Agent-Based Simulation—An Application to the New Electricity Trading Arrangements of England and Wales

Derek W. Bunn and Fernando S. Oliveira

***Abstract*—This paper presents a large-scale application of multiagent evolutionary modeling to the proposed new electricity trading arrangements (NETA) in the U.K. This is a detailed plant-by-plant model with an active specification of the demand side of the market. NETA involves a bilateral forward market followed by a balancing mechanism and then an imbalance settlement process. This agent-based simulation model was able to provide pricing and strategic insights, ahead of NETA's actual introduction.**

***Index Terms*—Agent-based simulation, electricity markets, evolutionary economics, imperfect competition, market design.**

I. Introduction

CONVENTIONAL economic modeling approaches have shown only a limited ability to develop insights into the strategic behavior of firms competing in new restructured markets, such as electricity. These new electricity markets, which have emerged around the world since the early 1990s, tend to be characterized by an oligopoly of generators, very little demand-side elasticity in the short term, and complex administered market mechanisms, which are designed to facilitate both financial trading and physical real-time system balancing. It is not surprising, therefore, that theoretical economic analysis tends to oversimplify, overaggregate, and represent stylized versions of these markets. For example, analytical evaluation of the supply function equilibria for these markets has either assumed that the supply functions are continuous [1], whereas in practice, generating units are offered in discrete blocks or that the industry ownership consists of symmetric equally sized firms [2] or both [3]. From a different auction theory perspective, von der Fehr and Harbord [4] recognized that there is a need to model discretely, but for tractability, their analysis was still restricted to a duopoly. Rothkopf [5] observes that the daily repetition of these auctions means that all players are involved in repeated games rather than single auctions and, therefore, it is to be expected that the continuous process of daily experimentation and learning by all of the firms will lead to multiple transient equilibria in practice.

It is precisely these features of the problem which suggest that agent-based computational methods could perform a useful role. The development of a detailed simulation platform representing the agents, the markets, and the market clearing mechanisms, together with reinforcement learning to facilitate profit-seeking behavior by the agents, can, in principle, provide a computational framework to overcome the limitations of the analytical approaches. Furthermore, Roth and Erev [6], [7] have recognized that the intermediate behavior of a dynamic game maybe more important than its asymptotic properties. They have shown that reinforcement-learning models outperform the equilibrium predictions in certain games. In the broader context of management research, the strategy and organizational behavior literature [8] has also started to model organizations as adaptive systems that learn with experience by trial and error through exploiting and exploring their environments.

However, in practice, there have been very few large-scale applications to industry behavior. It is still, therefore, an open question as to how well agent-based simulation can provide useful insight at the firm level in such a complex market as electricity, not only with respect to the specification of appropriate learning, but also in terms of analysing the multiagent multiple equilibria output from the simulations. In this paper, we describe one such application of agent-based simulation to explore the economic implications of the proposed changes to the U.K. electricity market and thereby provide a timely case study in using evolutionary computation in practice.

This study was actually motivated early in 2000 by a major U.K. electricity company who was seriously concerned about the impending radical reforms of the electricity market, which were scheduled to be introduced in November 2000. The new electricity trading arrangements (NETA) would replace the mandatory daily uniform price auction (the "Pool"), which had operated to provide a competitive wholesale market since 1990. The new market would mainly be based upon continuous bilateral trading up to "gate closure" at 3.5 h ahead of real time. Following gate closure, the system operator (SO) would operate a kind of balancing market, the design of which was deliberately intended to financially reward flexible plant and discourage players on both the generation and demand side of the market from being out of balance. The risks involved in this new system appeared, therefore, to be considerable and any model-based insights would have to capture subtle details of the interrelationship of the bilateral trading to the balancing market and the relative plant economics that would ensue. With no realistic analogies from electricity markets elsewhere and only some limited simulations from an experimental role-playing game [9], agent-based simulation appeared to offer a real possibility to develop detailed insights into the potential market ahead of its introduction (which actually happened in March 2001).

Manuscript received August 30, 2000. This work was supported by the Fundação para a Ciência e a Tecnologia, Portugal, under Grant Praxis XXI/BD/21769/99.

The authors are with the Decision Sciences Department, London Business School, London NW1 4SA, U.K. (email: dbunn@london.edu; foliveira@london.edu).

Publisher Item Identifier S 1089-778X(01)09040-3.

1089–778X/01$10.00 © 2001 IEEE

Previously, Bower and Bunn [10], [11] applied agent-based simulation to look at one aspect of the auction design involved in NETA, namely, the switch from uniform to discriminatory pricing. Their model did not capture the interaction between the bilateral trading and the balancing market nor did it incorporate any sophistication in the agents' learning abilities. The simulation platform developed here is a much more detailed representation of how NETA was designed to function.

1) It actively models the demand side (suppliers).
2) It models the interactions between two different markets [the bilateral market and the balancing mechanism (BM)] and also models the settlement process (SP).
3) It takes into accounts the daily dynamic constraints and it assumes different marginal costs for each generation technology.

II. Specification of the NETA Simulation Platform

A. Overview of NETA

The conceptualised agents in this model represent generators (i.e., generation companies possibly owning several plants with different generation technologies), suppliers (i.e., the agents purchasing from wholesale market in order to "supply" end-use customers), and the SO. In the following sections, we describe how we have modeled the bilateral trading, the BM, the SP, and the strategic behavior of suppliers and generators. The basic structure of the model followed [9], [12], and [13], although the details of the BM reflect later revisions [14], [15].

NETA is based upon unadministered bilateral trading between generators, suppliers, traders, and customers, which takes place in the forward markets before gate closure. At the introduction of NETA, four organizations had set up power exchanges (PXs) to facilitate this, although it is clear that market forces will cause liquidity to gravitate to one or two of them. The BM works as a market where the SO buys and sells increments (*incs*) or decrements (*decs*) of electricity in order to balance the system as a whole. However, individual generators and suppliers may be out of balance. During the SP, the SO will compare the contract positions (quantities contracted), plus whatever is bought or sold in the BM with the actual position (quantities generated or consumed) for each one of the suppliers and generators (plant by plant) to calculate the *imbalances*. The imbalance may be a *spillage* (if a plant is generating more than it has contracted or if a supplier is consuming less than it has contracted) or a *top-up* (if a plant is generating less than it has contracted or if a supplier is consuming more than it has contracted). For both types of imbalances, there is a price: if an agent is spilling, it will receive as payment for the electricity generated the system sell price (SSP); if an agent is topping up, it will pay the system buy price (SBP). The spread between the two prices is intended to provide a penalty for being out of balance: the SSP (SBP) is expected to be considerably lower (higher) than the prices in the forward markets.

Let us define the following variables.

QD_t Total quantity of demand at time t.
QPX_t Total quantity of demand at time t in the bilateral markets.
QTP_t Total quantity of top-up at time t.
$QBBM_t$ Total quantity bought in the BM at time t.
QG_t Total quantity of generation at time t.
QS_t Total quantity of spillage at time t.
$QSBM_t$ Total quantity sold in the BM at time t

Then, the two markets together are described by the following identities:

demand identity
$$QD_t = QPX_t + QTP_t + QBBM_t$$
and generation identity
$$QG_t = QPX_t + QS_t + QSBM_t.$$

The task of the SO is to buy or sell in the BM enough electricity to ensure that the system is always in balance
$$QSBM_t - QBBM_t = QTP_t - QS_t.$$

The day-ahead and within-day balancing schedules are as follows [15, p. 160].

SO Day-Ahead Balancing Process:

1) By 09:00, publishes the day-ahead demand forecast.
2) By 11:00, receives the initial physical notifications (IPNs).
3) Calculates the available national plant margin or shortfall.
4) Verifies system security given demand predictions, the submitted IPN's, and planned transmission outage.
5) By 12:00, issues the total system plant margin data to the market for the day ahead.
6) Forecast constraint costs based on the estimated final physical notifications (FPNs) and bid (offer) prices and volumes.
7) If necessary, calls the most economic balancing services contracts to ensure the system security.
8) During the following 11 hours, receives updates of the physical notifications (PNs).
9) By 16:00, publishes the revised national plant margin and zonal margin.

SO Within-Day Balancing Process:

1) Publish half hourly averaged demand forecasts for a defined period until gate closure.
2) As participants become aware of changes to their physical position, they will advise the SO.
3) At defined times, the zonal and national margins will be reassessed and provided to the market.
4) Undertake security analysis and reassess the requirements of balancing services contracts.
5) At gate closure, the PNs will become FPNs and the SO will have received bids (offers) of the prices and volumes from the participants in the BM.
6) During the BM, the SO will balance the system, taking into account technical constraints, dynamic operating characteristics of generation and demand balancing services, and uncertainty in demand.

B. Modeling NETA

Multiagent simulations can have various design objectives, e.g., to replicate the time series properties of a real market [16] or to search for consensus [17]. Our model was designed primarily to search for possible market equilibria, given all the parameters defining the agents and the market structure.

Modeling the balancing process is not a straightforward task and demands some simplifications to be adopted in order to make the problem tractable. These simplifications are much less restrictive than the ones usually presented in the literature.

1) The transmission system was not modeled. This implies that the model does not capture regional imbalances or transmission constraints.
2) We model only a typical day, taking into account some plant dynamics. The model is aimed at analysing the process of finding the equilibrium solution for the specific daily profile simulated.
3) We have simplified the continuous nature of trading in NETA and represented both the forward market and the BM as two sequential one-shot markets. This implies that the flows of information are much simpler than reality: the players only submit their offer (bids) to the bilateral market knowing the SO's demand forecasts and submit their offers (bids) to the BM knowing the expected position of the system.
4) Finally, our model assumes independence between generators and suppliers, although vertical integration is a reality in the industry. We adopt this simplification since the regulator imposes a condition that all the trading between the suppliers and generators belonging to the same company has to be subject to separate imbalance accounts.

In formulating the market models, the "natural" approach would seem to be the continuous double auction [18], [19], but its implementation with computer agents has several problems. It requires direct communication and negotiation between agents, the development of a multicriteria algorithm when the quantity is variable, and finally the computational time could be an issue in a large model such as ours.[1] Hence, we used the single call market (SCM) developed by Cason and Friedman [20], but we have adapted it to reflect the NETA trading principles with the agents paying the price bid or receiving the price offered instead of paying (receiving) the clearing price. It should be noted that the same type of auction was also adopted by [4], where they refer to it as a sealed-bid multiple-unit auction. Nicolaisen *et al.* [21] adopted an auction very similar to ours with the main difference being that their agents pay (receive) the midpoint of the bid-ask spread. By adopting an SCM, we also avoid explicitly modeling beliefs about opponents [22].

Another challenging task was to model the suppliers and generators. We intended to build bounded rational agents that learn about their environment and improve their behavior with experience, avoiding the perfect foresight paradigm. But at the same time, following Marcet and Nicollini [23],[2] the agents' behavior should not be too nonrational (we have to impose some lower bounds on the agents' level of rationality). The internal consistency requirement is the most demanding one and we had to make sure that the agents do not choose completely unreasonable actions, even in the early stages of the learning process.

Another problem facing the agents is the crossing of information. The agents take actions in two different markets—the PX and the BM—and receive feedback from these two markets. In order to facilitate the association between an action and its results, the feedback process should be as close to the action as possible. In order to accommodate this requirement, the agents have operational objectives. Keeping in mind that the final goals of the agent are the strategic objectives, we need to make sure that there is an interrelation between these two types of objectives. To ensure this link, we define operational rules.

Each supplier is characterized by the following prespecified parameters: market share, BM exposure (contract cover), retail price, prediction error, and search propensity.

Market share refers to the retail market in terms of the relative quantity of electricity sold. *BM exposure* reflects the percentage of forecast demand it intends to purchase in the PX. The *retail price* is not a crucial variable; since this model is aimed at short-term analysis, it just allows the suppliers to be modeled as profit maximizing. *Prediction error* reflects, as mean average prediction error, the capability of the agent to predict its own demand. *Search propensity* is an integer ranging from one to ten that defines a heuristic controlling how the agents search for the best payoff and transform past experience into future policies. A low *search propensity* parameter will create stable policies that change slowly with experience, whereas a high parameter will create reactive policies that tend to follow short run experience).

Each supplier has the following instruments.

1) *Markup in the PX (learned by the agent):* This is relative to the PX price (PXP) in the previous day.
2) *Markup in the BM (learned by the agent):* This is relative to the PXP in the same day.

Each supplier has two strategic objectives:

1) to maximize *total daily profits*, given the market structure and given its market share;
2) to minimize the difference between its *objective* for the BM exposure and the *actual* BM exposure.

The two operational objectives are to maximize daily profits in the PX and in the BM, respectively. They enable each agent to associate the outcomes (such as prices, quantities, and profits) to each instrument used and to the strategic objectives.

Suppliers also have an operational rule based upon *adaptive expectations*, namely, never to bid (offer) more (less) in the BM than the previous day's SBP (SSP) since those are the expected imbalance prices.[3]

Each generator is characterized by the following parameters: plants, cycles, capacity, availability, BM exposure (contract cover), and search propensity.

[1]In the experiments presented in this paper, we have simulated 200 trading days. In each trading day, there are 24 auctions for each hour in the bilateral market and 24 auctions for the BM (one for each hour).

[2]They define three concepts that maybe used to establish lower bounds of rationality. In asymptotic rationality, the agents' behavior will asymptotically converge to the optimum. In epsilon-delta rationality, agents may have some resistance to change and may be satisfied within a solution close enough to the optimal behavior. With internal consistency, agents' behavior has to have some lower rational bounds in the short run, i.e., agents try to do the best they can over a limited horizon.

[3]This operational rule, also adopted for the generators, establishes lower bounds for rationality. One of the agents' strategic objectives is BM exposure. Agents want to have a low BM exposure because of high BM price and volume uncertainty. If suppliers (generators) choose to pay (receive) a higher (lower) price in the PX than in the BM, their behavior would be irrational since this behavior would contradict there BM exposure objective.

TABLE I
RELATION BETWEEN CYCLES (MAXIMUM NUMBER OF STARTUPS WITHIN A DAY) AND TYPE OF PLANT IN THE U.K.

Type	Cycle
Gas Turbine	3
Oil	3
Pumped Storage	3
Small coal	2
CCGT	1
Large coal	1
Inter-connector	0
Nuclear	0

Peak technologies (used to supply peak demand), gas turbine, oil, and pumped storage were classified as having three daily cycles. Midmerit technologies were classified with one and two cycles, which include CCGT and coal. Finally, the base-load plants (running in a nonstop regime) are the nuclear stations and the interconnectors.

Plants owned by each generator are specified at the generating set level. Plants of the same type are assumed to have similar marginal costs, startup costs, and no-load costs. The agent's decision-making does not take into account explicitly the startup and no-load costs, although it will implicitly learn to do so. Thus, each agent has an objective for the position of each plant in the load duration curve (we identify for each plant the maximum number of cycles per day that it can operate) and the plant's profit it will be penalized if it does not meet that objective. This is a reasonable way of incorporating some consideration of dynamic plant constraints. Thus, we define the parameter *cycles* for each type of plant (see Table I).

Thus, base-load plants with high startups or inflexible technology need to run continuously and specify zero or just one cycle. Flexible plant with low startup cost can have a higher number of cycles. The installed *capacity* is assumed to be available with a probability specified by its *availability* parameter, which reflects outage rates for each plant. *BM exposure* and *search propensity* are the same as for the suppliers.

Each generator has the same instruments as the suppliers, namely, it learns to mark up in the PX from the previous day and to mark up in the BM from the PX outcome. Its strategic objectives are:

1) to maximize *daily profits*, given the market structure;
2) to minimize the difference between its *objective* and the *actual* BM Exposure for each plant.

The two operational objectives are to maximize the *daily profits* in the PX and in the BM, respectively, for each plant.

In order to avoid inconsistent behavior during the learning process, we impose some lower bounds of rationality through operational rules.

1) *Portfolio Management:* A plant with higher or equal number of cycles will never undercut the offers of another plant with equal or less number of cycles.
2) *Noninterruption:* Plants that have to run continuously or plants with one cycle may run without profit in certain hours of the day.
3) *No Loss Leading:* Plants with one cycle do not run without profit at the beginning or at the end of the day. They prefer not to run at all if the price is to low.
4) *Peak Premia:* Peak plant never offer prices below marginal cost.

Together with the following two extra operational rules for the BM.

1) *Adaptive Expectation:* Never bid (offer) above (below) the previous SBP (SSP).
2) *Avoidable Cost:* Never pay more than the marginal cost for "speculative" *decs.*

The model is organized into trading days or iterations and runs in a day-ahead mode for a set of 24-h periods. A trading day starts with the agents buying (selling) electricity in the PX. In the PX, the suppliers try to buy, at a price as low as possible, the amount of electricity needed to fulfil their contract cover objectives. The generators will try to sell at a price as high as possible, given their portfolio, and given the amount of electricity they want to save to sell in the BM. At gate closure, each agent will know exactly how much it has sold or bought and provides the SO with its FPN.[4]

Then the trading in the BM begins. The SO's total demand forecasts are common knowledge in the industry, period by period (and we have assumed for these experiments that the forecasts are accurate). Nevertheless, each one of the suppliers will have some uncertainty predicting its own demand. Thus, using their FPN's and its demand *forecast*, the SO calculates the total system surplus or shortfall for each period in the day ahead. Given this total system position, the SO will accept either *incs* or *decs* in the BM. The trades in the BM are done between the SO and each one of the generators and suppliers offering (bidding) the *incs* or *decs* into the BM. We assume that the SO will, first of all, in the interests of efficiency, clear all possible *arbitrage opportunities* (ARB_t). Then

$$\text{Excess}_t = QD_t - FPNs_t - ARB_t - QS_t + QTP_t.$$

If $\text{Excess}_t > 0 (< 0)$, the SO will be accepting *extra-incs* (*extra-decs*), above the arbitrage level ARB_t, from generators (only).[5]

After all trading in the BM has occurred and the SO has bought or sold whatever energy was needed to balance the system, the imbalances of each generator and supplier are calculated. The imbalance prices and costs are computed. If the SO accepts *incs*, the SBP will be defined as the weighted average of the offers accepted in the BM. Otherwise, if the SO accepts *decs*, the SSP will be defined as the weighted-average of the bids accepted in the BM. Hence, in our model, at each hour, only one imbalance price will be defined (SBP or SSP). Thus, if an agent is long (short) when the system is short (long), there will be no imbalance price defined for its case. We adopted a rule that the SO has indicated it may have to use if

[4]*Good behavior assumption:* In this model, we assume that agents always communicate the true FPN's and that they always deliver what was contracted in the PX and in the BM.

[5]It should be noted that a supplier without load management will not be influencing the net position of the system and so its bids in the BM are only to cover its own uncertainty to avoid the imbalance charges. The bids (offers) of these players will only be accepted if there is an arbitrage opportunity.

there are insufficient bids (offers),[6] which is to take the average of past SBP (SSP) values for that particular hour for the SBP (SSP) not defined.

III. Learning Algorithm

Although there are several agent-based concepts coming from the field of multiagent systems [24], almost all assume that an agent must communicate or be physically separated from the environment. The agents used in our model have no communication capabilities. The agents are conceptual identities representing "economic agents" in the market, having capacity to "receive" information, learn from the interactions, and act on the simulated environment. They are *autonomous agents*.[7]

The agents in this model learn online (given the information set, an agent modifies its actions in order to maximize its own profit). A "natural" way to model the learning process in a way that "captures" human and organizational learning process is reinforcement learning [6]–[8]. In the artificial intelligence community, this technique has been used for a long time to make computer programs that can learn to play games [25]–[27].

A. Defining the Markups

The learning problem is defined as follows: each player has a tuple of instruments $< mk\_px\_b(mk\_px\_o), mk\_bm\_b(mk\_bm\_o) >$, where $mk\_px\_b(mk\_px\_o)$ is the ratio between the price bid (offered) in the PX and the PXP in the previous day. The $mk\_bm\_b(mk\_bm\_o)$ is the ratio between the price bid (offered) in the BM and the PXP in the same day (the agents already have information on the PXP when they bid or offer in the BM). The attraction of this model is its simplicity: the agents only learn a ratio for the whole day, the margin revision in the PX or BM is the same for each different hour of the day. This avoids the adoption of different learning behavior for the different levels of demand [4]. Another strong point about this learning algorithm is the assumption that the information set only includes the prices in the previous day. Given the present state of the world, the agent only tries to improve its position given the knowledge accumulated in the past, assuming that the present contains all the relevant information (i.e., the Markov property).

Let t represent an iteration (K is the maximum number of iterations) $t = 1, \ldots, K$ and i represent the time of the day $i = 1, \ldots, 24$. Let also PB_{t_i} $(PBbm_{t_i})$ represent the price bid at iteration t and hour i in the PX (BM) and PO_{t_i} $(PObm_{t_i})$ represent the price offered at iteration t and hour i in the PX (BM). Then

$$\begin{aligned} PB_{ti} &= PXP_{t-1_i} * mk\_px\_b_t \\ PO_{ti} &= PXP_{t-1_i} * mk\_px\_o_t \\ PBbm_{ti} &= PXP_{t_i} * mk\_bm\_b_t \\ PObm_{ti} &= PXP_{t_i} * mk\_bm\_o_t. \end{aligned}$$

[6]In practice, this will not occur as frequently as in the model. The dynamic nature of buying and selling by the SO throughout the 3.5-h BM window, and the locational nature of buying and selling to balance the system at the nodal level, means that the SO will be more active than our model reflects.

[7]An autonomous agent is a system situated within a part of an environment that senses that environment and acts on it, over time, in pursuit of its own agenda and so as to affect what it senses in the future [24].

Each agent learns a different policy for each one of the four markups. This is a very important point about this learning algorithm. Agents do not learn how to choose prices they learn how to choose the markups on the previous day prices: the result is a model in which prices are *unbounded*.

The learning process is the same for each one of the four markups. Each one of the markups was partitioned into ten discrete intervals. In the experiments described in this paper, the ratio between the price bid (offered) assumes different ranges for suppliers and generators in different markets.

Suppliers learn different markups for the PX and the BM (one for *incs* and another one for *decs*).[8]

1) *Bids in the PX from 0.95 to 1.2:* This allows for decreases and increases in the price offered.
2) *Incs in the BM from* (-0.2) *to 3:* This allows for incs with negative prices.
3) *Decs in the BM from* (-0.2) *to 3:* This allows for decs with negative prices.

Generators also learn different markups for the PX and the BM (one for incs and another one for decs).

1) *Offers in the PX from* (-0.15) *to 1.15:* This allows for decreases and increases in the price offered. It also allows the generators to buy in the PX (by offering a negative price).
2) *Incs in the BM from 0.6 to 3:* This allows for incs higher or lower than the PXP.
3) *Decs in the BM from* (-0.25) *to 1.1:* This allows for decs with negative prices.

The larger range defined for the generators allows them to bid above or below the clearing price and allows some plants to receive a negative price in order to make sure some low cycle plant runs through short price troughs.

B. Learning a Policy

At each instance, the agent calculates the *expected daily profit* and the *expected acceptance rate* for each one of the markups used at that specific iteration.

1) The expected daily profit is calculated using exponential smoothing of the profits earned on past trading days.
2) The expected acceptance rate is calculated using exponential smoothing of the number of hours that a bid (offer) was accepted in the past trading days.

Thus, given the expected acceptance rate and the expected daily profit, each player calculates the *expected reward* for each markup. Then the agent constructs a *utility function* over the markups. In order to construct this utility function, the agent ranks the markups by decreasing rank of expected reward. The markup with the highest expected reward will receive a higher perceived utility value. This transformation also takes into account the search propensity parameter, such that a low parameter will be associated with a conservative utility function. A high search parameter, on the other hand, defines a more adventurous agent in trying different markups. Finally, the agent

[8]The intervals are constructed in a fairly *ad-hoc* way. They were defined in order to allow enough freedom in the possible choices (allowing for increases and decreases in offers and bids or infinitely negative or infinitely positive bids or offers).

will transform the utility function into a policy—an association between each markup and the probability of bidding (offering) that markup. The agent's policy will be used to choose the price in the following day. The basic ideas behind the construction of the utility function and the derivation of the policy are inspired by the fitness function and selection mechanisms that have been used in genetic algorithms [28], [29].

Let $j = 1, \ldots, 10$ represent the interval index (the markup number) and $t = 1, \ldots, K$ represent the iteration. Then, $\mathrm{Prf}_{tj}(\mathrm{Art}_{tj})$ represents the daily profit (acceptance rate) at iteration t using markup j. $\mathrm{Exp}(\mathrm{Prf}_{tj})$[9] represents the expected time t daily profit (conditional on acceptance) of markup j. $\mathrm{Exp}(\mathrm{Art}_{tj})$ represents the expected time t acceptance rate using markup j. $\mathrm{Exp}(\mathrm{Rwd}_{tj})$ represents the expected time t reward of markup j. $\mathrm{Rank}(j)$ stands for the rank of the markup j. Util_j and Pol_j stand for the perceived utility and the probability of using a markup j. The policy is calculated with the following algorithm. At the end of the day, after receiving the feedback with the prices and quantities traded in each hour:

Step 1) Calculate the new expected daily profit and acceptance rate for the markups used.

Let Profit_{tj}^{i} represent the profit at hour i and iteration t of the markup j, then the daily profit and the acceptance rate will be, respectively, $\mathrm{Prf}_{tj} = \sum_{i=1}^{24} \mathrm{Profit}_{tj}^{i}$ and

$$\mathrm{Art}_{tj} = \frac{\text{Number of bids (offers) accepted}_{tj}}{24}.$$

Then, for each used markup j

$$\mathrm{Exp}(\mathrm{Prf}_{tj}) = \mathrm{Exp}(\mathrm{Prf}_{t-1j}) + \alpha \cdot [\mathrm{Prf}_{t-1j} - \mathrm{Exp}(\mathrm{Prf}_{t-1j})]$$
$$\mathrm{Exp}(\mathrm{Art}_{tj}) = \mathrm{Exp}(\mathrm{Art}_{t-1j}) + \alpha \cdot [\mathrm{Art}_{t-1j} - \mathrm{Exp}(\mathrm{Art}_{t-1j})].$$

Step 2) Recalculate the expected reward for each markup. For every markup j

$$\mathrm{Exp}(\mathrm{Rwd}_{tj}) = \mathrm{Exp}(\mathrm{Prf}_{tj}).\mathrm{Exp}(\mathrm{Art}_{tj}).$$

Step 3) Rank the markups j by descending value of the expected rewards. The ranking idea is used in the "genetic algorithms" literature to avoid the fast convergence of the selection process, possibly to a second best solution.

Step 4) Calculate the perceived utility of each markup j

$$\mathrm{Util}_j = U \cdot \left(\frac{\text{Search Propensity} - \text{n}}{\text{Search Propensity}} \right)^{\mathrm{Rank}(j)-1}$$

where, for each agent, U, *Search Propensity*, and n equal 1000, 4, and 3, respectively. This approach is quite flexible enabling the construction of a wide variety of utility functions. After calculating the perceived utility from each markup, the agent transforms this utility function into a policy.

Step 5) Calculate the "policy," i.e., the probability of using each markup j. For this purpose, we have used the rule of proportionality: the probability of choosing a certain markup is directly proportional to the weight of that markup perceived utility in the sum of perceived utilities of all markups

$$\mathrm{Pol}_j = \frac{\mathrm{Util}_j}{\sum_k \mathrm{Util}_k}.$$

This algorithm differs from Erev and Roth's work [6], [7] on three main points.

1) *Our model is a pure strategy stochastic game*. Our agents try to learn the best pure strategy and the learning process is organized in order to find pure strategy equilibria
2) *Our model takes into account the exploration vs exploitation problem*. In Erev–Roth's model, these concepts are not present. In the reinforcement learning problem [25]–[27], the agent has to define how much exploration and exploitation to undertake. In order to allow agents to explore new strategies, we used the utility function model. Thus, agents only use, with a high probability, the best strategy or the first three best strategies and only with a small probability do they use the other ones.
3) *The players always have the same probabilities of exploitation or exploration, independently of the iteration of the game*. This approach keeps the agents' capacity of reaction to changes constant along the simulation: this makes the model suitable to deal with nonstationary environments. In Roth–Erev's basic model [6], the probability of choosing a certain strategy is directly proportional to the expected reward. The implication of this is that Roth–Erev's algorithm tends to converge to local equilibria with insufficient search of the solution space.

Table II illustrates the learning process.

Overall, the simulation model can be summarized by the behavioral pseudocode outlined in the Appendix.

IV. Initial Simulations of the New Electricity Trading Arrangements

The NETA simulation model, as described above, was applied to the full system of England and Wales as it existed in summer 2000 with 80 generating plants owned by 24 generators who sell power to 13 suppliers. All the experiments have simulated 200 iterations (trading days) of the algorithm based upon the winter demand profile, shown in Fig. 1, with an available capacity of around 56 GW.

In the experiments presented, we use a 0.5 smoothing parameter (learning rate). The retail price was fixed at £80. The plant availability was defined as a function of the technology: base-load technology is available 99% of the days, while the flexible technology is available 95% of the days.

The estimated *marginal generation costs* for each plant ranged from £3 MWh to £94 MWh. The low marginal costs are associated with base-load plants, nuclear, combined cycle gas turbine (CCGT), and some large coal plants. The high marginal costs are associated with gas turbines, pumped storage, and oil plants. The estimates used are consistent with those used in other published studies on the U.K. generation market, as well as with known data on plant efficiencies and fuel costs. The

[9]In our notation, $\mathrm{Exp}(X_{tj})$ stands for the expected time t value of the variable X at time $t-1$.

TABLE II
POLICY DERIVATION EXAMPLE

Mark-up Categories	1	2	3	4	5	6	7	8	9	10
Exp_{tj}(Prf)	500	400	600	300	1000	700	800	850	750	900
Exp_{tj}(Art) (%)	100	94	98	80	85	70	65	70	60	55
Exp_{tj}(Rwd)	500	376	588	240	850	490	520	595	450	495
Rank(j)	5	9	3	10	1	7	4	2	8	6
Util_j	3.9	0	62.5	0	1000	0.3	15.6	250	0.1	1
Pol_j (%)	0.3	0	4.7	0	75	0	1.2	18.8	0	0.1

Agent has ten different possible markups and it builds a policy on them. $\mathrm{Exp}(\mathrm{Prf}_{tj})$, $\mathrm{Exp}(\mathrm{Rwd}_{tj})$, and $\mathrm{Exp}(\mathrm{Art}_{tj})$ represent, respectively, the expected time t profit and reward and the expected time t acceptance rate of each markup j. $\mathrm{Rank}(j)$ orders the rewards from the highest to the lowest expected reward. Util_j represents the perceived utility an agent receives from a certain markup. Pol_j represents the probability of using a certain markup when bidding (offering).

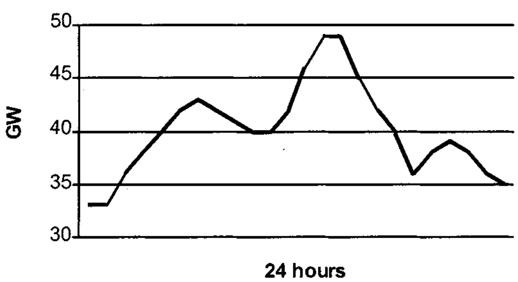

Fig. 1. Standard winter daily demand profile: the demand profile used in the experiments presented in this paper has three peaks of demand in order to make the generators task more difficult (they have to decide when to start a plant given the demand cycles).

relation between cycles and the plants' positions in the supply curve is presented in Fig. 2.

A. Market Results

Initial results give an overview of how the model is working and the evolution of prices and quantities traded in the PX and BM. In Table III, we analyze the impact of suppliers' prediction error on the quantities sold and bought in the BM and on the imbalances (spillage and top-up).

The quantity traded (bought or sold) in the BM is less than 2.5% of the total trade in the PX. As one might expect, the prediction error is a determinant factor in the amount traded and these results point to the economic gain, which could be achieved through better forecasting, by the suppliers (these agents are the ones that will be paying the imbalance charges resulting from prediction errors). Whilst forecasting load at the national level has become very accurate (2% mean absolute percentage error), at the regional level, where metering is less frequent and where there had not been a need to forecast accurately in such a short time scale, errors of between 5%–10% are thought to be the norm. Imbalances and trade in the BM are different concepts. The spillage and top-up represent from 2.5% to 2.9% of total trade in the PX, with a low prediction error and represent 5.0%–6.1% of total trade in the PX for the higher prediction error. The reason why imbalances represent a higher percentage than trades in the BM is that some positive imbalances can cancel the negative ones. Overall, these results show that the risks of NETA are much greater for suppliers. The

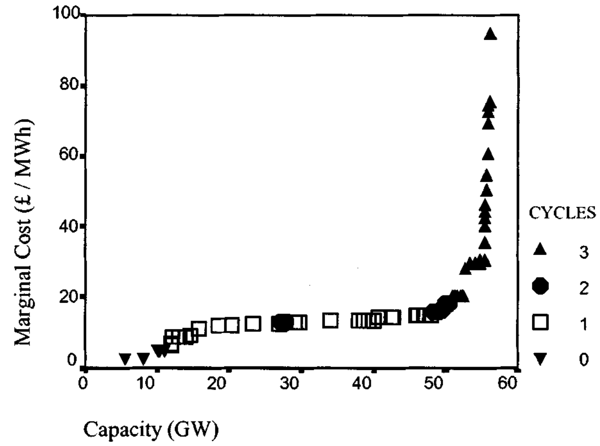

Fig. 2. Plant cycles in the supply curve: Plots the type of cycle defined for each plant along the supply curve. On average flexible plant have higher marginal cost.

TABLE III
QUANTITIES TRADED IN THE BM AND TOTAL IMBALANCES

Average Prediction Error	Sell BM	Buy BM	Spill	Top-up
10.0	2.4	1.3	5.1	6.2
5.0	1.3	0.9	2.5	2.9

Sell (buy) BM stands for the quantity sold (bought) in the BM. Spill (top-up) is the total electricity that suppliers consumed below (above) the FPNs or generators produced above (below) the FPNs. All quantities are presented as a percentage of the total trade in the PX.

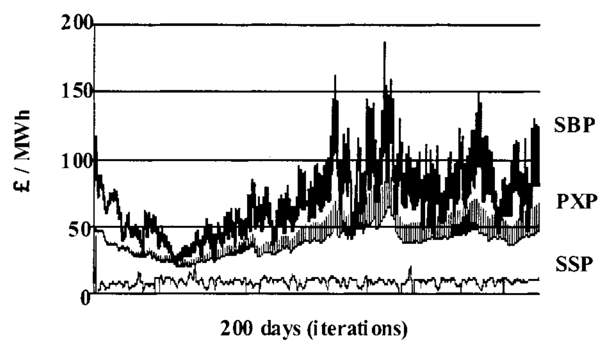

Fig. 3. Average daily prices (moving 24 hours window): PXP, SBP and SSP. Price evolution in the PX and in the BM for the 200 iterations of the experiment.

generators can control their imbalances with the exception of unplanned technical outages, while the suppliers are completely dependent on industry prediction capabilities.

An example of how the model learns can be seen in Fig. 3, which displays a representation of the average daily prices (24-h moving window) in the PX and in the BM (PXP, SBP, and SSP) during the 200 iterations of the baseline experiment. Note that market prices do emerge to create a wide spread between SBP and SSP and that PXP is centrally located between them. This is what the advocates of NETA hoped would occur in that out-of-balance players would regret they had not traded forward at PXP. The level of prices that have emerged around £50 MWh for winter days is rather high because in this experiment, as will be seen below, agents learn to exercise market power and create price spikes at the two peak periods. Notice that the SBP is much

TABLE IV
FINAL ITERATION FOR THE WINTER DAY BASELINE SCENARIO

Demand (GW)	Marginal Cost (MC)	Capacity Used (%)	PXP	SBP	Mark-up PXP/MC (%)	Ratio SBP/PXP	Mark-up SBP/MC (%)
33	13.25	58.8	14.55	16.57	9.8	1.14	25
33	13.25	58.8	14.53	18.00	9.7	1.24	36
36	13.28	64.2	14.89	17.79	12.2	1.19	34
38	13.36	67.7	15.44	19.00	15.6	1.23	42
40	14.32	71.3	15.71	20.00	9.7	1.27	40
42	14.36	74.8	16.99	34.32	18.3	2.02	139
43	14.60	76.6	17.72	34.16	21.4	1.93	134
42	14.36	74.8	16.88	30.85	17.6	1.83	115
41	14.36	73.1	16.03	29.27	11.6	1.83	104
40	14.32	71.3	15.72	20.00	9.8	1.27	40
40	14.32	71.3	15.78	26.50	10.2	1.68	85
42	14.36	74.8	16.71	33.28	16.4	1.99	132
46	14.62	82.0	29.46	63.00	101.5	2.14	331
49	15.87	87.3	318.72	603.70	1908.0	1.89	3703
49	15.87	87.3	509.18	1100.76	3108.0	2.15	6830
45	14.42	80.2	23.65	51.49	64.1	2.18	257
42	14.36	74.8	16.29	20.00	13.5	1.23	39
40	14.32	71.3	15.67	26.60	9.4	1.70	86
36	13.28	64.2	15.15	16.62	14.1	1.10	25
38	13.36	67.7	15.47	18.55	15.8	1.20	39
39	13.36	69.5	15.60	26.80	16.8	1.72	101
38	13.36	67.7	15.49	24.78	16.0	1.60	86
36	13.28	64.2	14.96	19.00	12.7	1.27	43
35	13.28	62.4	14.83	18.00	11.7	1.21	36

It describes the quantity demanded, the marginal costs, and prices (price in the PX and the SBP) during the last iteration for each hour of the day. Markups represent the percent increase of the PXP and of the SBP compared to the marginal cost.

more volatile than the PXP or the SSP. Notice also the emergence of daily price cycles (with high prices at the peaks) as agents learn from experience.

To look at the daily price profile in more detail, the actual prices on the last iteration are shown in Table IV.

The peak prices spikes are noteworthy, as the PXP ranges from £14.5 to £29.5 in the offpeak hours to £318 and £509 in the two peak hours. The markups PXP/MC above 60% are associated with *capacity used* above 80% and the "price explosion" only happens with *capacity used* above 85%.

The prices can be divided into three periods:

1) a low demand period (below 80% of *available capacity*), where the prices tend to be close to marginal costs;
2) an average demand period (between 80% and 85% of *available capacity*), where the prices are at least 60% above the marginal costs;
3) a high demand period (above 85% of *available capacity*), where the prices rise to at least 38 times the marginal cost.

Notice the small difference in *capacity used* between the high demand and the average demand period: this is due to the supply function's shape.

These results are stronger than the ones in [4]. The experiments show that even if the demand is lower than the total capacity, prices may rise well above the marginal cost. Thus, there is a threshold above which collusion is possible and sustainable and there is another above which only the regulator can stop prices from unbounded spiking.

In the rest of the analysis, we disregard these two hours with demand of 49 GW, since the equilibrium was not reached. It is also interesting to note that the offpeak prices (without these two hours) show a profile of PXP's consistent[10] with the 6-mo forward prices, which were emerging in summer 2000, ahead of the planned NETA introduction during the winter (and indeed close to what did occur in the early months of NETA after March 2001).

B. Strategic Implications for Suppliers

The 13 suppliers used in the experiments had market shares of between 3.6% and 12.2%. The analysis of the suppliers initially focused on two main issues: demand prediction capabilities and contract cover (how much of there expected demand they want to buy in the PX). In Table V, we specify the five experiments on this theme.

We used regression analysis to test the influence of the separate effects. The regression for the offpeak hours, where the dependent variable is the suppliers' *profit per unit sold*, is reported in Table VI.

The regression in Table VI shows that the lower the prediction error, the higher the profit per unit sold of the industry as a whole, given an objective for contract cover of 100%. The experiments have also shown that a contract cover of 115% (85%) has a negative (positive) impact on the suppliers' profit per unit sold. This is a result that surprised the supply industry since con-

[10]In August 2000, forward prices for offpeak winter 2000/2001 were averaging £22/MWh.

TABLE V
EXPERIMENTS WITH SUPPLIERS' PARAMETERS

Exp. Number	Prediction Error	Contract Cover
1	10	100
2	10	115
3	5	115
4	5	100
5	5	85

In each experiment, a different combination of demand prediction capabilities and contract cover objective is tried.

TABLE VI
DEPENDENT VARIABLE—SUPPLIERS PROFIT PER UNIT SOLD

		Two	Three	Four	Five
Coef.	60.8	-5.79	-11.87	0.71	2.4
t-stat	(396)	(-26.7)	(-54.8)	(3.2)	(11.2)

"Two," "Three," "Four," and "Five" are dummy variables representing experiments 2, 3, 4, and 5, respectively (i.e., "Two" is 1 to indicate experiment two or zero, otherwise). Experiment one is our base case. t statistics are presented in parentheses.

ventional wisdom has been that they would be wise to be risk averse to imbalances and, if anything, be slightly overcontracted in their FPNs. In our model, however, the intuition is that if all suppliers undercontract, the PXP falls given the reduction in demand. Following that, the generators expect a lower price in the BM and also have extra capacity available to sell, so the BM price falls as well.

C. Strategic Implications for Generation

In order to study generators' behavior, we analyze two dependent variables: profit per unit of capacity available and price offered in the PX (price-offered PX). Note these variables are defined at the plant level. The objective is to identify a relationship between generators' behavior and plant ownership structure. The regression parameters for the profits per unit of capacity available in the offpeak hours (for the accepted offers in the PX) are presented in Table VII (Model 1). The regression parameters for price-offered PX (for the accepted offers in the offpeak hours) are presented in the Table VII (Model 2).[11]

Applying hierarchical clustering to Model 1 and Model 2 parameters, we have identified the following strategic groups of generators.

1) *AES, MG, and PG:* These generators have lower profits per unit and lower prices than average. These players own base-load plants that they use intensively with low profits per unit of capacity available.
2) *BE and EDF:* These generators have a low price and continuous running policy.
3) *EP:* This generator has a portfolio based on some base-load plants (large coals) with "high" marginal costs and a few very flexible and "low" cost plants, the pumped storage. EP had a very good performance on the pumped storage and a below average performance on the large coals.

[11]Although some of the estimated regression parameters are not statistically significant, the overall results do provide an indication for the informal clustering into strategic groups.

TABLE VII
DUMMY VARIABLES

Var.	Model 1	Model 2
	5.9	16.7
	(57.2)	(168)
C1	-2.8	0.3
	(-27.6)	(2.9)
C2	0.16	1.3
	(0.87)	(7.5)
C3	4.3	4.8
	(14)	(16.8)
AES	-1.4	-0.7
	(-12.2)	(-6.4)
BE	1.23	-0.3
	(9.8)	(-2.6)
EDF	1.9	-0.4
	(15.5)	(-3.3)
EP	-0.05	-0.3
	(-0.34)	(-2.2)
MG	-0.4	-0.2
	(-2.1)	(-1.4)
NP	-0.15	0.9
	(-0.8)	(5.6)
PG	-2.1	-0.3
	(-17.2)	(-2.9)
TXU	0.5	2.6
	(2.4)	(12.5)

AES, BE, EP, MG, NP, PG, and TXU identify the seven main owners of plant portfolios. C1, C2, and C3 identify plants with one cycle, two cycles, and three cycles, respectively. Dependent variables in model 1 and model 2 are the plant profit per unit of capacity available and the plant prices offered PX (for the accepted offers only), respectively. t statistics are presented in parentheses.

4) *NP and TXU:* These generators belong to the same structural group, characterized as "diverse portfolios with dominant positions." These agents have a policy of high pricing (prices above average).

The flexible technology (two and three cycles) tends to price and have profits per unit above average. The fact that a plant belongs to a certain portfolio of plant owned by different generators also has a significant effect in its pricing behavior. This suggests that the interrelatedness of plant ownership and profits will continue to promote the active buying and selling of plant through the capital markets, as generators seek to reposition their portfolios of plant.

V. CONCLUSION

Agent-based computational methods can provide insights into pricing and strategic behavior in complex new markets such as electricity. Imperfect competition, exercised through the daily repetition of a competitive market with administered market rules, creates a process of continuous experimentation and gaming that agent-based simulation is able to imitate. Overall, the results obtained in this study were plausible to the industry, the prices reasonably well calibrated, and the model represented the first detailed study of the interrelationship of the bilateral trading to the balancing market in the proposed NETA for the U.K.

In summary, the insights into NETA which came out of this modeling were the following.

1) The amount of trading in the BM may be less than 2% of the total amount sold in the bilateral markets, but the imbalances will be much higher (how high depends on the prediction errors by the suppliers).
2) There is no incentive for the SO to publish more accurate predictions, yet this would improve the efficiency of the industry.
3) NETA will have a greater risk impact on suppliers than on generators, even though one of the ideals of the reform was to make the wholesale electricity market less of a "generators' market."
4) The capacity margin (demand as a percentage of total capacity available) seems to be the most important factor behind the possibility of collusive behavior on the generation side. There is a threshold above which collusion is possible and sustainable and there is another above which only the regulator can restrain unbounded spikes. Other experiments not reported in this paper on the parameter *availability* have shown that withholding capacity may have an extreme impact on the peak prices (but, in the experiments reported, a tightening of the capacity margin by 5% had a important impact on prices).
5) The experiments have also shown that a contract cover of 115% (85%) has a negative (positive) impact on the suppliers' profit per unit sold. If they collectively learn to undercontract in this way, they will exert some market power on the supply side.
6) Flexible plant will be relatively more valuable in the new market mechanism, yet its value will depend upon the portfolio of ownership. This interrelatedness will ensure that there will be continued activity in buying and selling plant amongst generators as they seek to reposition themselves in the market.

Finally, there is always a caveat in the use of models such as these. They embed a number of behavioral assumptions and are constrained by exogenous parameters. For example, the learning model is based upon adaptive expectations between the PX, the BM, and the next day's PX and BM. It is plausible to think of more sophisticated agents, but that may raise computational limitations. The model simulated here is already very large and complex by multiagent standards. Similarly, it would be attractive if some of the parameters were endogenous variables. Contract cover is one, in particular, where we would like to see some emergent behavior. Nevertheless, this application has represented a major innovative application of agent-based simulation to a real issue—that of understanding the implications NETA, where there had previously been a shortage of reliable analytical, empirical, and experimental insight.

APPENDIX

```
Pseudocode for the NETA Simulation System
Behavior
For I = 1:Number of iterations
    Step 1) Suppliers predict demand for each
        hour
    Step 2) Generators define which plants can
        run
    Step 3) Generators offer in the PX
    Step 4) Suppliers bid in the PX
    Step 5) Trading in the PX and calculation
        of System Position in each one of the
        hours
    Step 6) Generators and Suppliers offer
        (bid) into the BM
    Step 7) Trading in the BM and calculation
        of imbalance prices
    Step 8) Settlement Process: calculation of
        imbalances for each one of the suppliers
        and generators (plant by plant)
    Step 9) Learning
End
```

```
Subcode for Suppliers' Learning Behavior
    Step 1) Verify if the objective for the BM
        Exposure was achieved. If it was not,
        penalize the profit obtained in the PX.
    Step 2) Given the daily profits in the PX
        and in the BM revise, for the markup
        used, the:
        - Expected profit in the PX and in the
        BM
        - Expected acceptance rate in the PX and
        in the BM
    Step 3) Define the new bidding policy for
        the next day.
    In the beginning of the next day:
    Step 4) After calculating predictions for
        demand, define the quantities and prices
        bid, making sure that the operational
        rule of behavior is respected.
```

```
Subcode for Generators' Learning Behavior
For each plant in the portfolio:
    Step 1) Verify if the objectives for the BM
        Exposure and the Position in the Load
        Duration Curve were achieved. If they
        were not achieved penalise the profit
        obtained.
    Step 2) Given the daily profits in the PX
        and in the BM revise, for the markup
        used, the:
        - Expected profit in the PX and in the
        BM.
        - Expected acceptance rate in the PX and
        in the BM.
    Step 3) Define the new offering policy for
        the next day.
    In the beginning of the next day:
    Step 4) Define the quantities and prices
        offered for each plant, after knowing
        which plants are available, making sure
        that all the operational rules of
        behavior are respected.
```

REFERENCES

[1] R. Green, “Increasing competition in the British electricity spot market,” *J. Ind. Econom.*, vol. 44, no. 2, pp. 201–216, 1996.

[2] A. Rudkevitch, M. Duckworth, and R. Rosen, “Modeling electricity pricing in a deregulated generation industry: The potential for oligopoly pricing in a pool,” *Energy J.*, vol. 19, no. 3, pp. 19–48, 1998.

[3] R. Green and D. Newbery, “Competition in the british electricity spot market,” *J. Political Econ.*, vol. 100, no. 5, pp. 929–953, 1992.

[4] N.-H. M. von der Fehr and D. Harbord, “Spot market competition in the U.K. electricity industry,” *Econom. J.*, vol. 103, no. 418, pp. 531–546, May 1993.

[5] M. H. Rothkopf, “Daily repetition: A neglected factor in the analysis of electricity auctions,” *Electricity J.*, vol. 12, no. 3, pp. 61–70, Apr. 1999.

[6] A. E. Roth and I. Erev, “Learning in extensive-form games: Experimental data and simple dynamic models in the intermediate term,” *Games Econom. Behav.*, vol. 8, no. 1, pp. 164–212, Jan. 1995.

[7] I. Erev and A. E. Roth, “Predicting how people play games: Reinforcement learning in experimental games with unique, mixed strategy equilibria,” *Amer. Econom. Rev.*, vol. 88, no. 4, pp. 848–881, Sept. 1998.

[8] R. A. Bettis and C. K. Prahalad, “The dominant logic: Retrospective and extension,” *Strategic Manage. J.*, vol. 16, no. 1, pp. 5–14, Jan. 1995.

[9] (1999) Role playing simulations of the new electricity trading arrangements, a report to the Reta Programme. London Economics, London, U.K. [Online]. Available: http://www.londecon.co.uk/

[10] J. Bower and D. W. Bunn, “A model-based comparison of pool and bilateral market mechanisms for electricity trading,” *Energy J.*, vol. 21, no. 3, pp. 1–29, 2000.

[11] ——, “Experimental analysis of the efficiency of uniform-Price v discriminatory auctions in the england and wales electricity market,” *J. Econom. Dyn. Control*, vol. 25, no. 3–4, pp. 561–592, 2001.

[12] (1999, July) The new electricity trading arrangements. Office of Gas and Electricity Markets, London, U.K. [Online]. Available: http://www.ofgem.gov.uk/docs/reta3.pdf

[13] (1999, Oct.) The new electricity trading arrangements, Ofgem DTI conclusions document. Office of Gas and Electricity Markets/DTI, London, U.K. [Online]. Available: http://www.ofgem.gov.uk/docs/netaoct.pdf

[14] (2000, Mar.) Settlement administration agent user requirements specification. Office of Gas and Electricity Markets, London, U.K. [Online]. Available: http://www.ofgem.gov.uk/elarch/retadocs/mitbpm08.pdf

[15] (2000, Apr.) NGC incentives under NETA. Office of Gas and Electricity Markets, London, U.K. [Online]. Available: http://www.ofgem.gov.uk/docs/ngcfeb2.pdf

[16] B. LeBaron, W. B. Arthur, and R. Palmer, “Time series properties of an artificial stock market,” *J. Econom. Dyn. Control*, vol. 23, no. 9–10, pp. 1487–1516, Sept. 1999.

[17] E. Ephrati and J. S. Rosenschein, “Deriving consensus in multiagent systems,” *Artif. Intell.*, vol. 87, no. 1–2, pp. 21–74, Nov. 1996.

[18] P. R. Milgrom, “Auction Theory,” in *Advances in Economic Theory: Fifth World Congress*, T. Bewley, Ed. Cambridge, U.K.: Cambridge Univ. Press, 1987, pp. 1–32.

[19] R. Wilson, “Game-theoretic analysis of trading processes,” in *Advances in Economic Theory: Fifth World Congress*, T. Bewley, Ed. Cambridge, U.K.: Cambridge Univ. Press, 1987, pp. 33–70.

[20] T. N. Cason and D. Friedman, “Price formation in single call markets,” *J. Econometrica*, vol. 65, no. 2, pp. 311–345, Mar. 1997.

[21] J. Nicolaisen, V. Petrov, and L. Tesfatsion. (2000, Aug.) Market power and efficiency in a computational electricity market with discriminatory double-auction pricing. Dept. Economics, Iowa State Univ., Ames, IA. [Online]. Available: http://www.econ.iastate.edu/tesfatsi/mpeieee.pdf

[22] D. Carmel and S. Markovitch, “Learning models of intelligent agents,” in *Proc. 13th Nat. Conf. Artificial Intelligence*, Aug. 1996, pp. 62–67.

[23] A. Marcet and J. P. Nicollini. (1997) Recurrent hyperinflations and learning. Dept. Econ. and Business, Univ. Pompeu Fabra, Barcelona, Spain. [Online]. Available: http://www.econ.upf.es/

[24] S. Franklin and A. Graesser, “Is it an agent, or just a program?: A taxonomy for autonomous agents,” in *Proc. 3rd Int. Workshop Agent Theories, Architectures, and Languages*, Aug. 1997, pp. 21–35.

[25] G. Weiss, “Adaptation and learning in multi-agent systems: Some remarks and a bibliography,” in *Adaptation and Learning in Multiagent Systems*, G. Weiss and S. Sen, Eds. Berlin, Germany: Springer-Verlag, 1995, pp. 1–21.

[26] L. P. Kaelbling, M. L. Littman, and A. W. Moore, “Reinforcement learning: A survey,” *J. Artif. Intell. Res.*, vol. 4, pp. 237–285, Jan. 1996.

[27] R. S. Sutton and A. G. Barto, *Reinforcement Learning: an Introduction*. Cambridge, MA: MIT Press, 1998.

[28] Z. Michalewicz, *Genetic Algorithms + DataStructures = Evolution Programs*, 2nd ed. Berlin, Germany: Springer-Verlag, 1992.

[29] D. Whitley and J. Kauth, “Genitor: a different genetic algorithm,” Dept. Comput. Sci., Colorado State Univ., Fort Collins, CO, Tech. Rep. CS-88–101, 1988.

Derek W. Bunn received the M.A. degree from Cambridge University, Cambridge, U.K., and M.Sc. degree from Imperial College, London University, U.K., and the Ph.D. degree from the London Business School, London Univeristy, U.K.

He is currently a Professor of Decision Sciences at the London Business School, London, U.K. Previously, he was with Stanford University and Oxford University. He is the Editor-in-Chief of the *Journal of Forecasting* and *Energy Economics* and has been on the editorial boards of ten other journals

Fernando S. Oliveira received the B.S. degree in economics and the M.Sc. degree in artificial intelligence and computation from the University of Porto, Porto, Portugal, in 1995 and 1999, respectively. He is currently working toward the Ph.D. degree in decision sciences at the London Business School, London, U.K.