RICE UNIVERSITY

Essays in Power System Economics

by

Özgür İnal

A THESIS SUBMITTED
IN PARTIAL FULFILLMENT OF THE
REQUIREMENTS FOR THE DEGREE

Doctor of Philosophy

APPROVED, THESIS COMMITTEE:

Peter R. Hartley (signature)
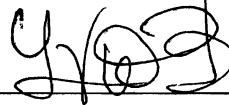
Peter R. Hartley, Chair
George and Cynthia Mitchell Chair in
Sustainable Development and
Environmental Economics, Professor of
Economics

(signature)
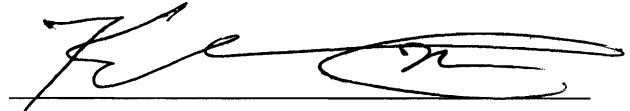
Simon Grant
Lay Family Professor of Economics

(signature)

Leonardo Dueñas-Osorio
Assistant Professor of Civil and
Environmental Engineering

(signature)

Kenneth B. Medlock, III
James A. Baker, III, and Susan G. Baker
Fellow in Energy and Resource
Economics, James A. Baker III Institute
for Public Policy
Adjunct Professor of Economics

Houston, Texas
April, 2011

Abstract

In the first chapter, we propose a new method for modeling competition in electricity spot markets, namely, by approximating the supply functions of the competitors with cubic splines. We argue that this method is preferable to approximation by linear or piecewise-affine functions, which have been the main approaches to date. We apply our method to the firms competing in the Texas market. We also show that, more often than not, we will observe that the marginal revenue functions of the firms will have increasing segments which may lead to multiple profit-maximizing optima for a firm.

In the second chapter, we model the effects of forward contracting on power prices in wholesale electricity markets. In contrast to most of the previous literature, we explicitly model power retailers, and introduce risk aversion. As expected, increasing the number of players have pro-competitive effects on the spot price of electricity. We also find that as the generators bid more competitively, spot and forward prices converge. Our model also captures the effects of level and variability of power demand on the players' contracting decisions.

In the final chapter, we depart from equilibrium approach and utilizing agent-based modeling, analyze the effects of increased power demand price sensitivity on the level and volatility of power prices. We find that as the price sensitivity increases at the demand side, power price as well as its volatility decrease significantly. We also argue that the celebrated Herfindahl-Hirschman Index to measure market concentration is not a suitable metric for power markets.

Acknowledgements

I could not have completed this dissertation, at least not sanely, without the help and encouragement of many people.

I feel immensely fortunate to have Dr. Peter Hartley as my advisor and my heartfelt thanks go to him. His support, generosity, attentiveness and patience far exceeded any reasonable limits I can think of.

I started working with Dr. Hartley following Dr. Ronald Soligo's advice. I am thankful to him for being a great friend and mentor since I came to Rice University.

I am grateful to my thesis committee members, Dr. Simon Grant, Dr. Leonardo Dueñas-Osorio and Dr. Kenneth Medlock, for their time and invaluable advice. I feel privileged to be Dr. Grant's student and assistant, and the opportunity to discuss my work with him. Leonardo was always available whenever I needed help with understanding issues related to power systems. Ken has been a source of much good advice, both for research and for my career path in general.

I am also grateful to Rice University Shell Center for Sustainability and Rice University Office of Graduate and Postdoctoral Studies for their financial support, and James A. Baker III Institute for Public Policy for providing me with a very comfortable work environment.

For someone in academia, it is not easy to find detailed information on the mechanics of market operations. Whenever I needed such information, Mr. Oğuz Orhon was a phone call away. If he was not able to answer my questions, he directed me to the right person. Among those were his colleagues Cesar Seymour, Chase Allen and Bill Jordan from GDF Suez. Alex Elsik and Tim Bowling from Keystone Energy Partners and Art Deller from ERCOT also generously shared their institutional knowledge.

I used Maple extensively in all three chapters of this dissertation. There is an amazing community of Maple users at MaplePrimes.com. I especially thank Robert Israel, acer, Joe Riel and Doug Meade for their answers to the questions I posted.

I thank Dr. Deian Tabakov for his friendship and help with LaTeX, Dr. Min Ouyang for helping me understand power flows, and Ms. Burcu Ciğerli, for her fresh pair of eyes.

I do not know how I could overcome various administrative issues without the help of our graduate program coordinator, Mrs. Altha Rodgers.

My girlfriend Özge Güleç went through this difficult period of my life with me. I cannot thank her enough.

And finally, for their unlimited patience and support, I would like to thank my parents and my two beautiful sisters, to whom this work is dedicated.

Contents

Illustrations

Tables

Chapter 1

A New Tool for Modeling the Electricity Spot Markets and its Application to the ERCOT Market

1.1 Introduction

The two competing models in the industrial organization literature for explaining how firms compete in electricity markets are the celebrated Cournot model and the supply function equilibrium model of Klemperer and Meyer (1989, [51]).[1]

A typical electricity wholesaler owns several generation facilities whose marginal costs differ depending mainly on the facilities' fuel source. As a result, rather than submitting a single price which corresponds to the unique profit maximizing quantity given a realization of its residual demand, a firm[2] may do better by submitting a schedule of price-quantity pairs taking into account various possible realizations of its residual demand. This is essentially the motivation behind the supply function equilibrium model that Klemperer and Meyer (1989) analyze. Green and Newbery (1992, [36]) and Bolle (1992, [10]) are the first authors to realize the appropriateness of Klemperer and Meyer's approach for modeling electricity spot markets.

Although some authors still prefer the Cournot model ([12], [16], [48]), mainly for its mathematical convenience, many others prefer the supply function equilibrium

---

[1]Some examples are [11], [12], [16], [42] and [48] for the Cournot model, and [3], [7], [8], [36], [37] and [46] for the supply function equilibrium model.

[2]We use the terms "electricity wholesaler", "generator", "bidder", "player" and "firm" interchangeably.

model since it better represents the everyday operations of the firms competing in electricity markets. We also take this approach. After developing our model, we use data on the firms from the Electric Reliability Council of Texas (henceforth ERCOT) area to test whether the firms are behaving optimally.

The rest of this paper is organized as follows: We will present our motivation in section 1.2, the institutional setting of the ERCOT market in section 1.3, our model in section 1.4. Section 1.4.3 analyzes the increasing marginal revenue case and section 1.5 talks about applying our model to the firms competing in the ERCOT market. Section 1.6 concludes.

1.2 Motivation

As mentioned above, supply function equilibrium models represent the operations in electricity markets more accurately than Cournot models do. However this accuracy comes at a cost. To be specific, letting the strategy space of the players consist of supply functions entails solving a set of differential equations to find the equilibrium (or equilibria) of the model. However, this is a cumbersome task unless the modeler makes some restrictive assumptions.

The first is assuming that there are two symmetric firms, as in Green and Newbery (1992, [36]). While this was realistic for the then recently deregulated British electricity market analyzed in Green and Newbery (1992), in today's markets there are many asymmetric firms. Hence, this assumption needs to be dropped.

The second is assuming that the firms possess linear supply and marginal cost functions.[3] There are very successful applications of this approach, most notably by

---

[3]In the literature, it is customary to call lines passing through the origin "linear function" and lines with non-zero intercepts "affine". We will also use this terminology.

Rudkevich (1999, 2002, 2005; [78, 79, 80]) and Baldick et al. (2000, 2004; [6, 7]). By assuming linearity, Rudkevich (1999, 2005) can find a closed form solution for the equilibrium in his stylized model with n firms. In that model, just by observing what had happened in the market in the previous period, players can attain a profit-maximizing equilibrium and convergence to this equilibrium happens very fast. Baldick et al. (2000, 2004; [6, 7]) relax the linearity assumption of Rudkevich (1999, 2005; [78, 80]) and assume that firms possess piecewise-affine supply functions as in figure 1.1. The piecewise-affine assumption allows the authors to model more than two asymmetric firms with capacity constraints. They find that their model fits the data from the British market better than the linear case.
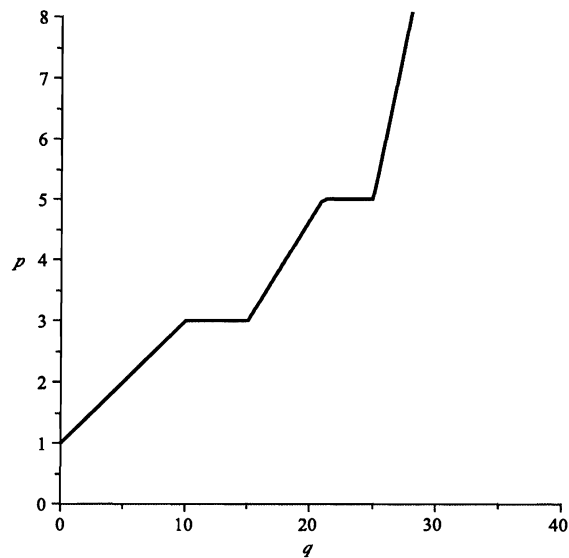
Figure 1.1 : Piecewise-affine supply function.

The approach that Baldick et al. take is not free from problems, however. First and most important, if the firms possess piecewise affine supply functions such as the one in figure 1.1, the capacity constraints will cause the aggregate supply function to

have jumps - a simple example of which is shown in figure 1.2. In the event that the market demand intersects the aggregate supply at both of its pieces, it is not clear what price-quantity pair will clear the market. Baldick et al. (2000, 2004) overcome this type of discontinuity problem by removing some parts of the aggregate supply function from both pieces, and interpolating the function at the jump. A second drawback of this approach is that we are deprived of the convenience of calculus methods. Finally, we lose the idea that the response of any one firm can be modelled as profit-maximizing. This is problematic if we believe, for example, that financial market pressures on privately owned firms would encourage profit maximization as a goal.
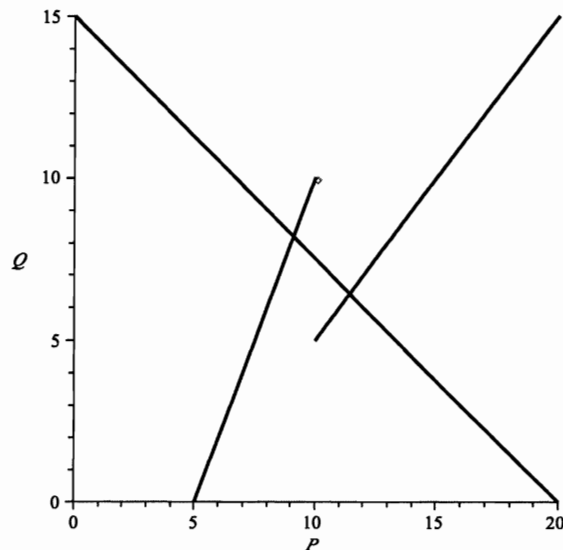
Figure 1.2 : Market clearing price-quantity pair is not well-defined.

These shortcomings motivated us to interpolate the supply and cost functions using shape-preserving cubic splines, which are continuous and differentiable. Next, we present a brief summary of how the ERCOT market operates followed by a stylized

model and the details of our approach.

1.3 Structure of the ERCOT Market

After the Texas Legislature amended the Public Utility Regulatory Act to deregulate the wholesale generation market in 1995, ERCOT became the first independent system operator in US. Its main missions are to direct the operation of the electric grid, which currently covers 75 percent of the land area in Texas and 85 percent of the state's electric load ([103]), and administer the power market.

In Texas, most wholesale electricity is traded via bilateral agreements. In order to balance supply (i.e., generation) and demand (i.e., load) in real time, ERCOT also administers a secondary (auction) market, called the Balancing Electricity Services (BES) market (or the "spot market"), where an average of 5% of the total transactions occur. On each day, market participants submit to ERCOT their production and obligation schedules for the following day, through their qualified scheduling entities (QSEs). However, actual production and consumption of electricity may vary due to unanticipated circumstances such as weather conditions, unplanned plant outages or transmission problems. As a result, companies are allowed to increase or decrease their real time production relative to the schedule they submitted one day ago. For each hour, generators offer bid schedules, composed of up to 40 price-quantity pairs, 20 monotonically increasing price-quantity pairs for augmenting (UBES) and 20 monotonically decreasing price-quantity pairs for decreasing (DBES) their production relative to their day-ahed schedule.[4] While DBES bids are mandatory, the firms are not obliged to submit UBES bids. ERCOT aggregates these bid schedules

---

[4]In ERCOT's jargon, these are called "Up Balancing Energy Services Bids - UBES" and "Down Balancing Energy Services Bids - DBES" respectively

into a single supply function for each type of service at each congestion zone.[5] Every fifteen minutes, fourteen minutes before the operating interval, ERCOT intersects the market demand with the aggregate supply to determine the "market clearing price of energy". If there is no congestion on the transmission lines, the ERCOT region becomes a single market. In case of congestion, the market is separated into zones, each having different market clearing prices. ERCOT notifies the QSEs of the market clearing price and gives them deployment instructions ten minutes before the operating interval. QSEs start to run their units accordingly, five minutes before the operating interval. For a complete description of the market's operations, see the ERCOT protocols, particularly section 6 in [23] and the very insightful article by Teng et al. (2004, [91]).

As mentioned in Hortaçsu and Puller (2005, 2008; [45, 46]), the market participants have a great deal of information on their competitors. Most plants in Texas have similar production technology[6] and their fuel efficiency data is publicly available. Also, traders seem to know which rival generators are producing at any point in time. Furthermore, it is possible to purchase real time data on the generation of large competitors.[7] Finally, every bidder has access to the aggregate bid data, which is released by ERCOT with a 2-day lag. Since each generator knows its own bids, this helps bidders infer their residual demand, assuming that the aggregate bids two days ago are similar to the ones today or are similar functions of other publicly observable data such as weather statistics.

---

[5]ERCOT organizes sources of supply and demand into "Congestion Zones" based on the likelihood of transmission constraints between these market areas. The ERCOT region is currently divided into four zones. For additional details see §7 in [23].

[6]Most plants in Texas use natural gas.

[7]This can also be inferred from the public real-time data on flows through a large number of nodes on the system.

1.4 A Stylized Model

We denote by $N = \{1, 2, ..., n\}$ the set of firms in the market. $C_i(q_i)$ is firm i's cost function and it is assumed to be quadratic, convex and twice continuously differentiable. $D_t(p)$ is the aggregate demand for electricity at time t. It is assumed to be differentiable and its slope satisfies $-\infty < D_p < 0$.

At each period t, each firm simultaneously submits a supply function $S_{it}(p)$. The residual demand faced by firm i is then $RD_i(p) = D_t(p) - \sum_{j \neq i} S_{jt}(p)$. The system administrator calculates the market clearing price, denoted as $p_t^\star$ using the market clearing condition

$$\sum_{i \in N} S_{it}(p_t^\star) = D(p_t^\star) \tag{1.1}$$

We are interested in the noncooperative equilibrium of the game

$$\Gamma = (N, (S_{it})_{i \in N}, (\pi_{it})_{i \in N}),$$

where N is the set of firms, S_{it} is player i's strategy and $\pi_{it} : \Pi_{j \in N} S_{jt} \times D_t(p) \rightarrow \mathbb{R}$ is firm i's expected profit.

Following Klemperer and Meyer (1989), we assume that the profit maximizing price-quantity pairs can be related to each other by a supply function for each firm: $q_i = S_i(p)$. That is, at any time t, each q_i corresponds to a specific price p. Then the problem of firm i becomes to maximize, with respect to p, the profit

$$\pi_{it}(p) = p \left(D_t(p) - \sum_{j \neq i} S_{jt}(p) \right) - C_i \left(D_t(p) - \sum_{j \neq i} S_{jt}(p) \right) \tag{1.2}$$

with the first order condition

$$\sum_{j\neq i} S'_{jt}(p) = \frac{S_{it}(p)}{p - C'(S_{it}(p))} + D'_t(p), \tag{1.3}$$

for all $i \in N$, where $f'(x)$ is the first derivative of $f(x)$.

Any set of nondecreasing supply functions which solve the set of equations (1.3) is an equilibrium of the game Γ.[8]

1.4.1 A Simple 2-Firm Example

The following 2-firm example will make our motivation clearer:

Let's suppose that the market demand is $D(p) = 80 - p$ and firm j observes that firm i bid[9] price-quantity pairs $\{(p, q) : (0, 0), (5, 20), (10, 35), (15, 45),$ $(20, 50), (40, 60)\}$[10] and approximates i's supply by interpolating the observed bids in a linear fashion and gets the following piecewise linear function:

---

[8]Note that this formulation does not take into account the capacity constraints. As explained in Green and Newbery (1992, [36]) §IIB, under capacity constraints, a solution intermediate between Cournot and Bertrand outcomes will not be stable and firms will find it profitable to deviate to the Cournot solution.

[9]This is without loss of generality since what matters for a given firm is its residual demand function. A firm may aggregate all of its competitors' bids and subtract this from the (expected) market demand to find its residual demand. That is to say, firm j may represent the total of firm i's competitors.

[10]As mentioned in section 1.3, ERCOT, the independent system operator, releases the aggregate bid data with a two-day lag. Hence, in our simple two-firm example, firm j can easily figure out what i bid.

$$S_i(p) = \begin{cases} 4p & \text{if } p \le 5 \\ 3p + 5 & \text{if } p \le 10 \\ 2p + 15 & \text{if } p \le 15 \\ p + 30 & \text{if } p \le 20 \\ \frac{p}{2} + 40 & \text{if } p \ge 20 \end{cases} \qquad (1.4)$$

Then, the residual demand faced by firm j will be

$$RD(p)_j = D(p) - S_i(p) = \begin{cases} 80 - 5p & \text{if } p \le 5 \\ 75 - 4p & \text{if } p \le 10 \\ 65 - 3p & \text{if } p \le 15 \\ 50 - 2p & \text{if } p \le 20 \\ 40 - \frac{3p}{2} & \text{if } p \le 26.67 \end{cases} \qquad (1.5)$$

Inverting (1.5), multiplying by q and then differentiating the whole expression with respect to q, we get the marginal revenue curve of firm j:

$$MR_j(q) = \begin{cases} \frac{80}{3} - \frac{4}{3}q & \text{if } q \le 10 \\ 25 - q & \text{if } q \ge 20 \\ \frac{65}{3} - \frac{2}{3}q & \text{if } q \ge 35 \\ \frac{75}{4} - \frac{1}{2}q & \text{if } q \ge 55 \\ 16 - \frac{2}{5}q & \text{if } q \ge 80 \end{cases} \qquad (1.6)$$

Figure 1.3 plots the residual demand and marginal revenue functions for firm j.
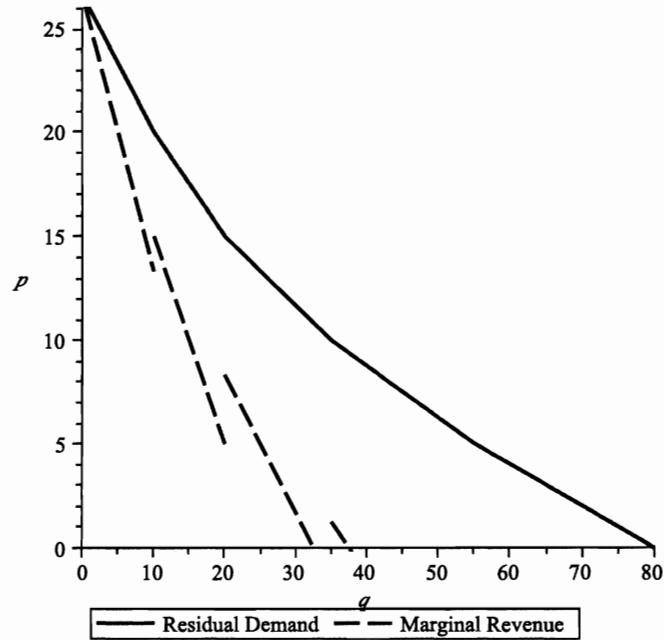
Figure 1.3 : Kinks in residual demand will lead to discontinuities in the marginal revenue.

The jumps in the marginal revenue curve are a potential problem. Instead of taking an ad-hoc approach such as "interpolating the quantities at the jump" ([6], [7]) it would be worthwhile to obtain a continuous marginal revenue curve, after the operation $D(p) - S_i(p)$. Given that the market demand is smooth, this, in turn, could be achieved if the supply function of the first firm was smooth. This shortcoming motivates us to try to fit a smooth function to a given set of (p, q) pairs. If firm j interpolated i's bid with a smooth function, its marginal revenue curve would also be smooth.

1.4.2 Approximating Competitor's Supply with a Spline

The aforementioned problem with piecewise linear approximation motivates us to approximate supply by interpolating a set of bid data with a smooth polynomial. To the best of our knowledge, no one has taken this approach before.

There is a huge literature on approximating functions and/or data with polynomial splines. One can consult [32, 33, 72, 82] and the references therein. A potential problem is that while market rules require firms to submit bids that are non-decreasing in price, polynomial splines generally do not respect this requirement. Thus, in addition to a smooth approximation, we also demand that our interpolant is monotone. Fortunately, we have algorithms at our disposal to achieve this task.

We will follow Fritsch and Carlson's (1980, [33]) method. Let $\pi = p_1 \leq p_2 \leq ... \leq p_n$ be a partition of the interval $[p_1, p_n]$. Given these n price values and the corresponding quantities, that is, n (p, q) pairs, the aim is to construct a piecewise defined cubic function $S(p)$ on π which is $\mathscr{C}^1$ and is such that $S(p_i) = q_i$. After noting that the set of all polynomials of the third degree, $\mathcal{P}_3 = \{s(p) = \sum_{i=1}^{3} c_i p^{i-1}, \quad c_i, p \in \mathbb{R}\}$, forms a 3-dimensional vector space with a finite basis (e.g., $\{1, p - a, (p - a)^2\}$), we can, on each interval $I_i = [p_i, p_{i+1}]$ represent $S(p)$ as a cubic polynomial:

$$S(p) = q_i H_1(p) + q_{i+1} H_2(p) + d_i H_3(p) + d_{i+1} H_4(p) \qquad (1.7)$$

where $d_j = S'(p_j)$, $j = i, i+1$ and the $H_k(p)$ are the Hermite basis functions for I_i, namely, $H_1(p) = \phi((p_{i+1}-p)/h_i)$, $H_2(p) = \phi((p-p_i)/h_i)$, $H_3(p) = -h_i \psi((p_{i+1}-p)/h_i)$ and $H_4(p) = h_i \psi((p - p_i)/h_i)$, where $h_i = p_{i+1} - p_i$, $\phi(x) = 3x^2 - 2x^3$ and $\psi(x) = x^3 - x^2$.

Hence, in order to interpolate $\{(p_i, q_i) : i = 1, 2, ..., n\}$ we essentially need an

algorithm which calculates the derivative values d_i such that they all have the same sign with the slope of the secant connecting two consecutive points:

$$sgn(d_i) = sgn(d_{i+1}) = sgn(\Delta_i) \qquad (1.8)$$

where $\Delta_i = (q_{i+1} - q_i)/h_i$.

The following lemmata will give the necessary and sufficient conditions for a cubic polynomial to be monotone on an interval:

Lemma 1.4.1. *Let $\alpha_i = d_i/\Delta_i$, $\beta_i = d_{i+1}/\Delta_i$ and $\alpha_i + \beta_i \leq 2$. Then $S(p)$ is monotone on I_i iff (1.8) is satisfied.*

Proof. Suppose $S(p)$ is monotone increasing on I_i. Then $\Delta_i > 0$ and for all $p \in I_i$, $S'(p) > 0$. In particular, $sgn(S'(p_i)) = sgn(d_i)$ and $sgn(S'(p_{i+1})) = sgn(d_{i+1})$ are both positive and (1.8) is satisfied. The case when $S(p)$ is monotone decreasing is similar.

Conversely, assume (1.8) holds. First, using the formula $S(p) = \sum_{i=0}^{n} \frac{S^{(i)} p_i}{i!} (p - p_i)^i$, expand $S(p)$ around p_i to get

$$S(p) = \frac{d_i + d_{i+1} - 2\Delta_i}{h_i^2}(p - p_i)^3 + \frac{-2d_i - d_{i+1} + 3\Delta_i}{h_i}(p - p_i)^2 + d_i(p - p_i) + q_i. \quad (1.9)$$

Differentiating (1.9) with respect to p gives

$$S'(p) = \frac{3(d_i + d_{i+1} - 2\Delta_i)}{h_i^2}(p - p_i)^2 + \frac{2(-2d_i - d_{i+1} + 3\Delta_i)}{h_i}(p - p_i) + d_i \qquad (1.10)$$

and

$$S''(p) = \frac{6(d_i + d_{i+1} - 2\Delta_i)}{h_i^2}(p - p_i) + \frac{2(-2d_i - d_{i+1} + 3\Delta_i)}{h_i}. \qquad (1.11)$$

Note that $d_i + d_{i+1} - 2\Delta_i = (\alpha_i + \beta_i - 2)\Delta_i$. So, if $\alpha_i + \beta_i = 2$ then $S(p)$ is quadratic or linear, hence $S'(p)$ is linear or constant. Then, (1.8) characterizes monotonicity of $S(p)$ since $min\{d_i, d_{i+1}\} \leq S'(p) \leq max\{d_i, d_{i+1}\}$. If, on the other hand, $\alpha_i + \beta_i - 2 < 0$, there are two cases to consider:

1. If $q_i < q_{i+1}$ then $d_i + d_{i+1} - 2\Delta_i < 0$, therefore $S'(p)$ is concave. This, together with (1.8), implies that $0 \leq min\{d_i, d_{i+1}\} \leq S'(p)$ hence $S(p)$ is monotone increasing.

2. If $q_i > q_{i+1}$ then $d_i + d_{i+1} - 2\Delta_i > 0$, therefore $S'(p)$ is convex. This, together with (1.8), implies that $0 \geq min\{d_i, d_{i+1}\} \geq S'(p)$ hence $S(p)$ is monotone decreasing.

$\square$

Lemma 1.4.2. *Let* $\alpha_i + \beta_i > 2$ *and (1.8) holds. Then* $S(p)$ *is monotone on* I_i *if and only if* $2\alpha_i + \beta_i \leq 3$, *or* $\alpha_i + 2\beta_i \leq 3$, *or* $\varphi(\alpha_i, \beta_i) \geq 0$, *where* $\varphi(\alpha, \beta) = \alpha - (2\alpha + \beta - 3)^2/3(\alpha + \beta - 2)$.

Proof. First, we observe from (1.11) that $S'(p)$ attains its unique extremum at

$$p^\star = p_i + \frac{h_i}{3}\left[\frac{2\alpha_i + \beta_i - 3}{\alpha_i + \beta_i - 2}\right] \tag{1.12}$$

and

$$S'(p^\star) = \varphi(\alpha_i, \beta_i)\Delta_i. \tag{1.13}$$

To prove the claim, note that conditions (1.12) and (1.13) imply that $S(p)$ is monotone on I_i if and only if either $p^\star \notin (p_i, p_{i+1})$, or, $p^\star \in (p_i, p_{i+1})$ and $sgn(S'(p^\star)) =$

$sgn(\Delta_i)$. Simple algebra shows that $p^\star < p_i$ implies $2\alpha_i + \beta_i \leq 3$, $p^\star \geq p_{i+1}$ implies $\alpha_i + 2\beta_i \leq 3$ and the condition $\varphi(\alpha_i, \beta_i) \geq 0$ is equivalent to saying $p^\star \in (p_i, p_{i+1})$ and $sgn(S'(p^\star)) = sgn(\Delta_i)$.

$\square$

These results suggest the following algorithm:

Let the data to be interpolated be (p_k, q_k) for $k = 1, 2, ..., n$.

1. Initialize the slope parameters.

   - Using three point difference formula[11], initialize d_i for $i = 2, ..., n - 1$.

   - For d_1 and d_n, let $d_1 = (q_2 - q_1)/(p_2 - p_1)$ and $d_n = (q_n - q_{(n-1)})/(p_n - p_{(n-1)})$.

2. Check whether the slope paremeters satisfy the necessary and sufficient conditions for monotonicity. If they do not satisfy those conditions, update their values accordingly:

   - Let $\Delta_k = (q_{k+1} - q_k)/(p_{k+1} - p_k)$ for $k = 1, 2, ..., n - 1$.

   - If $\Delta_k = 0$ then do the update $d_k = d_{k+1} = 0$, for $k = 1, 2, ..., n - 1$. Else

     - Set $\alpha_k = d_k/\Delta_k$ and $\beta_k = d_{k+1}/\Delta_k$.

     - If $\alpha_k^2 + \beta_k^2 > 9$ then do the updates $d_k = \tau_k d_k$ and $d_{k+1} = \tau_k d_{k+1}$ where $\tau_k = 3(\alpha_k^2 + \beta_k^2)^{-1/2}$.

We implemented this algorithm with Maple ([66, 67]). Let's turn back to our 2-firm example and see what difference this makes:

If firm j approximated i's supply with a monotone cubic spline instead, he would get

---

[11]Given three points, (p_i, q_i), $i = k - 1, k, k + 1$, the formula is $\frac{q_k - q_{k-1}}{2(p_k - p_{k-1})} + \frac{q_{k+1} - q_k}{2(p_{k+1} - p_k)}$.

$$S_i(p) = \begin{cases} \frac{1}{10}p^2 - \frac{1}{50}p^3 + 4p & \text{if } p \le 5 \\[2mm] \frac{9}{2}p - \frac{1}{10}p^2 & \text{if } p \le 10 \\[2mm] \frac{9}{2}p - \frac{1}{10}p^2 & \text{if } p \le 15 \\[2mm] -45 + \frac{51}{4}p - \frac{3}{5}p^2 + \frac{1}{100}p^3 & \text{if } p \le 20 \\[2mm] 20 + \frac{5}{2}p - \frac{1}{16}p^2 + \frac{1}{1600}p^3 & \text{if } p \le 40 \end{cases} \qquad (1.14)$$

By evaluating the derivative of (1.14) at the critical points one can easily see that it is smooth. Figure 1.4 plots the linear and the spline interpolants to the given (p, q) pairs.
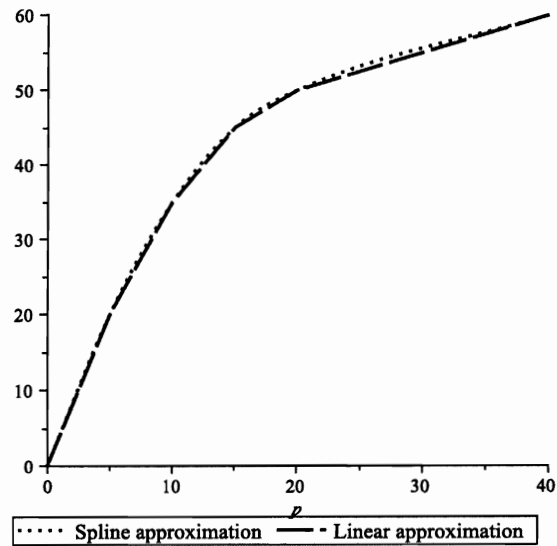
Figure 1.4 : Linear vs. Spline Inperpolation

This will give the following expression for the residual demand, which is necessarily smooth:

$$RD_j(q) = \begin{cases} 80 - 5p - \frac{1}{10}p^2 + \frac{1}{50}p^3 & \text{if } p \leq 5 \\[2ex] 80 - \frac{11}{2}p + \frac{1}{10}p^2 & \text{if } p \leq 10 \\[2ex] 80 - \frac{11}{2}p + \frac{1}{10}p^2 & \text{if } p \leq 15 \\[2ex] 125 - \frac{55}{4}p + \frac{3}{5}p^2 - \frac{1}{100}p^3 & \text{if } p \leq 20 \\[2ex] 60 - \frac{7}{2}p + \frac{1}{16}p^2 - \frac{1}{1600}p^3 & \text{if } p \leq 26.175 \end{cases} \tag{1.15}$$

Again, inverting $(1.15)^{12}$, multiplying by q and then differentiating the whole expression with respect to quantity, we get the marginal revenue curve of firm j. We suppress the extremely complicated expressions and just plot the result in figure 1.5:
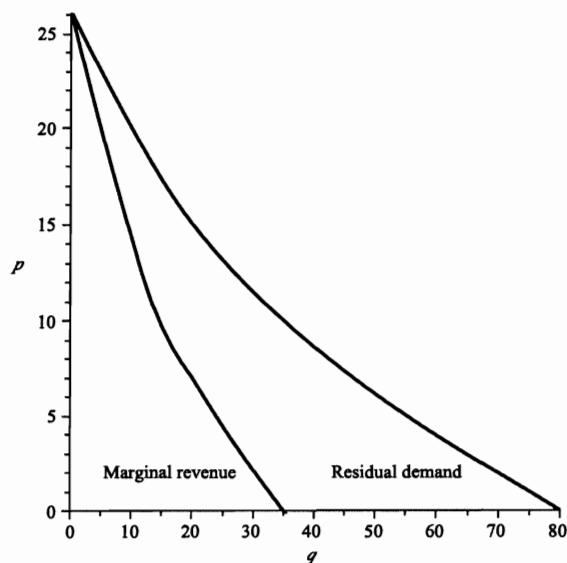
Figure 1.5 : Monotone spline interpolation results in differentiable residual demand and continuous marginal revenue functions.

---

[12]Not a trivial task. See appendix B for an alternative approach.

1.4.3 Increasing Marginal Revenue and the Possibility of Multiple Profit Maximizing Optima

One problem that has been ignored in modeling electricity spot markets is that the marginal revenue function of a generator can increase on some intervals. Some researchers previously noted this anomaly for a monopolist and analyzed the conditions under which we may expect a monopolist to have an increasing marginal revenue curve ([96], [30], [22]). Walters (1980, [96]) considers a monopolist facing a demand curve that yields such a marginal revenue curve and he argues that "[t]here is much evidence to show that such demand curves are characteristic of utility and service industries. For example, it is well known that an electric utility company, usually a statutory monopoly, has a very inelastic demand for domestic electricity for lighting purposes; but at lower prices the utility can break into the vast market for heating, air-conditioning, and industrial power. When not prohibited by law, and where technically feasible, electric utilities will price-discriminate in these markets."[13] Formby et al. (1982, [30]) "demonstrate that the conditions for a positively sloping marginal revenue curve are much less stringent than is generally recognized" and add that "positively sloping marginal revenue functions must be considered whenever convex demand functions are analyzed". They give examples of several functional forms frequently used in economic analysis which lead to upward sloping marginal revenue curves. Finally, Coughlin (1984, [22]) derives statements regarding the elasticities of a monopolist's demand and marginal revenue functions that are equivalent to the direction of change for the marginal revenue function.

---

[13]The common practice of using a different price schedule for different types of customers can most readily be rationalized as price discrimination. In particular, commercial customers typically purchase electricity mainly for lighting and other essentials and thus have relatively less elastic demand. They also typically face the highest charges.

None of these studies, however, consider an oligopolistic market structure and none of them suggest what to do in the case of multiple equilibria. We will show that, for electricity generating firms, ending up with a residual demand for which the marginal revenue is positively sloped is the rule, rather than the exception. As a result, multiple profit maximizing price-quantity pairs for a firm is not a very remote possibility. Furthermore, the set of ex-post profit maximizing points will not necessarily lie on a monotone path, a violation of ERCOT market rules. A reasonable hypothesis about what a firm is likely to do when faced with multiple profit maximizing optima, however, is still an open question.

First, we state some useful results.

Definition 1.4.3. *A real valued function f defined on $A \subset \mathbb{R}$ is convex on A if for all $x_1, x_2 \in A$ and $\alpha \in [0,1]$ we have $f\big(\alpha x_1 + (1-\alpha)x_2\big) \leq \alpha f(x_1) + (1-\alpha)f(x_2)$.*

Lemma 1.4.4. *f is convex if and only if for all x_1, x_2 and x_3 in A, such that $x_1 < x_2 < x_3$ we have*

$$\frac{f(x_2) - f(x_1)}{x_2 - x_1} \leq \frac{f(x_3) - f(x_2)}{x_3 - x_2} \tag{1.16}$$

Proof. Suppose f is convex and let x_2 be a convex combination of x_1 and x_3: $x_2 := \alpha x_1 + (1-\alpha)x_3$ for some $\alpha \in [0,1]$. Then we have:

$$\begin{aligned}
\frac{f(x_2) - f(x_1)}{x_2 - x_1} &\leq \frac{\alpha f(x_1) + (1-\alpha)f(x_3) - f(x_1)}{x_2 - x_1} \\
&= \frac{f(x_3) - f(x_1)}{x_3 - x_1} = \frac{f(x_3) - \big(\alpha f(x_1) + (1-\alpha)f(x_3)\big)}{x_3 - x_2} \\
&\leq \frac{f(x_3) - f(x_2)}{x_3 - x_2}
\end{aligned}$$

The inequalities in the first and third lines follow from the definition of convexity. In order to visualize the equalities, the figure below will be helpful.
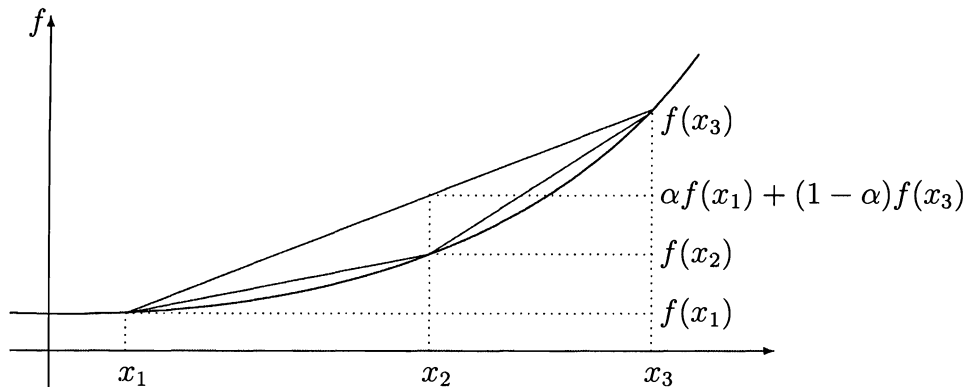
Conversely, assume that for all x_1, x_2 and x_3 (1.16) holds. Take any $\alpha \in (0,1)$ and without loss of generality let $x_2 = \alpha x_1 + (1 - \alpha)x_3$. Then,

$$
\begin{aligned}
\frac{f\big(\alpha x_1 + (1 - \alpha)x_3\big) - f(x_1)}{\alpha x_1 + (1 - \alpha)x_3 - x_1} &= \frac{f\big(\alpha x_1 + (1 - \alpha)x_3\big) - f(x_1)}{(1 - \alpha)(x_3 - x_1)} \\
&\leq \frac{f(x_3) - f\big(\alpha x_1 + (1 - \alpha)x_3\big)}{x_3 - \alpha x_1 - (1 - \alpha)x_3} \\
&= \frac{f(x_3) - f\big(\alpha x_1 + (1 - \alpha)x_3\big)}{\alpha(x_3 - x_1)}
\end{aligned}
$$

Canceling $(x_3 - x_1)$ gives

$$
\alpha\big(f\big(\alpha x_1 + (1 - \alpha)x_3\big) - f(x_1)\big) \leq (1 - \alpha)\big(f(x_3) - f(\alpha x_1 + (1 - \alpha)x_3)\big)
$$

upon rearranging we get

$$
f\big(\alpha x_1 + (1 - \alpha)x_3\big) \leq \alpha f(x_1) + (1 - \alpha)f(x_3).
$$

$\square$

Lemma 1.4.5. *Let $f : A \to B \subset \mathbb{R}$ be an invertible function and let f^{-1} be its inverse. If f is decreasing and convex, so is f^{-1}.*

Proof. Take any three points x_1, x_2 and x_3 from A and suppose, wlog, that $x_1 < x_2 < x_3$. Hence $y_3 := f(x_3) < y_2 := f(x_2) < y_1 := f(x_1)$ so f^{-1} is decreasing and by lemma 1.4.4

$$\frac{f(x_2) - f(x_1)}{x_2 - x_1} \leq \frac{f(x_3) - f(x_2)}{x_3 - x_2} < 0 \implies \frac{f(x_1) - f(x_2)}{x_2 - x_1} \geq \frac{f(x_2) - f(x_3)}{x_3 - x_2} > 0$$

$$\implies \frac{x_2 - x_1}{f(x_1) - f(x_2)} = \frac{f^{-1}(y_2) - f^{-1}(y_1)}{y_1 - y_2}$$

$$\leq \frac{x_3 - x_2}{f(x_2) - f(x_3)} = \frac{f^{-1}(y_3) - f^{-1}(y_2)}{y_2 - y_3}$$

$$\implies \frac{f^{-1}(y_1) - f^{-1}(y_2)}{y_1 - y_2} \geq \frac{f^{-1}(y_2) - f^{-1}(y_3)}{y_2 - y_3}$$

hence f^{-1} is convex. $\qquad\square$

Lemma 1.4.6. *Let $f : (a, b) \to \mathbb{R}$ be a $\mathscr{C}^2$ function. Then f is convex if and only if its second derivative is nonnegative on (a, b).*

Proof. See [75], page 26. $\qquad\square$

Now, suppose for simplicity there are only two firms and the market demand, $D(p)$, is perfectly inelastic.[14] Suppose firm 1's competitor bids a supply function

---

[14]Some large firms sign up for time-of-use metering of electricity and respond to spot prices. In particular, many co-generating firms can alter the amount of power they supply to, or take from, the grid and will do so in response to prices. As this last example illustrates, conceptually we can think of price-responsive demand as an alternative source of supply and take the non-responsive demand as the "demand" represented in the model. This view of price-responsive demand as a type of "supply resource" has even been formalized in many electricity markets. For example, ERCOT has a program called "Load acting as resource" (LaaR), in which a company with minimum load of 1 MWh can sign up to allow a portion or all of its energy demand to be responsive to market prices. In exchange for getting paid a fixed annual fee, the company has to curtail its energy demand immediately if asked by the system operator. It can be argued that existence of LaaRs provides some (imperfect) elasticity in real time, but they provide up to the 50% cap of ERCOT's Responsive Reserves procurement of 2300 MWs, which corresponds to only 1.8% of ERCOT peak demand, observed on August 16th, 2010. Between April 2006 and August 2010, LaaRs had been deployed only 13 times. See [24] for details.

$S_2(p) \in \mathscr{C}^2$ which is concave in p. Concavity is a realistic assumption here, given the capacity constraints. Then the residual demand function firm 1 faces will be $(D - S_2)(p)$. Since S_2 is concave, $-S_2$ and hence $D - S_2$ is convex. Consequently by lemma 1.4.5, the residual demand firm 1 faces, as a function of q, will be convex as well. Let $RD : [0, \bar{q}] \rightarrow \mathbb{R}_+$ denote that residual demand function. Then the expression for marginal revenue will be

$$MR(q) = \frac{d}{dq}\big(qRD(q)\big) = RD(q) + qRD'(q) \tag{1.17}$$

Now, consider the slope of the function defined by (1.17):

$$\frac{d}{dq}MR(q) = 2RD'(q) + qRD''(q). \tag{1.18}$$

The first term on the right hand side of equation (1.18) is negative since the residual demand function is negatively sloped. But if we do not impose any restrictions on the shape of the residual demand, namely, that it is globally concave, the second term on the right hand side will be positive on the intervals where $RD(q)$ is convex (by lemma 1.4.6) and on those intervals, the second term may dominate the first. Then (1.18) will be positive and marginal revenue will be increasing on that range, i.e., $2RD'(q) + qRD''(q) > 0$. Following Coughlin's analysis (1984, [22]), by rearranging this last equation we get:

$$2RD'(q) + qRD''(q) > 0$$

$$\iff -2 > qRD''(q)/RD'(q)$$

$$\iff -2 > \frac{dRD'(q)}{dq} \bigg/ \frac{RD'(q)}{q} = \frac{dRD'(q)}{RD'(q)} \bigg/ \frac{dq}{q}$$

$$\iff -2 > \text{quantity elasticity of slope of residual demand.}$$

$$(1.19)$$

So we can state:

Proposition 1.4.7. *Marginal revenue will be increasing if quantity elasticity of the residual demand is less than -2.*

Increasing marginal revenue creates the possibility of multiple profit maximizing quantities and prices for the firm. Multiplicity of optima, in turn, will create uncertainty as to what price-quantity pairs the firm will bid to the market. This is problematic not only because it creates or enhances circumstances under which the firm may manipulate the bidding process. It also creates uncertainty from the point of view of the system administrator who is responsible to keep the system balanced at all times.[15] How much quantity is dispatched from each generator is crucial since electricity, to a large extent, is not storable.

Consider the following figure where we used hypothetical bid data:[16]

---

[15]And this is a serious problem because the most distinctive feature of the electricity markets is that supply and demand should be in balance at any second. Supply-demand imbalance will lead to blackouts and/or deviations from the target system frequency. Lineweber and McNulty (2001, [57]) estimate that the US economy across all business sectors is losing between \$119 billion and \$188 billion annually due to power outages and power quality issues.

[16]Even though we used hypothetical bid data, the shape of the resulting residual demand function closely resembles that of figures 4.2 and 4.3 in Wolak (2003, [100]), where the author used real data from the Australian market.
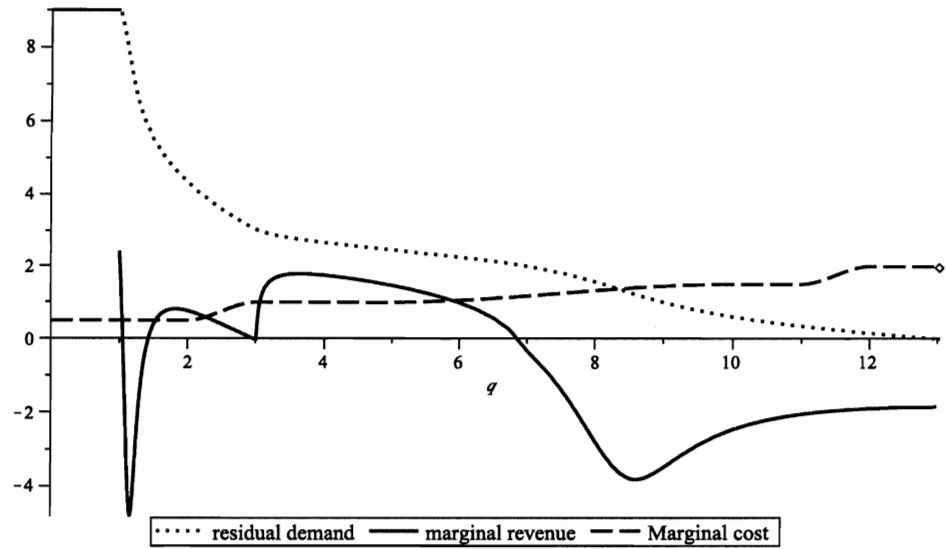
Figure 1.6 : Multiple profit-maximizing quantities.

Both $q = 1.04$ and $q = 5.91$ are profit-maximizing quantities, and at both of these quantities, marginal revenue and marginal cost curves intersect, and at both quantities, the profit of the firm is 8.402. The price corresponding to $q = 1.04$ is 8.73 and the price corresponding to $q = 5.91$ is 2.27. Given this realization of the residual demand, it is not clear which quantity the firm will produce. Even though one may think that the firm itself is indifferent, the same cannot be said for the system operator or for consumers. One profit-maximizing output level is almost 6 times the other. In addition, to maintain system balance, this uncertainty has to be resolved. Finally, from a market design perspective, the difference between the two profit maximizing prices gives this firm the opportunity to manipulate the bidding process.

1.5   Application to ERCOT Market

In this section, we use data on the firms' marginal costs and actual bids[17] to construct their ex-post profit maximizing supply functions. Figure 1.7 illustrates our method.[18] We start with a given realization of the residual demand function and calculate the corresponding marginal revenue. Intersection of marginal revenue with the marginal cost (the black, dashed curve) will give the profit maximizing quantity for that specific realization of the market demand. We shift market demand to get separate residual demand and marginal revenue functions and obtain other profit maximizing points. Connecting these gives us the ex-post optimal supply function for the firm (the dot-dashed blue curve).

Note that all four realizations of the residual demand and the corresponding marginal revenue curves intersect at a point to the right of the origin. This quantity is the amount of energy the firm sold in bilateral contracts. To the right of this point, the firm is a net seller in the BES market and hence acts as a monopolist on its residual demand curve. To the left of that point, the firm is a net buyer of electricity and acts as a monopsonist. While the forward contract position is private information of the generators, there is an intuitive method to estimate it using a firm's marginal cost function and its BES bids: If the firm is short in the market, it needs to buy power to cover its contractual obligations. Hence, up to this contract quantity, the firm will want to bid less than its marginal cost function and try to decrease the market price

---

[17]In the ERCOT region, each company is represented by a qualified scheduling entity (QSE) and ERCOT, prior to January 31, 2011, released QSE-specific bid data for the past two years, sixty days after the operating day. The data were available at ERCOT.com for the zonal market. ERCOT implemented a new market design and switched to the "nodal system" on December 1, 2010, for which there is no publicly available bid data.

[18]This is identical to Hortaçsu and Puller's (2005, 2008) "empirical strategy".
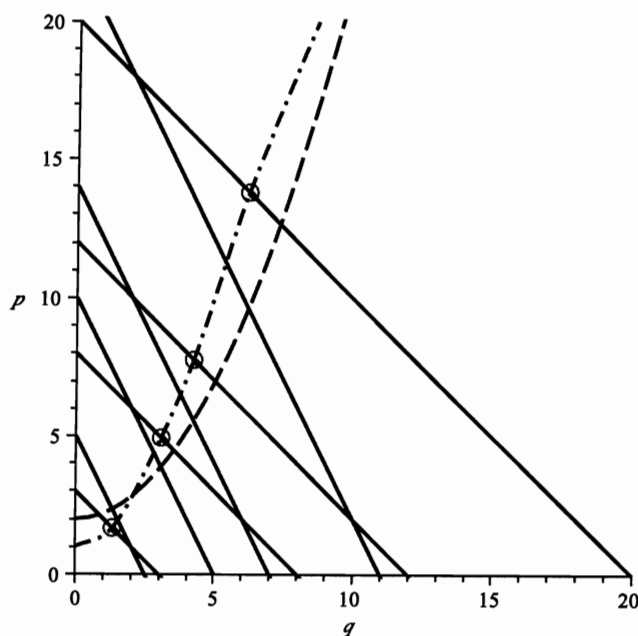
Figure 1.7 : Construction of the ex-post optimal supply function. The blue dots are the optimum quantity-price pairs.

and to the right of it the firm will want to bid more than its marginal cost[19].

1.5.1 Comparison of Actual BES Bids to Ex-post Optimal Bids

We apply our method for analyzing five firms; Austin Energy, Brazos Electric Power Cooperative, Calpine, Luminant and NRG from March 2009 through August 2009. Figures 1.8, 1.9, 1.10, 1.11 and 1.12 are from March 18, 2009 between 8 and 9 PM, a time period during which there was no zonal congestion and the whole ERCOT region had a single market clearing price.

As mentioned in section 1.3, the DBES bids are monotonically decreasing in price.

---

[19]Hortaçsu and Puller (2005, 2008; [45, 46]) say that "[t]his practice was acknowledged by all of the firms that they interviewed during their research." Our conversations with market participants also verified the validity of this approach.

In order to be able to present both the up and the down bids on the same figure, we first normalize the down bids by taking their mirror image with respect to the price axis. Then we analyze the DBES and UBES cases separately since the aggregate down and the aggregate up bids usually overlap and the period we chose was not an exception. After we calculate the ex-post optimum bids separately for the two cases, we combine the results for each firm to come up with its ex-post optimal bid stack. Occasionally, we had some optimum bid points lying outside the monotone path combining the others. Since such points violate the market rules, we eliminate those.

We start from a balancing market demand of -5000 MWs and go up to 4000 MWs by adding 100 MW increments.

Shioshansi and Oren (2007, [84]) and Hortaçsu and Puller (2008, [46]) analyzed the market in 2002. They found that the incremental bids of TXU, the biggest wholesaler in Texas in terms of installed capacity, were very close to its ex-post optimal supply function. However, TXU's decremental bids were much lower than ex-post optimum. We examined the recent behavior of TXU, now operating under the name Luminant. We found both their decremental and incremental bids to be fairly close to their ex-post optimum supply function. However, especially on the incremental side, they use a very coarse bidding strategy. We have difficulty explaining this because the high and low operating limits, high and low sustainable limits and ramp rates of their power plants seem to allow for a more refined bidding strategy. This is particularly the case between 0 and 700 MWs for the period shown in figure 1.8.

According to Hortaçsu and Puller's (2008) analysis, the second biggest producer Reliant, which now operates under the name NRG, was bidding remarkably close to its ex-post optimum bids. This result was supported by Shioshansi and Oren (2007).
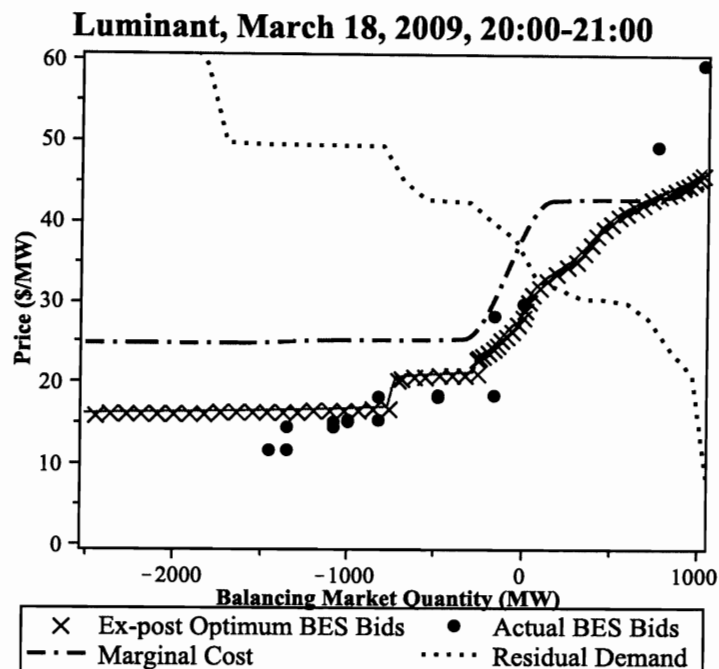
Figure 1.8 : Analysis for Luminant. Estimated contracted quantity is 700 MWs .

During the sample period we analyzed however, this is not the case (figure 1.9). NRG bids significantly below what would be ex-post optimum and most of the time does not submit any incremental bids. Some small wholesalers usually avoid the balancing market. It has been suggested that this may be because they do not want to modify their production schedules.[20] Hortaçsu and Puller (2008, [46]) suggest that the high cost of setting up and operating a real time bidding desk is another reason for smaller wholesalers' avoiding incremental bids. However, we do not think either reason holds for NRG and find their lack of incremental bids rather surprising.

Calpine is one of the biggest electricity producers in the ERCOT region and they own the most efficient generation fleet. It was also one of the firms Hortaçsu and

---

[20]Private conversations with real time power traders. Also see [45] and [46].

NRG, March 18, 2009, 20:00-21:00
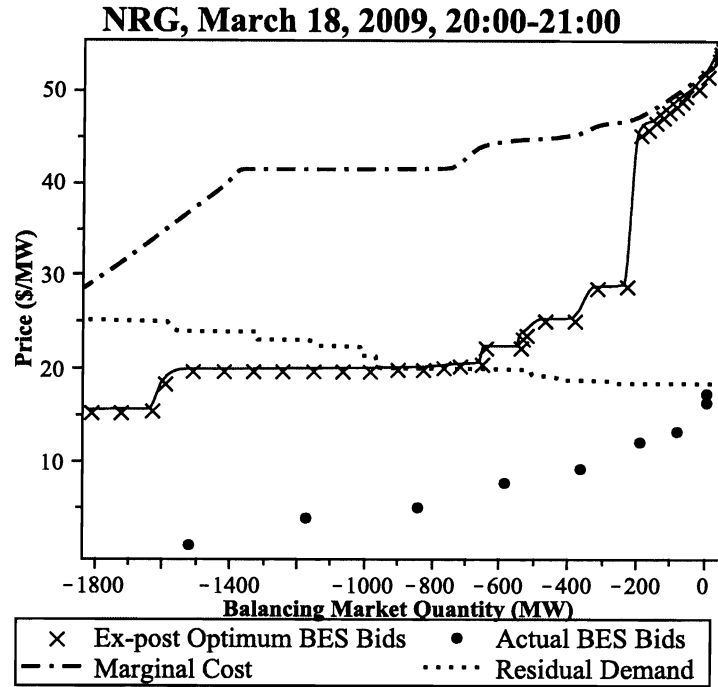
Figure 1.9 : Analysis for NRG. Estimated contracted quantity is 0 MWs.

Puller (2005, 2008) analyzed for the year 2002. Parallel to their results, we found that, on the decremental side, Calpine's bids are far from being optimal, although, we can say that they bid very close to optimum on the incremental side. A general tendency of Calpine is to use too few bid points (see figure 1.10).

Brazos Electric Power Cooperative (figure 1.11) is one of the smaller wholesalers. They are using a more refined bidding strategy compared to their bigger rivals and seem to bid close to their ex-post optimum supply function between 0 and 100 MWs, but outside that region, their bids are far from being optimal.

The final wholesaler we analyze is Austin Energy, another small producer (figure 1.12). Like Brazos Electric Power Cooperative, they also use a more refined bidding

Calpine, March 18, 2009, 20:00-21:00
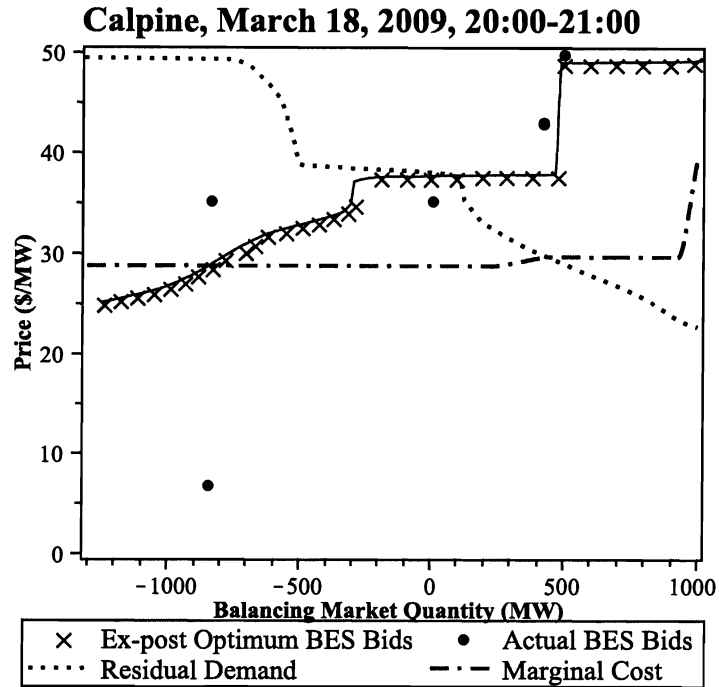
Figure 1.10 : Analysis for Calpine. Estimated contracted quantity is -844 MWs.

strategy with the incremental bids closer to their optima compared to the decremental bids.

Before concluding this section, it is important to note that we had to make many simplifying assumptions to arrive at these results. First of all, while we did our best to estimate the cost functions as precisely as possible, the real marginal costs may depart from our estimates since we assumed that ramprate constraints, upper and lower operating limits and upper and lower sustainable limits of the generators are not binding. In addition, we do not explicitly take into account start-up and shut-down costs, which may be significant.[21] Also, the heatrates we use for calculating the

---

[21]Another reason for choosing this specific time period is that the plants running between 20:00 to 21:00 were also running before 20:00 and after 21:00.

**Brazos Electric Power Cooperative,
March 18, 2009, 20:00-21:00**
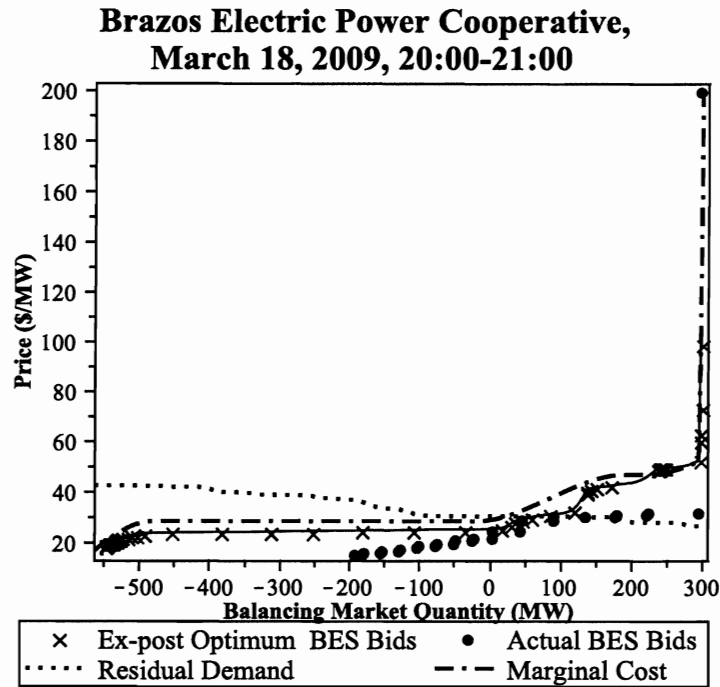
Figure 1.11 : Analysis for Brazos Electric Power Cooperative. Estimated contracted quantity is 292 MWs.

marginal cost functions are average figures. In particular, one implication of this last assumptions is that the marginal cost of a given plant will be a constant function of quantity, which is not the case in reality. An excellent report by Hirst (2001, [41]) explains all these concepts and their implications in detail.

1.6 Conclusions and Future Research

We presented an alternative tool for modeling the bidding process in electricity spot markets. We approximate bidders' supply functions by interpolating their bid data as well as their marginal cost functions with piecewise cubic splines. This method gives us differentiable residual demand, and thus, continuous marginal revenue curves.

Austin Energy, March 18, 2009, 20:00-21:00

X	Ex-post Optimum BES Bids	●	Actual BES Bids
·····	Residual Demand	— · —	Marginal Cost
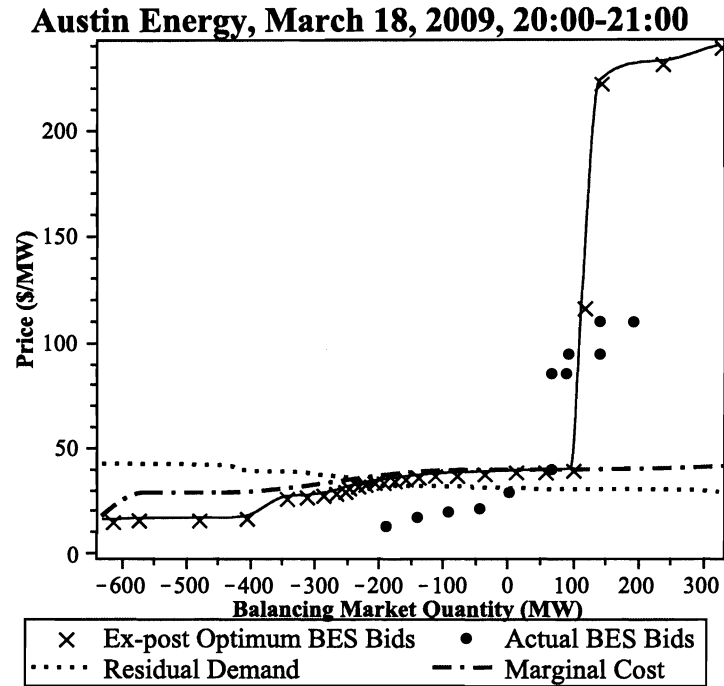
Figure 1.12 : Analysis for Austin Energy. Estimated contracted quantity is 35 MWs.

Furthermore, as mentioned in Teng et al. (2004, [91]), rather than working with step functions, ERCOT divides two adjacent bids into up to 100 segments (see the next figure) in order to calculate the market clearing price. Our method is in essence dividing two adjacent bids into infinitely many segments.
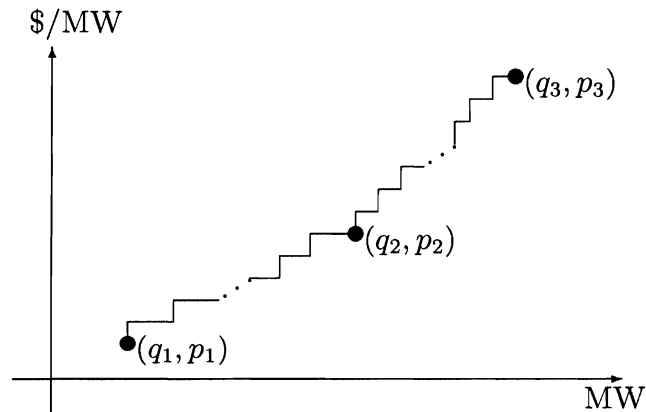
Also, having such a flexible tool to calculate the optimal bid curves efficiently facilitates analyzing many more time periods and circumstances, and allow us to investigate a wider range of hypotheses about firm behavior and market outcomes.

However, these advantages come at a cost; namely, now each supply function has up to $3n$ parameters, where n is the number of polynomial pieces in a given supply function. This makes it hard to find a closed form solution for the equilibrium conditions. It is also hard now to analyze whether and how learning in this environment will take place (see [78] and [80] for the linear supply function equilibrium case).

We also showed that, in electricity markets where the competitors' residual demand functions will necessarily have convex sections, their marginal revenue curves will have increasing segments. This anomaly may lead to multiple profit maximizing optima for some players, which, in turn, causes additional uncertainty about the amount that player will want to supply to the balancing market. We are currently working on further implications of this anomaly and how to deal with the uncertainty it creates.

We applied our technique to firms competing in the Texas wholesale market and found that while the firms with the biggest stakes in the market generally bid closer to what economic theory predicts,[22] they do not fully take advantage of their market power. Smaller firms still have some chance to bid strategically, even though their residual demand is much more elastic. However, they do not appear to fully exploit this opportunity.

---

[22]Another explanation for Luminant's and Reliant's success may be that these firms were the incumbent utilities prior to deregulation. Arguably, they are familiar with the whole system, especially the transmission grid, more than any of their competitors. This certainly is a comparative advantage for these two firms. We thank Martin Lin for pointing this out.

For our analysis, we confined ourselves to those hours during which there are no transmission constraints. However, the existence of transmission constraints, or expectation of them, can alter the bidding behavior of the firms. In future research, we will also incorporate these constraints. Another question we will investigate as an extension to our analysis is whether allowing the firms to bid smooth functions rather than step functions would alter their bidding strategies significantly.

Chapter 2

Electricity Forward Markets and Competition in Supply Functions: The Case of Risk-Averse Agents

2.1 Introduction

The interaction between forward/futures contracts and the spot price of the under-lying commodity has attracted the attention of researchers from several fields such as economics, finance, industrial organization and power systems engineering. In their seminal work, Allaz and Vila (1993, [2]) show that existence of future trad-ing in an oligopolistic market, where the oligopolists act non-cooperatively and have Cournot conjectures, promotes competition and as the frequency of trading increases, we achieve the competitive outcome.[1] Contrary to this encouraging result, Ferreira (2003, [28]) provides a counter-argument to the previous literature on the positive effects of trading in forward/futures markets on competition, where he shows that the introduction of futures market may have an anti-competitive effect. Mahenc and Salanié (2004, [64]) find that under Bertrand competition, equilibrium prices are lower if there is no forward contracting. Liski and Montero (2006, [58]) show that if the game is played infinitely, it does not matter if the oligopolists compete in price or quantity; the possibility of forward trading allows firms to sustain collusive profits

---

[1]Even though it is not stated by Allaz and Vila (1993, [2]) and other researchers building on their model (see, for example, [34] and [15]), it turns out that symmetry of cost functions and number of players are critical assumptions for this result to hold. In [86], Su provides a three-player counterexample to the Allaz-Vila result and the existence results for the Allaz-Vila model under asymmetric cost functions.

that would not be possible without the existence of forward markets. Green and Coq (2010, [35]) focus on the length of contracts and find out that the pro-competitive effect of forward contracting is not guaranteed and by selling "the right amount of contracts", firms can sustain prices above marginal cost.[2] Hughes and Kao (1997 [47]) reconsider the validity of another assumption in the Allaz-Vila model, namely that the contracts are publicly observable.[3] They find that if the contracts are not observable, producers prefer not to participate in the forward market and the Allaz-Vila result fails to hold. Ferreira (2006, [29]) also investigates the effects of observability on market efficiency but finds that imperfect observability may induce even more competitive outcomes that those of Allaz-Vila model.[4]

Endowed with this theoretical background, many researchers applied these ideas to wholesale power markets. Powell (1993, [73]) is the first to model the producers and the retailers separately. The risk-neutral producers engage in Cournot competition and set the quantity in the spot market but the forward price in the contract market. The retailers are risk-averse and are endowed with mean-variance utility functions. Under the assumption that the generators do not collude, the competitive outcome is attained. Green (1999 [34]) analyzes a symmetric duopoly, competing in supply functions in the spot market, and finds that competition in the contract market leads to lower spot market prices. Chung et al. (2004, [18]) have a very similar

---

[2]In a remotely related model, Aghion and Bolton (1987, [1]) show that an incumbent firm can block new entry by signing forward contracts. Newbery (1998, [68]) also investigates the relationship between forward contracts and entry, and relates these ideas to the competition in the British wholesale power market. He finds that "if entry remains contestable and the contract market is reasonably liquid and active ... then the inefficiencies of market power caused by too few generators are much reduced."

[3]In power markets these data are strictly confidential.

[4]In the aforementioned work by Ferreira (2003, [28]), contracts are observable. In this literature, researchers almost unanimously agree on the policy recommendation that authorities should make contracts observable. On the other hand, see an extensive survey by Madhavan (2000, [63]) where he points out that more transparency does not necessarily mean increased efficiency.

setup to Green's (1999 [34]) but they extend Green's model to multiple asymmetric firms. Their results are also similar to Green's. Wolak (2000, [99]) develops a bidding model with forward contracts, and applies his model to the National Electricity Market in Australia, for which he has confidential forward contract data. He finds that forward contracts are effective in mitigating market power of generators. Kamat and Oren (2004, [49]), Yao et al. (2007, [106] and 2008, [107]) use a Cournot model over an electricity network and hence can incorporate transmission constraints into their model. These papers formulate the model as an equilibrium problem with equilibrium constraints.[5] The players are risk-neutral. They find that existence of transmission constraints, or even a small probability of congestion will result in substantially reduced forward contracting but agree with the result that forward contracting mitigates market power. Bushnell (2007, [15]) also uses a Cournot model but extends it to multiple players. Given the assumptions, his corollary 3 states that "the impact of one round of forward contracting on the Lerner index is equivalent to an increase in firms to a number equal to the square of the number of firms in the market". Wang et al. (2008, [97]) compare the effect of financial options contracts under both Cournot competition and competition in linear supply functions and find that the mitigation effect of contracting is higher under linear supply function competition. Niu et al. (2005, [70]) develop players' optimal spot market bids given their forward positions, and test their model with data from ERCOT. Their model is useful in the sense that it can be used to analyze the effects of different forward contract levels on real-time

---

[5]In these models, each generator solves a mathematical problem with equilibrium constraints (MPEC). In their manuscript, Luo et. al (2006, [62]) describe an MPEC as "an optimization problem in which the essential constraints are defined by a parametric variational inequality or complementarity system." In the studies cited above, the parameter is the competitors' forward positions. These models look very promising and in [106] and [107], the authors develop effective algorithms for solving relatively large systems.

market prices. They agree with the conventional view that forward contracts mitigate market power. In a very recent paper, Holmberg (2011, [43]) models an asymmetric duopoly competing in supply functions, and he also incorporates price caps. His model also implies that forward contracting has pro-competitive effects.

Coq and Orzen (2006, [21]) test the Allaz-Vila model in an experimental setting. Their results support the view that forward contracting is an effective means of mitigating market power, but suggest that entry is more effective than introducing forward markets. Brandts et al. (2008, [14]), in addition to analyzing power markets modeled with Cournot players, extend these experiments to a supply function competition setting with quadratic marginal cost functions. In their setting with no demand uncertainty, they find that under both settings efficiency is improved if forward markets are introduced. They agree with Coq and Orzen's (2006, [21]) results in that availability of forward contracts leads to efficiency gains, but find that the addition of another producer increases production more than does forward contracting, and efficiency gains from forward contracting are not guaranteed.

A negative result comes from Sánchez et al. (2009, [81]). They develop an agent-based simulation model of the Spanish power market. The results of their simulations, where the agents are risk-neutral, suggest that if forward contracting is voluntary, only the small players will want to do so, while the dominant players will exercise market power in the spot market rather than trade forward.

In all of the aforementioned models, which imply that forward contracting increases efficiency, the intuition is that locking in prices by signing forward contracts decreases market power and spot prices by shrinking the size of the spot market that (dominant) firms can manipulate. But then a natural question arises which is never asked in these models (Harvey and Hogan (2000, [38]) are the only ones to explicitly

raise this question): Why would the dominant firms voluntarily engage in forward contracting, if it decreases their potential profits? In the Allaz-Vila type settings, clearly producers do not care about competition; they are trying to achieve the first-mover advantage by selling forward, but since everyone does the same, a prisoners' dilemma type result occurs. Others implicitly answer the question by appealing to the "strategic incentive" of the players:[6] By increasing its forward sales, a producer lowers the forward price, and hence its competitors' forward sales. Left with this unsold capacity, the competitors have to bid more aggressively in the spot market. Again, since each player will want to behave the same way, they collectively end up worse-off.

Leaving Powell's work (1993) aside, none of these papers models the buyers in the contract market. Another common attribute in these models is that the players are risk-neutral. In our opinion, the risk-neutrality assumption should be dropped. Power generation technologies are extremely capital intensive investments. In a report prepared for members and committees of the US Congress, Kaplan (2008, [50]) estimates that power plants cost above $2,100 per kilowatt, with the exception of combined cycle natural gas plants ($1,200 per kilowatt). In particular, a kilowatt capacity of a nuclear plant costs well above $3,000. In light of these data, when modeling wholesale power markets it is only natural to think that generators are risk-averse agents, who want to recover their investment costs without too much exposure to volatile power prices. While generators' risk preferences may change over time, that is, as they recover their fixed costs, our conversations with professionals from the industry reinforces our intuition that even then, generators are risk-averse players. The retailers, on the other hand, operate on very narrow margins, hence, arguably, they are even

---

[6]See, for example, [34], [18] and [43].

ERCOT real time prices, 02/02/2011
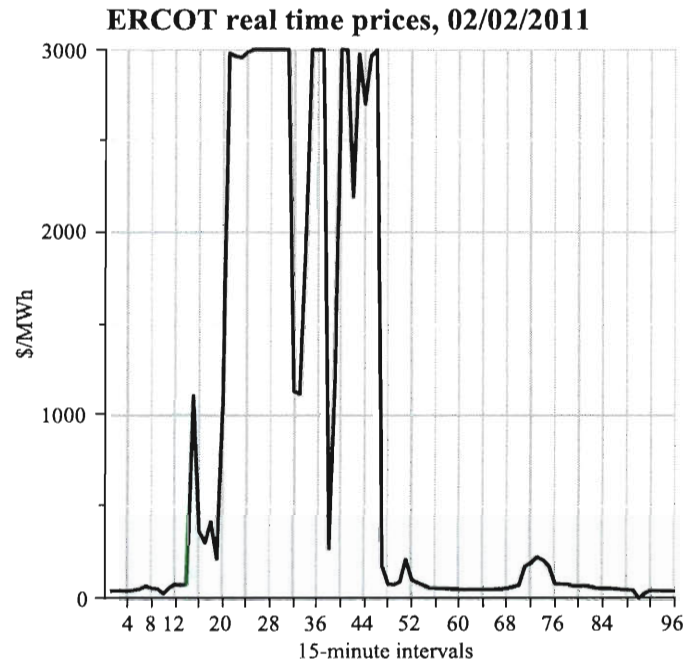
Figure 2.1 : ERCOT power prices under extreme weather conditions. Mean power price: $783.22, range: $3,004.74, standard deviation: $1188.96. *Source: www.ercot.com*

more risk-averse than the generators. They try to mitigate the effects of power price volatility, and as can ben seen in figure 2.1, this volatility may be extreme, especially under severe weather conditions.[7] Bessembinder and Lemmon (2002, [9]) model both the generators and retailers as risk-averse agents. While they assume the generators choose quantities (rather than a supply function) in the spot market, their model is more general in the sense that cost functions range from quadratic to quintic. The advantage of this approach is that the authors can account for the fact that when demand is very high, inefficient peaking plants start running and we observe a supply stack that resembles a "hockey stick". The results of their analyses imply that the

---

[7]Currently, offers are capped at $3000/MWh in ERCOT.

forward power price is a biased forecast of the expected spot price. They also come up with equations describing producers' and retailers' optimal forward positions.

Another assumption in many of these models (see, for example [34] and [18]) that we find disturbing is that "a sufficiently large proportion of the buyers in the contract market are risk-neutral with rational expectations, and will therefore drive the contract price to equal the expected spot price", that is to say, there is no risk premium. In light of the empirical evidence[8], we also drop this assumption.

A brief description of wholesale power markets

Before we go into the details of our model, we give a brief description of the nature of competition in a wholesale power market.

Producers generate electricity from many different sources with varying marginal costs. The overall marginal cost of a power producer is a non-decreasing function of quantity ("merit order"). Since the main determinants of production costs, the fuel used and the heat rates,[9] are common knowledge, each producer can estimate its competitors' cost functions with a great deal of accuracy.[10]

Power producers sell electricity in the wholesale market to large industrial customers and retailers. Retailers add a mark-up and sell electricity to their customers. Since electricity cannot be stored in a cost effective way to alleviate price spikes, market participants trade forward contracts to smooth out their income streams.

---

[8]See, for example, Longstaff and Wang (2004, [61]).

[9]Heat rate is an input-output measure of efficiency of a power plant and tells how much heat energy is required (measured in Btu's-British thermal units) to produce one kilowatt-hour of energy. It implicitly assumes a linear short-run production technology.

[10]There is no heat rate associated with hydroelectric plants, wind turbines and solar cells but the short run marginal costs of these sources are essentially zero - except for hydroelectricity based on stored water, where the short run cost is essentially the opportunity cost of the water. In many systems, such as ERCOT, however, there is little such capacity.

In the next section we develop a stylized model of a wholesale power market, where the players as risk-averse agents. While this complicates the algebra quite a bit, we believe it is an indispensable feature of the players in power markets. Since electricity is not storable at the wholesale level, and supply and demand should balance at all times, buy-hold-sell type of strategy is not an option and hence basing our arguments on the expectation that forward and spot power prices would converge is not very logical[11]. In an effort to represent the bidding behavior in today's wholesale power markets more realistically, we also model the spot market as one where the players compete with supply functions.

2.2   Model

We model a wholesale market with set $M = \{G_1, G_2, ..., G_m\}$ of risk averse generators and set $N = \{R_1, R_2, ..., R_n\}$ of risk averse retailers endowed with the utility function

$$u_k(\pi_k) = -e^{-\gamma_k \pi_k}, \quad k \in M \cup N, \tag{2.1}$$

where γ_k is the risk aversion parameter of k and π_k is its profit function:

$$\pi_{G_i}(p) = p \cdot q_{G_i}(p) + (f - p) \cdot x_{G_i} - C_{G_i}(q_{G_i}(p)), \tag{2.2}$$

if k is a generator, and

$$\pi_{R_i}(x_{R_i}) = q_{R_i} \cdot (p_R - p) + x_{R_i} \cdot (p - f), \tag{2.3}$$

---

[11]A party which has a contract to purchase power also has to arrange a buyer ("sink") for that power because of suppy-demand and non-storability constraints. As also pointed out in [9], occasionally some power marketers default on their power purchase agreements, not because they cannot afford to honor their contracts, but because they cannot arrange a sink.

if k is a retailer.

Generator i's aim is to maximize its expected utility of profits where the term $(f - p)x_{G_i}$ in its profit function represents the "two-way contracts for differences": a generator needs to refund its customers if the spot price p exceeds the contract price f and vice versa. Retailers serve their demand (q_{R_i}) at a fixed price (p_R), hedge their positions in the contract market and try to maximize their expected utility of profit.

The generators are characterized by quadratic cost functions given by $C_i(q_{G_i}) = 0.5c_iq_{G_i}^2 + a_iq_{G_i}$ for all $i = 1, 2, ..., n$, which leads to an affine marginal cost function for each generator:

$$MC_i = \frac{dC_i(q_{G_i})}{dq_{G_i}} = c_iq_{G_i} + a_i, \quad i = 1, 2, ..., n. \tag{2.4}$$

In order to ensure strict convexity of the cost functions, we assume that $c_i > 0$ for all i.

Generators compete with nondecreasing and affine supply functions

$$q_{G_i}(p) = \begin{cases} 0 & \text{if} \quad p < max\{0, -\alpha_i/\beta_i\} \\ \alpha_i + \beta_ip & \text{if} \quad p \geq max\{0, -\alpha_i/\beta_i\} \end{cases} \quad i = 1, 2, ..., n, \tag{2.5}$$

by choosing α_i and β_i[12].

Generators also compete in the forward contract market by selling x_{G_i}, $i = 1, ..., m$ units of contracts to the retailers. Retailers' demand for the forward contracts is denoted by x_{R_i}, $i = 1, ..., n$.

---

[12]Our model is more general than it actually looks since we can interpret supply as net of baseload generation (hydro, nuclear and large coal plants) and variable and intermittent generation sources such as wind and solar. Analysis of generators' marginal cost curves also shows that over a very wide range of output (again, net of baseload and peak demand), marginal costs are approximately linear. This point is also stressed in [34] and [101].

We assume π_{G_i} and π_{R_i} are distributed normally, where randomness is coming from the inelastic market demand, denoted by Q. Let μ denote the expected profit of player $i \in M \cup N$ and σ^2 denote its variance. Then:

Lemma 2.2.1. *For player $i \in M \cup N$, maximizing its expected utility of profit boils down to maximizing the expression*

$$\mu - \frac{1}{2}\gamma\sigma^2.$$

Proof. Expected utility of profit is equal to

$$\int -e^{-\gamma\pi_i}dN(\pi_i) = \frac{1}{\sigma\sqrt{2\pi}}\int -e^{-\gamma\pi_i}e^{-\frac{1}{2}\left(\frac{\pi_i-\mu}{\sigma}\right)^2}d\pi_i = \frac{1}{\sigma\sqrt{2\pi}}\int -e^{-\frac{1}{2\sigma^2}\left(\gamma\pi_i2\sigma^2+(\pi_i-\mu)^2\right)}d\pi_i$$

Noting that $\left(\gamma\pi_i2\sigma^2 + (\pi_i - \mu)^2\right) = \left((\pi_i - \mu + \gamma\sigma^2)\right)^2 - \gamma\sigma^2\left(\gamma\sigma^2 - 2\mu\right)$, we can write the above expression as

$$-e^{-\gamma\left(\mu-\frac{\gamma\sigma^2}{2}\right)}\left[\frac{1}{\sigma\sqrt{2\pi}}\int -e^{-\frac{1}{2\sigma^2}\left(\pi_i-(\mu-\gamma\sigma^2)\right)^2}d\pi_i\right]$$

But the quantity inside the square brackets is the integral of the pdf of a normal random variable with mean $\mu - \gamma\sigma^2$ and variance σ^2 over its full support, hence it is equal to 1. It follows that, the objective of i is to maximize $\mu - \frac{1}{2}\gamma\sigma^2$. $\square$

2.2.1 Equilibrium in the Spot Market

We start solving our model at the spot market stage, taking the contract positions of the players as given.

Empirical evidence suggests that, as the spot market approaches, the system operator and the wholesalers can determine the system load with a great deal of accuracy.

This implies that generator i's problem at the spot market stage can be written as [13]

$$\max_p \pi_{G_i}(p) = pq_{G_i}(p) + (f - p)x_{G_i} - C_i(q_{G_i}(p)). \tag{2.6}$$

Market demand, Q, which is equal to the total retail demand $\sum_{i=1}^{n} q_{R_i}$ will be satisfied by total generation, given by $\sum_{i=1}^{m} q_{G_i}$. This implies that $q_{G_i} = Q - \sum_{j \neq i}^{m} q_{G_j}$. Then we can write (2.6) as [14]

$$\max_p \quad p\left(Q - \sum_{j \neq i} q_{G_j}(p) - x_{G_i}\right) + f \cdot x_{G_i} - \frac{1}{2}c_i\left(Q - \sum_{j \neq i} q_{G_j}(p)\right)^2 - a_i\left(Q - \sum_{j \neq i} q_{G_j}(p)\right) \tag{2.7}$$

with the first-order condition

$$q_{G_i}(p) - x_{G_i} - \left(p - a_i - c_i q_{G_i}(p)\right)\left(\sum_{j \neq i} \beta_j\right) = 0, \quad i = 1, 2, ..., m. \text{[15]} \tag{2.8}$$

By (2.5), we can write (2.8) as

$$\alpha_i + \beta_i p = x_{G_i} - (a_i + c_i\alpha_i)\left(\sum_{j \neq i} \beta_j\right) + (1 - c_i\beta_i)\left(\sum_{j \neq i} \beta_j\right)p \tag{2.9}$$

---

[13]G_i can estimate the generation stack and the (inelastic) market demand. Hence, given its own bids, it also knows whether it is going to be dispatched or not (q_{G_i}). Here, it is implicitly assumed that G_i has an idea as to what G_j bid for that day.

[14]Where no confusion arises, we drop the limits from the summation sign. So $\sum a_i$ means $\sum_{i=1}^{n} a_i$ and so on.

[15]By equation (2.5), the second-order condition is

$$\frac{\partial^2 \pi_i}{\partial p^2} = -\left(\sum_{j \neq i} \beta_j\right)\left(2 + c_i \sum_{j \neq i} \beta_j\right),$$

which is non-positive since for all i, c_i and β_i are positive. Hence, the solution to (2.8) is indeed the maximizer.

which should hold for any value of p, hence we have

$$\alpha_i = x_{G_i} - (a_i + c_i\alpha_i)\left(\sum_{j\neq i}\beta_j\right) \qquad (2.10)$$

and

$$\beta_i p = (1 - c_i\beta_i)\left(\sum_{j\neq i}\beta_j\right)p \qquad (2.11)$$

which together give

$$\alpha_i = \frac{x_{G_i} - a_i\sum_{j\neq i}\beta_j}{1 + c_i\sum_{j\neq i}\beta_j} \qquad (2.12)$$

and

$$\beta_i = \frac{\sum_{j\neq i}\beta_j}{1 + c_i\sum_{j\neq i}\beta_j} \qquad (2.13)$$

From this point on, we assume that the producers use identical generating technologies and treat them as symmetric players.

Given the symmetry assumption, total physical production is

$$\sum_{i=1}^{m} q_{G_i} = \sum_{i=1}^{m}(\alpha_i + \beta_i p) = \frac{1}{1 + c(m-1)\beta}\left[\sum x_{G_i} - am(m-1)\beta + m(m-1)\beta p\right].$$

$$(2.14)$$

Equating total physical production to market demand $Q = \sum q_{R_i}$ we get

$$p = \frac{\left(1 + c(m-1)\beta\right)Q + am(m-1)\beta - X}{m(m-1)\beta}, \qquad (2.15)$$

where $X = \sum x_{G_i} = \sum x_{R_i}$ is the total forward position. It is clear from equations (2.13) and (2.15) that forward positions of generators have no effect on the equilibrium supply functions but the equilibrium spot price is decreasing in the total forward

position.

We can express generator i's production as

$$q_{G_i} = \alpha_i + \beta_i p \tag{2.16}$$

$$= \frac{x_{G_i} - a(m-1)\beta}{1 + c(m-1)\beta} + \frac{(m-1)\beta}{1 + c(m-1)\beta} \frac{\left(1 + c(m-1)\beta\right)Q + am(m-1)\beta - X}{m(m-1)\beta}$$

$$= \frac{mx_{G_i} - am(m-1)\beta + \left(1 + c(m-1)\beta\right)Q + am(m-1)\beta - X}{m[1 + c(m-1)\beta]} \tag{2.17}$$

To simplify the notation, let $z := m(m-1)\beta$ and $t := 1 + c(m-1)\beta$. Then

$$p = \frac{az - X + tQ}{z}, \tag{2.18}$$

and

$$q_{G_i} = \frac{mx_{G_i} - X + tQ}{tm}. \tag{2.19}$$

2.2.2 Equilibrium in the Contract Market

Following Bessembinder and Lemmon (2002, [9]), we model the contract market as a closed system. Given the spot price and quantity calculated above, we can now solve for the players' optimal forward positions.

Generators' Problem

Generator i's problem at the forward contracting stage is to

$$\max_{x_{G_i}} E(u_{G_i}) = \max_{x_{G_i}} \left(E(\pi_{G_i}) - \frac{\lambda}{2} var(\pi_{G_i}) \right)$$

with the following first order condition:

$$\frac{\partial E(u_{G_i})}{\partial x_{G_i}} = \frac{\partial E(pq_{G_i})}{\partial x_{G_i}} + f - E(p) - x_{G_i}\frac{\partial E(p)}{\partial x_{G_i}} - \frac{\partial}{\partial x_{G_i}}(\frac{c}{2}E(q_{G_i}^2) + aE(q_{G_i}))$$

$$- \frac{\lambda}{2}\left[\frac{\partial \text{var}(pq_{G_i})}{\partial x} + \frac{\partial}{\partial x_{G_i}}x_{G_i}^2 \text{var}(p) + \frac{c^2}{4}\frac{\partial}{\partial x_{G_i}}\text{var}(q_{G_i}^2) + a^2\frac{\partial}{\partial x_{G_i}}\text{var}(q_{G_i})\right]$$

$$+ \lambda\left[\frac{\partial}{\partial x_{G_i}}x_{G_i}\text{cov}(pq_{G_i},p) + \frac{\partial}{\partial x_{G_i}}\frac{c}{2}\text{cov}(pq_{G_i},q_{G_i}^2) + \frac{\partial}{\partial x_{G_i}}a\text{cov}(pq_{G_i},q_{G_i})\right.$$

$$\left. - \frac{\partial}{\partial x_{G_i}}\frac{x_{G_i}c}{2}\text{cov}(p,q_{G_i}^2) - \frac{\partial}{\partial x_{G_i}}ax_{G_i}\text{cov}(p,q_{G_i}) - \frac{\partial}{\partial x_{G_i}}\frac{ac}{2}\text{cov}(q_{G_i}^2,q_{G_i})\right] = 0$$

$$(2.20)$$

Substituting for p and q_{G_i} from the first stage, various variance and covariance terms from appendix C.1 and letting V stand for $\sum_{j \neq i}^{m} \frac{\partial x_{G_j}}{\partial x_{G_i}}$ we can express this first-order condition as[16]

---

[16]The second derivative of this expression with respect to the generator's forward position will not to be negative over all of its domain, parameterized by its six parameters, $\beta, c, m, var(Q), \lambda$ and $\sum_{j \neq i}^{m} \frac{\partial x_{G_j}}{\partial x_{G_i}}$. We will run our simulations only on the range where the second derivative is guaranteed to be negative. See appendix D.1 for details.

$$\frac{\partial E(u_{G_i})}{\partial x_{G_i}} = \frac{m-2-2V}{zm}E(Q) + \frac{a\left(m-1-V\right)}{tm}$$

$$+ \frac{2x_{G_i}(1-m)\left(1+V\right) + X_{-i}(2-m+2V)}{mtz} + f - \frac{az - x_{G_i} - X_{-i}}{z} - \frac{tE(Q)}{z} + \frac{x_{G_i}(1+V)}{z}$$

$$- a\left(\frac{m-1-V}{tm}\right) - c\left(\frac{m-1-V}{tm^2}E(Q) + \frac{x_{G_i}m^2 - m(2x_{G_i} + X_{-i} + x_{G_i}V) + (x_{G_i} + X_{-i})(1+V)}{m^2t^2}\right)$$

$$- \frac{\lambda}{2}\left[\frac{2(m-2-2V)}{m^2z^2}\left(x_{G_i}(m-2) - 2X_{-i} + az\right)\text{var}(Q) + \frac{2t(m-2-2V)}{m^2z^2}\text{cov}(Q^2, Q) + 2x_{G_i}\frac{t^2}{z^2}\text{var}(Q)\right.$$

$$\left. + \frac{c^2}{m^4t}\left(m-1-V\right)\left(\frac{2}{t}(x_{G_i}(m-1) - X_{-i})\text{var}(Q) + \text{cov}(Q^2, Q)\right)\right]$$

$$+ \lambda\left[\frac{t^2}{mz^2}\text{cov}(Q^2, Q) + \frac{t}{z}\left(\frac{x_{G_i}(m-2) - X_{-i} + az}{mz}\right)\text{var}(Q) + x_{G_i}\frac{t(m-2-2V)}{mz^2}\text{var}(Q)\right.$$

$$+ \frac{c}{2}\left(\frac{3m-4-4V}{m^3z}\text{cov}(Q, Q^2) + (m^3tz)^{-1}(4x_{G_i}(m-2)(m-1) + X_{-i}(8-6m))\right.$$

$$\left. + \frac{2az(m-1-V) + x_{G_i}(8-6m)V + 8X_{-i}V}{m^3tz}\text{var}(Q)\right)$$

$$+ \frac{a\text{var}(Q)}{m^2z}(m-2-2V) - \frac{c}{2}\left(\frac{t}{m^2z}\text{cov}(Q, Q^2) + \frac{2}{m^2z}(x_{G_i}(m-1) - X_{-i})\text{var}(Q)\right)$$

$$\left. - \frac{cx_{G_i}\text{var}(Q)}{m^2z}(m-1-V) - \frac{at}{mz}\text{var}(Q) - \frac{ac\text{var}(Q)}{m^3t}(m-1-V)\right] = 0. \qquad (2.21)$$

Factoring out x_{G_i}, X_{-i} and $\sum_{j\neq i}^{m}\frac{\partial x_{G_j}}{\partial x_{G_i}}$, we obtain the following system of equations:

$$\begin{bmatrix} A_1 & B_1 & \dots & B_1 \\ B_2 & A_2 & \dots & B_2 \\ \vdots & \vdots & \ddots & \vdots \\ B_m & B_m & \dots & A_m \end{bmatrix} \begin{bmatrix} x_{G_1} \\ x_{G_2} \\ \vdots \\ x_{G_m} \end{bmatrix} = \begin{bmatrix} C_1 \\ C_2 \\ \vdots \\ C_m \end{bmatrix},$$

where

$$A_i = \frac{2(1-m)}{mtz} + \frac{2}{z} - \frac{c(m-1)^2}{m^2t^2}$$
$$- \lambda \text{var}(Q)\left[\frac{(m-2)^2}{m^2z^2} + \frac{t^2}{z^2} + \frac{c^2(m-1)^2}{m^4t^2} - \frac{2t(m-2)}{mz^2} - \frac{2c(m-1)(m-2)}{m^3tz} + \frac{2c(m-1)}{m^2z}\right]$$
$$+ \sum_{j\neq i}^{m} \frac{\partial x_{G_j}}{\partial x_{G_i}}\left[\frac{2+m(t-1)}{mtz} + \frac{c(m-1)}{m^2t^2}\right.$$
$$\left. + \lambda \text{var}(Q)\left(\frac{2(m-2)}{m^2z^2} + \frac{c^2(m-1)}{m^4t^2} - \frac{2t}{mz^2} + \frac{2c(3m-4)}{m^3tz} + \frac{c}{m^2z}\right)\right], \qquad (2.22)$$

$$B_i = \frac{2+m(t-1)}{mtz} + \frac{c(m-1)}{m^2t^2} + \lambda \text{var}(Q)\left[\frac{2(m-2)}{m^2z^2} + \frac{c^2(m-1)}{m^4t^2} - \frac{2t}{mz^2} + \frac{c(4-3m)}{m^3tz} + \frac{c}{m^2z}\right]$$
$$+ \sum_{j\neq i}^{m} \frac{\partial x_{G_j}}{\partial x_{G_i}}\left[\frac{2}{mtz} - \frac{c}{m^2t^2} + \lambda \text{var}(Q)\left(\frac{4c}{m^3tz} - \frac{4}{m^2z^2} - \frac{c^2}{2m^4t}\right)\right], \qquad (2.23)$$

and,

$$C_i = E(Q)\left[\frac{2-m}{mz} + \frac{t}{z} + \frac{c(m-1)}{m^2t}\right] - f + a$$
$$- \lambda \text{cov}(Q^2, Q)\left[\frac{t(2-m)}{m^2z^2} - \frac{c^2(m-1)}{2m^4t} - \frac{ct}{2m^2z} + \frac{t^2}{mz^2} + \frac{c(3m-4)}{2m^3z}\right]$$
$$- \sum_{j\neq i}^{m} \frac{\partial x_{G_j}}{\partial x_{G_i}}\left[\lambda \text{cov}(Q^2, Q)\left(\frac{2t}{m^2z^2} + \frac{c^2}{2m^4t} - \frac{2c}{m^3z}\right) + E(Q)\left(\frac{c}{tm^2} - \frac{2}{zm}\right)\right] \qquad (2.24)$$

Solution to this system will give the optimal forward positions of the generators.

Retailers' Problem

We assume that retailers sell power to their customers with fixed price contracts. Let this fixed price be $1 + r$ times the wholesale spot price, where $r > 0$. Given this

assumption, the expression for retailer i's profit reduces to

$$\pi_{R_i} = q_{R_i} rp + x_{R_i}(p - f).$$ (2.25)

and retailer i wants to maximize

$$
\begin{aligned}
E(u_{R_i}) &= E(\pi_{R_i}) - \frac{\lambda}{2}\text{var}(\pi_{R_i}) \\
&= rE(pq_{R_i}) - fx_{R_i} + x_{R_i}E(p) \\
&\quad - \frac{\lambda}{2}\left[r^2\text{var}(pq_{R_i}) + x_{R_i}^2\text{var}(p) + 2rx_{R_i}\text{cov}(pq_{R_i}, p)\right]
\end{aligned}
$$ (2.26)

by choosing x_{R_i}. The first order condition for this problem is[17]

$$
\begin{aligned}
\frac{\partial E(u_{R_i})}{\partial x_{R_i}} &= r\frac{\partial E(pq_{R_i})}{\partial x_{R_i}} - f + E(p) + x_{R_i}\frac{\partial E(p)}{\partial x_{R_i}} \\
&\quad - \frac{\lambda}{2}\left[r^2\frac{\partial \text{var}(pq_{R_i})}{\partial x_{R_i}} + 2x_{R_i}\text{var}(p) + x_{R_i}^2\frac{\partial \text{var}(p)}{\partial x_{R_i}}\right. \\
&\quad \left. + 2r\text{cov}(pq_{R_i}, p) + 2rx_{R_i}\frac{\partial \text{cov}(pq_{R_i}, p)}{\partial x_{R_i}}\right] \\
&= 0.
\end{aligned}
$$ (2.27)

Substituting for p and various variance-covariance terms from appendix C.1, we can write

$$A_i \cdot x_{R_i} + B_i \cdot X_{-i} = C_i \quad \text{for all } i \in N,$$ (2.28)

where

$$A_i = \frac{2}{z} + \lambda\text{var}(Q)\left[\left(\frac{t - rs_{R_i}}{z}\right)^2 - \frac{rs_{R_i}}{z^2}(t - rs_{R_i})\sum_{j\neq i}^{n}\frac{\partial x_{R_j}}{\partial x_{R_i}}\right]$$

---

[17]Second order condition shows that x_{R_i} is indeed the maximizer. See appendix D.2 for details.

$$B_i = \frac{1}{z} - \frac{\lambda r s_{R_i} \text{var}(Q)}{z} \left(\frac{t - r s_{R_i}}{z} \right) + \frac{\lambda r^2 s_{R_i}^2}{z^2} \sum_{j \neq i}^{n} \frac{\partial x_{R_j}}{\partial x_{R_i}},$$

$$C_i = a - f + \left(\frac{t - r s_{R_i}}{z} \right) \left[E(Q) - \frac{\lambda a r s_{R_i}}{z} \left(a z \text{var}(Q) + t \text{cov}(Q^2, Q) \right) \right]$$

$$+ \sum_{j \neq i}^{n} \frac{\partial x_{R_j}}{\partial x_{R_i}} \left(\frac{\lambda r^2 s_{R_i}^2 a \text{var} Q}{z} + \frac{\lambda r^2 s_{R_i}^2 t \text{cov}(Q^2, Q)}{z^2} - \frac{r s_{R_i}}{z} E(Q) \right),$$

and s_{R_i} is the share of retailer i in the contract market, with $s_{R_1} + ... + s_{R_n} = 1$. The solution of the following system of equations will give x_{R_i}:

$$\begin{bmatrix} A_1 & B_1 & ... & B_1 \\ B_2 & A_2 & ... & B_2 \\ \vdots & \vdots & \ddots & \vdots \\ B_n & B_n & ... & A_n \end{bmatrix} \begin{bmatrix} x_{R_1} \\ x_{R_2} \\ \vdots \\ x_{R_n} \end{bmatrix} = \begin{bmatrix} C_1 \\ C_2 \\ \vdots \\ C_n \end{bmatrix}.$$

Both for the generators and the retailers, we have to solve a linear system of equations. To that end, we need to invert the "coefficient" matrix shown above. The following lemma tells us how to do that:

Lemma 2.2.2. *Let T be an $n \times n$ $(n > 2)$ matrix where the ith diagonal element is A_i and the off-diagonal elements on row i are B_i. Its inverse is another $n \times n$ matrix with the ith diagonal entry and the ij^{th} off-diagonal entries given respectively as*

$$\frac{\prod\limits_{j \neq i} A_j + \left[\sum\limits_{k=1}^{n-2} (-1)^k k \left(\sum\limits_{(S \setminus \{i\})_{n-k-2}} \prod\limits_{j \in (S \setminus \{i\})_{n-k-2}} A_j \prod\limits_{l \neq i,j} B_l \right) \right]}{\prod\limits_{j=1}^{n} A_j + \sum\limits_{k=2}^{n} \left[(-1)^{k-1}(k-1) \sum\limits_{S_{n-k}} \prod\limits_{i \in S_{n-k}} A_i \prod\limits_{j \neq i} B_j \right]} \tag{2.29}$$

$$\frac{-B_i \prod\limits_{k \neq \{i,j\}} (A_k - B_k)}{\prod\limits_{i=1}^{n} A_i + \sum\limits_{k=2}^{n} \left[(-1)^{k-1}(k-1) \sum\limits_{S_{n-k}} \prod\limits_{i \in S_{n-k}} A_i \prod\limits_{l \neq i} B_l \right]},$$ (2.30)

where S_m denotes the m-element subsets of $S = \{1, 2, 3, ..., n\}, m > 0$. If $m = 0$, $S_0 = \emptyset$, in which case $\prod_{k \neq i,j}$ and $\prod_{l \neq i}$ become $\prod_{k \neq i}$ and $\prod_l$, respectively. If $n = 2$, the solution is trivial.

Proof. As shown in appendix E, this follows from lemma E.0.1. $\qquad\square$

The expression

$$\sum_{(S \backslash \{i\})_{n-k-2}} \prod_{j \in (S \backslash \{i\})_{n-k-2}} A_j \prod_{l \neq i,j} B_l$$

means that the product

$$\prod_{j \in (S \backslash \{i\})_{n-k-2}} A_j \prod_{l \neq i,j} B_l$$

is going to be summed over all possible $(n - k - 2)$-element subsets of the set $S \backslash \{i\}$. An example will make it clearer:

Let $n = 4$ and

$$T = \begin{bmatrix} A_1 & B_1 & B_1 & B_1 \\ B_2 & A_2 & B_2 & B_2 \\ B_3 & B_3 & A_3 & B_3 \\ B_4 & B_4 & B_4 & A_4 \end{bmatrix}.$$

When we invert T, the first diagonal element and the off-diagonal element at the first row, second column will respectively be

$$\frac{A_2 A_3 A_4 - A_2 B_3 B_4 - A_3 B_2 B_4 - A_4 B_2 B_3 + 2 B_2 B_3 B_4}{denominator},$$ (2.31)

and

$$\frac{-B_1(A_3 - B_3)(A_4 - B_4)}{denominator}, \tag{2.32}$$

where *denominator* is

$$A_1A_2A_3A_4 - A_1A_2B_3B_4 - A_1A_3B_2B_4 - A_1A_4B_2B_3 - A_2A_3B_1B_4 - A_2A_4B_1B_3$$

$$- A_3A_4B_1B_2 + 2(A_1B_2B_3B_4 + A_2B_1B_3B_4 + A_3B_1B_2B_4 + A_4B_1B_2B_3) - 3B_1B_2B_3B_4. \tag{2.33}$$

Other diagonal and off-diagonal elements will be similar.

This lemma is useful for expressing the optimal forward positions analytically. However, the resulting expressions will be extremely long, and we see no benefit in presenting them here. Instead, we will implement this algorithm with the help of a computer and solve for the optimal forward positions numerically, as a function of the forward price, f. After we solve for the forward positions of both the generators and the retailers, we will obtain f by noting that, given our assumption that the system is closed, forward contracts will be in zero net supply:

$$\sum_{i=1}^{m} x_{G_i}(f) = \sum_{i=1}^{n} x_{R_i}(f).$$

There are two main types of models in the literature on forward contracts. The first one is the no-arbitrage type of models and we already argued that they are not suitable for analyzing power markets. In the second type of models, the forward prices are determined endogenously. French (1986, [31]), Fama and French (1987, [27]), Hirshleifer (1990, [39]) and Bessembinder and Lemmon (2002,[9])[18] are some examples of this approach. The main focus in these studies is the relationship between

---

[18]A more detailed literature review can be found in [61].

forward and spot prices, in particular the so-called "forward premium", $f - E(p)$, or the forward premium as a percentage of the spot price. From an investor's point of view, the investment decision may be a function of the forward premium. For example, in order to value a tolling agreement, a firm would want to know how the forward price it will lock in will compare to the spot price for a given level of power demand.

We also focus on this metric and present our results in the next section.

2.3 Results

Power demand varies significantly during the day and across seasons.[19] We ran some simulations in order to understand how the level and variation in power demand affect the direction of the forward bias. By playing with the levels of the several parameters of the model, we can make the bias arbitrarily large, but since we are not interested in the absolute values of the forward price and expected spot price in order to answer this particular question, the statistic we look at is the bias $f - E(p)$ as a percentage of expected spot price, $E(p)$. The result can be seen in figure 2.2, which is consistent with the intuition in the classical economics and finance literature: To hedge their long position in the underlying commodity, producers (i.e., power generators) get short in the forward/futures market. It is this hedging pressure that creates the downward bias in the contract price $(E(p) > f)$.[20] Risk-averse buyers (i.e., retailers), who take long positions in these contracts are being compensated by a positive expected profit for bearing risk. Our model implies that, given the level of expected demand, as

---

[19]In ERCOT, minimum demand observed during 2009 was 21,340 MWs on October 18 and the maximum was observed on July 13, at 63,516 MWs. The corresponding figures for 2010 were: 21,728 MWs on April 25, and 64,805 MWs on August 23rd.

[20]See [20], [19], [90], [39], [61] and the references therein for details.

the demand uncertainty increases, especially at lower levels of expected demand, this hedging pressure becomes even more dominant and retailers on the other side of the contract demand higher returns for bearing this risk.

Bias as a percentage of expected spot price
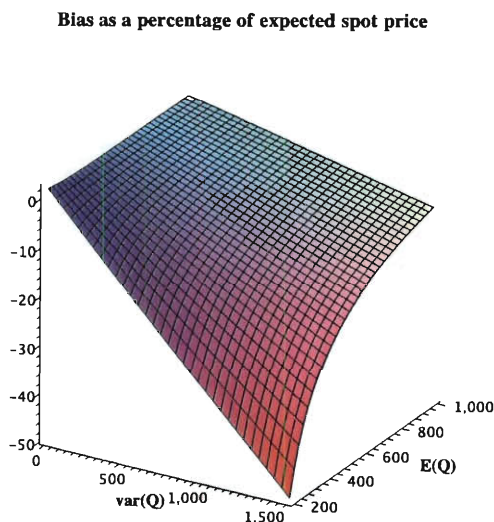
Figure 2.2 : There are 50 generators and 100 retailers with Bertrand conjectures, $a = 1$, $c = 5$, $\lambda_G = .2$, $\lambda_R = .5$, $\beta = .2$, $r = .2$.

To see what the setup in [9] would imply with quadratic costs, we reproduced their simulations with quadratic cost functions.[21] The general relationship between the variables and the percentage bias is the same, albeit at different levels.[22]

Next, we fix the risk aversion parameter of the retailers at $\lambda_R = 1$, expected demand at 100 and vary λ_G to see how this would affect the direction of bias. We

---

[21]This was necessary, since they only consider quartic cost functions in their simulations.

[22]These results seem to depend on how convex the cost function is. Hirshleifer (1990, [39]) shows that the equilibrium forward premium need not be negative, and in the Bessembinder-Lemmon setup with cubic, quartic and quintic cost functions the premium is strictly positive and increasing in demand variance. In their empirical work, Longstaff and Wang (2004, [61]) analyze the forward prices in the PJM market and find that average premia range from -$4.31 to $5.44.)

assume that generators are bidding competitively. As can be seen in figure 2.3, given the level of demand variance, as the generators become more risk averse, the bias increases in absolute value, conforming to our intuition. As the generators become less risk averse, maybe as they recover their fixed costs, the bias shrinks, again, consistent with our intuition.
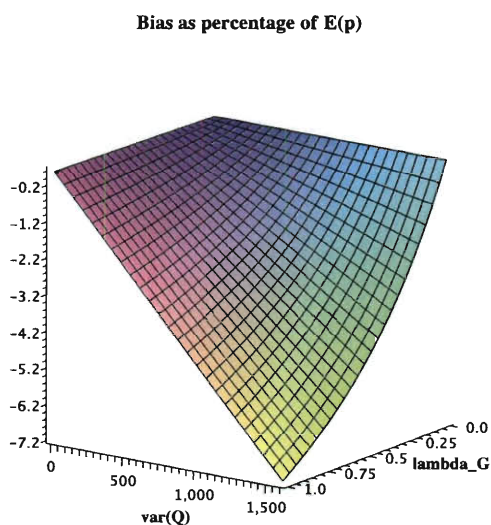
Figure 2.3 : Generators' risk preferences may change over time.

To assess the effect of supply function slope on the prices and the percentage bias, we calibrated our model with demand data from ERCOT. Generators expect demand to be 50,000 megawatts during a given hour, with a standard deviation of 2,000 megawatts. With 30 generators and 100 retailers, when the supply functions are very steep, we observe price spikes (and correspondingly high percentage bias in absolute value). As the generators start bidding more competitively, the spot price of power decreases dramatically and converges to the forward price (figures 2.4a and 2.4b). Here, $\beta = 50$ means they are bidding their marginal costs.

(a)   (b)
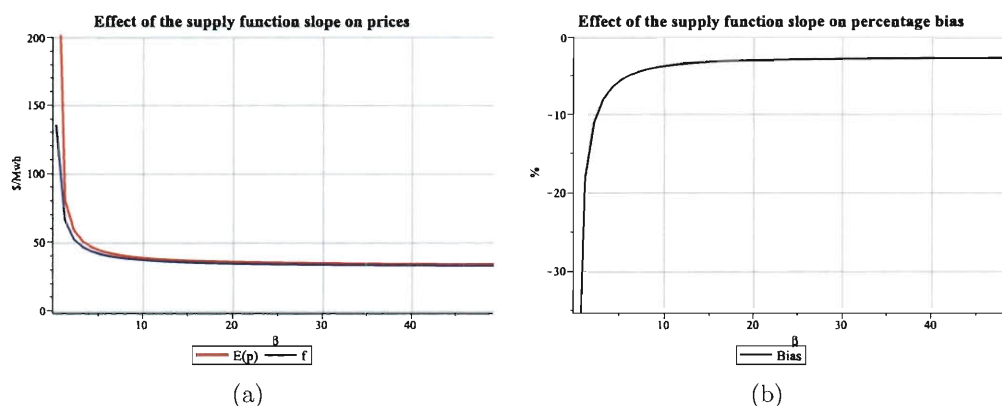
Figure 2.4 : $Q = 50,000$ MW, $\sigma_Q = 2000$ MW, $m = 30$, $n = 100$. Spot price converges to \$34.5/MWh

We observe the same trend by increasing the number of generators to 100 from 30, keeping the number of retailers constant. The spot price converges to \$10.1/MWh and the bias in forward price converges to less than one-half percent (figures 2.5a and 2.5b).

In order to see how increasing the number of market participants would affect the forward premium, we ran another set of simulations with these same parameters, by changing the number of generators and retailers. The bias is monotonically decreasing in the number of players. With 10 generators and 10 retailers, the bias is -33.9%. If we increase the number of generators to 100, the bias decreases to -0.5%. With 10 generators and 100 retailers, it is -3.2% and when we have 100 generators and 100 retailers, the bias as a percentage of expected spot price shrinks to -0.28%.[23]

[23]Even though it is virtually impossible to have 100 power generation companies in any market, bidding is done through qualified scheduling entities (QSEs) and for instance, in ERCOT there are about 160 registered QSE's operating as of March 2011 (see ercot.com/mktparticipants for details). While there is no way for us to know which QSE represents which company, except for generating companies running their own QSEs, a power generation company may use different QSEs for different generation assets it owns.
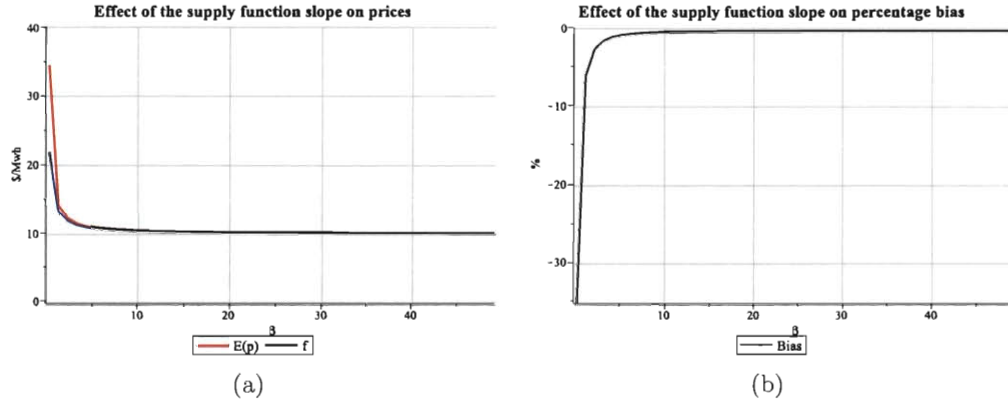
Figure 2.5 : $Q = 50,000$ MWh, $\sigma_Q = 2000$ MWh, $m = 30$, $n = 100$. Spot price converges to \$10.1/MWh

As a proxy to observability of contracts, we also check the effect of changes in generators' and retailers' conjectural variations on the amount of power sold forward. In the literature with risk-neutral agents, if the producers believe that their own contract decisions will not affect those of their competitors' (Cournot conjectures, $\sum_{j\neq i}^{m} \frac{\partial x_{G_j}}{\partial x_{G_i}} = 0$) generators choose not to sell any contracts. If, on the other hand, they believe that competitors' contracting decisions would completely offset a change in their own contract sales (Bertrand conjectures, $\sum_{j\neq i}^{m} \frac{\partial x_{G_j}}{\partial x_{G_i}} = -1$), they cover all of their expected output in the contract market and the competitive outcome will be achieved. Of course, $\sum_{j\neq i}^{m} \frac{\partial x_{G_j}}{\partial x_{G_i}}$ can take any value between -1 and 0, if we want to assess the effects of "imperfect observability".

We take $m = n = 30$ and fix demand at 1000 with a standard deviation of 100. Generators bid in their marginal costs. With Cournot conjectures, generators' total contract sales is 890, which increases only to 891 under Bertrand conjectures. Hence, our model is indecisive as to whether a change in conjectural variations would result

in more contract sales.

2.4   Conclusions and Future Research

We presented a 2-stage model of a wholesale power market where the generators compete with affine supply functions in the spot market and hedge their long positions by selling contracts in the forward market. We model the generators as risk-averse agents. Since electricity is not storable, we cannot assume the existence of a large number of risk-neutral power marketers to eliminate arbitrage opportunities. Therefore, we explicitly model the retailers, who take the opposite side of the forward contracts. Given the fact that the retailers are operating on very thin margins, we also model them as risk-averse agents.

We find that, as demand uncertainty increases, bias as a percentage of expected spot price becomes more negative. Another implication of our model is that as the number of market participants increases, and the generators bid closer to their marginal cost functions, the expected spot price of electricity and the forward price converge to each other.

We also analyze the effect of a change in generators' risk preferences over time. For a given level of expected demand, we consider how a change in the generators' risk parameter, together with a change in demand volatility, affect the forward bias. We find that, given the level of demand volatility, the bias increases in absolute value as the generators become more risk averse.

Risk-aversion is an essential characteristic of market participants in wholesale power markets. Even though it complicates the analysis quite a lot, we believe modeling the players as risk-averse agents is worth the effort. Again, in an effort to mimic the spot market operations in wholesale power markets, we modeled the spot market

as one where the generators' choice variables are affine supply functions. However, due to the technical limitations of the affine supply function equilibrium approach, we were not able to capture one of the most important characteristics of wholesale power markets, namely that the supply stack is approximately convex, and at high quantities it is very steep. Approximating the supply functions and the marginal costs with step functions would be worthwhile, but incorporating this into a model with risk-averse agents is an open question.

Another distinguishing characteristic of power prices is that they are usually highly skewed and may have high kurtosis.[24] Our model gives the skewness of the spot price

(a) $\mu = 24.2$, $\sigma = 19.7$, skewness=-8.1, kurtosis=714.15

(b) $\mu = 34.3$, $\sigma = 122.2$, skewness=9.3, kurtosis=89.3
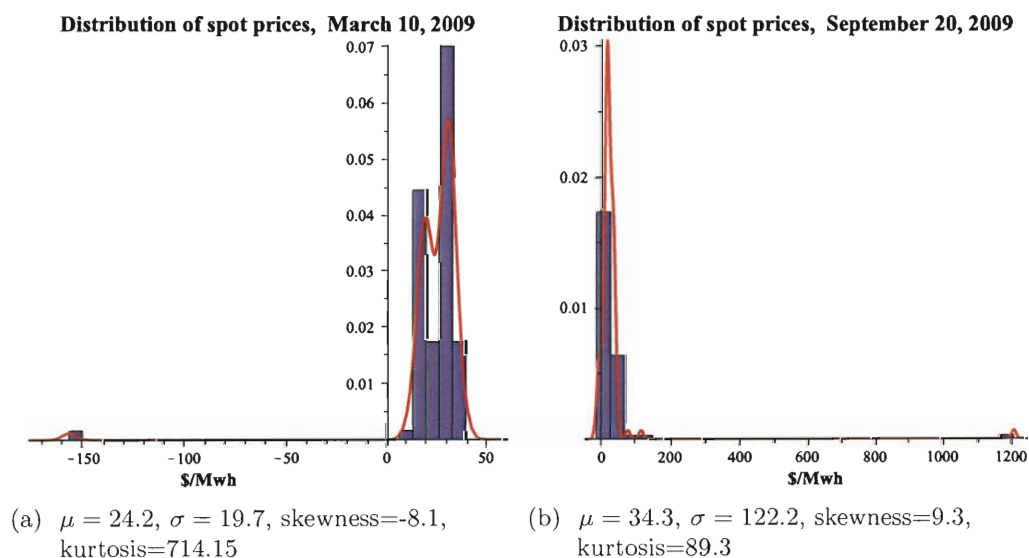
Figure 2.6 : Spot power prices are highly skewed. *Source: www.ercot.com*

---

[24]Skewness is the third, kurtosis is the fourth moment of a random variable around its mean. While skewness is mainly positive, we also frequently observe negative skewness in spot power prices. See figures 2.6a and 2.6b. Kurtosis measures the peakedness of a distribution and a higher kurtosis means more of the variance observed is the result of rare but extreme deviations, as opposed to frequent but modest deviations.

as $(t/z)^3 \cdot \text{skew}(Q)$, which is zero since we take demand to be normally distributed. However, there is strong empirical evidence which suggests otherwise, for both the spot and the forward markets (see, for example, Longstaff and Wang (2004, [61])). Bessembinder and Lemmon (2002, [9]) incorporate skewness into their model and express the premium as a linear combination of skewness and variance of spot power price (see their equation (13)). In their model, forward premia are positively related to skewness but negatively related to price volatility (which is a function of demand variance). Naturally, as the distribution of power prices becomes more skewed, these premia, which are always negative in our model, may become positive. Incorporating price skewness into a supply function competition model is thus another important improvement upon ours.

Chapter 3

An Analysis of Increased Power Demand Price Responsiveness on Locational Marginal Prices

3.1 Introduction

The electric utility sector originated in 1882, when Edison's Pearl Street Power Station began operating in Manhattan, New York City, and the first transmission line, a 2.4 Kv, 37-mile DC line was installed in Germany ([17]). Since then, generator and transmission line capacities have increased dramatically, but for more than a century we have not seen any major breakthroughs on the consumption side: The main concern in the industry had been to secure reliable and uninterrupted power supply. The industry did not focus on the potential efficiency gains from allowing demand side to enter the system more actively. As a result, the medium to long term demand for power, in virtually every wholesale power market had (and still has) very low price responsiveness.

In the past decade or so, some promising improvements in power grid management technology began to surface, the results of which we collectively refer to as the "smart grid". A smart grid can be summarized as a power system, which is capable of integrating the actions of all of its components (generation, transmission and distribution, consumption) by enabling two-way digital communication between these components. As a result of this two-way communication, a smart grid will be able to meet more demand without the need to add new generating capacity by increasing

efficiency of resource use.[1] Furthermore, the demand will be met by higher quality power; with fewer spikes in system voltage, more stable frequency and fewer blackouts. Greater price responsiveness of demand would also facilitate the integration of intermittent energy sources such as wind and solar.

For the purposes of this study, the most important feature of a smart grid is that it will enable the demand side become an active participant in a power market. This is going to be done by replacing our appliances, or retrofitting the existing ones with a smart control unit, communicating with the power grid. These devices will monitor the system frequency[2] in real time and will shift the time of electricity usage of appliances for which time is not critical. Clearly, this flexibility will increase the price responsiveness of demand.

Our aim in this paper is to assess the effects on the level and volatility of power prices of increased responsiveness of demand. We do so by simulating a hypothetical power market populated by autonomous agents. We will also argue that the widely-used market concentration index, the Herfindahl-Hirschman Index, is not a suitable metric for power markets.

3.2   Methodology

In our analysis, we will depart from an equilibrium approach and specifically look at how the agents in our hypothetical power system behave out of equilibrium. This

---

[1]Hirst (2008, [40]), the chief technologist at a British energy solutions company, estimates that if all refrigerators in U.K. had or were retrofitted with his company's "smart control unit", UK could shut down an 800-MW power plant. The same study also estimates the level of response available from retrofitting one million refrigerators with the same device as 20 MWs. This means the elimination of 20 MWs of spinning reserve every hour from the grid.

[2]System frequency indicates whether the supply and demand are in balance in a power grid. The target frequency in the US electricity grid is 60Hz. Any deviation from this value means there is excess supply or demand.

methodology is called "agent-based computational economics" (ACE). Tesfatsion and Judd (2006, [93]) define ACE as " *the computational study of economic processes modeled as dynamic systems of interacting agents who do not necessarily possess perfect rationality and information.*"[3].

Agent-based modeling is a relatively new methodology. More abundant and cheaper computing power has allowed researchers to tackle problems which would be very hard, if not impossible, to solve analytically. One of the fields to benefit most from this modeling approach is power system economics. One of the first papers is Visudhiphan and Ilić (1999, [95]) where the authors model a wholesale power market as a dynamic system in which generators can learn from their previous actions. They find that the daily repetition aspect of real electricity markets plays an important role in market dynamics such as exercising market power. Later, Bower and Bunn (2000, [13]) find that pool-based daily auctions (e.g., PJM) can result in lower power prices than continuous bilateral trading (e.g., ERCOT, prior to December 2010). Nicolaisen et al. (2001, [69]) use a modified version of the learning algorithm developed by Erev and Roth (1995, 1998; [25, 26]) to analyze the efficiency consequences of firm concentration and capacity choices.[4] In a recent paper, Li and Tesfatsion (2010, [55]) assess the effects of strategic bidding on system operator's net surplus collection[5] and find that as the generators act more strategically, which results in less efficient dispatch, the ISO's net surplus increases.

For our analyses, we will use the AMES Wholesale Market Testbed, which is

---

[3]For a more detailed overview, see Tesfatsion (2006, [92]). For several other definitions and a comparison to the classical equilibrium approach, see Arthur (2006, [4]).

[4]For a more extensive survey of the applications of ACE to wholesale power markets, see Weidlich and Veit (2008, [98]).

[5]System operators, or ISOs, are supposed to be revenue-neutral entities. In today's power markets, any positive surplus they collect is -mostly- distributed to financial transmission rights holders.

developed by University of Iowa researchers for experimental analysis of wholesale power markets.[6] AMES encompasses some of the core features of the Federal Energy Regulatory Commission's recommendations for a standard power market design. It models a one-settlement power market (that is, it only has the day-ahead market and the spot market is not modeled) complete with a realistic transmission system, and incorporates the actions of an independent system operator (ISO), load-serving entities, (LSEs)[7] and generation companies. The program allows for a user-specified stopping rule, such as a maximum number of days. The ISO oversees the grid and is responsible for operating the market as efficiently as possible. The LSEs try to secure power for their retail customers. They enter the market by bidding in their demand, which can be price elastic. The generators try to maximize their daily profits by bidding in linear supply functions. During each day, generator i chooses a supply function subject to its capacity constraints and reports it to the ISO for use in all 24 hours of the following day. After receiving demand and supply bids from the market participants, the ISO solves hourly DC optimal power flow problems where the objective is to minimize dispatch costs subject to supply, demand and transmission constraints. As part of the solution process, the ISO also obtains the "locational marginal prices" (LMPs) for each node in the system.[8] After this problem is solved for each day, the ISO settles the day-ahead market and posts the market-clearing prices and power supply commitments for the next day. Given this feedback, each

---

[6]See Sun and Tesfatsion (2007, [87, 88]) for details. AMES is a free and open-source platform developed in Java, and can be downloaded at www2.econ.iastate.edu/tesfatsi/AMESMarketHome.htm.

[7]We use the terms LSE and retailer interchangeably.

[8]Very briefly, LMP at a node can be defined as the "least cost of serving 1 MWh of incremental demand at that node". More technically, it is the Lagrange multiplier at the LP problem associated with that specific constraint. For determining the LMPs for the ubiquitous 3-node network, see Lin (2005, [56]). This only requires Ohm's law and Kirchoff's current and voltage laws. Solution to systems with four or more nodes is more complicated, but Liu et al. (2009, [60, 59]) are very accessible resources.

generator calculates its profits and using a stochastic reinforcement learning algorithm (details to follow in the next section), chooses another supply function to bid in for the next day. It is assumed that there are no system disturbances such as generator failures, line outages and weather shocks. Hence there is no need for ancillary services and they are not modeled.

In the next section, we provide the details of the model incorporated into AMES, the learning algorithm and the specifics of the hypothetical market we will be working on.

3.3   Model and Experimental Design

We have a set of three retailers $\{L_1, L_2, L_3\}$ and five capacity-constrained generators $\{G_1, ..., G_5\}$ interacting on the following network:[9]
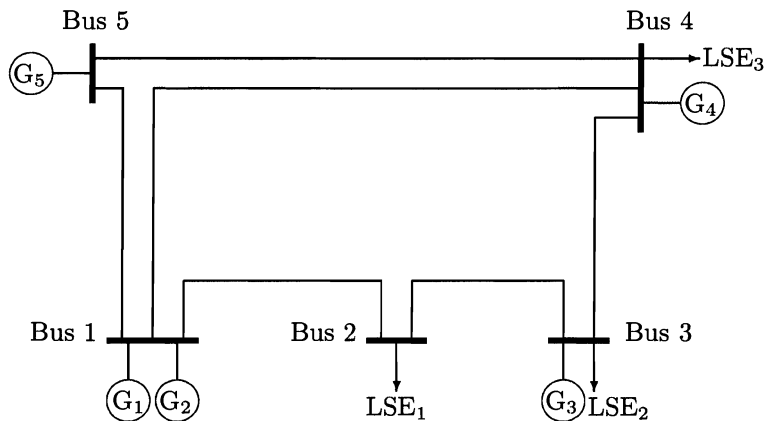
Generators' cost functions are given by

$$C_i(q_{G_i}) = F_i + a_i q_{G_i} + b_i q_{G_i}^2 \quad (\$/\text{MWh}),$$

[9]This network topology is based on the training documents used in the PJM market and ISO New England, hence it is safe to assume that it is a very good representation of a real market. One can also interpret each bus as a "zone", consisting of thousands of buses. See Lally (2002, [53]) for details.

where the parameters F_i, a_i, b_i and the upper capacity limits[10] are as given in table 3.1. Each generator is endowed with an *action domain*, denoted as AD_{G_i}, which

Table 3.1 : Generators' True Marginal Cost Functions and Capacities

	F_i	a	b	$q_{G_i}^{max}$
G_1	50	30	0.02	110
G_2	20	35	0.03	100
G_3	2000	15	0.008	520
G_4	200	25	0.009	200
G_5	1500	10	0.002	600

consists of a finite number of supply functions s_{G_i}, and includes G_i's true marginal cost function $a_i + 2b_i q_{G_i}$ (figure 3.1). The cardinalities of the sets AD_{G_i} are the same for all generators.

3.3.1 Generator Learning

In AMES, only the generators can learn from their past actions. The learning algorithm is based on the *stochastic reinforcement learning algorithm* developed by Ido Erev and Alvin Roth (1995, 1998; [25, 26]). According to this algorithm, a generator calculates his profit everyday after the day ahead market is settled. If the outcome is relatively good he increases the probability of choosing that particular supply function again for the next day, if the outcome is relatively bad, decreases the probability of choosing the same supply function. Let $m = 1, ..., M_i$ index the set AD_i and $\rho_{im}(1)$ denote the *initial propensity* of generator i to choose action m. Before day 1, each generator assigns equal probabilities to every action $m \in AD_i$. Let $\rho_{im}(d)$ stand for

---

[10]It is assumed that there are no binding "lower sustainable limits" hence $q_i^{min} = 0$ for all i.
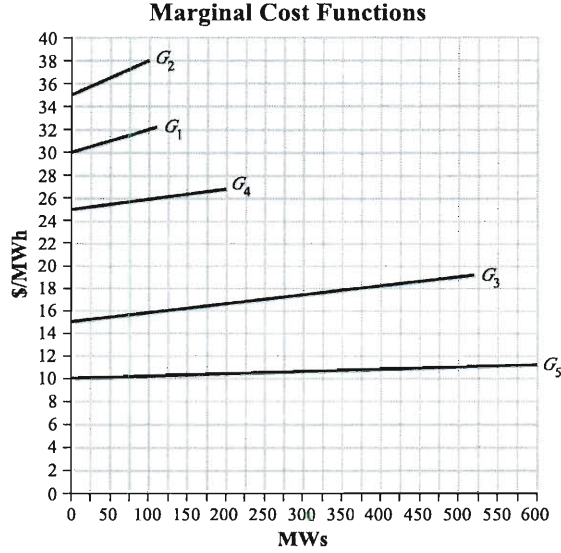
Figure 3.1 : Generators' true marginal cost functions.

the *current propensity* of generator i to choose $m \in AD_i$ on day d, for use in day $d + 1$. Then, the *choice probabilities* generator i uses to choose a supply function for day d are constructed by

$$\pi_{im}(d) = \frac{e^{\rho_{im}(d)/T_i}}{\sum_{j=1}^{M_i} e^{\rho_{ij}(d)/T_i}}, \tag{3.1}$$

where T_i is the *temperature parameter*, which determines how much weight the generator gives to propensity values in determining the choice probabilities. Note that as $T_i \to \infty$, $\pi_{im}(d) \to 1/M_i$, which means that the propensity values play no role in determining the probabilities. On the other extreme, as $T_i \to 0$, ρ_{im} become peaked around actions which have the highest propensity values.

At the end of each day, generator i updates the current propensities it assigns to action m to be used to calculate the action choice probabilities for day $d + 1$ as follows: Let m' be the action selected on day d and $P_{im'}(d)$ be the profit guaranteed

for day $d+1$ after the settlement. Then, for each $m \in AD_i$

$$\rho_{im}(d+1) = [1 - r_i]\rho_{im}(d) + \text{Response}_{im}(d), \tag{3.2}$$

where

$$\text{Response}_{im}(d) = \begin{cases} [1 - e_i] \cdot P_{im'}(d) & \text{if} \quad m = m' \\ e_i \cdot \rho_{im}(d)/[M_i - 1] & \text{if} \quad m \neq m', \end{cases} \tag{3.3}$$

Here $r_i \in [0, 1]$ is the *recency parameter*, which limits the growth of ρ over time; as r_i grows, past experience is slowly "forgotten". $e_i \in [0, 1)$ is the *experimentation parameter*, which facilitates experimentation with different actions while the generators are learning during the early stages. This parameter guarantees that the supply functions which resulted in good payoffs will not be the only ones chosen by a generator; he will bid in similar supply functions more often. Li and Tesfatsion (2010, [54]) carried out detailed analyses of the effects these parameters' values on generator earnings. They found the sweet-spot values of these learning parameters to be $r = 0.04$ and $e = 0.96$. We also use the same parameter values in our simulations. Choice of the initial propensity parameter and T_i are also critical: For the default values given in AMES, no learning occurs. In our simulations we set them such that $\rho_{im}(1)/T_i = 100$.

3.3.2 LSE bids

LSE bids are the combination of an hourly fixed quantity $q_{L_i}^F$ as given in figure 3.2, and a price-sensitive (inverse) demand function given by

$$p(q_{L_i}^S) = c_i(h, d) - 2d_i(h, d)q_{L_i}^2 \tag{3.4}$$
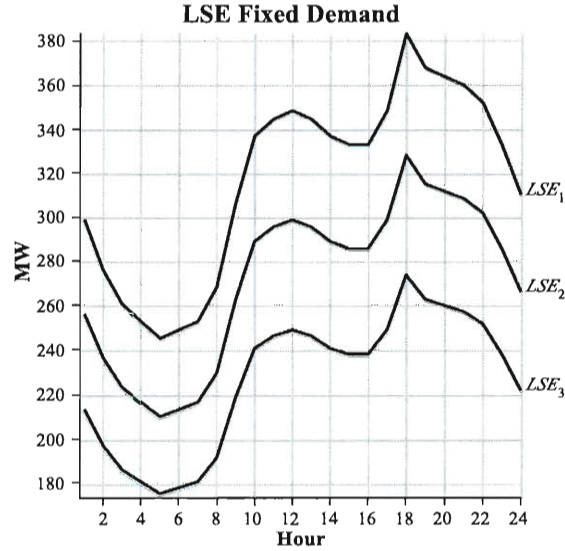
Figure 3.2 : Hourly fixed demand of the retailers.

for a given hour h and day d^{11}, where $q_{L_i}^S \in [0, q_{L_i}^{MAX}]$. Here, $q_{L_i}^{MAX}$ denotes the maximum potential price-sensitive demand and it is bounded from above by the benchmark-case value of $q_{L_i}^F$. These bids apply throughout the entire period of the simulation.

Following [55], we define the ratio R_i for retailer L_i to assess the effects of changes in demand sensitivity in LSE bids. For a given hour and day, R_i is defined as

$$R_i = \frac{q_{L_i}^{MAX}}{q_{L_i}^{MAX} + q_{L_i}^F},$$

(3.5)

where the numerator is as defined above, and the denominator is maximum total potential demand for L_i. For our analyses, we start with $q_{L_i}^{MAX} = 0$ and, keeping the denominator constant, we increase the numerator in steps, the first one being 10% of

---

[11] We drop h and d where no confusion arises.

$q_{L_i}^F$. We then keep increasing $q_{L_i}^{MAX}$ until we hit its upper limit, making it equal to 110% of the previous value at each step. At the end, we obtain a common set of R values for each retailer, ranging from 0 to 1. Clearly, as R increases, the retailers will be able to shed some of their demand in response to high prices. We set the stopping rule to 100 days.

The results of our simulations are in the next section.

3.4 Results

In our first set of simulations, we set the transmission line limits high enough to eliminate the possibility of transmission congestion.[12] Then, we ran the same set of simulations with the line capacities as given in table 3.2[13], which resulted in significant transmission constraints.

Table 3.2 : Transmission Line Limits

Line	From node	To node	Limit (MWs)	Reactance (Ω)
1	1	2	250	0.0281
2	1	4	150	0.0304
3	1	5	400	0.0064
4	2	3	350	0.0108
5	3	4	240	0.0297
6	4	5	240	0.0297

We first give the benchmark values of LMPs in figures B.1a and 3.3b, where the

---

[12]If there is no transmission congestion, LMPs at every node are going to be the same. With congestion, prices will differ, with the possibility of a unique price for each node. The outcomes may be quite counterintuitive: See Oren et. al (1995, [71]) and Wu et. al (1996, [102]) for some interesting examples.

[13]See Lally (2002, [53]), page 8. In table 3.2, *reactance* is the opposition to charge in an AC circuit. It is the complex part in the equation of the impedance.

generators do not bid strategically and the retailers only bid fixed quantities $(R = 0)$.

(a) No congestion.

(b) Network is congested.
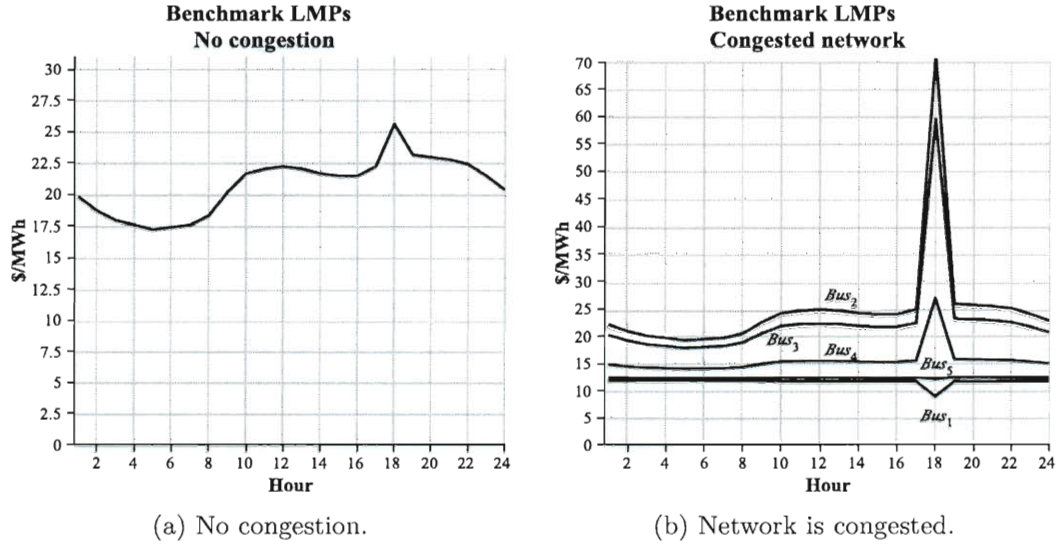
Figure 3.3 : Benchmark LMPs.

Next, we look at LMPs under both the transmission-constrained and unconstrained cases. We chose hour 18:00 since it corresponds to the time of the day when demand is highest. As expected, LMPs decrease as the retailers' ability to withhold demand, measured by R, increases (figures 3.4a and 3.4b).[14]

Figures 3.4a and 3.4b also suggest that the level of LMP is not a monotone function of the length of interaction in this market. Fixing R, we do not see a pattern of convergence in power prices as the length of interaction increases (figure 3.5). The variation in prices in figure 3.5 is a result of the learning algorithm: The generators continuously update their supply functions in an effort to increase their profits, which

---

[14]For figure 3.4b, we consider bus 2, since there is no generation located there and it hosts the largest retailer.

| (a) No congestion. | (b) Network is congested. |
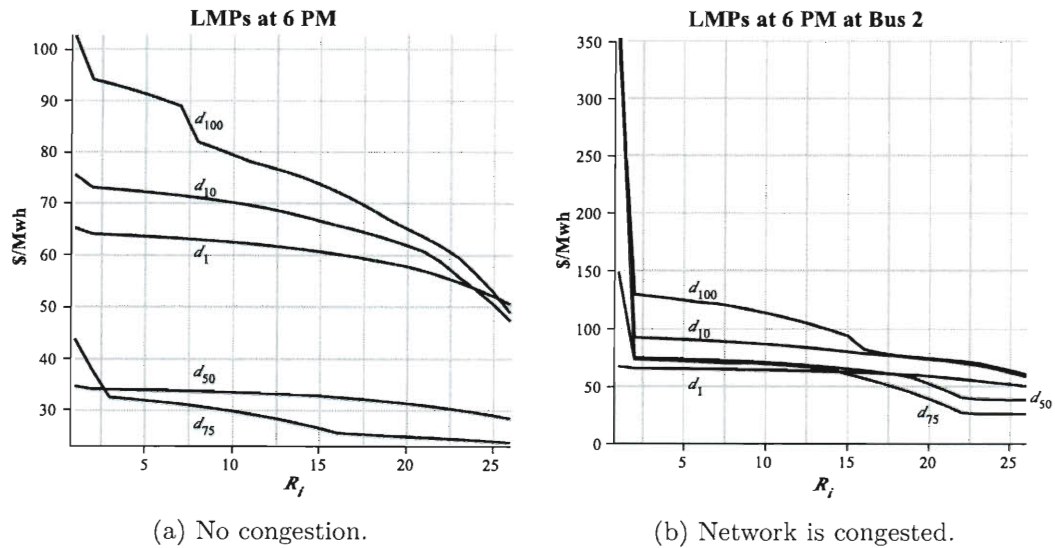
Figure 3.4 : LMPs as a function of R.

causes the spikes when R is low. As R increases, however, we see fewer spikes, but no convergence in prices. Figures 3.6a and 3.6b show that the prices are volatile, but volatility is decreasing significantly as R increases.

The United States Department of Justice, as well as many regulatory agencies in other countries, use the Herfindahl-Hirschman Index (HHI)[15] to measure market concentration, including concentration in power markets. Using the IEEE 30-bus network[16] with 9 generators and 21 retailers, we ran another set of simulations to show that the HHI is not an appropriate measure in power markets.

We first consider the retail side of the market and interpret the fixed demand of

---

[15]Given n firms with percentage market shares s_i, $i = 1, ..., n$, the HHI is $\sum_i s_i^2$. According to section §5.3 of the Horizontal Merger Guidelines of the Department of Justice, a market where the HHI is below 1,500 is considered unconcentrated and a market with HHI above 2500 is highly concentrated. For unconcentrated markets, "[m]ergers ... are unlikely to have adverse competitive effects and ordinarily require no further analysis." For HHI falling in between, the market is considered moderately concentrated. See [94] for details.

[16]Details of the network topology can be found at Shahidehpour et. al (2002, [83]) on page 478.
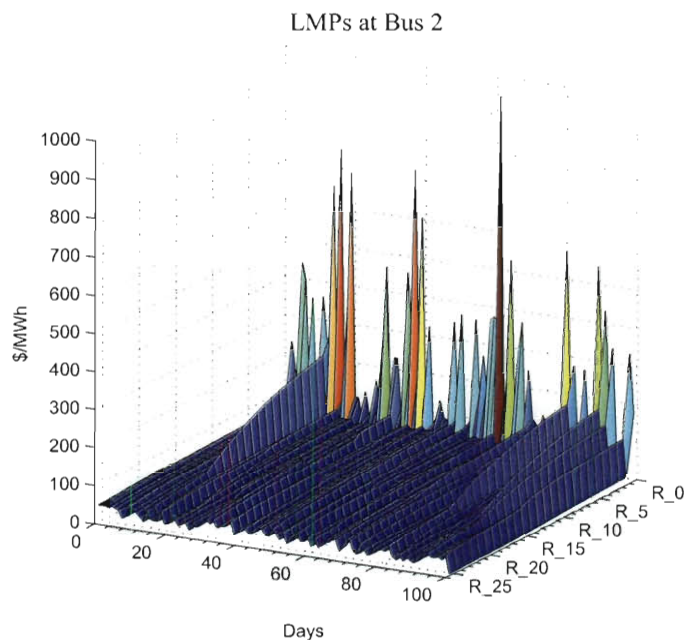
LMPs at Bus 2

Figure 3.5 : LMPs as function of R across days at bus 2 under transmission congestion.

the retailers as their capacities.

Under the benchmark case (Case 1), where the generators bid their true marginal cost functions and the demand is fixed, the total daily surplus of the retailers is $20.

We then took line 1-2 offline (Case 2). As a result of the increase in LMPs due to this constraint, the total retailer surplus jumped to 155. Retailers 2 and 3 are the ones who benefited the most from this line outage.

Given the data in table 3.3, the HHI is 1073.5, which is considered an unconcentrated market and a merger between retailers 2 and 3 would increase the HHI only by 40.3 points, to 1113.8, which is still considered unproblematic. Retailer 3 has about one-fourth of the market. Retailer 2 is tiny, but it may have talented traders who

LMP Volatility
No congestion

LMP Volatility at Bus 2
Contested network

(a) No congestion.

(b) Network is congested.
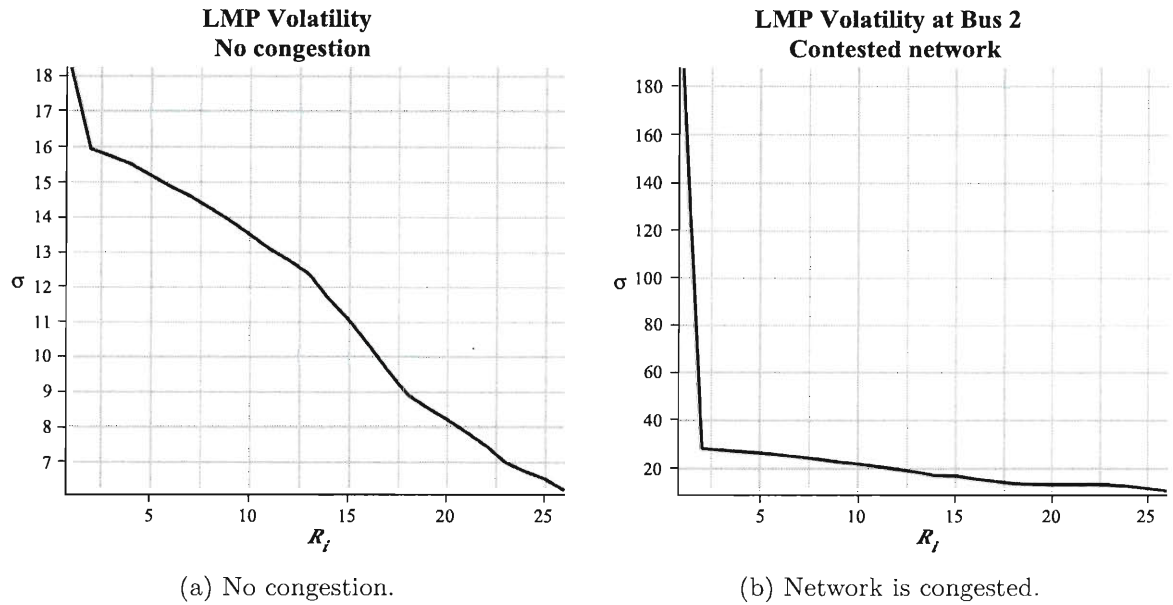
Figure 3.6 : LMP volatility as a function of R.

know how to game the market and use retailer 3's (demand) capacity as a leverage.[17] As a result, the acquisition of retailer 2 by retailer 3 may be a harmful combination. Under case 2, this merged firm gets almost all of the surplus, even though it has only 1/4th of the market.

The problem clearly lies in the fact that the HHI statistic does not take into account the locations of the players, which is a very critical variable in power markets. When we look at the generation side of the market, we see that even a small generator can exercise market power as a result of transmission congestion.

In the 30-bus test case, the HHI in the generation market (figure 3.4) is 1,408 and the HHI index again does not raise a red flag. However, under congestion, generators

---

[17]Traders can schedule trades to artificially create congestion. Before clearing the day-ahead market, the ISO foresees the problems these trades may create, and pays the traders to back up or modify their trades to relieve the congestion.

Table 3.3 : Retail Market

Retailer	Capacity	Case 1 surplus	Case 2 surplus	% change
1	430.99	0.00	0.00	-
2	47.65	9.28	121.15	1205%
3	1342.56	1.97	28.87	1366 %
4	679.22	0.00	0.00	-
5	452.8	0.00	0.00	-
6	595.8	0.14	0.00	-100%
7	115.2	0.41	0.07	-84%
8	222.44	0.96	1.54	61%
9	123.14	0.86	0.84	-2%
10	162.86	0.78	0.60	-23%
11	69.55	0.70	0.41	-41%
12	178.74	0.49	0.14	-72%
13	63.54	0.74	0.36	-51%
14	188.71	0.62	0.22	-64%
15	43.68	0.57	0.17	-69%
16	347.59	0.42	0.07	-83%
17	63.54	0.63	0.32	-49%
18	172.81	0.45	0.11	-77%
19	69.55	0.33	0.03	-92%
20	47.65	0.26	0.01	-97%
21	210.52	0.26	0.01	-97%
Total	5628.54	19.87	154.92	679%

8 and 9 can exercise market power and significantly increase their profits.

3.5 Conclusions and Future Research

We analyzed the effects of increased power demand price sensitivity on the level and volatility of power prices using agent-based modeling. We found that, as the price responsiveness of power demand increases, prices and their volatility decrease.

Even though we had to depart from equilibrium analysis, and hence the benefit of drawing general conclusions, this approach has many advantages. First of all, we can explicitly and easily incorporate a power grid, with arbitrary level of detail, into the model. Settlement rules specific to each market can also be incorporated

Table 3.4 : Generation Market

Generator	Capacity	Case 1 profit	Case 2 profit	% change
1	100	948,984	240,912	-74.6%
2	80	1,236,598	3,247,949	162.7%
3	50	1,722,341	2,599,050	50.9%
4	50	1,591,339	2,192,603	37.8%
5	20	109,009	220,522	102.3%
6	70	723,366	889,683	23%
7	60	1,008,295	1,242,993	23.3%
8	20	58,604	140,929	140.5%
9	20	16,269	54,889	237.4%
Total	470	7,414,803	10,829,529	46 %

without too much difficulty. We do not have to dispense with these just because of the computational burden and complexities they bring, thanks to cheap computing power. Second, we do not have to assume that the players are perfectly rational which, most likely, is not the case.[18] We believe that the observed behavior of the agents is a more realistic representation of the environment we are trying to model.

We did our analysis using the AMES testbed. While AMES is a flexible platform where the user can work with virtually any network topology, it has some limitations. First, only the generators can learn from their actions; the most the retailers can do is to bid in more price-sensitive functions. Considering the greater flexibility of consumption that will result from the widespread implementation of smart grids, extending the platform to one where the demand side can also learn would be worthwhile. Second, there is no justification for choosing the specific learning algorithm used. In the Roth-Erev learning algorithm, agents only update their own action choice probabilities. In the first paper of this dissertation, we conjectured that generators

[18]Erev and Roth (1998, [26]) discuss the implications of their learning algorithm "for developing a low-rationality, cognitive game theory".

can exercise market power by analyzing past market data and bid accordingly. Empirical evidence from the ERCOT market -partially- agreed with our conjecture. The model could be altered to incorporate this learning algorithm.

We also provided the results of 30-bus runs to show that the HHI is not a suitable metric to measure market power in power markets. The problem with the HHI is that it does not take the agents' locations into account. Any index to be used in power markets should take location into account, and it will probably require analyzing the participants on a node-by-node basis. We are currently working on this problem.

AMES implements locational marginal pricing as the congestion management mechanism. In their recent paper, Li and Tesfatsion (2010, [55]) analyze the distribution of surplus in a hypothetical power market using AMES. They find that the ISO surplus increases as generators behave less competitively. This result has the potential to create some incentive problems. As revenue-neutral entities, ISOs distribute their surplus, and the current trend is to allocate it to the transmission rights holders. The key conclusion of their paper is that this surplus "should be used pro-actively to mitigate the conditions encouraging generator capacity withholding". This surplus accrued to the ISO is a direct result of using LMP as the congestion management mechanism. While it makes perfect sense to implement LMP in AMES since LMP is the most widespread used mechanism today, even backed by the Federal Energy Regulatory Commission, there are serious criticisms for its use as a congestion management mechanism, see for example Rosenberg (2000 and 2004; [76, 77]). The most important one for our purposes is that the revenues collected (ISO's surplus) far exceed the actual redispatch costs to relieve transmission congestion. Rosenberg (2004, [77]) proposes what he calls the "Compensation/Charge Method", where the disparity between redispatch costs and congestion revenues disappears. Solution con-

cepts from cooperative microeconomics, such as the nucleolus and Shapley value can also be applied.[19] This is another problem we are currently working on.

---

[19]See for example Tan and Lie (2002, [89]), Yang et al. (2003, [105]), Du et al. (2006, [85]), Zolezzi and Rudnick (2007, [108]), Xiao and Li (2009, [104]), Azevedo et al. (2009, [5]).

Chapter 4

Epilogue

In this dissertation, we presented three essays on power system economics.

The first essay proposed a new method for modeling competition in wholesale power markets, namely, by approximating the supply functions of the competitors with cubic splines. We argued that this method is preferable to approximation by linear or piecewise-affine functions, which have been the main approaches to date. We applied our method to the firms competing in the Texas market and found that the firms with relatively big stakes in the market generally bid close to what economic theory would predict. We also showed that, more often than not, we will observe that the marginal revenue functions of the firms will have increasing segments which may lead to multiple profit-maximizing optima for a firm.

In the second essay, we focused on the relationship between the forward and spot prices of power. We departed from the general approach in the economics literature and modeled the players as risk-averse agents. We analyzed the effects of several variables such as the degree of risk aversion and power demand variability on the forward bias, defined as the difference between the forward and spot prices of power. The results of our simulations imply that as power demand becomes more variable, for a given level of expected demand, the bias increases in absolute value. This is consistent with the classical literature in finance. We obtained similar results as the agents became more risk-averse. We also showed that the two prices converge to each other as the generators bid closer to their marginal costs. In addition to this

convergence result, the bias converges to zero, as we increase the number of players.

The third essay analyzed, in an agent-based setting, the effects of increased power demand price sensitivity on the level and volatility of locational marginal prices. As intuition would suggest, we found that, as power demand becomes more price-responsive, both the level and the volatility of the prices decrease. We also argued that the widely-used index for measuring market concentration, the celebrated Herfindahl-Hirscman Index, is not a suitable metric for power markets.

We made many simplifying assumptions to do our analyses. In the first essay, we disregarded some physical constraints power generators face in real life; such as their high and low operating limits and ramp rate constraints. We also did not analyze bidding behavior when the transmission grid was congested. The effects of these assumptions remain to be investigated. Also, it is an open question as to whether the bidding behavior of the firms would change if they were allowed to bid continuous functions.

In the second essay, we assumed a linear functional form for the marginal cost functions for the sake of mathematical tractability. We also assumed that power demand was distributed normally. As a result of these assumptions, our model could not explain the fact that the forward bias can also be positive. How to relax these assumptions in a supply function equilibrium setting, however, is an open question.

As we argued in the third essay, an important open question is whether we can come up with a more suitable metric than the Herfindahl-Hirschman Index for measuring concentration in power markets. Any metric to replace the HHI should take into account the locations of the players. Tools from spatial statistics can be applied to investigate this issue further.

Appendix A

Construction of Marginal Cost Functions

Data on plants' capacity, fuel source and fuel efficiency[1] are publicly available at Energy Information Administration and Environmental Protection Agency websites[2]. We obtain the price data from EIA's website.

Each firm in our sample owns several plants. Marginal cost functions will be constructed by multiplying a firm's generation fleet's individual plant's heat rate with the relevant fuel price, and stacking them from the lowest to the highest cost generator after adding the relevant operating and management expenses to fuel cost[3]. ERCOT publishes data on which generators are available and online during a given hour, and only those generators are used in the marginal cost calculations.

We use EIA form 923 for natural gas and coal prices. For natural gas plants we assume 5 cents commodity charge per MMBtu and 1.5% pipeline fuel charge per MMBtu[4]. Coal plants need sulfur dioxide permits. SO_2 permit rates for each power plant in operation are given in the EPA's NEEDS database. We assume that the cost of a permit is $50/ton.

---

[1]Fuel efficiency is measured in (average) heat rate. Heat rate is defined as the number of British Thermal Units of heat required to produce a kilowatt-hour of energy.

[2]We thank Jennifer Rosthal for sharing her compiled dataset.

[3]More details are given in the following companion manual: www.owlnet.rice.edu/~inal/energy/ManualForEnergyEconStudent.pdf

[4]Private conversation with two natural gas traders. [45], [46] and [84] assume a charge of $0.10/MMBtu.

Appendix B

Reason for Choosing the Fritsch-Carlson Method and a Note on the Code

B.1 Why We Chose Fritsch-Carlson Method

Fritsch and Carlson (1980) argue that their method stands out "by its efficiency, in terms of the time required to determine the interpolant, storage required to represent it, and/or time required to evaluate it". While today's computers are much faster compared to the ones the authors had, we have more data points and efficiency of an algorithm is still desirable. Furthermore, Kvasov (2000, [52]) shows that the error of approximation with this method remains small (see his theorem 4.2). Moreover, the Fritsch-Carlson method is local, that is, a change in the data will only affect the relevant interval, and not the whole polynomial. This attribute of the Fritsch-Carlson method is especially useful, since we further smooth the residual demand function after fitting a spline to it.

Other methods, most notably the ones by Pruess (1979, [74]) and McAllister and Roulier (1981, [65]), could also be used. The first one uses (up to two) additional knots per interval of data and it requires, "a nonlinear iteration to determine the locations of the additional breakpoints" (see [33]). The latter, in addition to being monotonic, also preserves convexity of the data. It uses at most one additional breakpoint. Nevertheless, inserting additional data means increased storage requirements and more time for evaluation. Also, as described by the authors on page 340, their method

suffers from some pathological phenomena, in particular that "small changes in the slopes used for the determination of the various cases may cause radical variations in the resulting spline".

B.2   A Note on the Code for Simulations

As touched upon above, inverting the residual demand explicitly to get a function of quantity is the hardest and most time consuming part in our code. Instead of inverting the function explicitly, we could sample n (p, q) pairs from $RD : P \rightarrow Q$, store them in an $n \times 2$ matrix, transpose it and fit another spline to the resulting $2 \times n$ matrix of (q, p) val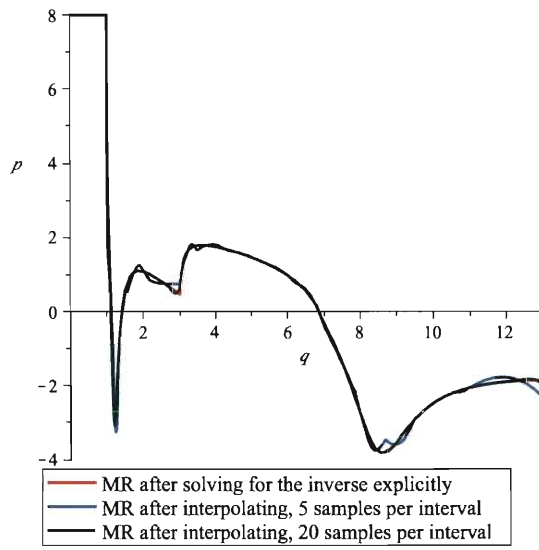ues. While "taking the inverse" using this method is much easier, it only exacerbates the increasing marginal revenue problem mentioned in section 1.4.3.

To illustrate this alternative approach, on the next page, 3 graphs are plotted, where, respectively, 5, 10 and 20 equally spaced sample points from each of $[p_i, p_{i+1}]$, $i = 1..4$ are taken. One can see that as the number of sampled points increases, the marginal revenue curve behaves nicer. On the following page, we superimpose the residual demand curves (left column) and the marginal revenue curves. Even though the residual demand curves produced by both methods almost perfectly match, even when 5 points are sampled, we cannot say the same for the marginal revenue curves (on the right). Despite the fact that the marginal revenue curves match perfectly as we sample more points, we will continue taking the inverse of the residual demand function explicitly because this second method is even more resource-consuming.
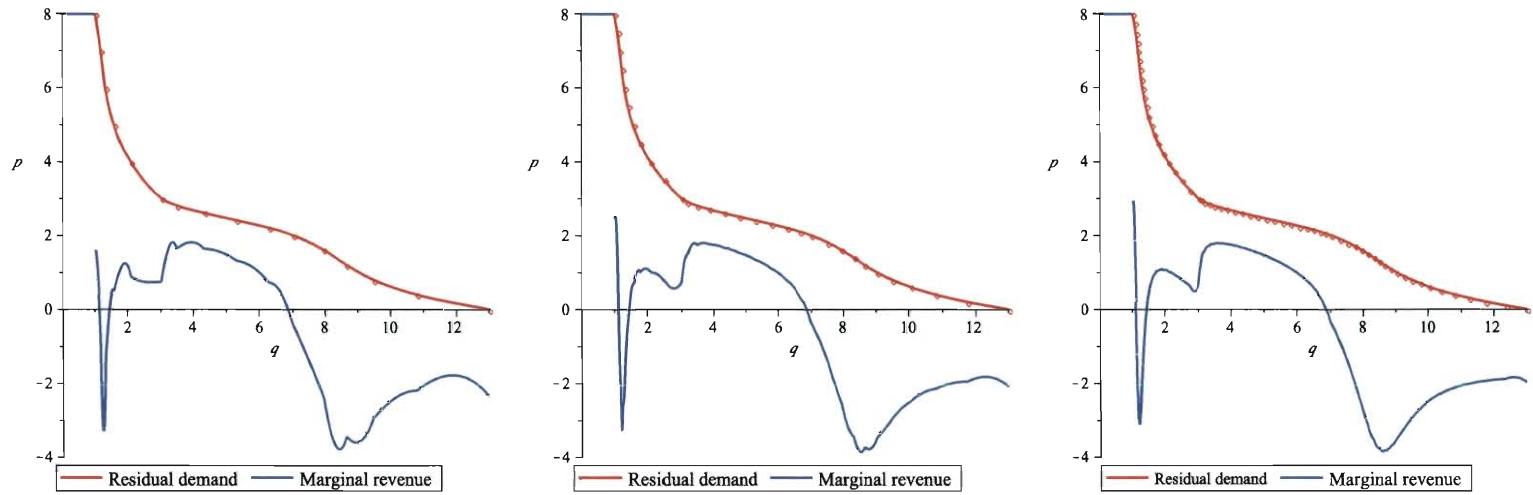
(a) 5 sample points taken on each interval (b) 10 sample points taken on each interval (c) 20 sample points taken on each interval

Figure B.1 : Interpolation by sampling points.

Inverse by explicitly solving

Inverse by interpolating, with 5 sample points per interval

MR after solving for the inverse explicitly

MR after interpolating, 5 samples per interval

MR after interpolating, 20 samples per interval
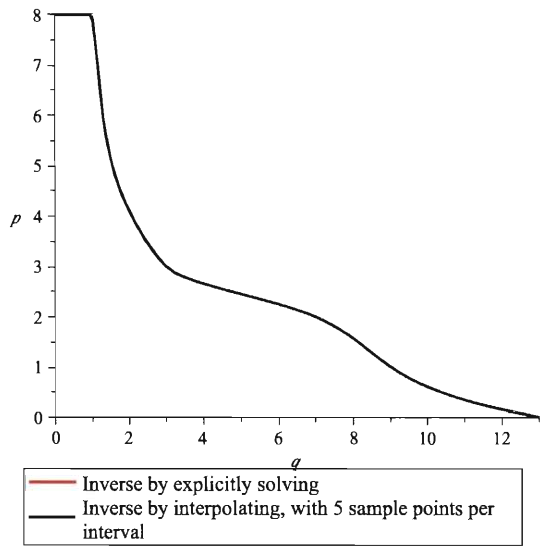
Appendix C

Variance-covariance terms and the related derivatives

Let X_{-i} denote $\sum_{j \neq i} x_{G_i}$ or $\sum_{j \neq i} x_{R_i}$ depending on the context (generator vs. retailer). We will assume that the retail price p_R is a constant multiple of the spot market price p: $p_R = (1 + r)p$

Let $a, b, c, d \in \mathbb{R}$ and W, X, Y and Z be random variables. Using the properties of covariance, namely, $\text{cov}(a + X, b + Y) = \text{cov}(X, Y)$ and

$$\text{cov}(aW + bX, cY + dZ) = ac\,\text{cov}(W, Y) + ad\,\text{cov}(W, Z) + bc\,\text{cov}(X, Y) + bd\,\text{cov}(X, Z)$$

we obtain the following, which we will use to solve for the optimum forward positions of the market participants:

C.1 Retailer

In what follows, we assume that $\partial Q / \partial x_{R_i} = 0$ since Q is the total retail (and hence market, in our case) demand which is not affected by the retailers' forward positions. Also, s_{R_i} denotes retailer i's share of the market demand Q, where $\sum_{i=1}^{n} s_{R_i} = 1$.

$$\mathrm{var}(pq_{R_i}) = \mathrm{var}(ps_{R_i}Q) = E\left[\frac{az - X + tQ}{z}s_{R_i}Q - E\left(\frac{az - X + tQ}{z}s_{R_i}Q\right)\right]^2$$

$$= E\left[\frac{s_{R_i}(az - X)}{z}(Q - E(Q)) + \frac{s_{R_i}t}{z}(Q^2 - E(Q^2))\right]^2$$

$$= \frac{s_{R_i}^2(az - X)^2}{z^2}\mathrm{var}(Q) + \left(\frac{s_{R_i}t}{z}\right)^2\mathrm{var}(Q^2) + \frac{2s_{R_i}^2t(az - X)}{z^2}\mathrm{cov}(Q, Q^2)$$

$$\implies \frac{\partial \mathrm{var}(pq_{R_i})}{\partial x_{R_i}} = -\frac{2s_{R_i}^2(az - X)\left(1 + \sum\limits_{j \neq i}^{n}\frac{\partial x_{R_j}}{\partial x_{R_i}}\right)}{z^2}\mathrm{var}(Q) - \frac{2s_{R_i}^2t\left(1 + \sum\limits_{j \neq i}^{n}\frac{\partial x_{R_j}}{\partial x_{R_i}}\right)}{z^2}\mathrm{cov}(Q^2, Q)$$

$$= -\frac{2s_{R_i}^2}{z^2}\left(1 + \sum\limits_{j \neq i}^{n}\frac{\partial x_{R_j}}{\partial x_{R_i}}\right)\left[(az - X)\mathrm{var}(Q) + t\mathrm{cov}(Q^2, Q)\right] \tag{C.1}$$

$$\mathrm{var}(p) = \frac{t^2}{z^2}var(Q) \implies \frac{\partial \mathrm{var}(p)}{\partial x_{R_i}} = 0 \tag{C.2}$$

$$\mathrm{cov}(pq_{R_i}, p) = E\left[\left(s_{R_i}pQ - E(s_{R_i}pQ)\right)\left(p - E(p)\right)\right]$$

$$= E\left[\left(s_{R_i}\frac{az - X}{z}(Q - E(Q)) + \frac{s_{R_i}t}{z}(Q^2 - E(Q^2))\right)\left(\frac{t}{z}(Q - E(Q))\right)\right]$$

$$= \frac{s_{R_i}t(az - X)}{z^2}\mathrm{var}(Q) + \frac{s_{R_i}t^2}{z^2}\mathrm{cov}(Q^2, Q)$$

$$\implies \frac{\partial \mathrm{cov}(pq_{R_i}, p)}{\partial x_{R_i}} = -\frac{s_{R_i}t}{z^2}\left(1 + \sum\limits_{j \neq i}^{n}\frac{\partial x_{R_j}}{\partial x_{R_i}}\right)\mathrm{var}(Q) \tag{C.3}$$

C.2 Generator

$$\text{cov}(pq_{G_i}, p) = \text{cov}\left(\frac{t}{mz}Q^2 + (\frac{x_{G_i}}{z} + \frac{a}{m} - \frac{2X}{mz})Q, \frac{t}{z}Q\right)$$

$$= \frac{t^2}{mz^2}\text{cov}(Q^2, Q) + \frac{t}{z}\left(\frac{x_{G_i}}{z} + \frac{a}{m} - \frac{2X}{mz}\right)\text{var}(Q)$$

$$\implies \frac{\partial\text{cov}(pq_{G_i}, p)}{\partial x_{G_i}} = \frac{t\left(m - 2 - 2\sum\limits_{\substack{j \neq i}}^{m}\frac{\partial x_{G_j}}{\partial x_{G_i}}\right)}{mz^2}\text{var}(Q) \qquad (\text{C.4})$$

$$\text{cov}(pq_{G_i}, q_{G_i}^2) = \text{cov}\left(\frac{t}{mz}Q^2 + (\frac{x_{G_i}}{z} + \frac{a}{m} - \frac{2X}{mz})Q, \frac{Q^2}{m^2} + (\frac{2x_{G_i}}{tm} - \frac{2X}{tm^2})Q\right)$$

$$= \frac{t}{m^3 z}\text{var}(Q^2) + \left[\frac{1}{mz}(\frac{2x_{G_i}}{m} - \frac{2X}{m^2}) + \frac{1}{m^2}(\frac{x_{G_i}}{z} + \frac{a}{m} - \frac{2X}{mz})\right]\text{cov}(Q, Q^2)$$

$$+ \left(\frac{x_{G_i}}{z} + \frac{a}{m} - \frac{2X}{mz}\right)\left(\frac{2x_{G_i}}{tm} - \frac{2X}{tm^2}\right)\text{var}(Q)$$

$$\implies \frac{\partial\text{cov}(pq_{G_i}, q_{G_i}^2)}{\partial x_{G_i}} = \left[\frac{3m - 4 - 4\sum\limits_{\substack{j \neq i}}^{n}\frac{\partial x_{R_j}}{\partial x_{R_i}}}{m^3 z}\right]\text{cov}(Q, Q^2)$$

$$+ \left[\left(\frac{m - 2 - 2\sum\limits_{\substack{j \neq i}}^{m}\frac{\partial x_{G_j}}{\partial x_{G_i}}}{mz}\right)\left(\frac{2x(m - 1) - 2X_{-i}}{tm^2}\right)\right.$$

$$+ \left.\left(\frac{x(m - 2) - 2X_{-i} + az}{mz}\right)\left(\frac{2(m - 1) - 2\sum\limits_{\substack{j \neq i}}^{m}\frac{\partial x_{G_j}}{\partial x_{G_i}}}{tm^2}\right)\right]\text{var}(Q)$$

$$= \frac{3m - 4 - 4\sum\limits_{\substack{j \neq i}}^{m}\frac{\partial x_{G_j}}{\partial x_{G_i}}}{m^3 z}\text{cov}(Q, Q^2)$$

$$+ \left[\frac{4x(m - 1)(m - 2) + (8 - 6m)X_{-i} + x(8 - 6m)\sum\limits_{\substack{j \neq i}}^{m}\frac{\partial x_{G_j}}{\partial x_{G_i}} + 8X_{-i}\sum\limits_{\substack{j \neq i}}^{m}\frac{\partial x_{G_j}}{\partial x_{G_i}}}{m^3 tz}\right.$$

$$+ \left.\frac{2az\left(m - 1 - \sum\limits_{\substack{j \neq i}}^{m}\frac{\partial x_{G_j}}{\partial x_{G_i}}\right)}{m^3 tz}\right]\text{var}(Q) \qquad (\text{C.5})$$

$$\text{cov}(pq_{G_i}, q_{G_i}) = \text{cov}\left[\frac{t}{mz}Q^2 + \left(\frac{x_{G_i}}{z} + \frac{a}{m} - \frac{2X}{mz}\right)Q, \frac{Q}{m}\right]$$

$$= \frac{t}{m^2 z}\text{cov}(Q^2, Q) + \frac{1}{m}\left(\frac{x_{G_i}}{z} + \frac{a}{m} - \frac{2X}{mz}\right)\text{var}(Q)$$

$$\implies \frac{\partial \text{cov}(pq_{G_i}, q_{G_i})}{\partial x_{G_i}} = \frac{\text{var}(Q)}{m^2 z}\left(m - 2 - 2\sum_{j \neq i}^{m}\frac{\partial x_{G_j}}{\partial x_{G_i}}\right) \tag{C.6}$$

$$\text{cov}(p, q_{G_i}^2) = \text{cov}\left(\frac{t}{z}Q, \frac{Q^2}{m^2} + (\frac{2x_{G_i}}{tm} - \frac{2X}{tm^2})Q\right)$$

$$= \frac{t}{zm^2}\text{cov}(Q, Q^2) + \frac{1}{z}(\frac{2x_{G_i}}{m} - \frac{2X}{m^2})\text{var}(Q)$$

$$\implies \frac{\partial \text{cov}(p, q_{G_i}^2)}{\partial x_{G_i}} = \frac{2\text{var}(Q)}{m^2 z}\left(m - 1 - \sum_{j \neq i}^{m}\frac{\partial x_{G_j}}{\partial x_{G_i}}\right) \tag{C.7}$$

$$\text{cov}(p, q_{G_i}) = \text{cov}(\frac{t}{z}Q, \frac{Q}{m}) = \frac{t}{zm}\text{var}(Q)$$

$$\implies \frac{\partial \text{cov}(p, q_{G_i})}{\partial x_{G_i}} = 0 \tag{C.8}$$

$$\text{cov}(q_{G_i}^2, q_{G_i}) = \text{cov}\left[\frac{Q^2}{m^2} + \left(\frac{2x_{G_i}}{tm} - \frac{2X}{tm^2}\right)Q, \frac{Q}{m}\right]$$

$$= \frac{1}{m^3}\text{cov}(Q^2, Q) + \frac{1}{m}\left(\frac{2x_{G_i}}{tm} - \frac{2X}{tm^2}\right)\text{var}(Q)$$

$$\implies \frac{\partial \text{cov}(q_{G_i}^2, q_{G_i})}{\partial x_{G_i}} = \frac{2\text{var}(Q)}{m^3 t}\left(m - 1 - \sum_{j \neq i}^{m}\frac{\partial x_{G_j}}{\partial x_{G_i}}\right) \tag{C.9}$$

$$\text{var}(pq_{G_i}) = \text{var}\left[\frac{t}{mz}Q^2 + \frac{mx_{G_i} - 2X + az}{mz}Q + \left(\frac{az - X}{z}\right)\left(\frac{mx_{G_i} - X}{tm}\right)\right]$$

$$= \frac{t^2}{m^2z^2}\text{var}(Q^2) + \left[\frac{mx_{G_i} - 2X + az}{mz}\right]^2\text{var}(Q) + 2\left(\frac{mx_{G_i} - 2X + az}{mz}\right)\frac{t}{mz}\text{cov}(Q^2, Q)$$

$$\implies \frac{\partial\text{var}(pq_{G_i})}{\partial x_{G_i}} = 2\left[\frac{mx_{G_i} + az - 2X}{m^2z^2}\right]\left(m - 2 - 2\sum_{j \neq i}^{m}\frac{\partial x_{G_j}}{\partial x_{G_i}}\right)\text{var}(Q)$$

$$+ \frac{2t}{m^2z^2}\left(m - 2 - 2\sum_{j \neq i}^{m}\frac{\partial x_{G_j}}{\partial x_{G_i}}\right)\text{cov}(Q^2, Q)$$

$$= \frac{2\left(m - 2 - 2\sum_{j \neq i}^{m}\frac{\partial x_{G_j}}{\partial x_{G_i}}\right)}{m^2z^2}\left(x_{G_i}(m - 2) - 2X_{-i} + az\right)\text{var}(Q)$$

$$+ \frac{2t\left(m - 2 - 2\sum_{j \neq i}^{m}\frac{\partial x_{G_j}}{\partial x_{G_i}}\right)}{m^2z^2}\text{cov}(Q^2, Q) \tag{C.10}$$

$$\text{var}(p) = \frac{t^2}{z^2}\text{var}(Q) \implies \frac{\partial\text{var}(p)}{\partial x_{G_i}} = 0 \tag{C.11}$$

$$\text{var}(q_{G_i}^2) = \frac{1}{m^4}\text{var}(Q^2) + 4\left(\frac{mx_{G_i} - X}{tm^2}\right)^2\text{var}(Q) + \frac{4}{m^2}\left(\frac{mx_{G_i} - X}{m^2t}\right)\text{cov}(Q^2, Q)$$

$$\implies \frac{\partial\text{var}(q_{G_i}^2)}{\partial x_{G_i}} = 8\left(\frac{mx_{G_i} - X}{m^4t^2}\right)\left(m - 1 - \sum_{j \neq i}^{m}\frac{\partial x_{G_j}}{\partial x_{G_i}}\right)\text{var}(Q) + \frac{4}{m^4t}\left(m - 1 - \sum_{j \neq i}^{m}\frac{\partial x_{G_j}}{\partial x_{G_i}}\right)\text{cov}(Q^2, Q)$$

$$= \frac{4}{m^4t}\left(m - 1 - \sum_{j \neq i}^{m}\frac{\partial x_{G_j}}{\partial x_{G_i}}\right)\left[\frac{2}{t}\left(x_{G_i}(m - 1) - X_{-i}\right)\text{var}(Q) + \text{cov}(Q^2, Q)\right] \tag{C.12}$$

$$\text{var}(q_{G_i}) = \frac{1}{m^2}\text{var}(Q) \implies \frac{\partial\text{var}(Q)}{\partial x_{G_i}} = 0 \tag{C.13}$$

Appendix D

Second order conditions

D.1   Generator

The second derivative of equation (2.21) with respect to the forward position generator i is

$$
\begin{aligned}
=&\frac{2(1-m)u}{mtz} + \frac{1+u}{z} + -\frac{c}{m^2t^2}\left(m^2 - m(1+u) + u\right) \\
&-\lambda var(Q)\left[\frac{(m-2u)(m-2)}{m^2z^2} + \frac{t^2}{z^2} + \frac{c^2(m-u)(m-1)}{m^4t^2} - \frac{t(m-2)}{mz^2}\right. \\
&\left. -\frac{t(m-2u)}{mz^2} - \frac{c\left(m(2m-3) + (4-3m)u\right)}{m^3tz} + \frac{c(2m-1-u)}{m^2z}\right]
\end{aligned}
\tag{D.1}
$$

where u stands for $1 + \sum_{j\neq i}\frac{\partial x_{G_j}}{\partial x_{G_i}}$ and since $\sum_{j\neq i}\frac{\partial x_{G_j}}{\partial x_{G_i}} \in [-1, 0]$, $u \in [0, 1]$. Also, $z = m(m-1)\beta$ and $t = 1 + c(m-1)\beta$, as before. It is safe to assume that the generators will not bid in a supply function less steep than their marginal cost; hence $c\beta \in (0, 1]$. Plugging in these, we get

$$
\frac{-1}{m^4(m-1)^2\beta^2(1+c\beta m - c\beta)^2}\varphi\left(c, m, u, \beta, \lambda, var(Q)\right)
$$

where $\varphi\big(c, m, u, \beta, \lambda, var(Q)\big)$ is given by

$$\lambda var(Q) \Bigg[\big(m^6 - 4m^3 - 4m^5 + m^2 + 6m^4\big)(c\beta)^4 + \big(- 4m - 4u + 4m^2 + 4um^2\big)c\beta + 4u$$

$$+ \big(- m^2 + 9m^3 + 3um^2 - 3um^3 - m - 11m^4 - mu + um^4 + 4m^5\big)(c\beta)^3$$

$$+ \big(- 4m^3 + 4um^3 - 4m^2 + 2mu + 4m - 7um^2 + 4m^4 + u\big)(c\beta)^2 \Bigg]$$

$$+ \beta \Bigg[\big(- m^6 - m^6u - 3m^4 + um^3 + 3m^5 + 3m^5u - 3um^4 + m^3\big)(c\beta)^2$$

$$+ \big(- um^2 + um^3 - 5m^5 - m^5u - 3m^3 + m^6 + 7m^4 + um^4\big)(c\beta)$$

$$+ um^4 + m^3 - 3um^3 - m^4 + 2um^2 \Bigg] \tag{D.2}$$

The expression to the left of $\varphi(\cdot)$ is clearly negative. Then we have to check over what ranges of its parameters $\varphi(\cdot)$ is positive. If we let $\gamma = \lambda \cdot var(Q)$, we have five parameters to deal with. Fixing three of those, we can check the sign of $\varphi(\cdot)$ by varying the other two. Unfortunately, there are too many variables to deal with and since $\varphi(\cdot)$ is not well-behaved we are not able to draw general conclusions as to what its sign would be. While we found it generally positive, for very small values of demand variance, risk aversion parameter *and m*, we were able to find some examples (see figures D.1a and D.1b) where $\varphi(\cdot)$ will be negative. Hence, one needs to be careful when carrying out simulations. For example, for $m = 20$, regardless of the values of other parameters, a sufficient condition for non-negativity of $\varphi(\cdot)$ is that $\lambda \cdot var(Q) \geq 1.3$. In our simulations m is large enough and we do not observe negative values for $\varphi(\cdot)$.
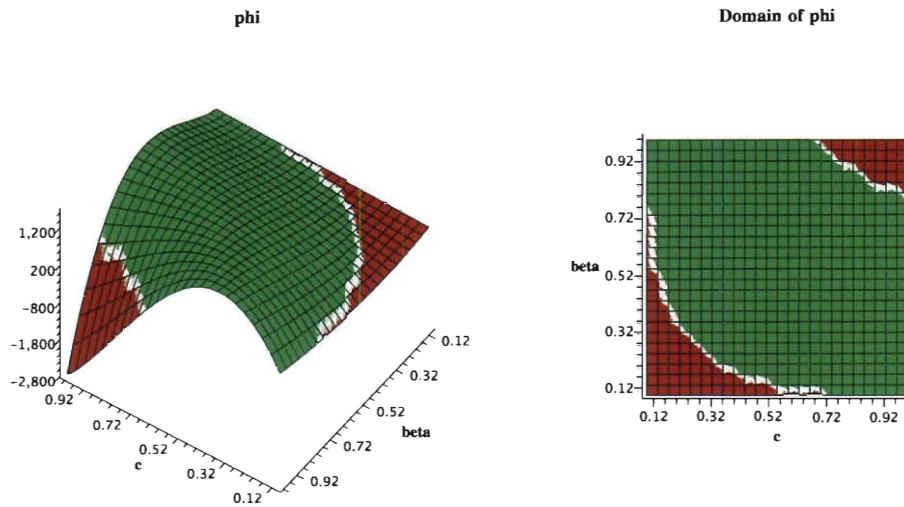
(a) Negative values are of $\varphi(\cdot)$ are in red.   (b) $\varphi(\cdot) > 0$ over the "green" domain.

Figure D.1 : Domain of $\varphi(\cdot)$. $\lambda var(Q) = .2$, $m = 6$, $u = 0$.

D.2   Retailer

For the second order condition to hold, we need to show that

$$-\frac{1}{z} - \frac{\lambda}{2}\left[r^2 A + 2t^2 var(Q) - 4rvar(Q)(1+W)\frac{st}{z^2}\right] \tag{D.3}$$

is negative, where $A = 2var(Q)(1+W)\frac{s^2}{z^2} + 2(1+W)var(Q)\frac{s^2}{z^2}$ and W is the conjectural variations. A sufficient condition is that $W = -1$ (Bertrand conjectures). Even under Cournot conjectures ($W = 0$), we can easily show that the condition holds by noting that

$$r^2 A + 2t^2 var(Q) - 4var(Q)\frac{rst}{z^2} = 4\left(\frac{rs}{z}\right)^2 + 2t^2 - r\frac{rst}{z^2}$$

$$= \left(\frac{rs - t^2}{z}\right) + \left(\frac{rs^2}{z} + \frac{1}{z^2}\right) \geq 0, \tag{D.4}$$

Hence, the whole expression (D.3) is negative, given $z = m(m-1)\beta > 0$ for $m > 1$.

Appendix E

Proof of lemma 2.2.2

Lemma E.0.1. *Let* $\begin{bmatrix} A & B \\ C & D \end{bmatrix}$ *be a block matrix. Its inverse is given by*

$$\begin{bmatrix} (A - BD^{-1}C)^{-1} & -A^{-1}B(D - CA^{-1}B)^{-1} \\ -D^{-1}C(A - BD^{-1}C)^{-1} & (D - CA^{-1}B)^{-1} \end{bmatrix}$$

Proof. See [44], §1.6. $\square$

We restate lemma 2.2.2 for convenience:

Lemma E.0.2. *Let T be an $n \times n$ ($n > 2$) matrix where the ith diagonal element is d_i and the off-diagonal elements on row i are c_i. Its inverse is another $n \times n$ matrix with the ith diagonal entry and the ij^{th} off-diagonal entries given respectively as*

$$\frac{\prod\limits_{j \neq i} d_j + [\sum\limits_{k=1}^{n-2} (-1)^k k (\sum\limits_{(S \setminus \{i\})_{n-k-2}} \prod\limits_{j \in (S \setminus \{i\})_{n-k-2}} d_j \prod\limits_{l \neq i,j} c_l)]}{\prod\limits_{j=1}^{n} d_j + \sum\limits_{k=2}^{n} [(-1)^{k-1}(k-1) \sum\limits_{S_{n-k}} \prod\limits_{i \in S_{n-k}} d_i \prod\limits_{j \neq i} c_j]} \tag{E.1}$$

$$\frac{-c_i \prod\limits_{k \neq \{i,j\}} (d_k - c_k)}{\prod\limits_{i=1}^{n} d_i + \sum\limits_{k=2}^{n} [(-1)^{k-1}(k-1) \sum\limits_{S_{n-k}} \prod\limits_{i \in S_{n-k}} d_i \prod\limits_{l \neq i} c_l]}, \tag{E.2}$$

where S_m denotes the m-element subsets of $S = \{1, 2, 3, ..., n\}, m > 0$. If $m = 0$, $S_0 = \emptyset$, in which case $\prod_{k \neq i,j}$ and $\prod_{l \neq i}$ become $\prod_{k \neq i}$ and $\prod_l$, respectively.

Proof. By induction. The claim trivially holds for $n = 3$. Suppose it holds for $n = r$. We have to show it is true for $n = r + 1$ as well. In order to do so, it is enough to note

97

that T is A, the column vector $[c_1, c_2, ..., c_{r-1}]$ is B, the row vector $[c_r \quad c_r \quad ... \quad c_r]$ is C and the scalar d_r is D in lemma E.0.1. $\square$

Bibliography

[1] P. AGHION AND P. BOLTON, *Contracts as a barrier to entry*, The American Economic Review, 77 (1987), pp. 388–401.

[2] B. ALLAZ AND J.-L. VILA, *Cournot competition, forward markets and efficiency*, Journal of Economic Theory, 59 (1993), pp. 1–16.

[3] E. J. ANDERSON AND A. B. PHILPOTT, *Using supply functions for offering generation into an electricity market*, Operations Research, 50 (2002), pp. 477–489.

[4] W. B. ARTHUR, *Out-of-equilibrium economics and agent-based modeling*, in Handbook of Computational Economics: Agent-Based Computational Economics, Volume 2, L. Tesfatsion and K. L. Judd, eds., North-Holland, 2006, pp. 1551–1564.

[5] F. AZEVEDO, H. KHODR, AND Z. VALE, *Transmission cost allocation using cooperative game theory: A comparative study*, in Energy Market, 2009. EEM 2009. 6th International Conference on the European, May 2009, pp. 1 –6.

[6] R. BALDICK, R. GRANT, AND E. KAHN, *Linear supply function equilibrium: Generalizations, application, and limitations*. POWER Working Paper Series, PWP-078, University of California Energy Institute, 2000.

[7] ———, *Theory and application of linear supply function equilibrium in electricity*

markets, Journal of Regulatory Economics, 25 (2004), pp. 143–167.

[8] R. BALDICK AND W. W. HOGAN, *Capacity constrained supply function equilibrium models of electricity markets: Stability, non-decreasing constraints, and function space iterations*. POWER Working paper, 2002.

[9] H. BESSEMBINDER AND M. L. LEMMON, *Equilibrium pricing and optimal hedging in electricity forward markets*, The Journal of Finance, 57 (2002), pp. 1347–1382.

[10] F. BOLLE, *Supply function equilibria and the danger of tacit collusion*, Energy Economics, 13 (1992), pp. 94–102.

[11] S. BORENSTEIN AND J. BUSHNELL, *An empirical analysis of the potential for market power in California's electricity industry*, Journal of Industrial Economics, 47 (1999), pp. 285–323.

[12] S. BORENSTEIN, J. BUSHNELL, AND S. STOFT, *The competitive effect of transmission capacity in a deregulated electricity market*, RAND Journal of Economics, 31 (2000), pp. 294–325.

[13] J. BOWER AND D. W. BUNN, *Model-based comparisons of pool and bilateral markets for electricity*, Energy Journal, 21 (2000), pp. 1–29.

[14] J. BRANDTS, P. PEZANIS-CHRISTOU, AND A. SCHRAM, *Competition with forward contracts: a laboratory analysis motivated by electricity market design*, The Economic Journal, 118 (2008), pp. 192–214.

[15] J. BUSHNELL, *Oligopoly equilibria in electricity contract markets*, Journal of Regulatory Economics, 32 (2007), pp. 225–245.

[16] J. B. Bushnell, E. T. Mansur, and C. Saravia, *Vertical arrangements, market structure, and competition: An analysis of restructured US electricity markets*, American Economic Review, 98 (2008), pp. 237–266.

[17] J. Casazza and F. Delea, *Understanding Electric Power Systems: An Overview of the Technology and the Marketplace*, IEEE Press, 2003.

[18] T. S. Chung, S. H. Zhang, K. P. Wong, C. W. Yu, and C. Y. Chung, *Strategic forward contracting in electricity markets: modeling and analysis by equilibrium method*, IEEE Proceedings on Transmission, Generation and Distribution, 151 (2004), pp. 141–149.

[19] P. H. Cootner, *Returns to speculators: Rejoinder*, The Journal of Political Economy, 68 (1960), pp. 415–418.

[20] ——, *Returns to speculators: Telser versus Keynes*, The Journal of Political Economy, 68 (1960), pp. 396–404.

[21] C. L. Coq and H. Orzen, *Do forward markets enhance competition? experimental evidence*, Journal of Economic Behavior & Organization, 61 (2006), pp. 415–431.

[22] P. J. Coughlin, *Changes in marginal revenue and related elasticities*, Southern Economic Journal, 51 (1984), pp. 568–573.

[23] ERCOT, *ERCOT Protocols*, July 2009. www.ercot.com/mktrules/protocols/library.

[24] ——, *ERCOT Emergency Interruptible Load Service*, August 2010. http://www.ercot.com/services/programs/load/eils/.

[25] I. EREV AND A. ROTH, *Learning in extensive form games: Experimental data and simple dynamic models in the intermediate term*, Games and Economic Behavior, 8 (1995), pp. 164–212.

[26] ——, *Predicting how people play games: Reinforcement learning in experimental games with unique, mixed strategy equilibria*, The American Economic Review, 88 (1998), pp. 848–881.

[27] E. F. FAMA AND K. R. FRENCH, *Commodity futures prices: Some evidence on forecast power, premiums, and the theory of storage*, The Business Journal, 60 (1987), pp. 55–73.

[28] J. L. FERREIRA, *Strategic interaction between futures and spot markets*, Journal of Economic Theory, 108 (2003), pp. 141–151.

[29] ——, *The role of observability in futures markets*, Topics in Theoretical Economics, 6 (2006), pp. 1–22.

[30] J. P. FORMBY, S. LAYSON, AND W. J. SMITH, *The law of demand, positive sloping marginal revenue, and multiple profit equilibria*, Economic Inquiry, 20 (1980), pp. 303–311.

[31] K. R. FRENCH, *Detecting spot price forecasts in futures prices*, The Business Journal, 59 (1986), pp. S39–S54.

[32] F. N. FRITSCH AND J. BUTLAND, *A method for constructing local monotone piecewise cubic interpolants*, SIAM Journal on Scientific and Statistical Computing, 5 (1984), pp. 300–304.

[33] F. N. FRITSCH AND R. E. CARLSON, *Monotone piecewise cubic interpolation*, SIAM Journal on Numerical Analysis, 17 (1980), pp. 238–246.

[34] R. GREEN, *The electricity contract market in England and Wales*, Journal of Industrial Economics, 47 (1999), pp. 107–124.

[35] R. GREEN AND C. L. COQ, *The length of contracts and collusion*, International Journal of Industrial Organization, 28 (2010), pp. 21–29.

[36] R. GREEN AND D. M. NEWBERY, *Competition in the British electricity spot market*, Journal of Political Economy, 100 (1992), pp. 929–953.

[37] ——, *Competition in the electricity industry in England and Wales*, Oxford Review of Economic Policy, 13 (1997), pp. 27–46.

[38] S. M. HARVEY AND W. W. HOGAN, *California electricity markets and forward market hedging*. last accessed on February 3, 2008, 2000.

[39] D. HIRSHLEIFER, *Hedging pressure and futures price movements in a general equilibrium model*, Econometrica, 58 (1990), pp. 411–428.

[40] D. HIRST, *Bridging ambient electrical supply with demand: A competitive market design proposal*. 31st IAEE International Conference, 18-20 June, Istanbul, Turkey, 2008.

[41] E. HIRST, *Real-time balancing operations and markets: Key to competitive wholesale electricity markets*, tech. rep., Edison Electric Institute and Project for Sustainable FERC Energy Policy, 2001.

[42] W. HOGAN, *A market power model with strategic interaction in electricity networks*, The Energy Journal, 18 (1997), pp. 107–142.

[43] P. HOLMBERG, *Strategic forward contracting in the wholesale electricity market*, Energy Journal, 32 (2011), pp. 169–202.

[44] R. A. HORN AND F. ZHANG, *Basic properties of the Schur complement*, in The Schur Complement and Its Applications, F. Zhang, ed., Springer, 2005, pp. 17–46.

[45] A. HORTAÇSU AND S. L. PULLER, *Understanding strategic bidding in restructured electricity markets: A case study of ERCOT*. NBER Working Paper 11123, 2005.

[46] ——, *Understanding strategic bidding in multi-unit auctions: A case study of the Texas electricity spot market*, RAND Journal of Economics, 39 (2008), pp. 86–114.

[47] J. S. HUGHES AND J. L. KAO, *Strategic forward contracting and observability*, International Journal of Industrial Organization, 16 (1997), pp. 121–133.

[48] W. JING-YUAN AND Y. SMEERS, *Spatial oligopolistic electricity models with Cournot generators and regulated transmission prices*, Operations Research, 47 (1999), pp. 102–112.

[49] R. KAMAT AND S. S. OREN, *Two-settlement systems for electricity markets under network uncertainty and market power*, Journal of Regulatory Economics, 25 (2004), pp. 5–37.

[50] S. KAPLAN, *Power plants: Characteristics and costs*, tech. rep., Congressional Research Service, 2008.

[51] P. D. KLEMPERER AND M. A. MEYER, *Supply function equilibria in oligopoly under uncertainty*, Econometrica, 57(6) (1989), pp. 1243–1277.

[52] B. I. KVASOV, *Methods of Shape-Preserving Spline Approximation*, World Scientific Publishing Co. Pte. Ltd., 2000.

[53] J. LALLY, *Financial transmission rights auction example*, tech. rep., ISO New England, 2002.

[54] H. LI AND L. TESFATSION, *Co-learning patterns as weakly emergent phenomena in wholesale electric power markets.* 2010.

[55] ——, Forthcoming: *ISO net surplus collection and allocation in wholesale power markets under locational marginal pricing*, IEEE Transactions on Power Systems, 26 (2010).

[56] M. LIN, *Management of transmission in the electricity markets*, in Energy Modeling: Advances in the Management of Uncertainty, V. Kaminski, ed., Risk Books, 2005, pp. 169–207.

[57] D. LINEWEBER AND S. MCNULTY, *The cost of power disturbances to industrial & digital economy companies*, 2001. Electric Power Research Institute.

[58] M. LISKI AND J.-P. MONTERO, *Forward trading and collusion in oligopoly*, Journal of Economic Theory, 131 (2006), pp. 212–230.

[59] H. LIU, L. TESFATSION, AND A. A. CHOWDHURY, *Derivation of locational marginal prices for restructured wholesale power markets*, The Journal of Energy Markets, 2 (2009), pp. 3–27.

[60] ——, *Locational marginal pricing basics for restructured wholesale power markets*, in IEEE Power & Energy Society General Meeting, July 2009.

[61] F. A. LONGSTAFF AND A. W. WANG, *Electricity forward prices: A high-frequency empirical analysis*, The Journal of Finance, 59 (2004), pp. 1877–1900.

[62] Z.-Q. LUO, JONG-SHI-PANG, AND D. RALPH, *Mathematical Programs with Equilibrium Constraints*, Cambridge University Press, 1996.

[63] A. MADHAVAN, *Market microstructure: A survey*, Journal of Financial Markets, 3 (2000), pp. 205–258.

[64] P. MAHENC AND F. SALANIÉ, *Softening competition through forward trading*, Journal of Economic Theory, 116 (2004), pp. 282–293.

[65] D. F. McALLISTER AND J. A. ROULIER, *An algorithm for computing a shape-preserving osculatory quadratic spline*, ACM Transactions on Mathematical Software, 7(3) (1981), pp. 331–347.

[66] M. B. MONAGAN, K. O. GEDDES, K. M. HEAL, G. LABAHN, S. M. VORKOETTER, AND J. M. P. DEMARCO, *Maple Advanced Programming Guide*, Maplesoft, 2005.

[67] ——, *Maple Introductory Programming Guide*, Maplesoft, Ontario, Canada, 2005.

[68] D. M. NEWBERY, *Competition, contracts and entry in the electricity spot market*, RAND Journal of Economics, 29(4) (1998), pp. 726–479.

[69] J. NICOLAISEN, V. PETROV, AND L. TESFATSION, *Market power and efficiency in a computational electricity market with discriminatory double-auction*

pricing, IEEE Transactions on Evolutionary Computation, 5 (2001), pp. 504–523.

[70] H. NIU, R. BALDICK, AND G. ZHU, *Supply function equilibrium bidding strategies with fixed forward contracts*, IEEE Transactions on Power Systems, 20 (2005), pp. 1859–1867.

[71] S. OREN, P. SPILLER, P. VARAIYA, AND F. WU, *Nodal prices and transmission rights: A critical appraisal*, The Electricity Journal, 8, pp. 24–35.

[72] E. PASSOW AND J. ROULIER, *Monotone and convex spline interpolant*, SIAM Journal on Numerical Analysis, 14 (1977), pp. 904–909.

[73] A. POWELL, *Trading forward in an imperfect market: The case of electricity in Britain*, The Economic Journal, 103 (1993), pp. 444–453.

[74] S. PRUESS, *Alternatives to the exponential spline in tension*, Mathematics of Computation, 33 (1979), pp. 1273–1281.

[75] R. T. ROCKAFELLAR, *Convex Analysis*, Princeton Landmarks in Mathematics, Princeton University Press, New Jersey, 1997.

[76] A. ROSENBERG, *Congestion pricing or monopoly pricing?*, The Electricity Journal, 13 (2000).

[77] ——, *A proposed alternative to locational marginal pricing and its impact on new generation*, The Electricity Journal, 17 (2004).

[78] A. RUDKEVICH, *Supply function equilibrium in power markets: Learning all the way*, Tech. Rep. TCA Technical Paper 1299-1702, Tabors Caramanis & Associates, 50 Church St. Cambridge, MA, 02138, 1999.

[79] ——, *Supply function equilibrium: Theory and applications*, in Proceedings of the 36th Hawaii International Conference on System Sciences, IEEE, 2002.

[80] ——, *On the supply function equilibrium and its applications in electricity markets*, Decision Support Systems, 40 (2005), pp. 409–425.

[81] J. J. SÁNCHEZ, D. W. BUNN, E. CENTENO, AND J. BARQUÍN, *Dynamics in forward and spot electricity markets*, IEEE Transactions on Power Systems, 24 (2009).

[82] L. L. SCHUMAKER, *Spline Functions: Basic Theory*, Cambridge University Press, Cambridge, UK, 3rd ed., 2007.

[83] M. SHAHIDEHPOUR, H. YAMIN, AND Z. LI, *Market Operations in Electric Power Systems*, IEEE Press, 2002.

[84] R. SIOSHANSI AND S. OREN, *How good are supply function equilibrium models: An empirical analysis of the ERCOT balancing market*, Journal of Regulatory Economics, 31 (2007), pp. 1–35.

[85] D. SONGHUAI, Z. XINGHUA, M. LU, AND X. HUI, *A novel nucleolus-based loss allocation method in bilateral electricity markets*, IEEE Transactions on Power Systems, 21 (2006), pp. 28–33.

[86] C.-L. SU, *Analysis on the forward market equilibrium model*, Operations Research Letters, 35 (2007), pp. 74–82.

[87] J. SUN AND L. TESFATSITON, *An agent-based computational laboratory for wholesale power market design*, in IEEE Power Engineering Society General Meeting, June 2007.

[88] ——, *Dynamic testing of wholesale power market designs: An open-source agent-based framework*, Computational Economics, 30 (2007), pp. 291–327.

[89] X. TAN AND T. LIE, *Application of the Shapley value on transmission cost allocation in the competitive power market environment*, Generation, Transmission and Distribution, IEE Proceedings-, 149 (2002), pp. 15 – 20.

[90] L. G. TELSER, *Returns to speculators: Reply*, The Journal of Political Economy, 68 (1960), pp. 404–415.

[91] S. TENG, J. YU, L. JOHNSON, AND J. MICKEY, *A study on balancing energy bid price overlap in ERCOT market*, in IEEE International Conference on Electric Utility Deregulation, Restructuring and Power Technologies, Hong Kong, April 2004.

[92] L. TESFATSION, *Agent-based computational economics: A constructive approach to economic theory*, in Handbook of Computational Economics: Agent-Based Computational Economics, Volume 2, L. Tesfatsion and K. L. Judd, eds., North-Holland, 2006, pp. 831–880.

[93] L. TESFATSION AND K. L. JUDD, *Agent-Based Computational Economics*, vol. 13 of Handbooks in Economics, North-Holland, 2006.

[94] THE UNITED STATES DEPARTMENT OF JUSTICE, *Horizontal merger guidelines*, 2011.

[95] P. VISUDHIPHAN AND M. D. ILIĆ, *Dynamic games-based modeling of electricity markets*, in IEEE Power Engineering Society 1999 Winter Meeting, 1999, pp. 274–281.

[96] A. A. WALTERS, *A note on monopoly equilibrium*, The Economic Journal, 90 (1980), pp. 161–162.

[97] R. WANG, Y. LI, AND S. ZHANG, *Analysis of supply function equilibrium in electricity markets with financial options contracts*, in Third International Conference on Electric Utility Deregulation and Restructuring and Power Technologies, DRPT 2008, April 2008.

[98] A. WEIDLICH AND D. VEIT, *A critical survey of agent-based wholesale electricity market models*, Energy Economics, 30 (2008), pp. 1728–1759.

[99] F. WOLAK, *An empirical analysis of the impact of hedge contracts on bidding behavior in a competitive electricity market*, International Economic Journal, 14 (2000), pp. 1–39.

[100] F. A. WOLAK, *Identication and estimation of cost functions using observed bid data: An application to electricity markets*, in Advances in Economics and Econometrics: Theory and Applications, Eighth World Congress, Volume 2, M. Dewatripont, L. P. Hansen, and S. J. Turnovsky, eds., Cambridge University Press, 2003, pp. 133–169.

[101] C. WOLFRAM, *Measuring duopoly power in the British electricity spot market*, American Economic Review, 89 (1999), pp. 805–826.

[102] F. WU, P. VARAIYA, P. SPILLER, AND S. OREN, *Folk theorems on transmission access: Proofs and counterexamples*, Journal of Regulatory Economics, 10 (1996), pp. 5–23.

[103] WWW.ERCOT.COM, *Company profile*, 2008.

[104] H. XIAO AND W. LI, *Allocation of congestion cost in a pool based model using Shapley value*, in Power and Energy Engineering Conference, 2009. APPEEC 2009. Asia-Pacific, 2009, pp. 1 –5.

[105] H. YANG, D. SHI, AND X. DUAN, *Congestion cost allocation method based on Aumann-Shapley value in bilateral model*, in IEEE Power Engineering Society General Meeting, 2003, vol. 2, 2003, pp. 1002–1006.

[106] J. YAO, S. S. OREN, AND I. ADLER, *Two-settlement electricity markets with price caps and cournot generation firms*, European Journal of Operations Research, 181 (2007), pp. 1279–1296.

[107] ——, *Modeling and computing two-settlement oligopolistic equilibrium in a congested electricity network*, Operations Research, 56 (2008), pp. 34–47.

[108] J. M. ZOLEZZI AND H. RUDNICK, *Transmission cost allocation by cooperative games and coalition formation*, IEEE Transactions on Power Systems, 17 (2002), pp. 1008–1015.