel Pakes, Michael Ostrovsky and Steven Berry. Simple estimators for the parameters of discrete dynamic games (with entry exit examples). *RAND Journal of Economics*, 38(2):373-399, 2007.

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

Elie Tamer. Incomplete simultaneous discrete response model with multiple equilibria. *Review of Economic Studies*, 70(1):147-165, 2003.

Matthew C. Weinberg and Daniel Hosken. Evidence on the accuracy of merger simulation. *Review of Economics and Statistics*, 95(5):1584-1600, 2013.

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

Catherine D. Wolfram. Strategic bidding in a multiunit auction: An empirical analysis of bids to supply electricity in England and Wales. *RAND Journal of Economics*, 29(4):703-725, 1998.

APPENDIX A

PROOFS

A.1 1st Order Condition of Optimality under CH

Suppressing the i and t indices, the maximization problem of player i with level k is given by,

$$\underset{\hat{S}(p)}{\text{Max}} \int_{\underline{p}}^{\overline{p}} \left(U \left(p \cdot S(p) - C \left(S(p) \right) - (p - PC)QC \right) \right) dH \left(p, S(p); k, QC \right)$$

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

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

Label the integrand:

$$F(p, S, S') = U'(pS(p) - C(S(p)) - (p - PC)QC)(pS'(p) + S(p) - C'(S(p))S'(p) - QC) \\ H\left(p, \hat{S}(p); k, QC\right)$$

$$F_S = - H_S^l U'(\cdot)(pS' + S - C'S' - QC) + H^l U''(\cdot)(p - C')(pS' + S - C'S' - QC) \\ + H^l U'(\cdot)(1 - C''S')$$

$$F_{s'} = -H^l U'(\cdot)(p - C')$$

$$\begin{aligned}\frac{d}{dp}F_{s'} =& - H^l_p U'(\cdot)(p - C') + H^l_S S' U'(\cdot)(p - C') \\ &+ H^l U''(\cdot)(p - C')(pS' + S - C'S' - QC) + H^l U'(\cdot)(1 - C''S')\end{aligned}$$

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

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

Therefore,

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

A.2 Proof of Proposition 2.2

Given additive separable form of bid function for bidders type 0 to type K from Proposition 2.1, uncertainty in rivals' contract obligation, QC_{-jt}, and the aggregate demand act a parallel shift in residual demand. Thus, all that matters is the distribution of a scalar random variable that is the sum of functions of rival contract position ($\sum_{j \neq i} QC_{jt}$) and total demand shock ($-\varepsilon_t$). Hence, for a bidder type k ($k = 0, 1, ..., K$), we can rewrite his believe about distribution of market clearing price as

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

Let θ_{it} denote $\sum_{j \neq i} QC_{jt} - \varepsilon_t$, $\Gamma(\cdot)$ be distribution of θ_{it} and $\Delta(l_{-i})$ denote the marginal distribution of the vector of rival firm types. Then, a bidder type k's believe about distribution of market clearing price involves two source of uncertainty, θ_{it} and rival type l_{-i}.

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

Take derivatives to find H_S and H_p

$$
\begin{aligned}
H_S &= \int_{l_{-i}} \gamma \left(D_t(p) - \sum_{j \neq i} \alpha_j^l(p) - \hat{S}_{it}^k(p) \right) \cdot \Delta(l_{-i}) \\
&= \int_{l_{-i}} \gamma^{l_{-i}}(p) \cdot \Delta(l_{-i})
\end{aligned}
$$

where $\gamma(\cdot)$ is derivative of $\Gamma(\cdot)$.

$$\begin{aligned} H_p &= -\int_{l_{-i}} \gamma\left(D_t(p) - \sum_{j\neq i}\alpha_j^l(p) - \hat{S}_{it}^k(p)\right)\left(D_t'(p) - \sum_{j\neq i}\alpha_j^{l\prime}(p)\right)\cdot\Delta(l_{-i}) \\ &= -\int_{l_{-i}} \gamma^{l_{-i}}(p)\left(D_t'(p) - \sum_{j\neq i}\alpha_j^{l\prime}(p)\right)\cdot\Delta(l_{-i}) \end{aligned}$$

Since the residual demand function faces by a bidder type k under a certain belief about rival type (l_{-i}) is given by

$$RD_{it}(p, \hat{S}_{it}^k(p); k_i, QC_{it}) = D_t(p) + \varepsilon_t - \sum_{j\neq i}\alpha_{jt}^l(p) - \sum_{j\neq i} QC_{jt}$$

with derivative

$$RD_{it}'(p) = D_t'(p) - \sum_{j\neq i}\alpha_{jt}^{l\prime}(p)$$

Therefore,

$$H_p = -\int_{l_{-i}} \gamma^{l_{-i}}(p) RD_{it}'(p)\cdot\Delta(l_{-i})$$

Replace H_s and H_p in equation (2.5) that represent first order condition of profit maximizing problem of a bidder type k, we get

$$p - C_{it}'\left(S_{it}^*(p)\right) = \frac{\int_{l_{-i}}\gamma^{l_{-i}}(p)\cdot\Delta(l_{-i})}{-\int_{l_{-i}}\gamma^{l_{-i}}(p)RD_{it}'(p)\cdot\Delta(l_{-i})}\left(S_{it}^*(p) - QC_{it}\right)$$

Then, under assumption 2.3, optimal bidding strategy for a type k bidder, $S_{it}^k(p)$, is traced out by

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