CHAPTER 4

Identification and Estimation of Cost Functions Using Observed Bid Data

An Application to Electricity Markets

Frank A. Wolak

1. INTRODUCTION

This paper presents several techniques for recovering cost function estimates for electricity generation from a model of optimal bidding behavior in a competitive electricity market. These procedures are applied to actual data from the Australian National Electricity Market (NEM1) to recover cost function estimates for a specific market participant. I find close agreement between the cost functions recovered from these procedures and those obtained from engineering estimates. The techniques developed in this paper for recovering cost function estimates are not limited to markets for electricity generation. They can be used to recover cost function estimates for a participant in any bid-based centralized market.

There are number of uses for the procedures developed in this paper. The primary use is to measure the extent of market power possessed by a market participant using only bid information and market-clearing prices and quantities. A major research effort in empirical industrial organization is the measurement of market power. Bresnahan (1989) summarizes much of this research, although there has been an explosion of recent research on this general topic. The techniques presented in this paper are a logical extension of the techniques described by Bresnahan (1989) to bid-based markets.

A major challenge for designers of competitive electricity markets is to devise market rules that limit the ability of generation unit owners to exercise market power. Market power is the ability of a firm owning generation assets to raise the market price by its bidding behavior and to profit from this price increase. Until the recent trend toward industry restructuring, electricity was supplied by vertically integrated geographic monopolies regulated by state public utilities commissions in the United States or by goverment-owned national or state monopolies in other countries around the world. All of the industry characteristics that resulted in these two market structures make wholesale markets for electricity generation ripe for the exercise of market power. Electricity is extremely costly to store, there are binding short-run capacity constraints on its production, and demand must equal supply throughout the electricity grid at every moment in time. In addition, because of the manner in which electricity was sold to final customers during the former vertically integrated regime, the retail demand for electricity is very price inelastic on hour-ahead, day-ahead, and even month-ahead time horizons. These features of the electricity production process and the insensitivity of retail demand to wholesale price fluctuations allow small defects in market design to enhance significantly the ability of generation unit owners to exercise market power.

For this same reason, seemingly innocuous changes in market rules can produce a large impact on market outcomes. Consequently, market design is an extremely important aspect of the ongoing industry restructuring process. The optimal market design problem can be thought of as a single-principal (the market designer), multiple-agent (many electricity generation unit owners and wholesale energy purchasers) problem. Although the market design problem fits into the general class of common agency problems, given the complexity of even an extremely simple transmission network, solving for the optimal market design is an immensely complex task. The set of feasible mechanisms for compensating generators for the energy and generation reserves they supply and charging wholesale consumers for the energy they demand is enormous. Consequently, for a given market structure there are many feasible market designs, but the optimal market design is unknown.

Fortunately, all markets that currently exist in the United States have in place a process whereby the market monitoring unit within the Independent System Operator (ISO), the entity that operates the wholesale market and transmission grid, studies all aspects of market performance in order to detect design flaws that degrade market performance and enhance the ability of firms to exercise market power.[1] The next step in this market design process is to devise and implement market rule changes that eliminate these design flaws and move closer to finding the optimal set of market rules for that market structure. Although economic theory plays a major role in this process, there are very few empirical methods with a firm foundation in economic theory for analyzing the vast volumes of bid and market outcomes data available to these market monitoring units. This paper develops these sorts of tools and illustrates their use in the market monitoring and design process.

The specific application I consider is the estimation of forward market energy positions from spot market bid functions. In virtually all competitive wholesale electricity markets generators and loads engage in forward financial or hedge contracts, which allow them to fix the price for a specified amount of energy delivered and consumed in real time.[2] As noted in Wolak (2000), even with

[1] For an across-country discussion of the institutions and performance of competitive electricity markets, see Wolak (1999).

[2] Hedge contracts are typically signed between a generating company and an electricity retailer. They are purely financial obligations that guarantee the price at which a fixed quantity of electricity will be sold at a mutually agreed-on time in the future to the purchaser of the forward contract. If the relevant spot market price exceeds the contract price, then the contract seller pays to the buyer the difference between these two prices times the contract quantity. If the market price is less than the contract price, the buyer pays the absolute value of this same price difference times the contract quantity to the seller.

knowledge of a firm's bidding behavior in a competitive electricity market, it is difficult, if not impossible, to determine if the firm is able to exercise market power without knowing the generation unit owner's forward contract position. For a specific bid function and marginal cost function, there is a portfolio of forward financial contracts that can rationalize that bid function as expected profit maximizing. Wolak (2000) also demonstrates the enormous influence a generation unit owner's financial contract position has on his or her incentive to bid to attempt to increase the market-clearing price in the spot electricity market. Consequently, a technique for estimating a market participant's hedge contract or forward market position from bids submitted to a spot market can allow a market monitor to determine more precisely when the generation unit owner is likely to possess significant market power.

The remainder of the paper proceeds as follows. Section 2 relates this research to the general literature in empirical industrial organization on measuring market power using data on market-clearing prices and quantities. This section discusses the gain in econometric identification of the underlying cost function that results from using bids in addition to market-clearing prices and quantities in the estimation process. Section 3 presents a model of optimal bidding behavior with hedge contracts for a generic competitive electricity market. This section defines a *best-response bidding strategy* as the set of daily bid prices and quantities that maximize expected daily variable profits given the strategies of other firms participating in the market. The section also defines the *best-response price* as the market-clearing price that maximizes the realized profits of the firm given the bids actually submitted by its competitors, the realized value of the stochastic shock to the price-setting process. Both of these concepts are used to derive estimates of the cost function for a bidder in a competitive electricity market using actual bid information, the firm's hedge contract position, and actual market outcomes.

Section 4 presents my estimation methodology based on the best-response price concept. Section 5 presents this methodology based on the best-response bidding strategy. Section 6 then describes the essential features of the Australian National Electricity Market and the data set used in my empirical analysis. Section 7 presents the results of this empirical analysis. Section 8 describes how these techniques might be used in the market design process and discusses directions for future research.

2. IDENTIFYING MARGINAL COST FUNCTIONS FROM BIDS AND MARKET PRICES AND QUANTITIES

Beginning with Rosse (1970), empirical industrial organization (IO) economists have devised estimation procedures to recover cost functions from data on market-clearing prices and quantities. Rosse used a sample of monopoly local newspapers and the assumption of profit maximization to estimate the underlying marginal cost function of the monopolists. Porter (1983) employed a related approach in his study of price wars in the U.S. railroad industry during the 1880s.

He assumes a firm-level, homogeneous product, quantity-setting conjectural variation oligopoly equilibrium. He aggregates the firm-level first-order conditions to produce an industrywide supply function, which he jointly estimates along with an industry-level demand function. Bresnahan (1981, 1987) quantifies the extent of market power possessed by each vehicle model in the U.S. automobile industry using a discrete choice, differentiated products model of individual demand with vertical product differentiation in the unobserved product quality dimension. Aggregating these discrete purchase decisions across U.S. households yields an aggregate demand system for all automobile models. Bresnahan assumes Nash–Bertrand competition among the automobile makers facing this aggregate demand system to estimate the implied marginal cost of producing automobiles of each quality level. More recently, Berry (1994), Berry, Levinsohn, and Pakes (1995), and Goldberg (1995) have extended the techniques pioneered by Bresnahan to discrete choice oligopoly models with horizontal product differentiation.

The basic idea of all the techniques just described can be illustrated using the following example, which follows from the intuition given in Rosse (1970). Let $P(q, W, \theta, \varepsilon)$ denote the inverse demand function facing a monopolist and $C(q, Z, \theta, \eta)$ its total cost function. The variables W and Z are demand and cost function shifters, respectively. Here θ is the vector of parameters to be estimated, and ε and η are unobserved, to the econometrician, stochastic shocks. These shocks are assumed to be observable to the monopolist. The profit function of the monopolist is

$$\pi(q) = P(q, W, \theta, \varepsilon)q - C(q, Z, \theta, \eta). \tag{2.1}$$

The first-order condition for profit maximization is

$$\pi'(q) = P'(q, W, \theta, \varepsilon)q + P(q, W, \theta, \varepsilon) - C'(q, Z, \theta, \eta) = 0. \tag{2.2}$$

The researcher is assumed to have only market-clearing price and quantity data and the values of the demand and supply shifters, W and Z, for a cross section of monopolists selling the same homogeneous product. The econometrician does not have information on production costs for any of the firms. The researcher could also have a time series of observations on the same information for one or a small number of monopolists over time. This lack of cost data is the standard case faced by empirical researchers studying unregulated industries, such as those for automobiles, airlines, and personal computers, which are a few of the industries in which these techniques have been applied. Industry associations or government regulators usually publicly disclose information on market-clearing prices and quantities, but they give little information on production costs.

For the econometrician to make any statement about the extent of market power exercised in this market, she or he must have an estimate of the marginal cost function, $C'(q, Z, \theta, \eta)$. This estimate is constructed in the following manner. The econometrician first specifies a functional form for the inverse demand function. Suppose she or he selects $P(q, W, \theta, \varepsilon) = \mathrm{a} + \mathrm{b}q + \mathrm{c}W + \varepsilon$, where a, b, and c are elements of θ. The parameters of a, b, and c must be estimated by standard instrumental variables techniques to account for the fact that observed q and unobserved ε are correlated. This correlation occurs because the observed market-clearing quantity is determined by solving the first-order condition for profit maximization given in (2.2). This implies that q^E, the equilibrium quantity, is a function of η and ε and the demand and supply shifters, W and Z, so that $q^E = f(W, Z, \eta, \varepsilon)$. The market-clearing price is then determined by substituting q^E into the inverse demand function.

Given these estimates for a, b, and c, the econometrician can then solve for the value of $C'(q, Z, \theta, \eta)$ implied by the first-order conditions for profit maximization given in (2.2), using the observed market-clearing prices and quantities. Rearranging (2.2) for the assumed parametric inverse demand function yields

$$C'(q^E, Z, \theta, \eta) = P'(q^E, W, \theta, \varepsilon)q + P(q^E, W, \theta, \varepsilon) = \mathrm{b}q^E + p^E. \tag{2.3}$$

For each value of p^E and q^E, the market-clearing prices and quantities, compute an estimate of the marginal cost, $C'(q^E, Z, \theta, \eta)$, using the right-hand side of (2.3) and an estimate of the demand parameter b. This marginal cost estimate can then be used to compute an estimate of the amount of market power possessed by the firm in each market, by computing the Lerner index:

$$L = [p^E - C'(q^E, Z, \theta, \eta)]/p^E. \tag{2.4}$$

The assumption of firm-level profit maximization implies that estimates of only the parameters of the demand function are needed to compute an estimate of the Lerner index of market power.

Researchers often select a functional form for $C'(q^E, Z, \theta, \eta)$ and use the implied marginal costs derived from (2.3) to estimate the elements of θ contained in the cost function. An alternative approach, beginning with Rosse (1970), estimates the parameters of the inverse demand and cost function jointly using the assumption of profit maximization to identify the marginal cost function from observed market-clearing prices and quantities.

The intuition embodied in this example is used in all of the papers described thus far. Porter (1983) estimates the aggregate demand function facing the oligopoly that he studies. He makes assumptions on the functional form of costs for each individual firm and the nature of the strategic interaction among firms – cartel or perfect competition – to deliver an aggregate supply function for the industry under each of these two behavioral assumptions. Then he jointly estimates these aggregate supply and demand equations as a switching regression model, using the assumption of profit maximization to identify parameters of the underlying individual cost functions from time series data on market-clearing prices and quantities.

Bresnahan (1987) specifies a discrete-choice demand structure in which each individual decides whether to purchase an automobile, and if so, which model.

He aggregates this discrete-choice demand structure across all consumers to derive a system of aggregate demand equations. Using various assumptions about the nature of strategic interaction – specifically, Nash–Bertrand competition or collusion among automobile producers – he estimates the parameters of this aggregate demand system along with the parameters of the marginal cost functions implied by the first-order conditions associated with profit maximization. Bresnahan (1981) allows for a richer stochastic specification in the aggregate demand system, but follows the same basic procedure to recover estimates of the marginal cost function.

Berry et al. (1995) allow for a multinomial logit discrete-choice demand structure at the consumer level and assume unobservable (to the econometrician) stochastic consumer-level marginal utilities of product attributes. These marginal utilities are assumed to be independent and nonidentically normally distributed across product attributes and independent and identically distributed across consumers. Integrating individual-level purchase decisions with respect to these normal distributions yields a product-level aggregate demand system for automobiles. The authors assume that the conditional indirect utility functions for each consumer contain the same vector of unobservable (to the econometrician) product characteristics, and that these product characteristics are uncorrelated with all observable product characteristics. This stochastic structure induces correlation between equilibrium prices and the vector of unobserved random product characteristics in the aggregate demand system. Berry et al. propose and implement an instrumental variables estimation technique that exploits this lack of correlation between observed and unobserved product characteristics to estimate the demand system jointly with the marginal cost function under the assumption of Nash–Bertrand competition among automobile producers. In contrast, Bresnahan (1981, 1987) relies on maximum likelihood techniques.

Goldberg (1995) uses individual household-level data to estimate a general discrete-choice model for automobile purchases at the household level. She then uses weights giving the representativeness of each of these households in the population of U.S. households to produce a system of aggregate demand functions for automobiles based on the choice probabilities implied by her model of household-level automobile demand. Using the assumption of Nash–Bertrand competition among automobile producers, she then computes implied marginal cost estimates similar to those given in (2.3), which she uses to estimate a marginal cost function for each automobile model.

The most important conclusion to draw from this line of research is that all marginal cost estimates are the direct result of the combination of the assumed functional form for the aggregate demand for the products under consideration and the assumed model of competition among firms. Similar to the example given here, the first-order conditions for profit maximization and the demand function for each product determine the implied marginal cost for that product. Consequently, a major focus of this research has been on increasing the flexibility and credibility of the aggregate demand system used. However, because supply and demand functions are the econometrician's creation for describing the observed joint distribution of market-clearing prices and quantities across markets, for any finite data set, a functional form for the demand curves faced by the oligopolists or the monopoly must be assumed in order to estimate any demand function. Although it may be possible to apply the nonparametric and semiparametric identification and estimation strategies described in Blundell and Powell (2003) and in Florens (2003) to this economic environment, all existing work on modeling oligopoly equilibrium has relied on functional form restrictions and models of firm-level profit-maximizing behavior to identify the underlying demand and cost functions.

Rosse (1970) and Porter (1983) explicitly make this functional form assumption for aggregate demand. Bresnahan (1981, 1987) and Berry et al. (1995) assume a functional form for the probabilities that determine individual purchase decisions. They derive the aggregate demand system actually estimated by summing these individual choice probabilities across consumers. Goldberg (1995) specifies a household-level choice model, which she estimates using household-level data. The aggregate demand functions entering into her oligopoly model are an appropriately weighted sum of these estimated household-level demand systems across all U.S. households.

3. MODELS OF BEST-RESPONSE BIDDING AND BEST-RESPONSE PRICING

This section shows how the techniques described herein can be extended to estimate underlying marginal cost functions using data on bids and market-clearing prices and quantities from competitive electricity markets. Specifically, I demonstrate how the availability of bids allows the econometrician to identify the underlying firm-level cost function purely through an assumption about firm behavior. A functional form assumption for aggregate demand is no longer necessary. I consider two models of optimizing behavior by the firm that recover an estimate of the firm's marginal cost function. The first model makes the unrealistic but simplifying assumption that the firm is able to choose the market-clearing price that maximizes its profits given the bids submitted by its competitors. The second model is more realistic, but entails a significantly greater computation burden. It imposes all of the constraints implied by the market rules on the bids used by the firm to set the market-clearing price. The firm is assumed to bid according to the rules of the competitive electricity market to maximize its expected profits. The second approach explicitly imposes the reality that the only way the firm is able to influence the market-clearing price is through the bids it submits.

A first step in describing both of these methodologies is a description of the payoff functions and strategy space for participants in a generic competitive electricity market. Specifically, I describe the structure of bids and how these bids are translated into the payoffs that a market participant receives for supplying energy to the wholesale electricity market.

A competitive electricity market is an extremely complicated noncooperative game with a very high-dimensional strategy space. A firm owning a single generating set (genset) competing in a market with half-hourly prices must, at a minimum, decide how to set the daily bid price for the unit and the quantity bid for forty-eight half-hours during the day.[3] In all existing electricity markets, firms have much more flexibility in how they bid their generating facilities. For instance, in NEM1, firms are allowed to bid daily prices and half-hourly quantities for ten bid increments per genset. For a single genset, this amounts to a 490-dimensional strategy space (ten prices and 480 half-hourly quantities). Bid prices can range from −9,999.99 $AU to 5,000.00 $AU, which is the maximum possible market price. Each of the quantity increments must be greater than or equal to zero and their sum less than or equal to the capacity of the genset. Most of the participants in this market own multiple gensets, so the dimension of the strategy space for these firms is even larger. The England and Wales electricity market imposes similar constraints on the bid functions submitted by market participants. Each genset is allowed to bid three daily price increments and 144 half-hourly quantity increments. Genset owners also submit start-up and no-load costs as part of the day-ahead bidding process. Bidders in the California ISO's real-time electricity market bid eleven price and quantity increments, both of which can vary on an hourly basis.

To compute the profit function associated with any set of bids the firm might submit, I must have an accurate model of the process that translates the bids that generators submit into the actual market prices they are paid for the electricity for all possible bids submitted by them and their competitors and all possible market demand realizations. The construction of a model of the price-setting process in NEM1 that is able to replicate actual market prices with reasonable accuracy is a necessary first step in the process of estimating cost functions from generator bidding behavior and market outcomes. Wolak (2000) devotes significant attention to demonstrating that the model of the price-setting process used here accurately reflects the actual price-setting process in NEM1.

In preparation for the empirical portion of the paper, I describe the two procedures for cost function estimation for NEM1 in Australia, although the modifications necessary to apply these methods to other competitive electricity markets and other bid-based markets are straightforward. In NEM1, each day of the market, d, is divided into the half-hour load periods i beginning with 4:00 A.M. to 4:30 A.M. and ending with 3:30 A.M. to 4:00 A.M. the following day. Let Firm A denote the generator whose bidding strategy is being computed. Define

Q_{id},	Total market demand in load period i of day d;
$\mathrm{SO}_{id}(p)$,	Amount of capacity bid by all other firms besides Firm A into the market in load period i of day d at price p;

[3] Electricity-generating plants are usually divided into multiple gensets or units. For example, a 2-GW plant will usually be divided into four 500-MW gensets.

$\mathrm{DR}_{id}(p) = Q_{id} - \mathrm{SO}_{id}(p)$,	Residual demand faced by Firm A in load period i of day d, specifying the demand faced by Firm A at price p;
QC_{id},	Contract quantity for load period i of day d for Firm A;
PC_{id},	Quantity-weighted average (over all hedge contracts signed for that load period and day) contract price for load period i of day d for Firm A;
$\pi_{id}(p)$,	Variable profits to Firm A at price p, in load period i of day d;
MC,	Marginal cost of producing a megawatt hour by Firm A; and
$\mathrm{SA}_{id}(p)$,	Bid function of Firm A for load period i of day d giving the amount it is willing to supply as a function of the price p.

For ease of exposition, I assume that MC, the firm's marginal cost, does not depend on the level of output it produces. For the general case of recovering marginal cost function estimates, I relax this assumption.

The market-clearing price p is determined by solving for the smallest price such that the equation $\mathrm{SA}_{id}(p) = \mathrm{DR}_{id}(p)$ holds. The magnitudes QC_{id} and PC_{id} are usually set far in advance of the actual day-ahead bidding process. Generators sign hedge contracts with electricity suppliers or large consumers for a pattern of prices throughout the day, week, and month, for an entire year or for a number of years. There is some short-term activity in the hedge contract market for electricity purchasers requiring price certainty for a larger or smaller than planned quantity of electricity at some point during the year.

In terms of the aforementioned notation, I can define the variable profits[4] (profits excluding fixed costs) earned by Firm A for load period i during day d at price p as

$$\pi_{id}(p) = \mathrm{DR}_{id}(p)(p - \mathrm{MC}) - (p - \mathrm{PC}_{id})\mathrm{QC}_{id}. \tag{3.1}$$

The first term is the variable profits from selling electricity in the spot market. The second term captures the payoffs to the generator from buying and selling hedge contracts. Assuming $\mathrm{QC}_{id} > 0$ (the generator is a net seller of hedge contracts), if $p > \mathrm{PC}_{id}$, the second term is the total payments made to purchasers of hedge contracts during that half-hour by Firm A. If $p < \mathrm{PC}_{id}$, the second term is the total payments made by purchasers of hedge contracts to Firm A. Once the market-clearing price is determined for the period, Equation (3.1) can be used to compute the profits for load period i in day d.

Financial hedge contracts impose no requirement on the generator to deliver actual electricity. These contracts are merely a commitment between the seller

[4] For the remainder of the paper, I use variable profits and profits interchangeably, with the understanding that I am always referring to variable profits.

(usually a generator) and the purchaser (usually a large load or load-serving entity) to make the payment flows described herein contingent on the value of the spot market-clearing price relative to the contract price. However, as discussed in detail in Wolak (2000), a generator that has sold a significant quantity of financial hedge contracts will find it optimal to bid more aggressively (to sell a larger quantity of energy in the spot market) than one that has sold little or no hedge contracts. This point can be illustrated by computing the first-order conditions for maximizing (3.1) with respect to p:

$$\pi'_{id}(p) = \mathrm{DR}'_{id}(p)(p - \mathrm{MC}) - (\mathrm{DR}_{id}(p) - \mathrm{QC}_{id}) = 0. \tag{3.2}$$

Because the residual demand curve is downward sloping and the firm can produce only a nonnegative quantity of electricity ($\mathrm{DR}_{id}(p) \geq 0$), the price that solves (3.2) for $\mathrm{QC}_{id} > 0$ is smaller than the price that solves (3.2) for $\mathrm{QC}_{id} = 0$. This result implies that, for the same values of MC and $\mathrm{DR}_{id}(p)$, the firm finds it profit maximizing to produce a larger amount of energy for $\mathrm{QC}_{id} > 0$ than it does for $\mathrm{QC}_{id} = 0$. Figure 1 from Wolak (2000) gives a graphical presentation of this logic. Another implication of the first-order condition (3.2) is that the contract price, PC_{id}, has no effect on the firm's profit-maximizing market-clearing price or output quantity. The level of the contract price simply determines the magnitude of the transfers that flow between the buyer and seller of the hedge contract. These same incentives to participate aggressively in the spot electricity market are also valid for a firm that has a contract guaranteeing physical delivery of QC_{id} units of electricity at price PC_{id} during hour i of day d.

The expression for Firm A's profits given in (3.1) illustrates two very important aspects of competitive electricity markets. First, unless a firm is able to move the market-clearing price by its bidding strategy, its profits are independent of its bidding strategy for a given hedge contract quantity and price. Given the market-clearing price, all of the terms in (3.1), the firm's actual variable profit function for load period i in day d, depend on factors unrelated to the bids it submits into the electricity market. Second, the difference between Equation (3.1) and the usual oligopoly model profit function is that the residual demand function $\mathrm{DR}_{id}(p)$ faced by Firm A is ex post directly observable given the bids submitted by all other market participants besides Firm A. As shown herein, the residual demand curve faced by Firm A at each price, p, is simply the aggregate demand function less the aggregate bid curve of all market participants besides Firm A, $\mathrm{DR}_{id}(p) = Q_{id} - SO_{id}(p)$.

In the standard oligopoly context, the residual demand faced by each market participant is not directly observable, because the aggregate demand function is not observable ex post. For example, in the Cournot duopoly model, the residual demand curve faced by one firm is simply the market demand, $D(p)$, less the quantity made available by that firm's competitor: $\mathrm{DR}(p) = D(p) - q_c$, where q_c is the quantity made available by the firm's competitor. Different from the case of a competitive electricity market, $D(p)$ is not directly observable, so that an estimate of $\mathrm{DR}(p)$ cannot be constructed without first econometrically estimating $D(p)$, the market aggregate demand function. In the case of a bid-based market such as electricity, even if load-serving entities could submit price-responsive demands, as is currently possible in most competitive electricity markets, the residual demand curve facing any competitor in these markets can be directly computed using all of the bids submitted by all other market participants.

Because this residual demand function can be constructed by the econometrician using bid data, there is no need to make a functional form assumption for the demand curve the bidder faces in order to compute its implied marginal cost for any level of output. Given a model for the price-setting process in this market and a behavioral model for bidders, implied marginal costs can be constructed for each observed level of output by Firm A.

The ex post observability of each generator's residual demand function has important implications for designing a competitive electricity market. Because the price elasticity of the residual demand curve faced by a bidder determines the extent of market power that it is able to exercise, the goal of the market design process is to face all bidders with a perfectly price-elastic residual demand function. Under these circumstances, no generator possesses market power. However, the residual demand curve faced by one market participant depends on the bids submitted by all other market participants. Therefore, aggressive bidding (very price-elastic bid supply functions) by a firm's competitors will leave it with a very elastic residual demand. This will cause the firm to bid very aggressively. This aggressive bidding will leave its competitors with elastic residual demand curves, which will cause them to bid more aggressively. This sequence of self-reinforcing aggressive bidding also works in the opposite direction to reinforce less price-elastic bidding. Specifically, if a firm bids a steep supply curve, that increases the incentive for its competitors to bid steep supply curves, because they now face more inelastic residual demand curves. Consequently, a very important aspect of the market design process is putting in place very strong incentives for aggressive spot market bidding.

Active participation by wholesale demanders in the forward electricity market is crucial to providing strong incentives for aggressive spot market bidding by generation unit owners. If a firm that owns significant generating capacity does not have a large fraction of this capacity tied up in forward contracts, then given the extreme inelasticity of the demand for electricity in any hour, this firm will find it profit maximizing to bid substantially in excess of its variable costs into the spot market during any hour that it knows some of its capacity is needed to meet total demand. Forward market commitments for a significant fraction of its capacity make this strategy less attractive because the firm earns the spot price only on any spot market sales in excess of its forward market commitments, rather than on all of its sales. In addition, because the number of firms that can compete to supply forward contracts far in advance of the delivery date is significantly greater than the number of firms that compete to supply electricity on a month-ahead, day-ahead, or hour-ahead basis, the price of electricity for delivery during these high-demand hours purchased far in advance is significantly cheaper than electricity purchased in short-term markets.

In particular, at time horizons greater than two years in advance of delivery, both existing and new entrants can compete to supply electricity. In contrast, a few hours before delivery, only the large generating units that are operating are able to compete to deliver electricity. Significant quantities of forward contracts guarantee that enough of these large generating units will be operating a few hours before delivery to ensure a workably competitive spot market for electricity. Wolak (2002) discusses these issues and the essential role of retail competition in developing an active forward electricity market.

I now introduce notation necessary to present the two procedures for recovering marginal cost estimates from bid data. Suppose that there are stochastic demand shocks to the price-setting process each period, and that Firm A knows the distribution of these shocks. This uncertainty could be due to the fact that Firm A does not exactly know the form of SO(p) – this function has a stochastic component to it – or it does not know the market demand used in the price-setting process when it submits its bids – Q is known only up to an additive error. Because I am not solving for an equilibrium bidding strategy, I do not need to be specific about the sources of uncertainty in the residual demand that Firm A faces. Regardless of the source of this uncertainty, Firm A will attempt to maximize profits conditional on the value of this uncertainty if the firm can observe it. If Firm A cannot observe this uncertainty, it will then choose its bids to maximize expected profits given an assumed distribution for this uncertainty. The two procedures for recovering the firm's underlying cost function from bid data differ in terms of their assumptions about whether the firm is able to achieve prices that maximize profits given the realization of this uncertainty or achieve only market prices that maximize expected profits taken with respect to the distribution of this uncertainty.

Let ε_i equal this shock to Firm A's residual demand function in load period i ($i = 1, \ldots, 48$). Rewrite Firm A's residual demand in load period i, accounting for this demand shock as $\mathrm{DR}_i(p, \varepsilon_i)$. Define $\Theta = (p_{11}, \ldots, p_{JK}, q_{1,11}, \ldots, q_{11,JK}, q_{2,11}, \ldots, q_{2,JK}, \ldots, q_{48,11}, \ldots, q_{48,JK})$ as the vector of daily bid prices and quantities submitted by Firm A. There are K increments for each of the J gensets owned by firm A. The rules of the NEM1 market require that a single price, p_{kj}, is set for each of the $k = 1, \ldots, K$ bid increments for each of the $j = 1, \ldots, J$ gensets owned by Firm A for the entire day. However, the quantity q_{ikj} made available to produce electricity in load period i from each of the $k = 1, \ldots, K$ bid increments for the $j = 1, \ldots, J$ gensets owned by Firm A can vary across $i = 1, \ldots, 48$ load periods throughout the day. In NEM1, the value of K is 10, so the dimension of Θ is $10J + 48 \times 10J$. Firm A owns a number of gensets so the dimension of Θ is more than several thousand. Let $\mathrm{SA}_i(p,\Theta)$ equal Firm A's bid function in load period i as parameterized by Θ. Note that by the rules of the market, bid increments are dispatched based on the order of their bid prices, from lowest to highest. This means that $\mathrm{SA}_i(p, \Theta)$ must be nondecreasing in p.

Figure 4.1 gives an example of two bid functions for different half-hours of the same day that are consistent with the rules of the Australian electricity

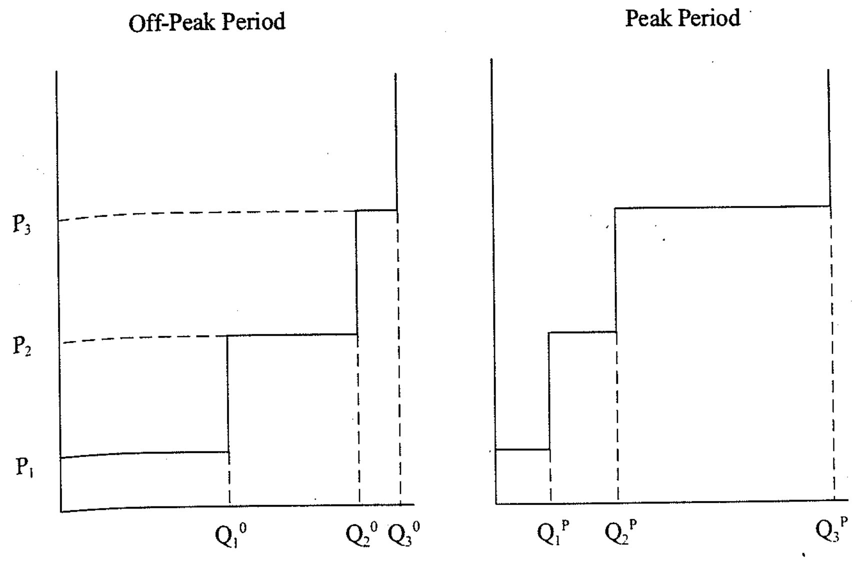

Figure 4.1. Sample bid functions for Australian electricity market.

market for the case of three bid price increments. Note that the only difference between the two bid curves is the quantity of energy that is made available at each price level during the two half-hours. Both the peak and off-peak period bid curves have the same price bids, as is required by NEM1 market rules, but the peak period bid curve assigns a large quantity of the capacity from the genset to the highest-price bid increment because the generator is reasonably confident that its highest-price bid will set the market-clearing price during this period. However, during the off-peak period that generator reduces the quantity that it bids at the highest price in order to be ensured that a significant fraction of its capacity will be sold at or above the intermediate bid price.

Let $p_i(\varepsilon_i, \Theta)$ denote the market-clearing price for load period i given the residual demand shock realization, ε_i, and daily bid vector Θ. It is defined as the solution in p to the equation $\mathrm{DR}_i(p, \varepsilon_i) = \mathrm{SA}_i(p, \Theta)$. Let $f(\varepsilon)$ denote the probability density function of $\varepsilon = (\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_{48})'$. Define

$$E(\Pi(\Theta)) = \int_0^\infty \cdots \int_0^\infty \sum_{i=1}^{48} [\mathrm{DR}_i(p_i(\varepsilon_i, \Theta), \varepsilon_i)(p_i(\varepsilon_i, \Theta) - \mathrm{MC}) - (p_i(\varepsilon_i, \Theta) - \mathrm{PC}_i)\,\mathrm{QC}_i] f(\varepsilon) d\,\varepsilon_1 \cdots d\varepsilon_{48} \tag{3.3}$$

as the expected profits to Firm A for the daily bid vector Θ.

Firm A's best-reply bidding strategy is the solution to the following optimization problem:

$$\max_{\Theta} E(\Pi(\Theta)) \text{ subject to } b_U \geq R\Theta \geq b_L\,. \tag{3.4}$$

Define Θ^* as the expected profit-maximizing value of Θ. Besides the extremely large dimension of Θ, there are several other reasons to expect this problem to be difficult to solve. First, in general, the realization of each residual demand function faced by Firm A is a nondecreasing, discontinuous step function, because the aggregate supply curve of all participants besides Firm A is a nondecreasing step function. Second, to compute the value of the objective function requires integrating with respect to a forty-eight-dimensional random vector ε. Most important, the dimension of Θ for Firm A is greater than 2,000. The linear inequality constraints represented by the matrix R and vectors of upper and lower bounds b_U and b_L imply that none of the q_{ik} can be negative and the sum of the q_{ik} relevant to a given genset cannot be greater than the capacity of the genset and that the prices for each bid increment cannot be smaller than −9,999.99 \$AU or larger than 5,000.00 \$AU. Wolak (2001a) computes this optimal bidding strategy for one market participant in the Australian electricity market and compares actual market outcomes with those that would exist under this optimal bidding strategy for a sample of days in NEM1.

At this point it is useful to compare the optimal bidding strategy problem given in (3.4) to the problem of computing an optimal supply function with demand uncertainty discussed in Klemperer and Meyer (1989) and applied to the electricity supply industry in England and Wales by Green and Newbery (1992). Rewrite Equation (3.1) with the residual demand function for load period i that includes the shock for period i as

$$\pi_{id}(p, \varepsilon_i) = \mathrm{DR}_{id}(p, \varepsilon_i)(p - \mathrm{MC}) - (p - \mathrm{PC}_{id})\mathrm{QC}_{id}. \tag{3.5}$$

Solving for the value of p that maximizes (3.5) yields $p_i{}^*(\varepsilon_i)$, which is the profit-maximizing market-clearing price given that Firm A's competitors bid to yield the residual demand curve, $\mathrm{DR}_{id}(p, \varepsilon_i)$, with demand shock realization, ε_i, for the hedge contract position, QC_{id} and PC_{id}. This optimal price also depends on QC_{id}, PC_{id}, and MC. I write it as $p_i{}^*(\varepsilon_i)$ to emphasize that it is Firm A's profit-maximizing price given the realization of ε_i. Because this price maximizes the ex post realized profits of Firm A, for the remainder of the paper I will refer to it as the *best-response price* for the residual demand curve $\mathrm{DR}_{id}(p, \varepsilon_i)$ with demand shock realization ε_i for the hedge contract position QC_{id} and PC_{id}. Substituting this value of p into the residual demand curve yields $\mathrm{DR}_{id}(p_i{}^*(\varepsilon_i), \varepsilon_i)$. This price and quantity combination yields Firm A the maximum profit that it can earn given the bidding behavior of its competitors and the demand shock realization, ε_i.

Klemperer and Meyer (1989) impose sufficient restrictions on the underlying economic environment – the demand function, cost functions, and distribution of demand shocks – so that tracing out the price–quantity pairs $(p_i{}^*(\varepsilon_i)$, $\mathrm{DR}_{id}(p_i{}^*(\varepsilon_i), \varepsilon_i))$ for all values of ε_i yields a continuous, strictly increasing equilibrium supply curve, $\mathrm{SA}_i(p)$, for their duopolists. For each demand shock realization, their supply curve yields the best-response price for each duopolist given the bidding strategies of its competitor. Because each realization of ε_i in the Klemperer and Meyer model is associated with a unique price–quantity pair, the symmetric equilibrium duopoly supply function in the Klemperer and Meyer model does not depend on the distribution of ε_i. For this same reason, the Klemperer and Meyer framework can allow ε_i to be only one dimensional.

However, the NEM1 market rules explicitly prohibit firms from submitting continuous, strictly increasing bid functions. They must submit step functions, where bid price remains constant for all forty-eight half-hours of the day, but the length of each price increment can change on a half-hourly basis. This market rule constrains the ability of generation unit owners to submit supply bids that set the ex post profit-maximizing price for all possible realizations of the ex post residual demand function they face. Therefore, the expected profit-maximizing step function bid function, $SA_i(p, \Theta)$, depends on the distribution of ε_i. For this reason, our best-reply bidding framework can allow for residual demand uncertainty, ε_i, that is multidimensional.

Because the market rules and market structure in NEM1 constrain the feasible set of price and quantity pairs that Firm A can bid in a given load period, it may be unable to achieve $p_i{}^*(\varepsilon_i)$ for all realizations of ε_i using its allowed bidding strategy. As noted herein, the allowed bidding strategy constrains Firm A to bid ten bid increments per genset, but, more importantly, the prices of these ten bid increments must be the same for all forty-eight load periods throughout the day. This may severely limit the ability of Firm A to achieve $p_i{}^*(\varepsilon_i)$. To the extent that it does, our best-response pricing procedure will yield unreliable estimates of the firm's underlying cost functions.

Best-response prices must yield the highest profits, followed by best-response bidding, because the former is based on the realization of ε_i as shown in (3.5), whereas the latter depends on the distribution of ε as shown in (3.3). The expected value of the generator's actual profits can only be less than or equal to the expected value of the best-response bidding profits. Recall that, by definition, the best-response price, $p_i{}^*(\varepsilon_i)$, yields the maximum profits possible given the bidding strategies of Firm A's competitors and the realized value of the residual demand shock, ε_i. The best-response bidding strategy that solves (3.3) for the expected profit-maximizing vector of allowable daily bid prices and quantities, Θ^*, yields the highest level of expected profits for Firm A within the set of allowable bidding strategies. Therefore, by definition, this bidding strategy should lead to average profits that are greater than or equal to Firm A's average profits from its current bidding strategy for the same set of competitors' bids and own hedge contract positions. The extent to which profits from a best-response bidding strategy lie below the maximum possible obtainable from best-response prices is not addressed here. Wolak (2001a) shows that a significant fraction of the difference between the actual variables profits earned by a firm in the Australian electricity market and the profits that it would earn from best-reply prices is due to the fact that the market rules constrain the ability of the firm to achieve $p_i{}^*(\varepsilon_i)$ for every realization of ε_i using a bidding strategy that respects the NEM1 market rules. In addition, given the high-dimensional strategy space available to Firm A, Wolak (2001a) also shows that a nonnegligible portion of the difference between the best-response pricing variable profits and variable profits under Firm A's current bidding strategy can be attributed to the use of bidding strategies that are not best response in the sense of not exactly solving the optimization problem (3.4).

Before both cost function estimation procedures are described in detail, it is useful to compare their properties. The best-response pricing approach has the advantage of computational simplicity and is broadly consistent with the approach used in the empirical IO literature, which uses a parametric model for demand and the assumption of profit-maximizing behavior to recover a cost function estimate. Similar to the empirical IO approach, this approach yields estimated marginal cost values for each observed level of output in the sample. The validity of the best-response pricing approach relies on the assumption that the firm is somehow able to achieve $p_i{}^*(\varepsilon_i)$ for every realization of ε_i. As discussed herein, this is unlikely to be strictly valid for NEM1. In contrast, the best-response bidding strategy approach respects all of the rules governing bidding behavior and market price determination in NEM1 and relies only on the assumption of bidding to maximize expected profits to recover cost function estimates. Because it imposes all of the restrictions on bidding behavior implied by the market rules, this approach is able to recover genset-level cost function estimates. If the assumptions necessary for the validity of the best-response pricing approach hold, then both approaches will yield valid cost function estimates, but the best-response bidding approach should yield more precise cost function estimates.

4. RECOVERING COST FUNCTION ESTIMATES FROM BEST-RESPONSE PRICES

This section describes a procedure for recovering marginal cost function estimates based on my model of best-response pricing. This procedure can also be used to recover estimates of a generator's forward hedge contract position. Recall that my assumption of best-response pricing does not impose any of the restrictions implied by the market rules on the behavior of the firm under consideration. This procedure assumes that the firm is able to observe the market demand and the bids submitted by all other market participants. It then constructs the realized value of its residual demand function implied by the market demand and these bids and then selects the profit-maximizing price associated with this residual demand given the firm's hedge contract position and marginal cost function. Because of its computational simplicity, this approach should be a useful diagnostic tool in recovering an estimate of a firm's marginal cost function or in diagnosing the extent of market power a firm might possess when the assumptions required for its validity are approximately true.

Let $C(q)$ denote the total variable cost associated with producing output level q. Rewrite the period-level profit function for Firm A in terms of this general variable cost function as

$$\pi(p) = \mathrm{DR}(p, \varepsilon)p - C(\mathrm{DR}(p, \varepsilon)) - (p - \mathrm{PC})\mathrm{QC}. \tag{4.1}$$

To compute the best-reply price associated with this realization of the residual demand function, $\mathrm{DR}(p, \varepsilon)$, differentiate (4.1) with respect to p and set the result equal to zero:

$$\pi'(p) = \mathrm{DR}'(p, \varepsilon)(p - C'(\mathrm{DR}(p, \varepsilon))) + (\mathrm{DR}(p, \varepsilon) - \mathrm{QC}) = 0. \tag{4.2}$$

This first-order condition can be used to compute an estimate of the marginal cost at the observed market-clearing price, p^E, as

$$C'(\mathrm{DR}(p^E, \varepsilon)) = p^E - (\mathrm{QC} - \mathrm{DR}(p^E, \varepsilon))/\mathrm{DR}'(p^E, \varepsilon). \tag{4.3}$$

$\mathrm{DR}(p^E, \varepsilon)$ can be directly computed by using the actual market demand and bid functions submitted by all other market participants besides Firm A. The market-clearing price, p^E, is directly observed. I also assume that QC is observed. Computing $\mathrm{DR}'(p^E, \varepsilon)$ is the only complication associated with applying (4.3) to obtain an estimate of the marginal cost of Firm A at $\mathrm{DR}(p^E, \varepsilon)$.

For most competitive electricity markets, bidders submit step functions rather than piecewise linear functions. Consequently, strictly speaking, $\mathrm{DR}'(p^E, \varepsilon)$ is everywhere equal to zero.[5] However, because of the large number of bid increments permitted for each generating facility in the Australian market – ten per generating unit – and the close to 100 generating units in the Australian electricity market, the number of steps in the residual demand curve facing any market participant is very large. In addition, because of the competition among generators to supply additional energy from their units, there are usually a large number of small steps in the neighborhood of the market-clearing price. Nevertheless, some smoothness assumption on $\mathrm{DR}(p, \varepsilon)$ is still needed to compute a value for $\mathrm{DR}'(p^E, \varepsilon)$ to use in Equation (4.3).

I experimented with a variety of techniques for computing $\mathrm{DR}'(p^E, \varepsilon)$ and found that the results obtained are largely invariant to the techniques used. One technique approximates $\mathrm{DR}'(p^E, \varepsilon)$ by $(\mathrm{DR}(p^E + \delta, \varepsilon) - \mathrm{DR}(p^E, \varepsilon))/\delta$, for values of δ ranging from ten Australian cents to one Australian dollar. Another technique approximates the residual demand function by

$$\mathrm{DR}(p, \varepsilon) = Q_d(\varepsilon) - \mathrm{SO}_h(p, \varepsilon), \tag{4.4}$$

where the aggregate bid supply function of all other market participants besides Firm A is equal to

$$\mathrm{SO}_h(p, \varepsilon) = \sum_{n=1}^{N} \sum_{k=1}^{10} qo_{nk}\, \Phi((p - po_{nk})/h). \tag{4.5}$$

[5] For the now-defunct California Power Exchange (PX), bidders submitted piecewise linear bid functions starting at point (0, 0) in price-quantity space and ending at $(2500, x)$ for any positive value of x. There were no limits on the number of these bid functions that any market participant could submit for a given hour. Therefore, the residual demand function facing any PX market participant was a piecewise linear function. Consequently, except at the points where two linear functions join, $\mathrm{DR}'(p^E, \varepsilon)$ is a well-defined concept.

Here qo_{nk} is the kth bid increment of genset n and po_{nk} is the bid price for increment k of genset n, where N is the total number of gensets in the market excluding those owned by Firm A. The function $\Phi(t)$ is the standard normal cumulative distribution function and h is a user-selected smoothing parameter. This parameter smooths the corners on the aggregate supply bid function of all other market participants besides Firm A. Smaller values of h introduce less smoothing at the cost of a value for $\mathrm{DR}'(p^E, \varepsilon)$ that may be at one of the smoothed corners. This second technique was adopted because it is very easy to adjust the degree of smoothing in the resulting residual demand function. Using this technique results in

$$\mathrm{DR}_h'(p, \varepsilon) = -\frac{1}{h} \sum_{n=1}^{N} \sum_{k=1}^{10} qo_{nk}\, \varphi((p - po_{nk})/h), \tag{4.6}$$

where $\varphi(t)$ is the standard normal density function. Using this method to compute $\mathrm{DR}'(p^E, \varepsilon)$, I can compute $C'(\mathrm{DR}(p^E, \varepsilon))$ by using Equation (4.3) for each market-clearing price.

There are variety of procedures to estimate the function $C'(q)$ given the $C'(q)$ and $q = \mathrm{DR}(p^E, \varepsilon)$ pairs implied by (4.3) applied to a sample of market-clearing prices and generator bids. In the empirical portion of the paper, I present a scatter plot of these $(C'(q), q)$ pairs and one estimate of $C'(q)$.

The first-order condition for best-reply pricing can also be used to compute an estimate of the value of QC for that half-hour for an assumed value for $C'(q)$ at that level of output. Rewriting (4.2) yields

$$\mathrm{QC} = (p^E - C'(\mathrm{DR}(p^E, \varepsilon)))\mathrm{DR}'(p^E, \varepsilon) + \mathrm{DR}(p^E, \varepsilon). \tag{4.7}$$

Different from the case of estimating the generator's marginal cost function, I expect QC to vary on a half-hourly basis both within and across days. Nevertheless, there are deterministic patterns in QC within the day and across days of the week. In the empirical portion of the paper, I quantify the extent to which the half-hourly implied values of QC follow the same deterministic patterns within the day and across days of the week as the actual values of QC.

In concluding this section, I find it important to emphasize that, strictly speaking, this procedure for estimating Firm A's marginal cost is valid only if the firm is somehow to able to obtain best-reply prices for all realizations of ε_i. As shown in Wolak (2000, 2001a), this is not possible because the market rules constrain the ability of expected profit-maximizing generators to set best-reply prices for all realizations from the joint distribution of forty-eight residual demand functions that Firm A faces each day. Nevertheless, as Section 7 shows, the deviation of actual prices from best-reply prices for Firm A is not so great as to make these calculations uninformative about Firm A's marginal cost function or its half-hourly hedge contract holdings. Given that these calculations are relatively straightforward to perform, I believe that they can be very useful diagnostic tools for computing marginal cost function estimates or forward contract position estimates that can be used in market power monitoring and analysis.

5. RECOVERING COST FUNCTION ESTIMATES FROM BEST-RESPONSE BIDDING

This section uses the assumption of best-response bidding, or, equivalently, bidding to maximize expected profits subject to the market rules on feasible bid functions, to recover estimates of genset-level marginal cost functions for Firm A. Imposing all of the bidding rules on the methodology used to recover the firm's marginal cost function will produce more accurate estimates than the methodology outlined in Section 4, even if the assumptions required for the validity of this simple approach hold. However, the procedure described here involves significantly more computational effort and econometric complexity. Specifically, I derive a generalized method of moments (GMM) estimation technique to recover genset-level cost functions for all of the units bid into the market by Firm A.

Deriving this estimation procedure requires additional notation. Define

$\mathrm{SA}_{ij}(p, \Theta)$,	The amount bid by genset j at price p during load period i;
$C_j(q, \beta_j)$,	The variable cost of producing output q from genset j;
β_j,	The vector of parameters of the cost function for genset j; and
$\mathrm{SA}_i(p, \Theta) = \sum_{j=1}^{J} \mathrm{SA}_{ij}(p, \Theta)$,	The total amount bid by Firm A at price p during load period i.

In terms of this notation, write the realized variable profit for Firm A during day d as

$$\begin{aligned}\Pi_d(\Theta, \varepsilon) = \sum_{i=1}^{48} \Big[& \mathrm{DR}_i(p_i(\varepsilon_i, \Theta), \varepsilon_i)\, p_i(\varepsilon_i, \Theta) \\ & - \sum_{j=1}^{J} C_j(\mathrm{SA}_{ij}(p_i(\varepsilon_i, \Theta), \Theta), \beta_j) \\ & - (p_i(\varepsilon_i, \Theta) - \mathrm{PC}_i)\,\mathrm{QC}_i \Big],\end{aligned}$$

where ε is the vector of realizations of ε_i for $i = 1, \ldots, 48$. As discussed herein, $p_i(\varepsilon_i, \Theta)$, the market-clearing price for load period i given the residual demand shock realization, ε_i, and daily bid vector Θ, is the solution in p to the equation $\mathrm{DR}_i(p, \varepsilon_i) = \mathrm{SA}_i(p, \Theta)$. As a way to economize on notation, in the development that follows I abbreviate $p_i(\varepsilon_i, \Theta)$ as p_i. The best-reply bidding strategy maximizes the expected value of $\Pi_d(\Theta, \varepsilon)$ with respect to Θ, subject to the constraints that all bid quantity increments, q_{ikj}, must be greater than or equal to zero for all load periods, i, bid increments, k, and gensets, j, and that for each genset the sum of bid quantity increments during each load period is less than the capacity, CAP_j, of genset j. As discussed earlier, there are also upper and lower bounds on the daily bid prices. However, Firm A's price bids for all bid increments, k, and gensets, j, and days, d, during my sample period are strictly below the upper bound and strictly above the lower bound.

This result allows me to use the first-order conditions for daily expected profit maximization with respect to Firm A's choice of the daily bid price increments to derive a GMM estimator for the genset-level cost function parameters. For all days, d, the moment restrictions implied by these first-order conditions are

$$E_{\varepsilon}\left(\frac{\partial \Pi_d(\Theta_d, \varepsilon)}{\partial p_{km}}\right) = 0 \tag{5.1}$$

for all gensets, m, and bid increments, k. I index Θ by d to denote the fact that there are different values of Θ for each day during the sample period. Equation (5.1) defines the $J \times K$ moment restrictions that I will use to estimate the parameters of the genset-level cost functions. The sample analog of this moment restriction is as follows:

$$\begin{aligned}\frac{\partial \Pi_d(\Theta_d, \varepsilon)}{\partial p_{km}} = \sum_{i=1}^{48}\Bigg[\Bigg(&\mathrm{DR}_i'(p_i(\varepsilon_i, \Theta), \varepsilon_i)\, p_i(\varepsilon_i, \Theta) \\ &+ (\mathrm{DR}_i(p_i(\varepsilon_i, \Theta), \varepsilon_i) - \mathrm{QC}_i) \\ &- \sum_{j=1}^{J} C_j'(\mathrm{SA}_{ij}(p_i(\varepsilon_i, \Theta)), \beta_j)\left(\frac{\partial\, \mathrm{SA}_{ij}}{\partial p_i}\right)\Bigg)\frac{\partial p_i}{\partial p_{km}} \\ &- \sum_{j=1}^{J} C_j'(\mathrm{SA}_{ij}(p_i(\varepsilon_i, \Theta)), \beta_j)\frac{\partial\, \mathrm{SA}_{ij}}{\partial p_{km}}\Bigg],\end{aligned} \tag{5.2}$$

where p_i is shorthand for the market-clearing price in load period i. Let $\ell_d(\beta)$ denote the $J \times K$ dimensional vector of partial derivatives given in (5.2), where β is the vector composed of β_j for $j = 1, \ldots, J$. Assuming that the functional form for $C_j(q, \beta_j)$ is correct, the first-order conditions for expected profit maximization with respect to daily bid prices imply that $E(\ell_d(\beta^0)) = 0$, where β^0 is the true value of β. Consequently, solving for the value of b that minimizes

$$\left[\frac{1}{D}\sum_{d=1}^{D}\ell_d(b)\right]'\left[\frac{1}{D}\sum_{d=1}^{D}\ell_d(b)\right] \tag{5.3}$$

will yield a consistent estimate of β. Let $b(I)$ denote this consistent estimate of β, where I denotes the fact that the identity matrix is used as the GMM weighting matrix. I can construct a consistent estimate of the optimal GMM weighting matrix using this consistent estimate of β as follows:

$$V_D(b(I)) = \frac{1}{D}\sum_{d=1}^{D}\ell_d(b(I))\,\ell_d(b(I))'. \tag{5.4}$$

The optimal GMM estimator finds the value of b that minimizes

$$\left[\frac{1}{D}\sum_{d=1}^{D}\ell_d(b)\right]' V_D(b(I))^{-1}\left[\frac{1}{D}\sum_{d=1}^{D}\ell_d(b)\right]. \tag{5.5}$$

Let $b(O)$ denote this estimator, where O denotes the fact this estimator is based on a consistent estimate of the optimal weighting matrix.

Operationalizing this estimation procedure requires computing values for the partial derivative of $\mathrm{SA}_{ij}(p, \Theta)$ with respect to p and p_{km} and the partial derivative of $p_i(\varepsilon_i, \Theta)$ with respect to p_{kj}. I use the same smoothing technique used in the previous section to compute the derivative of the residual demand function with respect to the market price to compute these partial derivatives. Define $\mathrm{SA}_{ij}^h(p, \Theta)$ as

$$\mathrm{SA}_{ij}^h(p, \Theta) = \sum_{k=1}^{10} q_{ikj}\,\Phi((p - p_{kj})/h), \tag{5.6}$$

which implies

$$\mathrm{SA}_i^h(p, \Theta) = \sum_{j=1}^{J}\sum_{k=1}^{10} q_{ikj}\,\Phi((p - p_{kj})/h). \tag{5.7}$$

This definition of $\mathrm{SA}_{ij}(p, \Theta)$ yields the following two partial derivatives:

$$\begin{aligned} \frac{\partial\,\mathrm{SA}_{ij}}{\partial p} &= \frac{1}{h}\sum_{k=1}^{10} q_{ikj}\,\phi((p - p_{kj})/h) \quad \mathit{and} \\ \frac{\partial\,\mathrm{SA}_{ij}}{\partial p_{kj}} &= -\frac{1}{h}\, q_{ikj}\,\phi((p - p_{kj})/h). \end{aligned} \tag{5.8}$$

The final partial derivative required to compute the sample analog of (5.1) can be computed by applying the implicit function theorem to the equation $\mathrm{DR}_i(p, \varepsilon_i) = \mathrm{SA}_i(p, \Theta)$. This yields the expression

$$\frac{\partial p_i(\varepsilon_i, \Theta))}{\partial p_{kj}} = \frac{\dfrac{\partial\,\mathrm{SA}_i(p_i(\varepsilon_i, \Theta), \Theta)}{\partial p_{kj}}}{\mathrm{DR}_i'(p_i(\varepsilon_i, \Theta), \varepsilon_i) - \mathrm{SA}_i'(p_i(\varepsilon_i, \Theta), \Theta)}, \tag{5.9}$$

where the derivative of the residual demand curve with respect to price used in this expression is given in Equation (4.6) and the other partial derivatives are given in (5.8). Given data on market-clearing prices and the bids for all market participants, I can compute all of the inputs into Equation (5.2). I only need to choose a value for h, the smoothing parameter that enters the smoothed residual demand function and the smoothed bid functions of Firm A. Once this smoothing parameter has been selected, the magnitudes given in (5.8) and (5.9) remain constant for the entire estimation procedure.

The final step necessary to implement this estimation technique is choosing the functional form for the marginal cost function for each genset. Firm A owns two power plants. One power plant has four identical gensets that the firm operates during the sample period. I refer to this facility as Plant 1. The gensets at Plant 1 have a maximum capacity of 660 MW and a lower operating limit of 200 MW. The other power plant has three identical gensets that the firm operates during the sample period. I refer to this facility as Plant 2. The gensets at Plant 2 have a maximum capacity of 500 MW and a lower operating limit of 180 MW. Because it is physically impossible for a genset to supply energy safely at a rate below its lowest operating limit, I specify a functional form for marginal cost to take this into account. Consequently, I assume the following parametric functional forms for the two unit-level marginal cost functions:

$$C_1'(q, \beta_1) = \beta_{10} + \beta_{11}(q - 200) + \beta_{12}(q - 200)^2, \tag{5.10}$$

$$C_2'(q, \beta_2) = \beta_{20} + \beta_{21}(q - 180) + \beta_{22}(q - 180)^2. \tag{5.11}$$

These functional forms are substituted into (5.2) to construct the sample moment restrictions necessary to construct the objective function I minimize to estimate

$$\beta = (\beta_{10}, \beta_{11}, \beta_{12}, \beta_{20}, \beta_{21}, \beta_{22})'.$$

Recall that for each genset the value of q entering (5.10) and (5.11) in the estimation procedure is the actual level of output produced by that unit during the half-hour period under consideration. I now turn to summarizing the relevant features of the NEM1 market in Australia necessary to understand the empirical work.

6. OVERVIEW OF NEM1

The Victoria Power Exchange (VPX) is the longest-running wholesale electricity market in Australia. It was established under the Electricity Industry (Amendment) Act of 1994 and formally began operation on July 1, 1994. The New South Wales (NSW) State Electricity Market (SEM) began operation May 10, 1996. NEM1 is the competitive electricity market established jointly by NSW and Victoria on May 4, 1997. It introduced unrestricted competition for generation dispatch across the two states; that is, the cheapest available generation, after allowing for transmission losses and constraints, is called on regardless of where it is located, and all wholesale energy is traded through the integrated market. The spot price in each state is determined with electricity flows in and between the state markets based on competitive bids or offers received in both markets.

The formation of NEM1 started the harmonization of the rules governing the operation of the two markets in Victoria and NSW. The market structures of the two electricity supply industries in Victoria and NSW are similar in terms of the relative sizes of the generation firms and the mix of generation capacity by fuel type, although the NSW industry is a little less than twice the size (as measured by installed capacity) of the Victoria industry and the largest three generators in NSW control a larger fraction of the total generation capacity in their state than the three largest generators in Victoria control in their state.

6.1. Market Structure in NEM1

Restructuring and privatization of the State Electricity Commission of Victoria (SECV) in 1994 took place at the power station level.[6] Each power station was formed into a separate entity to be sold. All former SECV generation capacity is now privately owned. The new owners are from within Australia and abroad. Currently there are eight generating companies competing in the VPX. The NSW–SEM has four generators competing to supply power. All generating assets are still owned by the NSW government. There are seven corporatized state-owned electricity distribution and supply companies serving NSW and the Australian Capital Territory (ACT). The eventual goal is to privatize both the generation and supply companies.

In both Victoria and NSW, there is an accounting separation within the distribution companies between their electricity distribution business and their electricity supply business. All other retailers have open and nondiscriminatory access to any of the other distribution companies wires. In NSW, the high-voltage transmission grid remains in government hands. In Victoria, the high-voltage transmission grid was initially owned by the government and was called PowerNet Victoria. It was subsequently sold to the New Jersey-based U.S. company, GPU, and renamed GPU-PowerNet. In NSW it is called TransGrid. Both the state markets operating under NEM1 – SEM in NSW and VPX in Victoria – were state-owned corporatized entities separate from the bulk transmission entities.

During 1997, the year of our sample, peak demand in Victoria was approximately 7.2 GW. The maximum amount of generating capacity that could be supplied to the market was approximately 9.5 GW. Because of this small peak demand, and despite the divestiture of generation to the station level, three of the largest baseload generators had sufficient generating capacity to supply at least 20 percent of this peak demand. More than 80 percent of the generating plant is coal fired, although some of this capacity does have fuel-switching capabilities. The remaining generating capacity is shared equally between gas turbines and hydroelectric power.

During 1997, the NSW market had a peak demand of approximately 10.7 GW and the maximum amount of generating capacity that could be supplied to the market was approximately 14 GW. There were two large generation companies, each of which controlled coal-fired capacity sufficient to supply more than 40 percent of NSW peak demand. The remaining large generators had hydroelectric, gas turbine, and coal-fired plants. The Victoria peak demand tends to occur during the summer month of January, whereas peak demand in NSW tends to occur in the winter month of July.

The full capability of the transmission link between the two states is nominally 1,100 MW from Victoria to NSW, and 1,500 MW in the opposite direction,

[6] Wolak (1999) provides a more detailed discussion of the operating history of the VPX and compares its market structure, market rules, and performance to the markets in England and Wales, Norway and Sweden, and New Zealand.

although this varies considerably, depending on temperature and systems conditions. If there are no constraints on the transfer between markets, then both states see the same market price at the common reference node. If a constraint limits the transfer, then prices in both markets diverge, with the importing market having a higher price than the exporting market.

6.2. Market Rules in NEM1

With a few minor exceptions, NEM1 standardized the price-setting process across the two markets. Generators are able to bid their units into the pool in ten price increments that cannot be changed for the entire trading day – the twenty-four-hour period beginning at 4:00 A.M. and ending at 4:00 A.M. the next day. The ten quantity increments for each genset can be changed on a half-hourly basis. Demanders can also submit their willingness to reduce their demand on a half-hourly basis as a function of price according to these same rules. Nevertheless, there is very little demand-side participation in the pool. A few pumped storage facilities and iron smelter facilities demand-side bid, but these sources total less than 500 MW of capacity across the two markets. All electricity is traded through the pool at the market price, and all generators are paid the market price for their energy.

7. RECOVERING IMPLIED MARGINAL COST FUNCTIONS AND HEDGE CONTRACT QUANTITIES

This section presents the results of applying the procedures described in Sections 4 and 5 to actual market outcomes. This requires collecting data on generator bids and market outcomes for a time period in which I also have information on the half-hourly values of QC, the quantity of the firm's forward contract obligations. I was able to obtain this information for a market participant in the Australian market for the period from May 15, 1997 to August 24, 1997. As discussed earlier, a major source of potential error in this analysis is the possibility that I have not adequately modeled the actual price-setting process in the Australian electricity market. Wolak (2000, Section 5) compares different models of the half-hourly price-setting procedure to determine which one does the best job of replicating observed half-hourly prices. This analysis found that the process I use in this paper – setting the half-hourly price equal to the price necessary to elicit sufficient supply from the aggregate half-hourly bid supply curve to meet the half-hourly market demand – replicates actual prices with sufficient accuracy.

I first compute implied marginal cost estimates using bid data submitted by Firm A's competitors, actual market prices, and total market demand. To give some idea of the range of residual demand curves faced by Firm A within the same day, Figures 4.2 and 4.3 plot the actual ex post residual demand curve faced by a firm in a representative off-peak demand period and on-peak demand period for July 28, 1997. These curves have been smoothed using

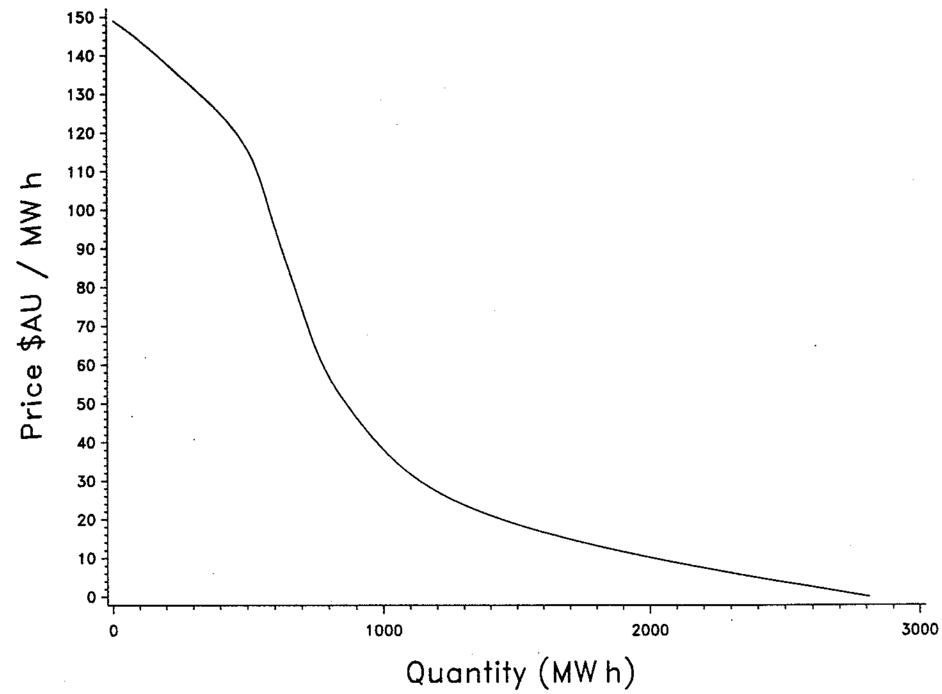

Figure 4.2. Residual demand curve for July 28, 1997 low demand.

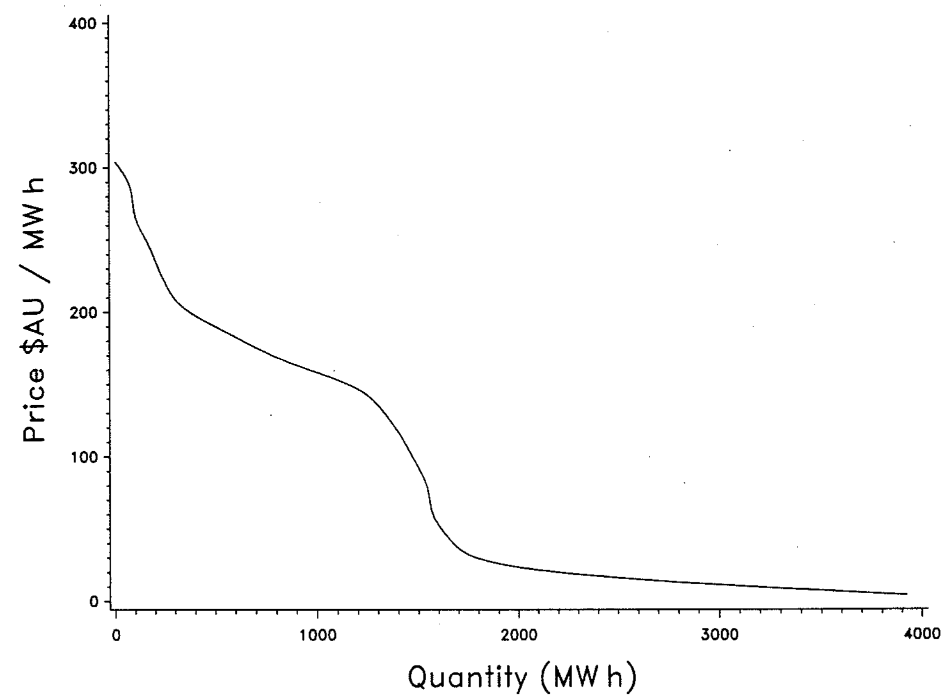

Figure 4.3. Residual demand curve for July 28, 1997 high demand.

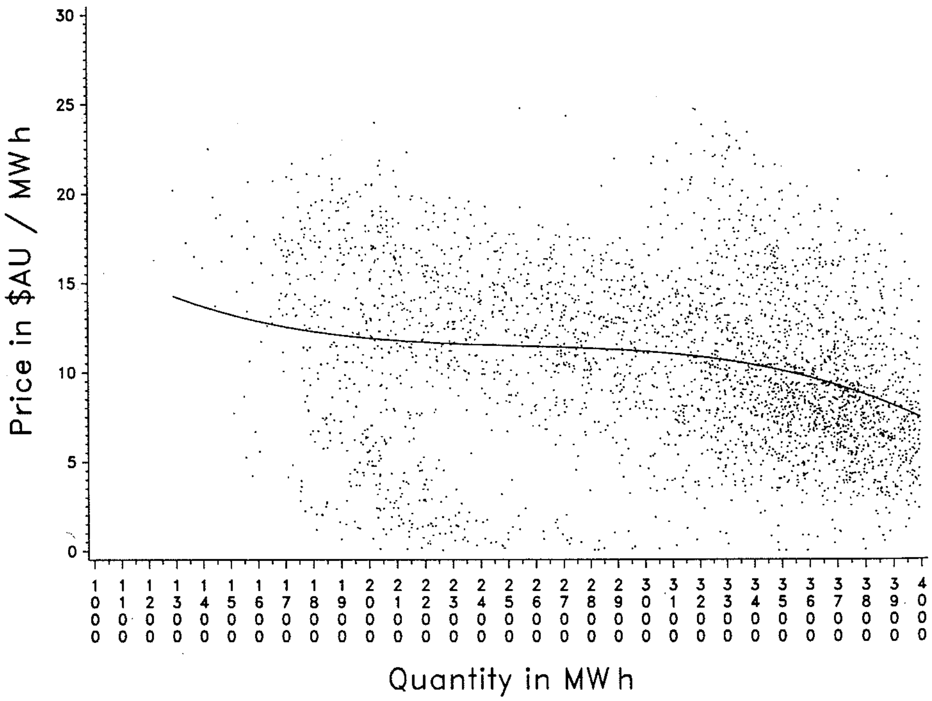

Figure 4.4. Implied marginal cost.

the expression for the residual demand curve given in (4.4) and (4.5), using a value of $h = 1$ \$AU. These curves confirm the standard intuition that Firm A possesses greater opportunities to exercise its market power during high market demand periods as opposed to low market demand periods. At every price level, Firm A faces a significantly higher residual demand during the high-demand load period than in the low-demand load period. I use the value of h employed to plot these figures for all of the results reported in this section. Repeating these results for values of $h = 10$ \$AU and 0.10 \$AU did not substantially alter any of the conclusions reported in the paragraphs that follow.

Figure 4.4 is a plot of the marginal cost and associated output demanded pairs, $(C'(\mathrm{DR}(p^E, \varepsilon)), \mathrm{DR}(p^E, \varepsilon))$, for all of the half-hourly market-clearing prices, p^E. The figure also plots the predicted values from the following cubic regression of the implied marginal cost, $C'(q)$, on q, the associated implied output of Firm A:

$$C'(q) = a + bq + cq^2 + dq^3 + \eta.$$

Table 4.1 gives the results of this implied marginal cost function regression. Although there is a considerable amount of noise in the sample of implied marginal cost and output pairs, the regression results are broadly consistent with the magnitudes of marginal costs implied by the heat rates and fuel prices of the facilities owned by Firm A. In particular, in discussions with the management of Firm A, they confirmed that their best estimates of their own marginal costs fluctuate between 15 \$AU/MWh and 10 \$AU/MWh.

Table 4.1. *Implied marginal cost regression: $C(q) = a + bq + cq^2 + dq^3$*

Parameter	Estimate	Standard Error	t-Stat
a	31	4.43	7.00
b	-2.21×10^{-2}	4.61×10^{-3}	−4.78
c	8.47×10^{-6}	1.54×10^{-6}	5.49
d	-1.11×10^{-9}	1.61×10^{-10}	−6.65

I now examine the accuracy of my procedure for estimating the half-hourly values of QC. This process requires selecting values for the marginal cost at any given output level. Consistent with the usual circumstances one is likely to face in market monitoring, I assume that the form of the generator's cost function is unknown. Therefore, I perform this procedure by assuming a rough estimate of the firm's marginal cost function. I assume a constant value for the marginal cost for all levels of output. More sophisticated marginal cost functions can be assumed, but the results with constant marginal cost should be the most informative about usefulness of this technique in market monitoring, because more accurate information on generation unit heat rates and output rates are generally unavailable. The two values of MC, the fixed level of marginal cost used in the procedure to recover estimates of QC for each half-hour, bound the fitted marginal cost curve given in Figure 4.4. These values are 10 \$AU/MWh and 15 \$AU/MWh. Figures 4.5 and 4.6 plot the implied half-hourly values of QC and the associated actual half-hourly values of QC for MC equal to 10 \$AU/MWh and 15 \$AU/MW h, respectively. Both of the figures also have a graph of the line QC(implied) = QC(actual) to illustrate visually how closely my procedure tracks the desired relationship between these two magnitudes. Both values of the marginal cost show a clear positive correlation between QC(implied) and QC(actual), although the consistency with the desired relationship seems greatest for an MC equal to 10 \$AU/MWh.

The values of QC(actual) vary on a half-hourly basis within the day and across days. However, there are still systematic patterns to these changes within the day and across days. On average, QC(actual) is higher on weekdays than weekends and higher during the peak hours of the day than the off-peak hours of the day. Consequently, another way to determine the usefulness of my procedure is to see if it captures the systematic variation in the values of QC(actual) within the day and across days of the week. To do this, estimate the following regression for both QC(actual) and QC(implied) for all load periods, $i = 1, \ldots, 48$, and days, $d = 1, \ldots, D$:

$$\mathrm{QC}(J)_{id} = \alpha + \sum_{m=1}^{3} \rho_m \,\mathrm{DMN}(m)_{id} + \sum_{p=1}^{6} \gamma_{p\mathrm{DWKD}}(p)_{id} + \sum_{r=1}^{47} \psi_r \,\mathrm{DPD}(r)_{i,d} + \nu_{id}, \tag{7.1}$$

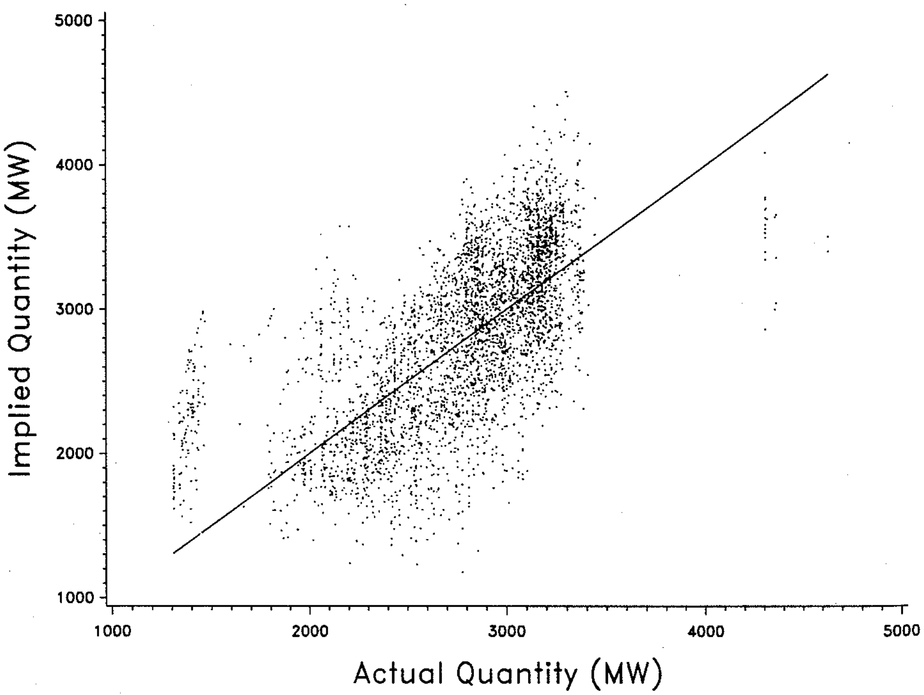

Figure 4.5. (MC = 10 \$AU/MW h) Implied vs. actual contract quantities.

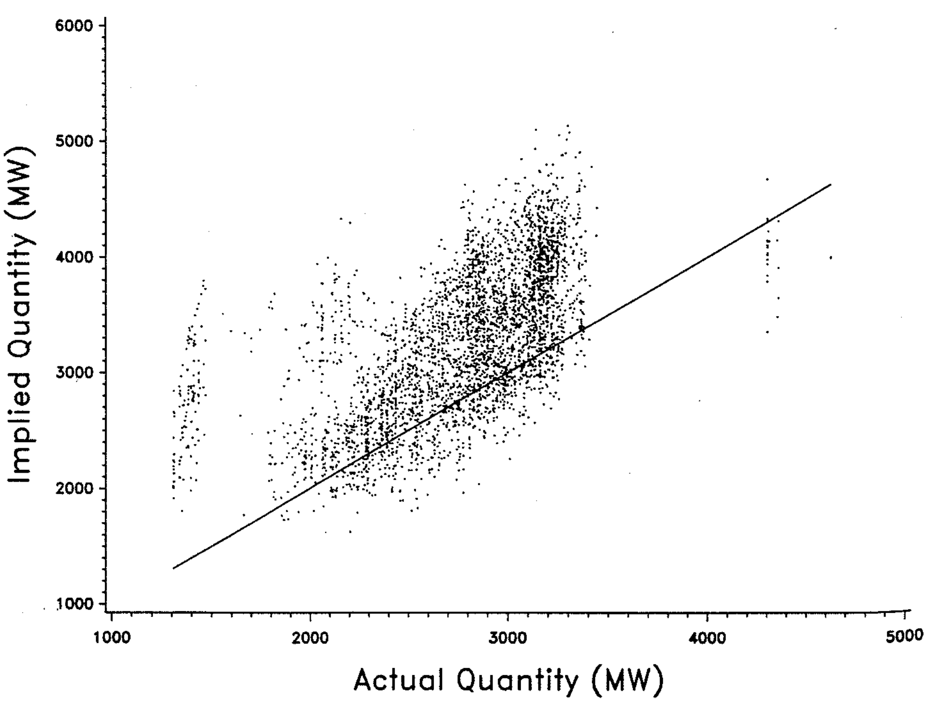

Figure 4.6. (MC = 15 \$AU/MW h) Implied vs. actual contract quantities.

where $\mathrm{DMN}(m)_{id}$ is a dummy variable equal to one when day d is in month m and zero otherwise, $\mathrm{DWKD}(p)_{id}$ is a dummy variable equal to one when day d is on day-of-the week p and zero otherwise, and $\mathrm{DPD}(r)_{id}$ is dummy variable equal to one when load period i is in load period-within-the-day r and zero otherwise. I compute estimates of the ρ_m, γ_p, and ψ_r for both QC(actual) and QC(implied), by estimating (7.1) by ordinary least squares. Table 4.2 reports the results of this estimation for QC(implied) with MC set equal to 10 \$AU/MW h. Figures 4.7 and 4.8 plot estimated values of γ_p ($p = 1, \ldots, 6$) and ψ_r ($r = 1, \ldots, 48$) for QC(implied) with MC equal to 10 \$AU/MWh and for QC(actual). The first value of γ_p is associated with Sunday and the excluded day of the week is Saturday. The first value of ψ_r is the half-hour beginning at 4:00 A.M. and ending at 4:30 A.M. and the excluded load period is the one beginning at 3:30 A.M. and ending at 4:00 A.M. the following day. These figures show a remarkable level of agreement between the deterministic part of the within-day and across-day variation in QC(implied) and QC(actual). These results provide strong evidence that even applying my procedure with this crude assumed marginal cost function yields a considerable amount of information about the level and pattern of forward contracting that a firm has undertaken.

At this point is it important to note that a generation unit owner's forward contract position is generally unknown to the market monitor. However, the analysis given here demonstrates that, with the use of the assumption of expected profit maximization with data on actual bidding behavior, something that is readily available to the market monitor, accurate estimates of the hourly levels of forward contract obligations can be obtained. Consequently, even this very rough procedure, which relies on best-response pricing, can be a very powerful tool for determining those instances when a market participant is likely to attempt to exercise market power in a spot electricity market. As shown in Wolak (2000), a generation unit owner's forward contract position is a very important determinant of the amount of market power that it is able to exercise in a spot electricity market.

I now examine the properties of my procedure for recovering genset-level marginal cost functions implied by best-reply bidding. As discussed in Section 5, Firm A has two types of identical gensets. Consequently, I estimate two genset-level marginal cost functions, applying the GMM estimation technique outlined in that section. I compute estimates of these unit-level marginal cost functions using both the identity matrix and a consistent estimate of the optimal weighting matrix. For all of the estimations reported, I assume $h = 1$ \$AU, although the results did not change appreciably for values of h ranging from 0.10 \$AU to 50 \$AU.

Table 4.3 reports the results of estimating $C_1'(q, \beta_1)$ and $C_2'(q, \beta_2)$, using the identity matrix as the weighting matrix and a consistent estimate of the optimal weighting matrix. Wolak (2001b) proves the consistency of these parameter estimates under the assumption that h tends to zero as the number of observations, D, tends to infinity. This paper also derives an expression for the variance

Table 4.2. *Contract quantity regression for QC(implied) for (MC = $10)*

Variable	Estimate	Standard Error	t-Stat
Constant	1882.37	59.70	31.53
DWKD1	−126.34	25.02	−5.05
DWKD2	215.72	25.95	8.31
DWKD3	282.29	25.64	11.01
DWKD4	330.84	25.54	12.95
DWKD5	391.34	25.01	15.65
DWKD6	397.24	25.45	15.61
DMN1	−12.46	40.29	−0.31
DMN2	34.90	39.19	0.89
DMN3	398.21	39.10	10.19
DPD1	−165.64	67.18	−2.47
DPD2	−147.10	67.42	−2.18
DPD3	−26.59	65.32	−0.41
DPD4	−54.94	65.11	−0.84
DPD5	113.07	65.90	1.72
DPD6	315.23	66.53	4.74
DPD7	602.74	65.90	9.15
DPD8	581.09	66.56	8.73
DPD9	654.61	67.07	9.76
DPD10	674.01	66.62	10.12
DPD11	743.62	66.84	11.13
DPD12	736.64	67.07	10.98
DPD13	777.06	66.18	11.74
DPD14	822.59	66.18	12.43
DPD15	898.80	66.18	13.58
DPD16	877.76	66.16	13.27
DPD17	820.91	66.39	12.37
DPD18	800.76	66.15	12.10
DPD19	757.67	66.15	11.45
DPD20	714.33	66.15	10.80
DPD21	706.38	65.93	10.71
DPD22	648.04	65.72	9.86
DPD23	663.86	65.93	10.07
DPD24	692.77	65.93	10.51
DPD25	741.97	66.14	11.22
DPD26	876.84	65.95	13.30
DPD27	866.17	65.91	13.14
DPD28	793.67	67.74	11.72
DPD29	739.79	67.68	10.93
DPD30	624.01	68.45	9.12
DPD31	695.82	66.32	10.49
DPD32	770.62	66.36	11.61
DPD33	879.34	66.82	13.16
DPD34	858.81	66.38	12.94
DPD35	848.25	67.27	12.61
DPD36	623.35	66.17	9.42
DPD37	739.35	67.07	11.02
DPD38	522.93	66.41	7.87
DPD39	462.97	67.04	6.91
DPD40	432.56	66.62	6.49
DPD41	421.45	67.06	6.28
DPD42	300.60	67.06	4.48
DPD43	145.39	67.06	2.17
DPD44	138.65	66.60	2.08
DPD45	64.44	66.60	0.97
DPD46	87.50	66.13	1.32
DPD47	55.02	65.68	0.84

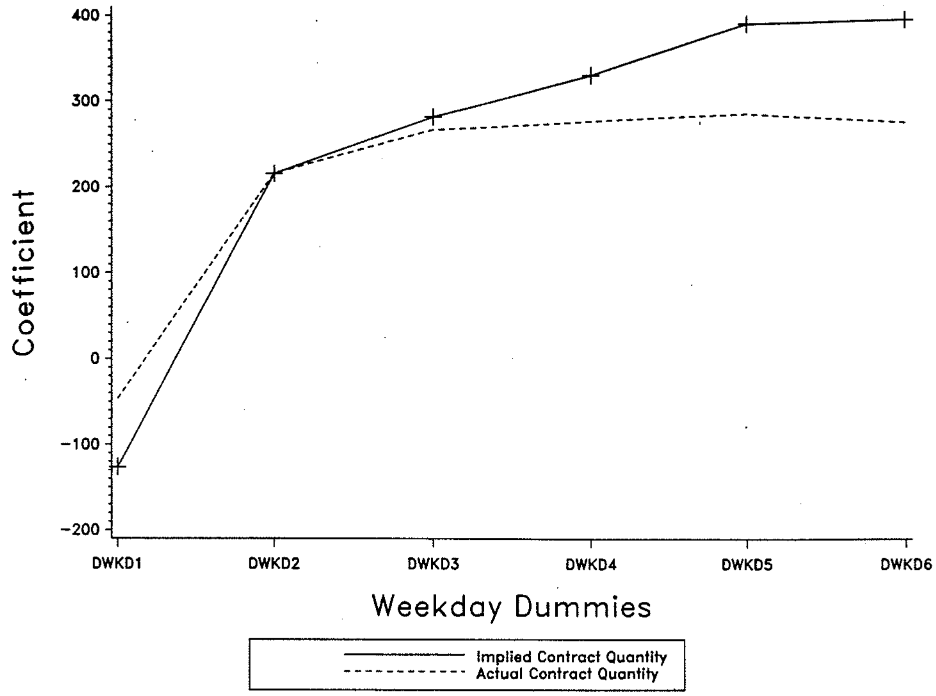

Figure 4.7. Contract quantity regression (MC = $AU10).

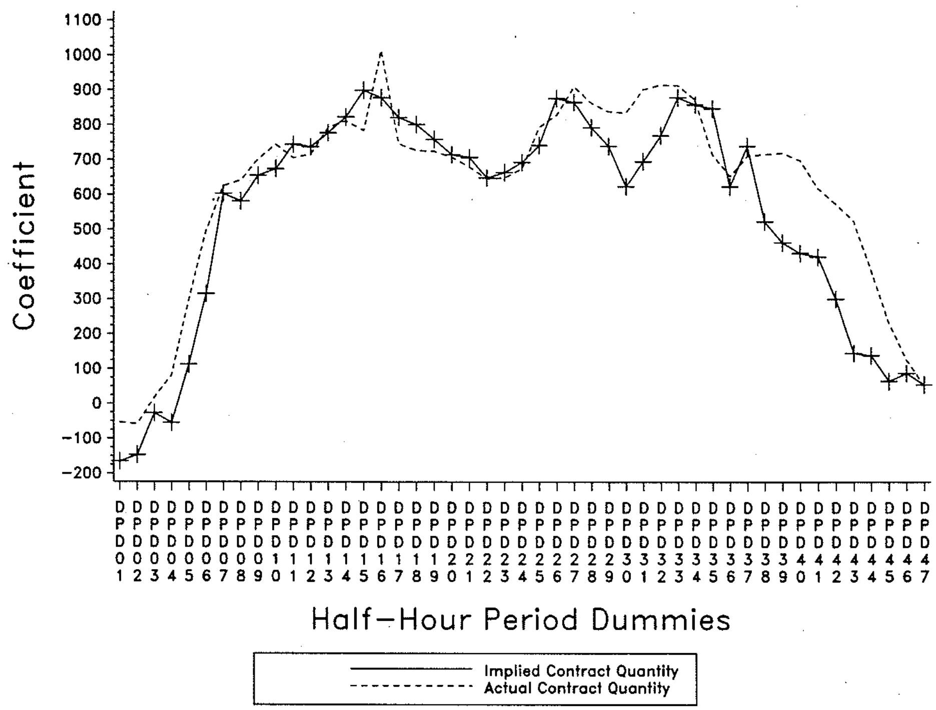

Figure 4.8. Contract quantity regression (MC = $AU10).

Table 4.3. *Genset-level marginal cost functions*

	Plant 1		Plant 2	
	Identity	Optimal	Identity	Optimal
β_{0k}	10.1	9.32	4.36	12.14
SE(β_{0k})	(1.23)	(1.14)	(1.53)	(0.74)
β_{1k}	−0.002	0.00103	−0.000448	0.0017
SE(β_{1k})	(0.006)	(0.000087)	(0.0041)	(0.000784)
β_{2k}	0.00000669	0.0000917	0.00031	0.0000686
SE(β_{2k})	(0.00001)	(0.00001)	(0.000085)	(0.00001)

Note: SE(β) = estimated standard error of the coefficient estimate, using the asymptotic covariance matrix given in Hansen (1982).

of the asymptotic normal distribution of these parameter estimates under the same assumptions.

The coefficient estimates are fairly precisely estimated across the four columns of results. As expected, the GMM estimates using a consistent estimate of the optimal weighting matrix appear to be more precisely estimated. The optimized value of the objective function from the GMM estimation with the consistent estimate of the optimal weighting matrix can be used to test the overidentifying restrictions implied by best-reply bidding. To estimate the six parameters of $C_1'(q, \beta_1)$ and $C_2'(q, \beta_2)$, I use seventy moment restrictions – ten bid increments for seven gensets. From the results of Hansen (1982), the optimized value of the objective function is asymptotically distributed as a chi-squared random variable with 64 degrees of freedom – the number of moment restrictions less the number of parameters estimated – under the null hypothesis that all of the moment restrictions imposed to estimate the parameters are valid. The optimized value of the objective function using a consistent estimate of the optimal weighting matrix is 75.40, which is less than the 0.05 critical value from a chi-squared random variable with 64 degrees of freedom. This implies that the null hypothesis of the validity of the moment restrictions given in (5.1) cannot be rejected by the actual bid data. This hypothesis test implies that given the parametric genset-unit cost functions in Equations (5.10) and (5.11), the overidentifying moment restrictions implied by the assumption of expected profit-maximizing behavior by Firm A cannot be rejected.

Figures 4.9 and 4.10 plot the estimated genset-level marginal cost functions for Plant 1 and Plant 2 along with pointwise 95 percent confidence intervals for the case of the consistent estimate of the optimal weighting matrix estimation results. Using the identity matrix as the GMM weighting matrix did not yield significantly different results. The confidence intervals indicate that the marginal cost curves are fairly precisely estimated. The results are broadly consistent with the results for the case of best-reply pricing. However, considerably more insight about the structure of Firm A's costs can be drawn from these results. Specifically, these results indicate the Plant 1 gensets are, for the same

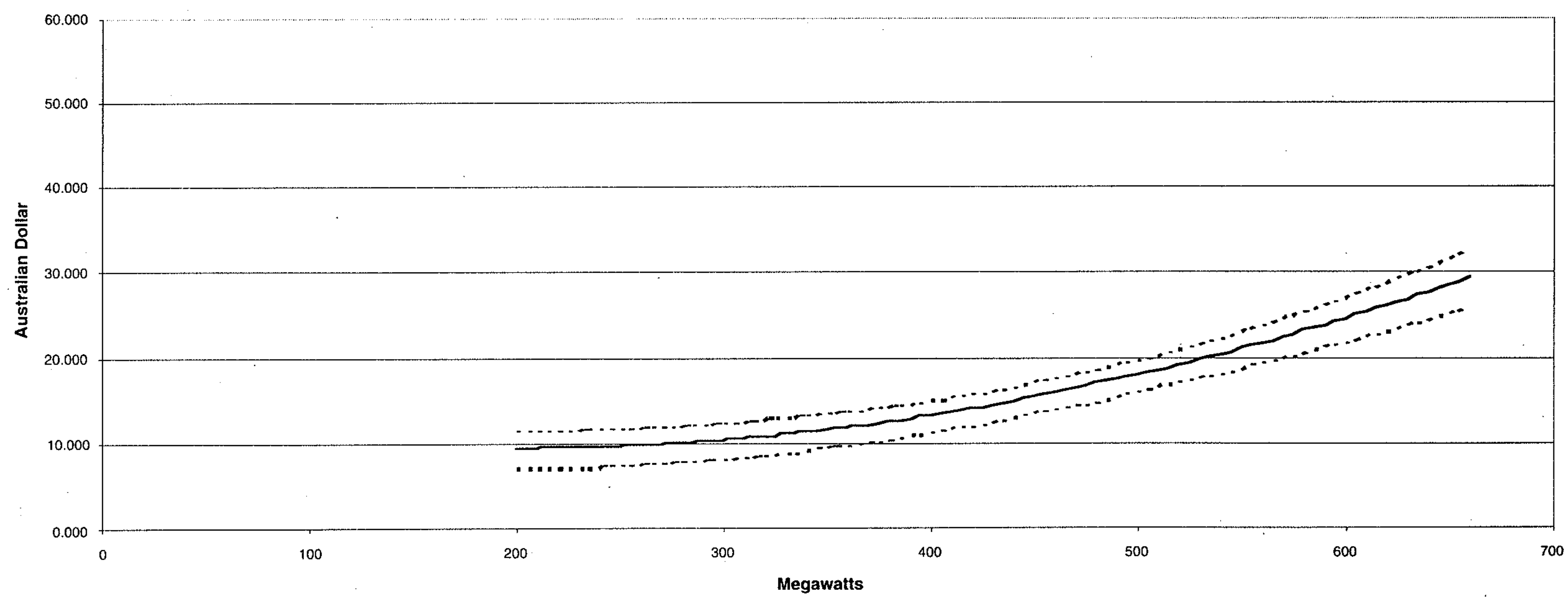

Figure 4.9. Marginal costs for Plant 1 of Firm A with optimal weighting matrix.

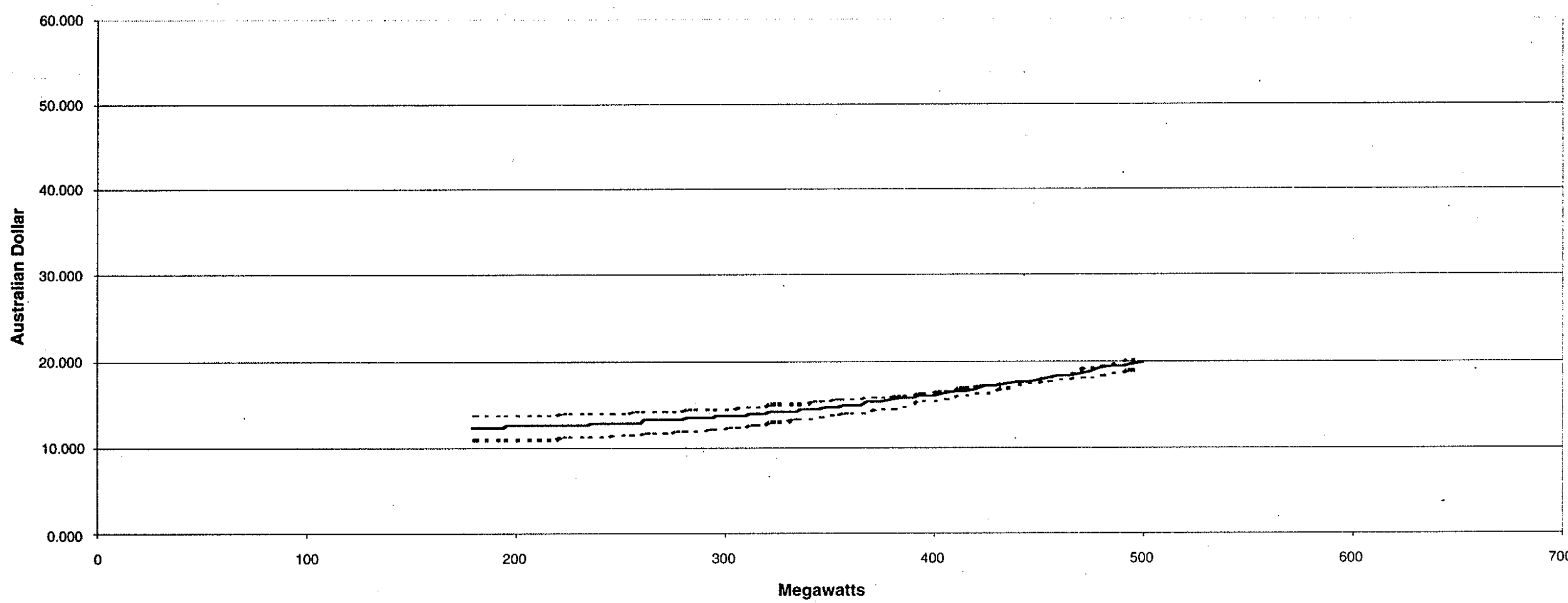

Figure 4.10. Marginal costs for Plant 2 of Firm A with optimal weighting matrix.

output levels, lower cost than Plant 2 gensets. This result was also confirmed by discussions with plant operators at Firm A and the fact that Plant 2 gensets are used less intensively by Firm A than are Plant 1 gensets.

The other result to emerge from this analysis is the increasing, convex marginal cost curves for all cases except the identity weighting matrix and the Plant 1 genset. One potential explanation for this result comes from discussions with market participants in wholesale electricity markets around the world. They argue that genset owners behave as if their marginal cost curves look like those in Figures 4.8 and 4.9 because they are hedging against the risk of unit outages when they have sold a significant amount of forward contracts. Because of the enormous financial risk associated with losing a genset in real time combined with the inability to quickly bring up another unit in time to meet this contingency, generation unit owners apply a large and increasing opportunity cost to the last one-third to one-quarter of the capacity of each genset. That way they will leave sufficient unloaded capacity on all of their units in the hours leading up to the daily peak so that they can be assured of meeting their forward financial commitments for the day even if one of their units is forced out.

This desire to use other units as physical hedges against the likelihood of a forced outage seems to be a very plausible explanation for the form of the marginal cost functions I recover, in light of the following facts about Firm A. First, during this time period, Firm A sold forward a large fraction of its expected output, and in some periods even more than its expected output. Second, all of Firm A's units are large coal-fired units, which can take close to 24 hours to start up and bring to their minimum operating level. Both of these facts argue in favor of Firm A's operating its units as if there were increasing marginal costs at an increasing rate as output approached the capacity of the unit.

8. IMPLICATIONS FOR MARKET MONITORING AND DIRECTIONS FOR FUTURE RESEARCH

There are a variety of uses for these results in market monitoring. Perhaps the most obvious is in constructing an estimate of the magnitude of variable profits earned by the firm over some time period. A major topic of debate among policymakers and market participants is the extent to which price spikes are needed for generation unit owners to earn an adequate return on the capital invested in each generating facility. This debate is particularly contentious with respect to units that supply energy only during peak periods. The methodology presented in this paper can be used to inform this debate.

Using these estimated marginal cost functions and actual market outcomes, one can compute an estimate of the magnitude of variable profits a generating unit earns over any time horizon. This information can then be used to determine whether the variable profit level earned on an annual basis from this unit is sufficient to adequately compensate the owner for the risk taken. This calculation should be performed on a unit-by-unit basis to determine the extent to which some units earn higher returns than other units. By comparing these variable profit levels to the annual fixed cost and capital costs of the unit, one can make a determination of the long-term profitability of each unit. Borenstein, Bushnell, and Wolak (2002) present a methodology for computing marketwide measures of the extent market power exercised in a wholesale electricity market. They apply their procedure to the California electricity market over the period from June 1998 to October 2000. These sorts of results should provide useful input into the regulatory decision-making process on the appropriate magnitude of allowable price spikes and the necessity of bid caps in competitive electricity markets. Determining the answers to these questions is particularly important in light of the events in all wholesale electricity markets throughout the United States during the summers of 1999 and 2000.

The framework outlined here can be extended in a number of directions. One extension involves using these methods in multisettlement electricity markets such as the California electricity supply industry. Here market participants make day-ahead commitments to supply or demand a fixed quantity of electricity and then purchase or sell any imbalance energy necessary to meet their actual supply and demand obligations in the ISO's real-time energy market. In this case the generator's profits from supplying electricity for the day are the result of selling into two sequential electricity markets. Consequently, one way to model this process is to assume best-reply pricing for the firm in both markets (and that the firm knows that it will attain best-reply prices in the real-time market when bidding into the PX market) and derive the implied marginal cost for the generator. Preliminary results from applying this procedure to California ISO and PX data are encouraging. Extending this procedure to the case of best-reply bidding in both markets is significantly more challenging.

A second direction for extensions is to specify Firm A's cost function as a multigenset cost function such as $C(q_1, \ldots, q_7, \beta)$. Assuming this functional form, I can examine the extent to which complementarities exist in the operation of different units. The marginal cost function for a given genset could also be generalized to allow dependence across periods within the day in the marginal cost of producing in a given hour, so that the variable cost for genset k might take the form $C_k(q_{1k}, \ldots, q_{48,k}, \beta_k)$, to quantify the impact of ramping constraints and other factors that prevent units from quickly moving from one output level to another.

ACKNOWLEDGMENTS

I thank Jeremy Fox, Jun Ishii, and Marshall Yan for outstanding research assistance. Bob Wilson and Lars Peter Hansen provided very helpful comments on previous drafts. Partial financial support for this research was provided by the National Science Foundation.

References

Berry, S. T. (1994), "Estimating Discrete-Choice Models of Product Differentiation," *RAND Journal of Economics*, 25(2), 242–262.

Berry, S. T., J. L. Levinsohn, and A. Pakes (1995), "Automobile Prices in Market Equilibrium," *Econometrica*, 63(4), 841–890.

Bresnahan, T. (1981), "Departures from Marginal-Cost Pricing in the American Automobile Industry," *Journal of Econometrics*, 17, 201–227.

Bresnahan, T. (1987), "Competition and Collusion in the American Automobile Market: The 1955 Price War," *Journal of Industrial Economics*, June, 45(4), 457–482.

Blundell, R. and J. L. Powell (2003), "Endogeneity in Nonparametric and Semiparametric Regression Models," in *Advances in Economics and Econometrics: Theory and Applications, Eighth World Congress*, Vol. 11 (ed. by Mathias Dewatripont, Lars Peter Hansen, and Stephen J. Turnovsky), Cambridge: Cambridge University Press.

Borenstein, S., J. Bushnell, and F. A. Wolak (2002), "Measuring Inefficiencies in California's Re-structured Electricity Market," forthcoming, *American Economic Review*, available from http://www. stanford.edu/~wolak.

Florens, J.-P. (2003), "Inverse Problems and Structural Econometrics: The Example of Instrumental Variables," in *Advances in Economics and Econometrics: Theory and Applications, Eighth World Congress*, Vol. 11 (ed. by Mathias Dewatripont, Lars Peter Hansen, and Stephen J. Turnovsky), Cambridge: Cambridge University Press.

Goldberg, P. K. (1995), "Product Differentiation & Oligopoly in International Markets: The Case of the U.S. Automobile Industry," *Econometrica*, 63(4), 891–952.

Green, R. J. and D. M. Newbery (1992), "Competition in the British Electricity Spot Market," *Journal of Political Economy*, 100(5), 929–953.

Hansen, L. P. (1982), "Large Sample Properties of Generalized Method of Moments Estimators," *Econometrica*, 50, 1029–1054.

Klemperer, P. D. and M. A. Meyer (1989), "Supply Function Equilibria in Oligopoly under Uncertainty," *Econometrica*, 57(6), 1243–1277.

Porter, R. H. (1983), "A Study of Cartel Stability: The Joint Executive Committee, 1880–1886," *Bell Journal of Economics*, Autumn, 14(2), 301–314.

Rosse, J. N. (1970), "Estimating Cost Function Parameters without Using Cost Data: Illustrated Methodology," *Econometrica*, 38(2), 256–275.

Wolak, F. A. (1999), "Market Design and Price Behavior in Restructured Electricity Markets: An International Comparison," in *Competition Policy in the Asia Pacific Region, EASE Vol. 8* (ed. by T. Ito and A. Krueger), Chicago: University of Chicago Press.

Wolak, F. A. (2000), "An Empirical Analysis of the Impact of Hedge Contracts on Bidding Behavior in a Competitive Electricity Market," *International Economic Journal*, 14(2), 1–40.

Wolak, F. A. (2001a), "A Model of Optimal Bidding Behavior in a Competitive Electricity Market," January, available from http://www.stanford.edu/~wolak.

Wolak, F. A. (2001b), "Identification, Estimation and Inference of Cost Functions in Multi-Unit Auction Models with an Application to Competitive Electricity Markets," February, available from http://www.stanford.edu/~wolak.

Wolak, F. A. (2002), "Designing a Competitive Wholesale Market that Benefits Consumers," available from http://www.stanford.edu/~wolak.