Dynamic Pricing & Learning in Electricity Markets*

Alfredo Garcia, Enrique Campos *and* James Reitzes

Systems & Information Engineering Department, University of Virginia, Charlottesville 22904

The Brattle Group, Washington DC 20036

April 18, 2002 (this version, November 20, 2003)

Abstract

We analyze the price-formation process in an infinite-horizon oligopoly model where hydroelectric generators engage in dynamic price-based competition. The analysis focuses on the role of "indifference" prices – i.e. prices that equate the gains from releasing or storing water. Strategies where players bid their indifference prices, and the marginal player undercuts the lowest-cost unsuccessful bidder, constitute a Markov Perfect Equilibrium (MPE) under appropriate conditions. These conditions involve symmetric production capacity and non-fractional (i.e., "all or nothing") output by successful bidders. Although the MPE solution represents an equilibrium consistent with dynamic strategic behavior, it requires computational sophistication by market participants. However, a basic "learning" procedure involving indifference prices converges to a Markov Perfect Equilibrium (MPE).

Subject Classification: Games: Stochastic, Noncooperative; Natural Resources: Energy, Water Resources; Economics: Restructured Electricity Markets, Dynamic Auctions.

---

*The first two authors gratefully acknowledge partial support from NSF grant ECS-0224747 and a University of Virginia Bankard Fellowship.

1 Introduction

With the extensive restructuring of power systems throughout the world, the analysis of strategic behavior in electricity markets has become an active area of research. This behavior typically has been examined in static models assessing "one-shot" equilibria in quantities, prices, or supply functions. The key assumption of such models is that firms' actions affect market performance only in the present period (see, for example, Green and Newbery (1992), Von der Fehr and Harbord (1993), Rudkevich, Duckworth, and Rosen (1998), Borenstein, Bushnell, and Knittel (1998), Borenstein and Bushnell (1999) and Anderson and Philpott (2002)). Besides allowing for more tractable analytics, this assumption is motivated by the fact that electricity, unlike other commodities, is not storable.

Nonetheless, many reasons suggest that electricity markets are characterized by dynamic strategic behavior. Dynamic strategies may arise when firms have flexibility in scheduling periods of plant maintenance, a key aspect of decision making for nearly all generators. Dynamic strategies also arise when hydroelectric (hydro) power is a significant source of production. This situation, which is the focus of our paper, describes market conditions in Norway and New Zealand, as well as several South American countries including Argentina, Brazil, Colombia, and Chile.

In markets with hydro power, generators can shift productive capacity from one period to another if they decide to hold onto water needed to produce electric power. A key determinant of the opportunity cost of supplying water in the present is the value that could be earned during periods of relative scarcity in the future. Interestingly, the likelihood of future scarcity depends not only on uncertain water inflows, but also on firms' valuation of their current water resources and their associated bidding strategies. The inherent interdependencies in this line of reasoning illustrate the complexity of analyzing dynamic strategic behavior.

This paper examines the price-formation process in an infinite-horizon oligopoly model where hydro generators engage in dynamic price-based competition. Players adjust their bids depending upon the "state" of the market where the "state" refers to the available capacity for generators in the market. The use of price as the control variable introduces certain analytical complexities, and it contrasts markedly with dynamic *Cournot* oligopoly models where output or capacity represents the control variable (see, for example, Crampes and Moreaux (2001), Scott and Read (1996), Bushnell (2003), Kelman et al. (2001) and Garcia and Arbelaez (2002)).

In general, a *Bertrand* model assumes that a firm determines the price at which it sells its output with firms being obligated to meet the resulting demand, while a *Cournot* model assumes that a firm determines its output or quantity while price is determined by some unspecified process so that market demand equals the total amount offered. Prices (quantities) form a *Bertrand* (*Cournot*) equilibrium if no single firm can profit from unilaterally changing its prices (quantities). Typically, these two models lead to very different predictions about market performance (see Fudenberg and Tirole (1991)).

In this paper, we build upon the work of Garcia, Reitzes and Stacchetti (2001) in which closed-form dynamic equilibrium solutions are obtained for simple examples using the notion of an "indifference" price. This is the price level where a player is indifferent between producing power with its water resources or holding onto those resources for future use. The relevant equilibrium solution concept is the *Markov Perfect Equilibrium (MPE).* Under *Markovian* pricing strategies, prices are a function of the state of *all* water reservoirs at the beginning of each period. The evolution of the "state" of the system is then determined by the corresponding pricing strategies of each player, the resulting market dispatch of water resources, and the realization of the random process governing water inflows. Under the equilibrium strategies, no participant in the market can increase its profits by unilaterally switching to a different pricing strategy.

In Garcia, Reitzes and Stacchetti (2001) a simple model with two players and limited reservoirs is used to analyze a number of issues of interest to regulatory economists, such as the likelihood of collusion and the effect of market behavior on the probability of outages relative to a centralized dispatch of resources by a benevolent planner.

By contrast, this paper extends the notion of "indifference" prices to a more general setting (finite number of players, integer-valued random inflows and demand) to provide a very simple characterization of Markov perfect equilibrium. Indeed, we find that a strategy of "*always bidding your indifference price*" represents an *MPE* under specified assumptions. In particular, this result arises when production capacities are symmetric and fractional dispatch cannot occur. In this MPE, all successful bidders put in a bid equal to their "indifference" price, except the marginal bidder may find it profitable to undercut the indifference price of the lowest (opportunity-)cost unsuccessful bidder.

We also study the behavior of a learning algorithm motivated by the notion of "indifference" prices. At each stage, players construct (unbiased) estimates of the future value of water under the current pricing strategy, and then adjust their pricing strategies to equal the "indifference" prices associated with the current strategy. We demonstrate this *this learning behavior converges to an MPE.*

While the concept of "opportunity cost" under centralized minimum-cost dispatch is well understood by decision makers in hydro-electric systems, the estimation of "opportunity cost" under bid-based market dispatch is not well understood. Due to computational difficulties involved in solving complex stochastic dynamic programming problems, market participants instead rely on heuristic procedures that incorporate some degree of error, or they encounter difficulties in making reasonable forecasts of other players' behavior.

However, a "learning" algorithm provides a reasonable description of interaction by boundedly rational participants in hydro-electric markets. Convergence of this "learning" algorithm to the MPE not only reinforces the descriptive value of the equilibrium solution but it supports the use of such an algorithm as a computational tool.

The structure of the paper is as follows. In section 2, we begin by introducing our relevant solution concept, the *Markov Perfect Equilibrium*, in an illustrative example.

In section 3, we proceed to formalize the notion of indifference prices for a general setting with a finite number of players, integer-valued random inflows and demand. The main result of sections 3 and 4 is a characterization of an MPE in which players simply bid their "indifference" prices for every “state” of the system.

In section 5, we show that a simple “learning” procedure converges to the *MPE* characterized. The objective of this section is two-fold. First, the convergence of a simple “learning” procedure shows how boundedly rational players may come to play the rather sophisticated *MPE* characterized earlier in the paper. Second, it motivates the use of a decentralized procedure as a means of computing the equilibrium solution. Similar to other learning models, we incorporate a notion of “inertia” in the process of forming estimates of future value, using empirical averages to construct these estimates. Thus, changes to the estimates of future value become progressively smaller over time. In the proof of convergence, we make use of a monotonicity property of indifference prices with respect to Markovian pricing policies. Intuitively, from a given player’s perspective, the price at which he/she is indifferent between selling or withholding should increase if all other players bid higher prices.

2 Illustrative Example

In this section we present a simple example illustrating the features of dynamic pricing equilibria in predominantly hydroelectric markets.

Two hydro generators, say i and j, drawing upon reservoirs of equal size, use water to produce electric power assuming an infinite planning horizon. Water reservoir levels are observed perfectly at the beginning of each period, after which both hydro generators simultaneously set prices for electric power. Both firms have a maximum reservoir capacity of 2 units and a maximum output capability equal to 1 unit per period. It is assumed that 1 unit of water produces 1 unit of output.

Demand also equals 1 unit in each period, and is perfectly inelastic. Hence, each player can satisfy the entire market demand when its reservoir has at least 1 unit of water.

During each period, generators make price bids for offering electric power up to their productive capability. The market then clears, with consumers purchasing from the low-priced firm. When both firms set the same price, it is assumed initially that one of the two firms supplies the entire market demand, where the probability of serving the market equals one half for each firm.

The "state" of the system is denoted by the two-tuple (x, y), where $x, y \in S = \{0, 1, 2\}$, that tracks reservoir levels for each generator. The state space is then $S \times S$. In strategically setting prices, each player faces uncertainty regarding the size of future water resources for itself and its rival. Water inflows, w, follow a simple random process. During each period, $w = 1$ with probability q, and $w = 0$ with probability $1 - q$. In other words, either it rains one unit (with probability q), or it does not. Water inflows occur at the end of the period. We further assume that the realization of w is the same for both players. Finally, marginal costs are identical for both players and normalized to equal zero.

When both reservoirs are empty, no production takes place. If one generator has a full reservoir and the other generator has an empty reservoir, the generator with water is effectively a monopolist. Its optimal strategy is to charge the maximum allowed price, c^*, which is either a reservation price (i.e., no transaction takes place if the price strictly exceeds the reservation price) or a price cap set by the regulator.

2.1 Markov Perfect Equilibria (MPE)

In light of the structure of the example, we posit that equilibrium strategies are symmetric, i.e. in equilibrium, given state (x, y), where $x \in S$ denotes generator i's reservoir level, and $y \in S$ denotes generator j's reservoir level, the players' equilibrium bidding strategies $\pi_i : S \times S \rightarrow [0, c^*]$ and $\pi_j : S \times S \rightarrow [0, c^*]$ are such that $\pi_i(x, y) = \pi_j(y, x)$. Let $V_{x,y}$ denote generator i's value function in equilibrium for the state $(x, y) \in S \times S$, i.e. the discounted value of the infinite stream of payoffs generated under strategy combination (π_i, π_j) starting from the initial state (x, y).

We begin by analyzing the simplest situation in which one of the reservoirs is empty and the other is not. As stated above, the generator with available capacity will earn c^* in the current period by simply bidding accordingly. In states where only generator i has available capacity,

$\pi_i(1,0) = \pi_i(2,0) = c^*$ and equilibrium value functions take on the following recursive form,

$$V_{1,0} = c^* + \beta[(1-q)V_{0,0} + qV_{1,1}], \tag{1}$$
$$V_{2,0} = c^* + \beta[(1-q)V_{1,0} + qV_{2,1}],$$

where $\beta \in (0,1)$ is the discount factor. If only generator j has available capacity, the value functions for generator i are as follows:

$$V_{0,1} = \beta[(1-q)V_{0,0} + qV_{1,1}], \tag{2}$$
$$V_{0,2} = \beta[(1-q)V_{0,1} + qV_{1,2}].$$

Finally, for the "rationing" state $(0,0)$ where reservoirs are depleted, the value functions take on the recursive form,

$$V_{0,0} = \beta[(1-q)V_{0,0} + qV_{1,1}]. \tag{3}$$

Our attention now shifts to reservoir states where both firms are capable of satisfying market demand. In these states, a generator setting its price above its rival's price may find itself in a monopoly position in the near future. This situation arises because the lower-priced rival reduces its reservoir, and that reservoir may not refill fast enough. An "opportunity cost" thus exists in selling electric power in the current period, since future monopoly opportunities may be forsaken. Consequently, when a price cap binds behavior in the monopoly state, the opportunity cost of selling power in competitive states is affected necessarily.

With the above trade-off in mind, we identify prices where each firm is indifferent between using water to sell electric power in the present period and holding that water for future use. For the $(1,1)$ and $(2,2)$ states, respectively, the "indifference prices" $p^*_{1,1}$ and $p^*_{2,2}$ must satisfy the following conditions:

$$p^*_{1,1} + \beta[(1-q)V_{0,1} + qV_{1,2}] = \beta[(1-q)V_{1,0} + qV_{2,1}], \tag{4}$$
$$p^*_{2,2} + \beta[(1-q)V_{1,2} + qV_{2,2}] = \beta[(1-q)V_{2,1} + qV_{2,2}].$$

Given the symmetry of our model, the prices $p^*_{1,1}$ and $p^*_{2,2}$ also capture the notion of indifference between releasing water or retaining water for generator j. Using the random tie breaking rule, in addition to (4) we have

$$V_{1,1} = \frac{p^*_{1,1}}{2} + \beta\left[\frac{1}{2}(1-q)V_{0,1} + \frac{1}{2}(1-q)V_{1,0} + \frac{1}{2}qV_{1,2} + \frac{1}{2}qV_{2,1}\right], \quad (4^*)$$

$$V_{2,2} = \frac{p^*_{2,2}}{2} + \beta\left[\frac{1}{2}(1-q)V_{1,2} + \frac{1}{2}(1-q)V_{2,1} + qV_{2,2}\right].$$

Now, when states $(1,2)$ and $(2,1)$ are reached, prices also exist for which generator i is indifferent between releasing water and receiving payments of $p^*_{1,2}$ and $p^*_{2,1}$, respectively, or retaining water for future use. These indifference prices satisfy the following conditions:

$$p^*_{1,2} + \beta[(1-q)V_{0,2} + qV_{1,2}] = \beta[(1-q)V_{1,1} + qV_{2,2}], \quad (5)$$

$$p^*_{2,1} + \beta[(1-q)V_{1,1} + qV_{2,2}] = \beta[(1-q)V_{2,0} + qV_{2,1}]. \quad (6)$$

Based on the preceding analysis, three different equilibrium outcomes are, in principle, possible. These outcomes are:

Case 1 $p^*_{1,2} > p^*_{2,1}$. In this case, the opportunity cost of releasing (the last unit of) water in state $(1,2)$ is *greater* than the opportunity cost of releasing water in the state $(2,1)$. Facing a lower opportunity cost of releasing water, firm j will undercut slightly firm i's bid of $p^*_{1,2}$. To compute the numerical value of $p^*_{1,2}$, one would solve the system of equations $(1)-(4^*)$, together with (5) and

$$V_{1,2} = \beta[(1-q)V_{1,1} + qV_{2,2}],$$

$$V_{2,1} = p^*_{1,2} - \delta + \beta[(1-q)V_{1,1} + qV_{2,2}].$$

Equation (6) is not binding since the player with 2 units undercuts the player with 1 unit by bidding $p^*_{1,2} - \delta$, $\delta > 0$.

Case 2 $p^*_{1,2} = p^*_{2,1} = p^*$. In this case, the opportunity cost of releasing water is the same even though the generators face differing reservoir levels. We obtain closed-form solutions for the *MPE* equilibrium by solving the system of equations $(1)-(4^*)$, (5) and (6) along with the following equations

$$V_{1,2} = \frac{p^*}{2} + \beta \left[\frac{1}{2}(1-q)V_{0,2} + \frac{1}{2}(1-q)V_{1,1} + \frac{1}{2}qV_{1,2} + \frac{1}{2}qV_{2,2} \right],$$
$$V_{2,1} = \frac{p^*}{2} + \beta \left[\frac{1}{2}(1-q)V_{1,1} + \frac{1}{2}(1-q)V_{2,0} + \frac{1}{2}qV_{2,2} + \frac{1}{2}qV_{2,1} \right].$$

Case 3 $p^*_{1,2} < p^*_{2,1}$. In this case, the opportunity cost of releasing water in state $(1,2)$ is *less* than the opportunity cost of releasing water in the state $(2,1)$. In equilibrium, generator i slightly undercuts generator j's bid of $p^*_{2,1}$, still receiving a price in excess of its own opportunity cost of releasing water. To compute the numerical value of $p^*_{2,1}$, one would solve the system of equations $(1) - (4^*)$, together with (6) and

$$V_{1,2} = p^*_{2,1} - \delta + \beta[(1-q)V_{0,2} + qV_{1,2}],$$
$$V_{2,1} = \beta[(1-q)V_{2,0} + qV_{2,1}].$$

Equation (5) is not binding since the player with 1 units undercuts the player with 2 units by bidding $p^*_{2,1} - \delta$, $\delta > 0$.

To illustrate, when $q = \frac{1}{3}$, $\beta = 0.999$ and $c^* = 1$, the solution to Case 1 is:

$$\begin{aligned}
p^*_{1,1} &= 0.896651 & V_{0,0} &= 269.516 \\
p^*_{2,2} &= 0.461302 & V_{0,1} &= 269.516 \\
p^*_{1,2} &= 0.692646 & V_{0,2} &= 269.429 \\
V_{1,0} &= 270.516 & V_{1,1} &= 270.325 \\
V_{1,2} &= 270.063 & V_{2,0} &= 271.325 \\
V_{2,1} &= 270.756 & V_{2,2} &= 270.35
\end{aligned}$$

In order to check the equilibrium conditions we compute $p^*_{2,1}$ (the equation that is not binding)

$$p^*_{2,1} = \beta[(1-q)V_{2,0} + qV_{2,1}] - \beta[(1-q)V_{1,1} + qV_{2,2}] = 0.801198$$

However, this case relies on the assumption that $p^*_{2,1} < p^*_{1,2}$. Selling at a price $p^*_{1,2}$ is not optimal for the firm with two units.

The system of equations implied by Case 2 has no solution, whereas the solution to Case 3 is:

$$\begin{aligned} p^*_{1,1} &= 0.89006, & V_{0,0} &= 284.98, \\ p^*_{2,2} &= 0.44812, & V_{0,1} &= 284.98, \\ p^*_{2,1} &= 0.80899, & V_{0,2} &= 284.95, \\ V_{1,0} &= 285.98, & V_{1,1} &= 285.84, \\ V_{1,2} &= 285.73, & V_{2,0} &= 286.84, \\ V_{2,1} &= 286.41, & V_{2,2} &= 285.98. \end{aligned}$$

In order to check the equilibrium conditions, we compute $p^*_{1,2}$ (the equation that is not binding)

$$p^*_{1,2} = \beta[(1-q)V_{1,1} + qV_{2,2}] - \beta[(1-q)V_{0,2} + qV_{1,2}] = 0.67385.$$

Note that $p^*_{2,1} > p^*_{1,2}$ as originally posited in Case 3. Therefore, the symmetric strategy π, with $\pi_1(x,0) = \pi_2(0,x) = c^*$ when $x > 0$ and

$$\begin{aligned} \pi_1(1,1) &= p^*_{1,1} &= \pi_2(1,1), \\ \pi_1(2,2) &= p^*_{2,2} &= \pi_2(2,2), \\ \pi_1(1,2) &= p^*_{2,1} - \delta &= \pi_2(2,1), \\ \pi_1(2,1) &= p^*_{2,1}, &= \pi_2(1,2), \end{aligned}$$

is a Markov Perfect Equilibrium.

3 A General Formulation

Now, we will examine equilibrium behavior under more general assumptions regarding reservoir sizes and replenishment rates. Let K_i be a positive integer that denotes the storage capacity of the reservoir owned by player $i \in \{1, 2, 3, ..., n\}$. The state space is $S = \prod_i \{0, 1, 2, ..., K_i\} \subseteq \mathcal{R}^n$. Player i's production capacity is denoted by x_i (also a positive integer) in each period. We shall assume that players can only produce electricity whenever the water stored exceeds productive capacity. In each period, demand for electricity is denoted by D (a positive integer) and is assumed to be perfectly inelastic with respect to price.

By assumption, bids must be linear. With this assumption, we gain analytical tractability at the expense of generality (non-linear bidding). The per-unit bid can take on any value in $\mathcal{P}$ =$[0, c^*]$, where c^* represents a maximum bid or price cap imposed by the regulator.

Given the bids b_i for players $i = 1, 2, 3, ..., n$, we shall denote by $b_{[k]}$ the kth-lowest bid, and $[k]$ represent the index set of players submitting that bid. The index set $[m^*]$ correspond to the marginal bidder or bidders, and is defined by

$$m^* = \arg\min\{k : \sum_{i=1}^{k} x_{[i]} \geq D\}.$$

where $x_{[i]}$ is the aggregate capacity of the $i - th$ ranked players. The spot price of electricity $p^*(\mathbf{b})$, ($\mathbf{b}$ =$(b_1, b_2, ..., b_n) \in B = \Pi_i \mathcal{P}$), corresponds to the marginal generator's bid,

$$p^*(\mathbf{b}) = b_{[m^*]}.$$

Let us now construct player i's dispatch as a function of the bids $\mathbf{b}$, and the "state" of the reservoirs, s =$(s_1, s_2, ..., s_n) \in S$.

Player i's dispatch can be defined as follows

$$D_i(\mathbf{b}; \mathbf{s}) = \begin{cases} \tilde{x}_i & b_i = p^*(\mathbf{b}) \text{ and } s_i \geq x_i, \\ x_i & b_i < p^*(\mathbf{b}) \text{ and } s_i \geq x_i, \\ 0 & \text{otherwise.} \end{cases} \cdot$$

where $\tilde{x}_i$ is the market dispatch if player i is the "marginal" bidder. This quantity represents the difference between market demand and the total productive capability of all other successful bidders and depends upon the specific allocation rule implemented in the event of multiple marginal bidders. As in the illustrative example, we shall assume a random allocation rule (i.e. any one of the players at the margin is equally likely to be chosen).

Since marginal production costs are assumed to equal zero, the immediate payoff for player i can be expressed as

$$r_i(\mathbf{b}; \mathbf{s}) = p^*(\mathbf{b}) \cdot D_i(\mathbf{b}; \mathbf{s}).$$

The evolution of player i's reservoir is governed by a first-order stochastic difference equation

$$s_i' = \min[s_i - D_i(\mathbf{b}; \mathbf{s}) + \xi_i, K_i],$$

where ξ_i is the random (non-negative, integer valued) water inflows to player i's reservoir. After one period, the resulting level from dispatch $D_i(\mathbf{b}; \mathbf{s})$, inflow ξ_i and initial reservoir level s_i would be $s_i - D_i(\mathbf{b}; \mathbf{s}) + \xi_i$. However, the limited reservoir capacity may result in water spillage if $s_i - D_i(\mathbf{b}; \mathbf{s}) + \xi_i > K_i$.

To simplify notation, we hereafter refer to the probability of reaching state $\mathbf{s}'$ from state $\mathbf{s}$ as $f(\mathbf{s}'; \mathbf{b}, \mathbf{s})$, which depends, of course, on the bidding strategies. Similarly, we shall denote by $F(\mathbf{s}'; \mathbf{b}, \mathbf{s})$ the associated cumulative distribution function. Finally, future payoffs are discounted using the factor $\beta \in (0, 1)$.

3.1 Stationary Markovian Pricing Strategies

We shall restrict our attention to Markovian pricing strategies, i.e. strategies where bids are a function of the current state of reservoirs. Moreover, we are interested in stationary (i.e. time-invariant) strategies.

For each player i, a Markovian pricing strategy is denoted by the mapping $\pi_i : S \mapsto \mathcal{P}$. A Markovian strategy combination, $\boldsymbol{\pi} = (\pi_1, \pi_2, ..., \pi_n)$, is a vector of Markovian pricing strategies for each player. We let Π represent the set of all Markovian (pure) strategy combinations.

With discounting and bounded payoffs in each period, a "value" function exists for every Markovian strategy combination. The value function, denoted by the mapping $Q^\pi : S \mapsto \mathcal{R}^n$, is defined recursively as

$$Q_i^\pi(\mathbf{s}) = r_i(\pi(\mathbf{s}); \mathbf{s}) + \beta \sum_{s' \in S} Q_i^\pi(\mathbf{s}') \cdot f(\mathbf{s}'; \pi(\mathbf{s}), \mathbf{s}).$$

3.2 Stationary Markov Perfect Equilibrium (MPE)

We are interested in *Subgame Perfect Markovian* strategy combinations that have the following property: at every time period, for any given state, no player can do strictly better by choosing a different price than the one prescribed by the strategy combination under consideration. This concept formalizes a notion of recursive rationality, i.e. play prescribed by the strategies from any state off the equilibrium path must also be in equilibrium (see Fudenberg and Tirole (1991)).

Formally, a strategy combination π^* is a stationary MPE iff for every player i, and every state $\mathbf{s} \in S$,

$$Q_i^{\pi^*}(\mathbf{s}) \geq Q_i^{(\pi_i, \pi^*_{-i})}(\mathbf{s}),$$

for all $\pi_i \neq \pi_i^*$, where (π_i, π^*_{-i}) is the strategy combination with player i bidding according to π_i (instead of π_i^*).

4 Characterization of *MPE*

We present now a discussion on the proof of existence of a stationary *MPE*, illustrating some of the key concepts that are necessary later on to prove the convergence of the learning procedure discussed in Section 5.1. Let $M_i \subset S$ be defined as

$$M_i = \{\mathbf{s} \in S \mid x_i \leq s_i\}.$$

In the following discussion we restrict our attention to stationary Markovian pricing strategies π_i in the reduced domain M_i, i.e. $\pi_i : M_i \subset S \mapsto \mathcal{P}$. This is done with no loss of generality, since at every stage, players with not enough stored water will not take part of the resulting market dispatch, regardless of their bids (their dispatch is null).

4.1 Valuation

If players are assumed to follow a Markovian pricing policy π, from an individual perspective, each player is faced at each state with the choice between releasing at the given price, withholding or setting the price at the margin. In order to compute the optimal response, we solve the dynamic programming recursive equations for the marginal unit valuation. For $\mathbf{s} \in M_i$ and

$b_i \in \mathcal{P}$ we have

$$V_i^\pi(\mathbf{s}; b_i) = p^*((b_i, \pi_{-i}(\mathbf{s})) \cdot D_i(b_i, \pi_{-i}(\mathbf{s}); \mathbf{s}) + \beta \sum_{s' \in S} V_i^\pi(\mathbf{s}') \cdot f(\mathbf{s}'; (b_i, \pi_{-i}(\mathbf{s})), \mathbf{s}), \tag{7}$$

$$V_i^\pi(\mathbf{s}) = \sup_{b_i \in \mathcal{P}} [V_i^\pi(\mathbf{s}; b_i)], \tag{8}$$

where $(b_i, \pi_{-i}(s))$ stands for the strategy combination that equals π except at state $\mathbf{s}$ where player i bids b_i. Equation (7) determines the value of selling today at the given prices (by bidding b_i). Equation (8) summarizes the value of today's best decision.

4.2 To Release or Not to Release ? (that is the question)

In what follows, we shall refer to *infra*-marginal players as those players bidding below market price $p^*(\pi(\mathbf{s}))$, i.e. player i is inframarginal if $\pi_i(\mathbf{s}) < p^*(\pi(\mathbf{s}))$. Similarly, player i is *supra*-marginal if $\pi_i(\mathbf{s}) > p^*(\pi(\mathbf{s}))$. Finally, *marginal* players are defined as being neither infra- or supra- marginal.

To characterize the optimal decision we now introduce the notion of *indifference prices* at state $\mathbf{s}$ under strategy combination π. The *indifference price* for player i is the price that equates the value obtained by releasing x_i units today or withholding. Let us denote it by $\tilde{p}_i(\mathbf{s}, \pi)$. This notion captures the opportunity cost of releasing x_i units today, at the current market price, under the assumption that players will continue to play according to strategy combination π, given state $\mathbf{s}$. Formally,

$$\tilde{p}_i(\mathbf{s}, \pi) \cdot x_i + \beta \cdot \sum_{s' \in S} V_i^\pi(\mathbf{s}') \cdot f(\mathbf{s}'; (0, \pi_{-i}(\mathbf{s})), \mathbf{s}) = \beta \cdot \sum_{s' \in S} V_i^\pi(\mathbf{s}') \cdot f(\mathbf{s}'; (c^*, \pi_{-i}(\mathbf{s})), \mathbf{s}).$$

We remark that $\tilde{p}_i(\mathbf{s}, \pi)$ is non-negative. By withholding output and retaining water reserves, a firm ensures that its water reserve level in the next period will be greater than (or equal to) the reserve level attained if it had released water to produce electricity. Since withholding by a given firm causes its rivals to release more water to produce electricity, rival reserves in the next period will be less than (or equal to) the levels attained if that firm had instead released water. Since a firm's value rises as its reserves increase and rival reserves decrease, it must hold that the future value associated with withholding is greater than (or equal to) the future value associated with releasing water. Also, $\tilde{p}_i(\mathbf{s},\pi) \leq c^*$ since the difference in per unit value between bidding c^* and releasing today can not exceed c^*.

4.3 Characterization of *MPE*

Let us now make the following standing assumption:

Assumption 1: $x_i = x > 0$ *for all* $i \in \{1, 2, 3, ..., n\}$ *and* $D \bmod x = 0$.

The above assumption implies that fractional dispatch does not occur. In situations where players can operate at a fraction of their full capacity level, the per-unit "indifference price" associated with fractional production may differ from that associated with full production. Bidding resulting in fractional dispatch cannot be ruled out as potentially profitable deviations from bidding to achieve full dispatch. We assume, though, that players are dispatched on an "all or nothing" basis. This captures some of the technical constraints associated with electricity production, where a relatively narrow range of efficient production levels may exist. Allowing fractional dispatch involves an even more complex optimization problem, which will be left for future research.

We now describe a bidding strategy $\tilde{p}_i^*(\mathbf{s},\pi)$ for the *marginal* bidder(s), given state s and strategy π, that will be used in characterizing equilibrium bidding:

$$\tilde{p}_i^*(\mathbf{s},\pi) = \begin{cases} c^* & [m^*+1] = \varnothing \text{ and } [m^*] \text{ is a singleton,} \\ \tilde{p}_{[m^*+1]}(\mathbf{s},\pi) - \delta & [m^*+1] \neq \varnothing \text{ and } [m^*] \text{ is a singleton,} \\ \tilde{p}_i(s,\pi) & \text{otherwise,} \end{cases}$$

where $[m^*+1]$ is the index set associated with lowest *supra*-marginal bids under strategy combination π at the given state s, and $\delta > 0$. Also,

$$\tilde{p}_{[m^*+1]}(\mathbf{s},\pi) \quad = \quad \min_{i \in [m^*+1]} \tilde{p}_i(\mathbf{s},\pi).$$

The next result follows from the discussion above:

Theorem 1: *Under Assumption 1, a strategy combination of the form* $\pi_i(s) = \tilde{p}_i(s, \pi)$, $i \notin [m^*]$, *and* $\pi_i(s) = \tilde{p}_i^*(s, \pi)$, $i \in [m^*]$, *for all* $s \in S$, *is a Markov Perfect Equilibrium.*

Proof: Given state $\mathbf{s}$, for *infra-marginal* players we have

$$p^*(\pi(\mathbf{s})) \quad > \quad \tilde{p}_i(\mathbf{s}, \pi).$$

Let us consider a deviation in which *infra*-marginal player i bids b, where $b > p^*(\pi(\mathbf{s}))$. In light of Assumption 1, $D_i(b, \pi_{-i}(\mathbf{s}); \mathbf{s}) = 0$ and

$$\textstyle\sum_{s' \in S} V_i^\pi(\mathbf{s}') \cdot f(\mathbf{s}'; (c^*, \pi_{-i}(\mathbf{s})), \mathbf{s}) \quad = \quad \sum_{s' \in S} V_i^\pi(\mathbf{s}') \cdot f(\mathbf{s}'; (b, \pi_{-i}(\mathbf{s})), \mathbf{s}).$$

Moreover,

$$\sum_{s' \in S} V_i^\pi(\mathbf{s}') \cdot f(\mathbf{s}'; (0, \pi_{-i}(\mathbf{s})), \mathbf{s}) = \sum_{s' \in S} V_i^\pi(\mathbf{s}') \cdot f(\mathbf{s}'; \pi(\mathbf{s}), \mathbf{s}).$$

By definition of indifference prices,

$$p^*(\pi(\mathbf{s})) \cdot x_i + \beta \sum_{s' \in S} V_i^\pi(\mathbf{s}') \cdot f(\mathbf{s}'; \pi(\mathbf{s}), \mathbf{s}) > \beta \sum_{s' \in S} V_i^\pi(\mathbf{s}') \cdot f(\mathbf{s}'; (b, \pi_{-i}(\mathbf{s})), \mathbf{s}) = V_i^\pi(\mathbf{s}; b).$$

For *supra-marginal* players we have

$$p^*(\pi(\mathbf{s})) \quad < \quad \tilde{p}_i(\mathbf{s}, \pi).$$

For any bid b below the price $p^*(\pi(\mathbf{s}))$ we have in light of Assumption 1, $D_i(b, \pi_{-i}(\mathbf{s}); \mathbf{s}) = x_i$ and

$$\sum_{s' \in S} V_i^\pi(\mathbf{s}') \cdot f(\mathbf{s}'; (b, \pi_{-i}(\mathbf{s})), \mathbf{s}) = \sum_{s' \in S} V_i^\pi(\mathbf{s}') \cdot f(\mathbf{s}'; (0, \pi_{-i}(\mathbf{s})), \mathbf{s}).$$

Also, the resulting market price will not increase (i.e. $p^*((b, \pi_{-i}(\mathbf{s}))) \leq p^*(\pi(\mathbf{s})))$ thus by definition of indifference prices,

$$p^*((b, \pi_{-i}(\mathbf{s}))) \cdot x_i + \beta \cdot \sum_{s' \in S} V_i^\pi(\mathbf{s}') \cdot f(\mathbf{s}'; (b, \pi_{-i}(\mathbf{s})), \mathbf{s}) < \beta \cdot \sum_{s' \in S} V_i^\pi(\mathbf{s}') \cdot f(\mathbf{s}'; (c^*, \pi_{-i}(\mathbf{s})), \mathbf{s}) = V_i^\pi(\mathbf{s}).$$

Finally, for marginal player i, c^* is a best reply to π when $[m^* + 1] = \varnothing$ and $[m^*]$ is a singleton. Also, bidding $\tilde{p}_{[m^*+1]}(\mathbf{s}, \pi) - \delta$ is a best reply to π when $[m^* + 1] \neq \varnothing$ and $[m^*]$ is a singleton. Otherwise, the marginal player is indifferent between selling or withholding full capacity x_i at a market price $p^*(\pi(\mathbf{s})) = \tilde{p}_i(s, \pi)$. ■

Example 1 Assume that there are two identical players, where $K = x = 1$ and $D = 1$. In each period, each player receives a water inflow of 1 unit with probability q (otherwise, the inflow equals 0). Both firms have the same realized water inflow (i.e., water inflows are perfectly correlated). Four possible states therefore exist: $(0,0)$, $(0,1)$, $(1,0)$, and $(1,1)$. Consider the symmetric strategy $\pi(1,0) = \pi(0,1) = \pi(0,0) = c^*$ and $\pi(1,1) = \beta(1-q)c^*$. Then,

$$\begin{aligned} V^\pi_{0,0} &= \beta\left[(1-q)V^\pi_{0,0} + qV^\pi_{1,1}\right], \\ V^\pi_{0,1} &= \beta\left[(1-q)V^\pi_{0,0} + qV^\pi_{1,1}\right], \\ V^\pi_{1,0} &= c^* + \beta\left[(1-q)V^\pi_{0,0} + qV^\pi_{1,1}\right]. \end{aligned}$$

We note that for all $\mathbf{s}$

$$\pi_i(\mathbf{s}) = \tilde{p}_i(\mathbf{s}, \pi).$$

Thus, in light of Theorem 1, π is an MPE.

Example 2 Back to the illustrative example in section 2, the reader can verify that when $q = \frac{1}{3}$, $\beta = 0.999$ and $c^* = 1$, the symmetric strategy π, with $\pi_1(x,0) = \pi_2(0,x) = c^*$ when $x > 0$ and

$$\begin{aligned} \pi_1(1,1) &= p^*_{1,1} &&= \pi_2(1,1), \\ \pi_1(2,2) &= p^*_{2,2} &&= \pi_2(2,2), \\ \pi_1(1,2) &= p^*_{2,1} - \delta &&= \pi_2(2,1), \\ \pi_1(2,1) &= p^*_{2,1}, &&= \pi_2(1,2), \end{aligned}$$

satisfies all the conditions required for Theorem 1 to hold.

Example 3 The following example illustrates the importance of Assumption 1. Two identical players and $K = x = 2, D = 3$, and inflow is equal to 2 units every period. The state $s = (2,2)$ is the only reachable state. The indifference price as defined above is clearly equal to zero. Notice that bidding zero is not an equilibrium since a player may profitably deviate by bidding c^*. Nonetheless, there are asymmetric pure strategy equilibria, in which one player bids c^* and the other bids $\frac{c^*}{2}$.

In general, the indifference price associated with fractional production may differ substantially from the indifference price associated with full production. Therefore, whenever fractional dispatch is a possibility (e.g. when Assumption 1 does not hold) bidding deviations (from the strategy in which all players bid their indifference prices) resulting in fractional dispatch, can not be ruled out as potentially profitable deviations.

4.4 Incorporating Thermal Generators

The equilibrium characterized in Theorem 1 still arises when thermal (i.e., coal, gas, nuclear-powered) electricity production offers competition to hydro production in the market. In the simplest case, a thermal generator j has a deterministic capacity $s_j = x_j = K_j$ and incurs a marginal cost c_j. The resulting profit function for a given profile b is

$$r_j(b; s) = (p^*(b) - c_j) \cdot D_j(b; s).$$

Under the marginal pricing rule formalized above, bidding marginal costs is optimal regardless of other players' bids.

A more elaborate model of a power plant could account for the availability of say N independent units, each of capacity x_j and outage probability q. One step transition probabilities can be easily computed depending upon the current "state" $s_j \in \{0, x_j, 2 \cdot x_j, \ldots, N \cdot x_j\}$ (total available capacity) and how quickly units out of operation are brought back on-line.

5 Learning Dynamic Equilibrium

Even though the *MPE* solution concept formalizes a notion of strategic stability, it does require a great degree of computational sophistication on the part of market participants. This feature undermines the descriptive value of the *MPE* solution concept. While many learning models have been proposed in the game theoretic literature to explain how boundedly rational players may come to play an equilibrium, there are very limited results on learning in dynamic games.

Fudenberg and Levine (1998) provide a summary exposition of recent work that focuses on learning in "static" games or games in normal form. Perhaps, the oldest and best known learning algorithm is "fictitious play" in which at each step, players compute their best replies based on the assumption that opponents decisions follow a probability distribution in agreement with the historical frequency of their past decisions. Convergence of beliefs to Nash equilibrium has been established by Monderer and Shapley (1996) for potential games and for certain classes of supermodular games by Milgrom and Roberts (1991). Typically, learning results rely on the analysis of the differential equation defined by the expected motion of the process while imposing restrictions on the structure of the errors (see Kushner and Yin (1997)).

The application of many of these results to dynamic games is not straightforward since in dynamic games, players are to "learn" policies (i.e., rules for determining actions on the basis of a "state" variable) as opposed to simply actions. In order to learn equilibrium policies, players need to construct good approximations to the value functions. There is therefore a compromise on the speed of policy versus value updating (see Borkar (2002)). In our application however, assuming there exists π of the form, $\pi_i(s) \in \tilde{p}_i(s, \pi)$ for all i and $s \in S$, our characterization result (Theorem 1) enable us to avoid these technical difficulties, since the optimal policy has a simple representation in terms of the value function.

5.1 The Learning Algorithm

In what follows, we shall work under the assumption that demand is inelastic and assumption 1 holds. For a given Markovian pricing strategy π^0, the learning algorithm **(L)** works as follows; for $k = 0, 1, 2, \ldots$

(L.1) Estimate "Indifference Prices"

For supra-marginal and infra-marginal players

$$\hat{p}_i(\mathbf{s}, \pi^k) = \tilde{p}_i(\mathbf{s}, \pi^k) + \varepsilon_k,$$

For marginal players, we follow the definition given in section 4.3.1, which requires marginal players to estimate indifference prices for the current lowest supra-marginal bidder(s) to estimate $\tilde{p}_i^*(\mathbf{s}, \pi)$.

(L.2) Update Policy

$$\pi_i^{k+1}(s) = (1 - \alpha_k) \cdot \pi_i^k + \alpha_k \cdot \hat{p}_i^k(\mathbf{s}, \pi^k),$$

where ε_k stands for the estimation error at iteration k and α_k is the "gain" of the recursive estimation procedure (typically, $\alpha_k = \frac{1}{k}$). The structure of these errors will be discussed in section 5.3.

5.2 Monotone Indifference Prices

In the study of convergence of the proposed learning algorithm we shall use a monotonicity property of indifference prices with respect to Markovian pricing policies. Intuitively, from a given player's perspective, the price at which he/she is indifferent between selling or withholding should increase if all other players bid higher prices. Formally,

Assumption 2 (Monotone Indifference Prices): *If $\pi \geq \pi'$ then $\tilde{p}_i(\mathbf{s}, \pi) \geq \tilde{p}_i(\mathbf{s}, \pi')$, for every player i and any given state s.*

The last inequality is equivalent to

$$\int V_i^{\pi}(\mathbf{s}') \cdot dF(\mathbf{s}'; c^*, \pi_{-i}(\mathbf{s})) - \int V_i^{\pi}(\mathbf{s}') \cdot dF(\mathbf{s}'; 0, \pi_{-i}(\mathbf{s})) \geq \int V_i^{\pi'}(\mathbf{s}') \cdot dF(\mathbf{s}'; c^*, \pi'_{-i}(\mathbf{s})) - \int V_i^{\pi'}(\mathbf{s}') \cdot dF(\mathbf{s}'; 0, \pi'_{-i}(\mathbf{s})), \quad (9)$$

where we have rewritten the definition of $p_i(\mathbf{s}, \pi)$ using the distribution function F.

We first note that for a fixed bid b for player i, stage revenues are increasing in π, i.e., if $\pi \geq \pi'$ then

$$p^*((b, \pi_{-i}(\mathbf{s}) \cdot D_i(b, \pi_{-i}(\mathbf{s})))) \geq p^*((b, \pi'_{-i}(\mathbf{s})) \cdot D_i(b, \pi'_{-i}(\mathbf{s}))),$$

from standard arguments in dynamic programming the value function defined in (7) and (8) inherits this monotonicity, i.e.,

$$V_i^{\pi}(\mathbf{s}) \geq V_i^{\pi'}(\mathbf{s}). \tag{10}$$

A sufficient condition for monotone indifference prices is to require that all players have identical characteristics (reservoir capacity, production capacity, inflow probabilities).

Lemma 1 (Symmetry): *If* $K_i = K; x_i = x$, $D \bmod x = 0$ *and* $\xi_i = \xi$ *for* $i \in \{1, 2, 3, ..., n\}$, *then* $\tilde{p}_i(s, \pi)$ *is monotone.*

Proof: For symmetric states $\mathbf{s}, \mathbf{s}' \in S$ (i.e., $s_i = s_j$ and $s'_i = s'_j$; $i \neq j$),

$$F(\mathbf{s}'; c^*, \pi_{-i}(\mathbf{s})) = F(\mathbf{s}'; c^*, \pi'_{-i}(\mathbf{s})),$$

$$F(\mathbf{s}'; 0, \pi_{-i}(\mathbf{s})) = F(\mathbf{s}'; 0, \pi'_{-i}(\mathbf{s})). \tag{11}$$

By contradiction, if $\tilde{p}_i(\mathbf{s}, \pi) < \tilde{p}_i(\mathbf{s}, \pi')$ for $\pi \geq \pi'$, in light of (11) it must be the case that $V_i^{\pi}(\mathbf{s}') < V_i^{\pi'}(\mathbf{s}')$, but this contradicts (10).■

A symmetric dynamic game structure is not a necessary condition for monotonicity (see Garcia, Reitzes and Stacchetti (2001) section 2.4 for an example). Inequality (9) is also referred to as *supermodularity* or *increasing differences* of expected continuation values in strategy combination π. This property is exploited by Topkis (1978) and (1998) to develop an abstract framework for monotone comparative statics of parametrized optimization problems and stochastic games. The application of Topkis's results to determine more general conditions for monotonicity of indifference prices is the subject of future research.

5.3 Convergence of the Learning Algorithm

We introduce some standard notation used in stochastic approximation techniques. Let

$$H_i(\pi)(\cdot) = \hat{p}_i(\cdot, \pi) - \pi_i,$$

and let $h_i(\pi)(\cdot) = E\left[H_i(\pi)(\cdot)\right]$. Then we can rewrite the algorithm **(L)** as:

$$\pi_i^{k+1} = \pi_i^k + \alpha_k h_i(\pi^k) + \varepsilon_k,$$

where $\varepsilon_k = \alpha_k(H_i(\pi^k) - h_i(\pi^k))$, and $\sum_k \alpha_k = \infty$, and $\sum_k \alpha_k^2 < \infty$.

Let $\mathcal{F}_k = \sigma(\pi^0, \pi^1, \ldots, \pi^k)$ be the history of the algorithm up to stage k. We assume that the indifference price estimates are unbiased; namely, $E[\varepsilon_{k+1} \mid \mathcal{F}_k] = 0$ for all k. We note that since π, and $\tilde{p}_i(\cdot, \pi)$ are bounded by c^*, we have

$$E[\|\varepsilon_k\|^2 \mid \mathcal{F}_k] \leq (\alpha_k \cdot 2c^*)^2,$$

and hence $\sum_k \|\varepsilon_k\|^2$ converges with probability 1.

Lemma 2:

$$\sum_k \varepsilon_k < \infty,$$

and $\lim_{k\to\infty}(\pi^{k+1} - \pi^k) = 0$, *with probability 1.*

Proof: Consider the truncation,

$$w_n^{(c)} = \sum_{k=1}^{\min\{n, M(c)\}} \varepsilon_k,$$

where $M(c) = \min\{k \mid \|\varepsilon_k\|^2 > (\alpha_k \cdot 2c)^2\}$. As a consequence $w_n^{(c)}$ is a martingale with bounded second moment for all c, and therefore converges with probability 1. Moreover there exists c, namely c^*, for which $M(c^*) = \infty$, and hence $\sum_k \varepsilon_k < \infty$. A clear consequence is $\lim_{k\to\infty}(\pi^{k+1} - \pi^k) = 0$, w.p.1 ■

In order to introduce our convergence result, we shall briefly review the notion of "distance" to a set in a normed space. Let $(X, \|\cdot\|)$ be a normed space and let $A \subset X$ be some subset. The distance from a point $p \in X$ to the set A is defined as

$$d(p, A) = \inf_{a \in A} \|p - a\|.$$

We recall that under Assumption 2, $\tilde{p}$ is monotone on π. Moreover

$$\tilde{p}(\cdot, c^*) < c^* \quad \text{and} \quad \tilde{p}(\cdot, 0) > 0. \tag{11}$$

In words, if players are bidding c^*, any player will have an incentive to "undercut" as a best reply. Similarly, if electricity is currently sold for free (all players are bidding 0) it is clearly profitable for any player to withhold in order to capture future monopoly rents associated with the likely event when all other players will be unable to supply demand. Thus, monotonicity and (11) implies that there exists at least one fixed point, say π^*, such that for $\epsilon > 0$ and $\pi^* - \epsilon \leq \pi' < \pi^* < \pi'' \leq \pi^* + \epsilon$, then

$$\pi' < \tilde{p}(\cdot, \pi') \leq \pi^* \leq \tilde{p}(\cdot, \pi'') < \pi''. \tag{12}$$

Equivalently,

$$\begin{aligned} \|\pi' - \pi^*\| &> \|\tilde{p}(\cdot, \pi') - \pi^*\|, \\ \|\pi'' - \pi^*\| &> \|\tilde{p}(\cdot, \pi'') - \pi^*\|. \end{aligned}$$

In words, the fixed point π^* is a "local attractor" to any dynamic sequence generated with the use of the indifference price operator (i.e. the algorithm L). Let Π^* denote the set of all such fixed points π^*.

Theorem 2: *(Convergence to an MPE) Under Assumptions 1 and 2,*

$$d(\pi^k, \Pi^*) \to 0 \textit{ with probability 1.}$$

Proof: Let $\epsilon > 0$ be as in (12). By Lemma 2, there exists $k(\epsilon)$ such that $d(\pi^k, \Pi^*) \leq \epsilon$ with probability 1, for all $k \geq k(\epsilon)$. In light of (12)

$$\left\|\pi^{k+1} - \pi^*\right\| \leq (1 - \alpha_k) \left\|\pi^k - \pi^*\right\| + \alpha_k \left\|\tilde{p}(\pi^k) - \pi^*\right\| \leq \left\|\pi^k - \pi^*\right\| \tag{13}$$

Thus, π^k converges to some $\pi^* \in \Pi^*$. Finally, by Theorem 1, any element $\pi^* \in \Pi^*$ is a MPE. ■

5.4 Numerical Illustration

In this section we present examples that illustrate the application of the learning algorithm. In this algorithm, each player has access to a unbiased simulator, and to the history of pricing decisions by the other players.

5.4.1 Estimating the Value Functions

We start with a stationary policy

$$\pi^0 : S \rightarrow \{0, \frac{c^*}{M}, 2\frac{c^*}{M}, \ldots c^*\}^n,$$

that maps each state into the bids made by the players in the past. Note that the price interval $[0, c^*]$ is discretized to the set $\mathcal{P} = \{0, \frac{c^*}{M}, 2\frac{c^*}{M}, \ldots, c^*\}$, for some integer M.

Each player i has an initial estimate $\tilde{V}_i^{\pi^0}(\mathbf{s}; p)$ of the value functions for each state $\mathbf{s}$, and price $p \in \mathcal{P}$.

At stage k of the algorithm, each player i proceeds as follows.

1. Using simulation, the player obtains a raw estimate of the value function $\hat{V}_i^{\pi^k}(\mathbf{s}; p)$ for each state s, and price $p \in \mathcal{P}$.

2. It then corrects its estimate using previous estimates, i.e.,

$$\tilde{V}_i^{\pi^k}(\mathbf{s}; p) = (1 - \alpha_k) \cdot \tilde{V}_i^{\pi^{k-1}}(\mathbf{s}; p) + \alpha_k \cdot \hat{V}_i^{\pi^{k-1}}(\mathbf{s}; p),$$

   for all states s.

3. The estimate of the indifference prices is computed as follows, given π^k, for inframarginal and supramarginal players

$$\hat{p}_i(\mathbf{s}, \pi^k) = \left\lfloor \frac{1}{x_i} \cdot [\tilde{V}_i^{\pi^k}(\mathbf{s}; c^*) - \tilde{V}_i^{\pi^k}(\mathbf{s}; 0)] + p^*((0, \pi^k_{-i}(\mathbf{s})) \right\rfloor, \tag{14}$$

   for each state s where $\lfloor x \rfloor$ is the closest point in $\{0, \frac{c^*}{M}, 2\frac{c^*}{M}, \ldots, c^*\}$ to $x \in [0, c^*]$. For marginal players, we follow the definition given in in section 4.3.1.

The algorithm presented above, can be computed asynchronously, by randomly selecting a single player to perform the estimation procedure at a stage. If an unbiased estimate is available, convergence of the asynchronous algorithm can be proved by using a set of stepsizes α_k^i for each player i, and simply making $\alpha_k^i = 0$ for each playing not updating. A similar argument holds for the case a single state is randomly selected at each stage. This would correspond to an on-line implementation of the algorithm. Note that the assumptions on the stepsizes still hold, since each player will update infinitely often.

5.4.2 Back to Illustrative Example

The illustrative example of Section 2 is implemented here with the parameters $c^* = 1$, $q = \frac{1}{3}$, $\beta = 0.999$, $K = 2$ and $D = 1$ (grid size $M = 1000$). The evolution of the algorithm with initial condition $\pi^0 = c^*$, is shown in Figure 1. The prices computed in section 2 correspond to the dashed lines.

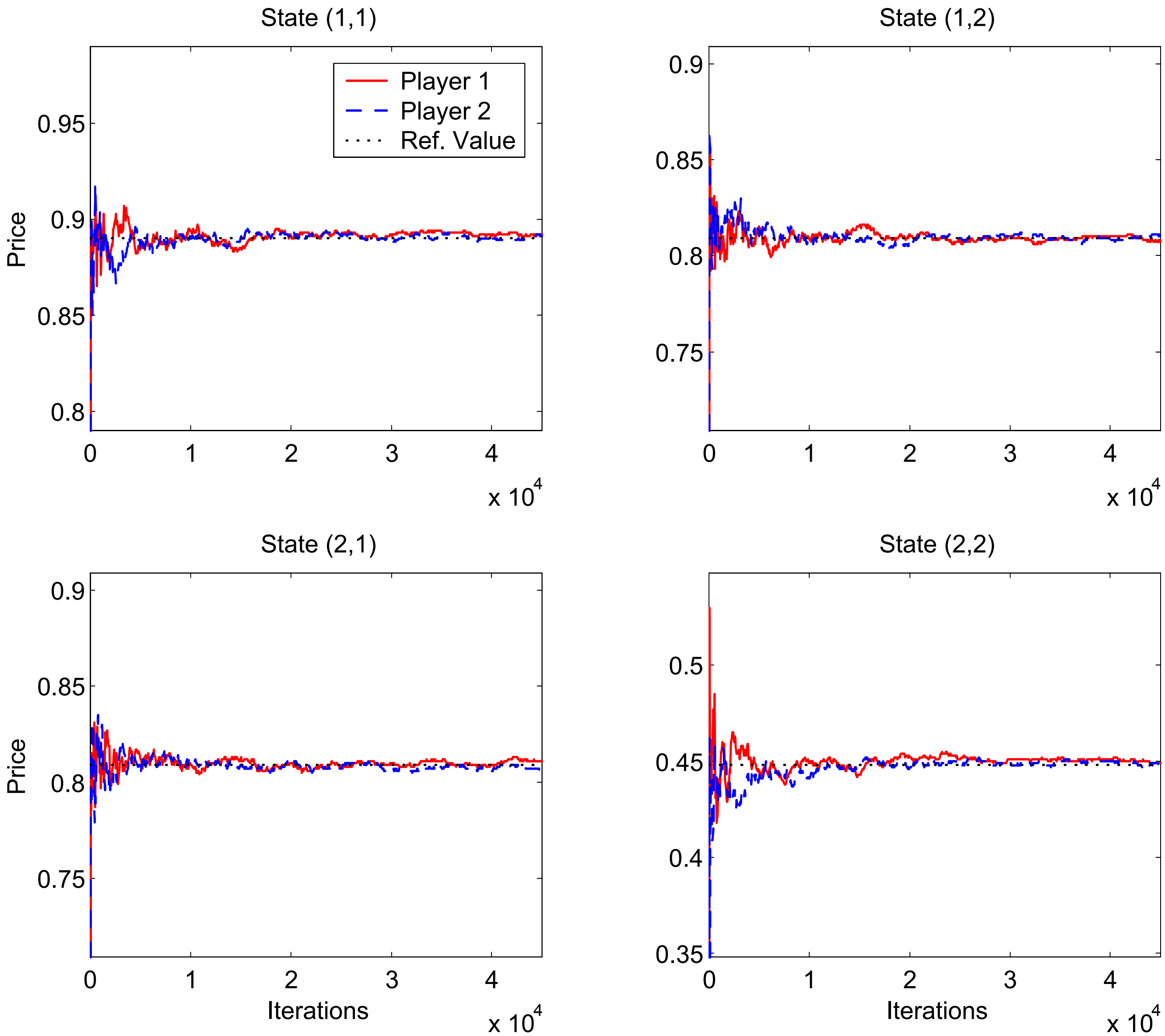

Figure 1: Sample path for learning algorithm.

6 Conclusions

In deregulated power markets with hydro-electric generation, market participants must analyze the strategic value of water resources in formulating their pricing strategies. In this paper, we have analyzed firm strategies in an infinite-horizon oligopoly model where hydro generators engage in dynamic price-based competition. We have provided a simple characterization of a *Markov Perfect Equilibrium (MPE)* in terms of "indifference" prices. These are the prices equating the gains from releasing or storing water under the considered strategy combination. Bidding indifference prices results in an *MPE* whenever production capacities are symmetric and fractional dispatch does not occur. The *MPE* solution concept formalizes a notion of equilibrium that is consistent with dynamic strategic behavior. However, it requires a great degree of computational sophistication on the part of market participants, a feature that undermines its descriptive value. Thus, we have shown that, as long as "indifference" prices are monotonically increasing in pricing strategies, a simple adaptive "learning" procedure converges to the characterized *Markov Perfect Equilibrium (MPE)*.

References

[1] ANDERSON E. J. AND A.B. PHILPOTT. 2002. Using Supply Functions for Offering Generation into an Electricity Market. Operations Research, 50 (3) pp. 477-489

[2] BORENSTEIN S. AND BUSHNELL J. 1999. An Empirical Analysis of the Potential for Market Power in California's Electricity Market. Journal of Industrial Economics, Vol 47, No. 3, September, pp. 285-323

[3] BORENSTEIN S., BUSHNELL J. AND KNITTEL C. 1999. Market Power in Electricity Markets: Beyond Concentration Measures. Energy Journal, 20 (4). pp 65-88

[4] BORKAR V. Reinforcement Learning in Markovian Evolutionary Games. 2002. Advances in Complex Systems, Vol. 5, No. 1 pp. 55-72

[5] BUSHNELL J. 2003. A Mixed Complementarity Model of Hydrothermal Electricity Competition in the Western United States. Operations Research, 51 (1) pp. 80-93

[6] CRAMPES C. AND MOREAUX M. 2001. Water Resource and Power Generation. International Journal of Industrial Organization 19, pp. 975-997

[7] FUDENBERG D. AND TIROLE J. 1991. Game Theory, MIT Press.

[8] FUDENBERG D. AND LEVINE D. 1998. The Theory of Learning in Games, MIT Press.

[9] GARCIA A. AND ARBELAEZ L. 2002. Market Power Analysis for the Colombian Electricity Wholesale Market. Energy Economics, Vol. 24 (3), pp. 217-229

[10] GARCIA A., REITZES J. AND STACCHETTI E. 2001. Strategic Pricing when Electricity is Storable. Journal of Regulatory Economics, Vol 20 No. 3 pp. 223-247

[11] GREEN R.J AND NEWBERY D. 1992. Competition in the British Electricity Spot Market. Journal of Political Economy, vol.100 pp. 929-953

[12] KELMAN R., BARROSO L. PEREIRA M. 2001 Market Power Assessment and Mitigation in Hydrothermal Systems. IEEE Transactions on Power Systems pp. 354 -359

[13] KUSHNER H. AND YIN G. 1997. Stochastic Approximation Algorithms and Applications, Springer Verlag

[14] MILGROM P. AND ROBERTS J. 1991. Adaptive and Sophisticated Learning in Normal Form Games. Games and Economic Behavior. Vol 3 pp. 82-100

[15] MONDERER D. AND SHAPLEY Ll. 1996. Fictitious Play Property for Games with Identical Interests. Journal of Economic Theory, Vol 68, pp 258-265

[16] RUDKEVICH A., DUCKWORTH M. AND ROSEN R. 1998. Modeling Electricity Pricing in a Deregulated Generation Industry: The Potential for Oligopoly Pricing in a PoolCo. The Energy Journal, Vol 19 No. 3. pp 19-48

[17] SCOTT T.J AND READ G. 1996. Modeling Hydro Reservoir Operation in a Deregulated Electricity Market. International Transactions on Operations Research, Vol 3, No3/4 1996, pp. 243-254

[18] TOPKIS D. 1978. Minimizing a Submodular Function on a Lattice. Operations Research 26, pp 305-321

[19] TOPKIS D. 1998. Supermodularity and Complementarity, Princeton University Press.

[20] VON DER FEHR N., HARBORD D. 1993. Spot market competition in the UK electricity market. The Economic Journal, May, pp. 531-546