

Multi-Agent Models for Consumer Choice and Retailer Strategies in the Competitive Electricity Market

Graziano ABRATE, Ettore BOMPARD, Roberto NAPOLI, Bo WAN



Working Paper 12, 2006

© HERMES
Fondazione Collegio Carlo Alberto
Via Real Collegio, 30
10024 - Moncalieri (TO)
Tel: 011 670 5250
Fax: 011 6705089
info@hermesricerche.it
<http://www.hermesricerche.it>

I diritti di riproduzione, di memorizzazione e di adattamento totale o parziale con qualsiasi mezzo (compresi microfilm e copie fotostatiche) sono riservati.

PRESIDENTE

Giovanni Fraquelli

SEGRETARIO

Cristina Piai

COMITATO DIRETTIVO

Giovanni Fraquelli (*Presidente*)
Cristina Piai (*Segretario*)
Guido Del Mese (ASSTRA)
Carla Ferrari (Compagnia di San Paolo)
Giancarlo Guiati (GTT S.p.A.)
Mario Rey (Università di Torino)

COMITATO SCIENTIFICO

Tiziano Treu (*Presidente*, Università "Cattolica del Sacro Cuore" di Milano e Senato della Repubblica)
Giuseppe Caia (Università di Bologna)
Roberto Cavallo Perin (Università di Torino)
Giovanni Corona (CTM S.p.A.)
Graziella Fornengo (Università di Torino)
Giovanni Fraquelli (Università del Piemonte Orientale "A. Avogadro")
Carlo Emanuele Gallo (Università di Torino)
Giovanni Guerra (Politecnico di Torino)
Marc Ivaldi (IDEI, Université des Sciences Sociales de Toulouse)
Carla Marchese (Università del Piemonte Orientale "A. Avogadro")
Luigi Prosperetti (Università di Milano "Bicocca")
Alberto Romano (Università di Roma "La Sapienza")
Paolo Tesauro (Università di Napoli "Federico" II)

Multi-Agent Models for Consumer Choice and Retailer Strategies in the Competitive Electricity Market

Abrate Graziano ^(*), Bompard Ettore ^(**), Napoli Roberto ^(**), Wan Bo ^(**)

Abstract: The benefit to end-users of electricity mainly in terms of low prices is strictly related to the interaction among the consumers themselves and the retailers. The literature on retail competition has highlighted possible distortions in the market outcomes due to unresponsive demand and boundedly rational consumers. Hence a model appropriate to represent this aspect of the market is needed. The paper proposes a general model of the interaction among retailers and consumers in the electricity market based on the theory of multi-agent systems. The objective is to incorporate different assumptions on the behavior of the players that correspond to different levels of interaction in determining the market outcomes. The model is validated interpretively by the results provided in terms of game theory equilibrium. The results of an extensive set of simulations on a population of 1000 consumers is presented and discussed.

Key words — Flat Rate, Time-of-Use (TOU) Pricing, Multi-agent, Retail competition.

Acknowledgement

We gratefully acknowledge financial support by the Lagrange project of Turin – supported by Institute for Scientific Interchange (ISI) Foundation and CRT Foundation – in addition to HERMES - Higher Education and Research on Mobility regulation and Economics of local Services – Turin.

* HERMES Research Center (Collegio Carlo Alberto – Moncalieri – TO).

Email address: g.abrate@hermesricerche.it

** Polytechnic of Turin, Department of Electrical Engineering (DELET).

1. INTRODUCTION

Deregulation aims at providing market-based efficient electricity prices to consumers. Wholesale competition is enhanced, on the supply side, by participation of several generations firms, and, on the demand side, by promoting consumer choice and retail competition. An extensive literature has grown up on the vertical relations between markets, discussing the opportunity to pass the Real Time Price (RTP) signals generated in the wholesale market on to the final consumers. Recently, Borenstein [1, 2] studied the RTP-induced wealth transfers from producers to consumers; Holland and Mansur [3] analyzed the short-run effects of time-varying prices on electricity load, prices, consumer surplus, generator profits, efficiency, and emissions. Other works have focused on the measurement of consumer response to real time rates (e.g., Faruqui and George, [6]).

Moreover, several works have addressed the issue of electricity market efficiency under retail competition, together with the problem of unresponsive consumers ([8, 11]). In particular, Borenstein and Holland [8] have shown that when a part of the demand is insensitive to wholesale price variations, retail competition would fail to achieve the second-best optimum, attainable by a regulator with perfect information on the demand curves. Joskow and Tirole [11] have stated that if retailers are able to measure real time consumption for each consumer, this distortion still holds if consumers are assumed to be heterogeneous and boundedly rational. Though, a much little attention has been given to the modeling of the retail market itself, considering the consumer choice between alternative pricing schemes, and the consequence of this behavior on retailers' strategies.

The most important data which characterizes consumer heterogeneity is the load profile, i.e. the hourly distribution of his consumption. In fact, a basic challenge for retailers is competing for the more convenient consumers from a load profile point of view. The identification of the *cheaper* consumers calls on the need of installing proper metering systems, in order to compute the actual average cost of serving consumers with different load profiles.¹ Retailers strategies are strictly related on whether it is optimal to pass on to the final consumers the wholesale price signal or to offer a different price schedule. To simplify the analysis, the question may be whether to offer RTP, Time of Use (TOU) or flat rates. then, it is important to understand how consumers will self-select to these two options.

RTP differs from TOU rates in that it is based on actual (as opposed to forecasted) prices which may fluctuate many times a day and are load-sensitive,

¹ Hunt (2002) is a good reference for the description of retail competition challenges.

rather than varying with a fixed schedule. Though RTP is more efficient, TOU have been more widely used and accepted, in part because it is easier and less costly to implement. The consumer choice between these two options is related to risk management and unexpected fluctuations in load profiles and prices. When it comes to the choice between TOU (or RTP) and flat rates, which is the focus of our work, the average load profile becomes the most important variable to take into account, together with the potential elasticity of substitution across different pricing periods.

Under the hypothesis of hourly wholesale pricing, we can start pointing out two basic prediction of the economic theory.

A) Within a perfect competitive framework, a flat rate option is never offered by the retail market if switching to time varying prices does not imply any additional transaction costs to the consumers [8]. This result holds because, under the zero profit assumption of the retail side, the flat rate would simply be a weighted average of real time wholesale costs of serving flat consumers. To better understand the statement, suppose there are only two time-periods. If there is no risk aversion, all consumers whose load is higher during the cheapest period would save money by selecting the real time option; but this would mean an increase in the flat rate since the relative weight of the cheapest period would decrease. Then, more consumers would prefer to switch to real time pricing. All consumers will end up by selecting real time price, except the one who has the heaviest consumption during the more expensive period, and the latter would be indifferent between the two options. This is also the unique Nash equilibrium of the game, obtained by iterative elimination of dominated strategies.

B) Suppose that there are only 2 retailers in the market. Since they would engage in price competition (they have no constraints on how much they can buy from the wholesale markets), the result under uniform prices would be the Bertrand paradox, so retailers will make no profit. The possibility of price discrimination among consumers with different load profiles offering different price schedules may however produce positive profits for retailers.

As stated before, limited work has been devoted to this issue. This is partly due to the fact that it is not easy to give a specific mathematic model for the complex interaction between different players in a competitive electricity market. Nowadays, multi-agent approaches have attracted a lot of research interest when studying the power market problems. In a multi-agent approach, the players of the studied case are modeled as intelligent agents, who have the ability to react to the environment according to their past experience. This kind of approach is attractive because it is more practical to capture the try-and-error behavior mechanism of the market players without the need of an explicit mathematic model.

Nowadays, a lot of electricity market simulations have been carried out based on a multi-agent model using reinforcement learning algorithm since the

autonomous agents have the ability to react rationally according to their past experiences. This kind of models is especially suitable to the electricity markets where there exist a lot of autonomous agents such as generators, retailers and customers. Most of the works have focused on the modeling of the power exchange and mainly on generators strategies. Bagnall and Smith [14] and Bunn and Oliveira [15] discussed alternative market structures in the context of the UK electricity market; Xiong et al. [5] compared alternative auction mechanisms, i.e. the uniform price and pay-as-bid electricity auction markets; Watanabe et al. [13] studied the problem of congestion management and on its effects on generators and distributors companies' strategies. Fujii et al. [16] analyzed the pricing processes modeling the strategic bidding of the demand side on the basis of the utility function of customers. In addition, some software based on multi-agent model is available, for example MASCES: a multi-agent system that simulates competitive electricity markets [17]. However, to our knowledge, no works focused on retail competition and in the modeling of the characteristics of end customers.

In this paper, we are interested in describing the behavior of retailers and heterogeneous consumers within a framework of retail competition, interpreting the prediction from the multi-agent model in terms of equilibrium, provided by game theory. This framework allows capturing the idea that agents act rationally in the market but need time to learn their optimal strategies from the experience.

The work is organized as follows. Next section introduces the multi-agent model for retail market studies and also describes in details the retail market structure and section 3 forms the models for consumers and retailers and section 4 illustrates the market analysis provided with simulation results. Finally, section 5 summarizes the main insights from the analysis. There is also a list of symbols and notations at the appendix section 6.

2. MAS FOR RETAIL MARKETS STUDIES AND RETAIL MARKET STRUCTURE

2.1 MAS FOR RETAIL MARKETS STUDIES

We resort to multi-agent approach for modeling the interaction among the sellers (the retailers) and the buyers (the consumers) in competitive electricity markets. An agent is a rational entity who knows how to make his choice based on his past experience. *Multi Agent System* (MAS) is a system composed of a population of agents simultaneously pursuing individual objectives. In order to develop their best choice among a set of actions, the agents adjust their behavior by reinforcement learning algorithms. The basic intuition underlying any reinforcement learning algorithm is that the tendency to implement an action

should be strengthened if it produces favorable results and weakened if it produced unfavorable results [4].

The appeal of using a multi-agent approach is related to the simplified assumptions on agent behavior. For example, when consumers have to make their choices among alternative price schedules, they may not know exactly their load profile and the price variations at any time, but they certainly know the total bill they have to pay. As to the retailers, they may choose their strategies by simply observing their effect on their profits. The idea is that agents can learn from their past strategies; sometimes they may choose a wrong option, but they are able to learn from their own errors. This idea can be computationally implemented by using a *Q-Learning algorithm*.

Q-Learning algorithm is a kind of reinforcement learning algorithm proposed by Watkins for solving the Markovian Decision Problems (MDPs) with incomplete information [4]. It does not need an explicit model of its environment and can be used on-line to find the optimal strategy through experience obtained from the direct interaction with its environment.

The *Q-Learning algorithm* can be mathematically formulated as follows.

Assume that an agent interacts with its environment at each of a sequence of discrete time points $t = t_0, t_1, \dots, t_n$. Let $S = \{s_1, s_2, s_3, \dots, s_n\}$ be the finite set of possible states of the environment and $A = \{a_1, a_2, a_3, \dots, a_m\}$ be the finite set of admissible actions the agent can take. At each time step t , the agent senses the current state $s_t = s \in S$ of its environment and on that basis selects an action $a_t = a \in A$ (Figure 1). As a result of its action, the agent receives an immediate reward r_{t+1} , and the environment's state changes to the new state $s_{t+1} = s' \in S$, with a transition probability $P_{ss'}(a)$.

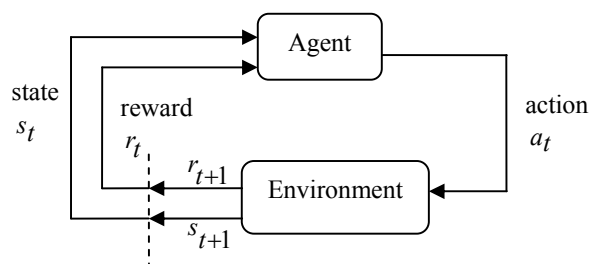


Figure 1. Agent's interaction with the environment

The objective of the agent is to find an optimal policy $\pi^*(s) \in A$ for each state s to maximize the total amount of reward it receives over the time span $T = t_n - t_0$.

Q-Learning algorithm provides an efficient real time approach to determine the optimal policy by estimating the optimal Q -values $Q^*(s,a)$ for pairs of states and admissible actions.

The Bellman optimality equation for $Q^*(s,a)$ is given as follows [4],

$$Q^*(s,a) = \sum_{s'} P_{ss'}(a) [R_{ss'}^a + \gamma \max_{a'} Q^*(s',a')] \quad (1)$$

where $R_{ss'}^a = r_{t+1}$ is the immediate reward from taking action a in the state s and transitioning from state s to s' , a' is the admissible action in the new state s' , and $\gamma(0 \leq \gamma \leq 1)$ is a scaling factor used to discount the future rewards. If γ is small, it means that the expected future rewards count for less.

Any policy selecting actions that are greedy (always select the action with the greatest reward) with respect to the optimal Q -values is an optimal policy [5]. Thus, the optimal policy is

$$\pi^*(s) = \underset{a}{\operatorname{arg\,max}}(Q^*(s,a)) \quad (2)$$

Without knowing the transition probability $P_{ss'}(a)$ the *Q-Learning algorithm* can find the $Q^*(s,a)$ in a recursive manner by using the available information s_t , a_t , s_{t+1} and r_{t+1} .

The updating rule for *Q-Learning* can be explained as:

$$Q_{t+1}(s,a) = \begin{cases} Q_t(s,a) + \alpha \Delta Q_t(s,a) & \text{if } s = s_t \text{ and } a = a_t \\ Q_t(s,a) & \text{otherwise} \end{cases} \quad (3)$$

where $\alpha(0 < \alpha \leq 1)$ is the learning rate, $Q_t(s,a)$ and $Q_{t+1}(s,a)$ are the Q -values for state-action pair (s,a) at time t and $t+1$, respectively, and $\Delta Q_t(s,a)$ is given by:

$$\Delta Q_t(s,a) = \{r_{t+1} + \gamma \max_{a'} [Q_t(s_{t+1},a')]\} - Q_t(s,a) \quad (4)$$

The learning rate α reflects the degree to which estimated Q -values are updated by new data. High values imply more rapid updates, with a risk of instability.

If the Q -values for each admissible state-action pair (s, a) is visited infinitely time, and the learning rate α decreases over the time step t in a suitable way, then as $t \rightarrow \infty$, $Q_t(s, a)$ converges with probability one for all admissible state action pairs.

2.2 RETAIL MARKET STRUCTURE

We consider that retailers buy electricity in the wholesale power exchange and we assume that there are no other costs in serving electricity to the final consumer other than the wholesale cost². To further simplify the analysis, even if of course the method proposed can be applied to more general context, we assume there are only two pricing periods for each day, for example daily hours (D) and night hours (N), while retail prices are assumed to be updated on a weekly basis. The average weekly wholesale prices during the day and during the night are defined as P_{DW} and P_{NW} respectively. We further define the everyday system real time price on the following basis.

The system real time prices (P_D, P_N) change around the weekly average and are related to the variations in the system load profile. Defining the load profiles L_D and L_N as the percentage of every consumption during the day and night respectively, we write the expression of real time prices as follows:

$$P_D = P_{DW} + (L_D - L_{DW}) \cdot P_{DW} + e_{PD} \quad (5)$$

$$P_N = P_{NW} - (L_N - L_{NW}) \cdot P_{NW} + e_{PN} \quad (6)$$

With the terms e_{PD} and e_{PN} , we introduce a random variation in prices. We assumed their mean values to be equal to 0, and their distribution to be uniform over the interval $U[-\delta_{PD(N)} \cdot (P_{D(N)W}); \delta_{PD(N)} \cdot (P_{D(N)W})]$, where $\delta_{PD(N)}$ is a parameter which express the degree of random volatility in the wholesale prices.

For representing the demand in the market, we consider 1000 consumers, defined as heterogeneous in terms of load profiles, but homogeneous in terms of total units of electricity consumption. To model load profiles heterogeneity and variability, we resorted to the following assumptions:

- 1) the individual load profile changes everyday, but the weekly average individual load profile is fixed;
- 2) since all customers have the same total level of consumption, the system load profile is simply the average of each consumer's load profile. Thus we modeled heterogeneity assuming that the weekly average individual

² Since the paper focuses on the comparison of different tariff structures (Flat, Time-of-Use, Real-Time pricing), this simplifying assumption – which is often used in the literature (see Borenstein and Holland, 2004) – does not affect our analysis unless other retailers' costs (such as billing) are themselves affected by the type of tariff used.

load profile is distributed around the system average load, such that $L_{iDW} = L_{DW} + e_W$ (with e_W satisfying uniform distribution $U[-\delta_W L_{DW}; \delta_W L_{DW}]$, where δ_W is a parameter which express the degree of random volatility in the average weekly load;

- 3) the i^{th} consumer load profile changes around the weekly average: $L_{iD} = L_{iDW} + e_D$ (with e_D satisfying uniform distribution $U[-\delta_D \cdot L_{iDW}; \delta_D \cdot L_{iDW}]$, where δ_D is a parameter which express the degree of random volatility in the average weekly load of the i^{th} customer.

The parameter e_W ensures consumers' heterogeneity around the system average, while the parameter e_D attributes a daily variability to the individual load profile.

3. MODELS FOR CONSUMERS AND RETAILERS

3.1 Model of the behavior of consumers

Consumers compare their bills under alternative pricing schemes, and make their choices according to what happened in the past. The basic idea is that when they choose their pricing scheme they only have some expectation on how much they will be paying, and when they receive the bill they will update these expectations and their future choices. In formulas, the (weekly) bill (negative reward) of each consumer is computed as follows:

$$\mathbf{R}_{con} = -1 \cdot 7 \cdot (P_{D_ret} \cdot \mathbf{L}_{DW} + P_{N_ret} \cdot (1 - \mathbf{L}_{DW})) \quad (7)$$

Where, \mathbf{R}_{con} is a 1000×1 dimension vector denoting the 1000 consumers reward and $P_{D(N)_ret}$ are the retail prices paid.³ The \mathbf{L}_{DW} is a 1000×1 dimension vector to denoting the 1000 consumers' individual average weekly daily load profile.

Another strategic possibility for a consumer when real time pricing is adopted would be modifying his load profile in response to price signals, shifting part of the consumption from the more expensive to the cheapest time period (i.e. from the day to the night). This elastic behavior is accounted for through a parameter d_{WM} as the maximum percentage of the daily load he/she may be willing to shift in order to save money (since in our model day price is always more expensive than night price).

³ Note that retail price is constant over a week, and that if $P_{D_ret} = P_{N_ret}$ the consumer is under flat rate. We should also note that, although a consumer selects a price option once a week, the reward is calculated everyday, because the load profile of each consumer and the system real time price change everyday. The aim of each agent is to maximize his reward (minimize his bill) during the long run (or one month in our study case).

3.2 Model of the behavior of retailers

In our model, we restricted the analysis to only 2 retailers for sake of simplicity. This assumption is fairly realistic when we consider a market where a local monopoly has been in place for a long time. Retailers have to decide prices (flat or time-of-use) once a week. We consider as the decision variable for the retailer the markup with reference to the wholesale cost.

Thus, if the retailer offers a flat rate, the price is calculated by:

$$P_{F_ret} = P_{FW} \cdot (1 + \Delta_F) \quad (8)$$

where Δ_F can assume a discrete number of values, which is the action set of the retailer.

When a retailer offers time-of-use prices, there are two decision variables, because he needs to fix the mark up for both day and night prices.⁴

$$P_{D_ret} = P_{DW} \cdot (1 + \Delta_D) \quad (9)$$

$$P_{N_ret} = P_{NW} \cdot (1 + \Delta_N) \quad (10)$$

In this case, the possible strategies are $S = \{\Delta_D \times \Delta_N\}$.

To compute the everyday profit of retailer i , we must consider how many consumers (and which consumers) will select his proposed tariff. The reward (R_{ret1}) can be written as follows:

$$R_{ret1} = (P_{D_ret1} - P_D) \cdot \left(\sum_{i \in ret1} L_{iD} \right) + (P_{N_ret1} - P_N) \cdot \left(\sum_{i \in ret1} L_{iN} \right) \quad (11)$$

When the tariff is flat, it means that P_{D_ret1} and P_{N_ret1} are equal, and the reward (R_{ret2}) can be rewritten as follows:

$$R_{ret2} = P_{F_ret2} \cdot \sum_{i \in ret2} (L_{iD} + L_{iN}) - P_D \cdot \left(\sum_{i \in ret2} L_{iD} \right) - P_N \cdot \left(\sum_{i \in ret2} L_{iN} \right) \quad (12)$$

4. RETAIL MARKET ACTIVITIES

In order to process our multi-agent simulation, we first need to define the system real time prices and consumers load profile.

Table 1 describes four different average load profile and prices scenarios,

⁴ In this case one of the make ups may even be negative: a higher (or lower) price difference between the two time periods may attract consumers with certain load profiles.

corresponding ideally to four types of weeks. We assumed to have a higher daily price in weeks corresponding to the highest daily consumption, as typically observed in real power exchange markets.⁵ When we assume perfect competition in the retail market (as described in the introduction section), we are able to compute the weekly flat rate charged to final consumers when *all* consumers select flat prices:

$$P_{FW} = P_{DW} \cdot L_{DW} + P_{NW} \cdot L_{NW} \quad (13)$$

Table 1. Typical weeks for the various scenarios

	System load profile day, night	Average flat rate if all consumers select flat rate
Week 1 average: ($P_{DW} = 12, P_{NW} = 2.4$)	$L_{DW} = 65\%$, $L_{NW} = 35\%$	8.64
Week 2 average: ($P_{DW} = 8, P_{NW} = 3$)	$L_{DW} = 55\%$, $L_{NW} = 45\%$	5.75
Week 3 average: ($P_{DW} = 14, P_{NW} = 2.5$)	$L_{DW} = 69\%$, $L_{NW} = 31\%$	10.435
Week 4 average: ($P_{DW} = 10, P_{NW} = 2$)	$L_{DW} = 60\%$, $L_{NW} = 40\%$	6.8

Figure 2 shows the consumers weekly average load profile under these 4 scenarios, and help to understand consumer heterogeneity. Figure 3 shows the system load profile and corresponding everyday real time prices when random terms (the random load fluctuation percentage e_W / e_D) are set to 0.2 and the maximum willingness (d_{WM}) to shift consumption across time periods is 20%.

⁵ It may be worth to recall that the choice of the week as the relevant time dimension for retail prices update is made for sake of simplicity, and here the idea of having different week with different load distribution should reflect the fact that typically the peak load problem may be more serious in specific periods or seasons.

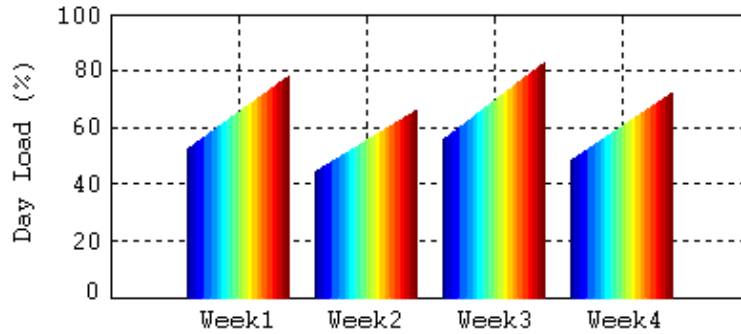


Figure 2. Consumers' average load profile for the four scenarios

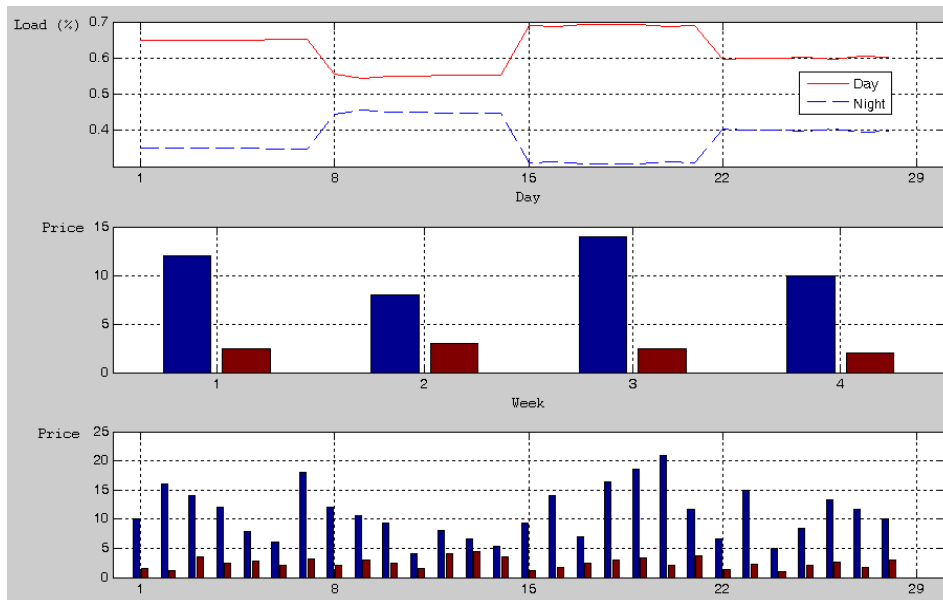


Figure 3. System average load profile and real time prices

We consider here three different situations in order to illustrate the behavior of consumers and retailers in a competitive electricity market. First, only the consumers are modeled as agents and the retail side is assumed to be zero profit. Then, only the retailers are modeled as agents and the consumers always make the “right” choice according to their load profile as an assumption. These two situations, though not completely, are useful as references cases for our last situation, in which, both consumers and retailers are considered simultaneously as agents.

4.1 Consumers are agents, retailers make zero profits (perfect competition)

The consumer's learning procedure can be articulated into three steps.

STEP 1: *State Identification*. In our model, the state of consumers has 2 components, the week type and the observation of the average (flat) price paid during the previous correspondent period.⁶ Actually, we assume the perception about flat price to be limited to binary information, i.e. the fact that it has been higher or lower than P_{FW} .

STEP 2: *Action Selection*. After having identified the current state, each agent looks at what his experience suggests to be the best action in this scenario (flat rate or time-of-use pricing). More technically, he selects the action with maximum Q -value corresponding to current state. To balance the exploration (to explore possible better strategies) and exploitation (to exploit the best strategies found by past experience), ϵ -greedy policy (with ϵ possibility to select randomly instead of selecting the best strategies up to now in order to explore possible better strategies) is used in the proposed action selection. That is, during the action selection process, an agent selects most of the time an action with maximum Q -value in the state s . But, with a small probability ϵ , he chooses randomly, independently from the Q -values. In other words, the agent has a relatively small probability to explore other choices even if they appear less profitable according to past experience. In our model, this probability is equal to 10%.

STEP 3: *Q-table update*. At the end of each week, agents have learnt something more on their optimal choice. Thus, Q -values are updated, using (1)-(2).

In the first situation, we assume the retailers make zero profit, and we focus on investigating consumers' behavior. As it has been stated in Section I, in such a situation the flat rate should disappear, in the absence of risk aversion and in a world without transaction costs. What will happen if we consider consumers as agents in the sense described in previous Sections? Basically, they will make sometimes a "wrong" choice due to the imperfect information of their own load profiles, although the probability to make "wrong" choice is very low under a long time of study because they learn from their past experiences.

Table 2 shows results of different simulations made to see how different parameters in our model may impact on the optimal choices of consumers. In each case it is shown how many consumers (over the total of 1000) select flat or TOU price option.

⁶ Note that the weekly average flat rate can not be obtained until all of the consumers have made their choices.

Table 2. Simulation results of different cases results under the zero profit assumption of retail side

Case 1	Case 2	Case 3	Case 4	Case 5	Case 6	Case 7
No heterogeneity in load profile without random.	Heterogeneity in load profile without random.	Increased heterogeneity in load profile	With random in RTP and load profile	With willingness to delay consumption	The same as case 5 but only 4 states	The same as case 4 but only 4 states
$\delta_{Per} = 1.0$ $\delta_W = .0$ $\delta_D = .0$ $\delta_{PD} = .0$ $\delta_{PN} = .0$ $d_{WM} = .0$	$\delta_{Per} = 1.0$ $\delta_W = .2$ $\delta_D = .0$ $\delta_{PD} = .0$ $\delta_{PN} = .0$ $d_{WM} = .0$	$\delta_{Per} = 1.0$ $\delta_W = .4$ $\delta_D = .0$ $\delta_{PD} = .0$ $\delta_{PN} = .0$ $d_{WM} = .0$	$\delta_{Per} = 1.0$ $\delta_W = .2$ $\delta_D = .2$ $\delta_{PD} = .2$ $\delta_{PN} = .2$ $d_{WM} = .0$	$\delta_{Per} = 1.0$ $\delta_W = .2$ $\delta_D = .2$ $\delta_{PD} = .2$ $\delta_{PN} = .2$ $d_{WM} = .2$	$\delta_{Per} = 1.9$ $\delta_W = .2$ $\delta_D = .2$ $\delta_{PD} = .2$ $\delta_{PN} = .2$ $d_{WM} = .2$	$\delta_{Per} = 1.9$ $\delta_W = .2$ $\delta_D = .2$ $\delta_{PD} = .2$ $\delta_{PN} = .2$ $d_{WM} = .0$
FL RTP	FL RTP	FL RTP	FL RTP	FL RTP	FL RTP	FL RTP
0 0	397 427	447 478	447 434	37 726	38 723	210 387
0 0	189 392	160 626	177 410	0 0	0 0	0 0
0 0	400 430	459 452	426 399	4 633	5 635	95 213
0 0	86 224	111 543	79 246	0 0	0 0	0 0
0 0	447 442	457 463	408 402	45 717	46 720	214 392
0 0	194 408	158 634	185 428	0 0	0 0	0 0
0 0	436 458	451 461	391 408	32 748	29 742	194 371
0 0	171 378	154 619	168 389	0 0	0 0	0 0

Case 1 just tests that when there is no consumer heterogeneity on load profiles, all consumers are indifferent between the two options (in other words, the consumers' probability to select flat or TOU price is the same, 50%). The other cases show the results when consumers have different load profiles under different assumptions concerning uncertainty, willingness to shift consumption across time periods, number of states identified by the consumers. The main insights from this analysis may be summarized as follows:

- Despite TOU selections prevail, a large proportion of consumers would be indifferent and several flat rate selections can be observed (at least 20% when consumption shifts across time periods is ignored); sometimes selections are equally split between the two options (even if this happens only when flat rate is observed to be unexpectedly low). The interesting thing in this result is that some consumers *rationaly* select the flat option, which should instead become sub-optimal for everyone according to theory. The reason is that theory would work if all consumers believe that other consumers are also rational. Instead, in our model at each time there is a little part of consumers being choosing their action randomly, and this affects the Q -values; this means that the multi-agent framework can capture the idea that, even if an agent behaves rationally, he must consider the probability that the others are not doing so.
- Since we do not model risk aversion, random variations around the average does not have substantial effects on the results. Also, the extent to which consumer heterogeneity is more severe does not change the simulation predictions. In the following Sections, we present simulations where the

values of these parameters are the following: $\delta_{PD} = \delta_{PN} = 0.1$ and $\delta_W = \delta_D = 0.2$ ⁷.

- When comparing results obtained with reference to different week (and therefore different load condition), they are quite similar, but it seems that the highest the difference in prices and loads between day and night, the highest is the percentage of flat selections. This trend is certainly not so evident, however it may be explained by the fact that “making an error” selecting TOU prices for consumers with high daily consumption may have an higher impact on Q -values when daily prices are higher.
- As expected, when consumers are able to shift consumption from the more expensive to the cheapest period, TOU rates become more appealing and flat rate selections decrease below the 5%.
- Finally, we consider a reduction of the number of states identified by the consumers, supposing he does not have (or does not care about) the information about the previous flat rate. It seems that if we reduce 2 states to 1, the lowest percentage will appear in flat rate selection and the highest in TOU). Thus, we may conclude that more detailed state definition does not help in leading to the theoretical best practice.

4.2 Retailers are agents, while consumers are assumed to just make the best choice at each time.

Similarly, we can define the learning steps of retailers as follows:

STEP 1: *State Identification*. In the case of retailers, the state is defined as the strategy chosen by the competitor in the previous time period. Note that in this second part of simulations we restrict the analysis to 1 typical week only.

STEP 2: *Action Selection*. Retailers have to choose the optimal mark-up on wholesale cost of electricity. Since the states are just the previous actions of the competitor, the optimal action has a nice interpretation in terms of “reaction” to his competitor’s behaviour. Considering the different state and the corresponding optimal action, this situation reflects the concept of reaction function typically used in oligopoly models.

STEP 3: *Q-table update*.

In this situation we analyze the results when there is not an assumption of perfect competition among retailers, but we endogenize the behavior of retailers in the model. Therefore we want to investigate the possibility of making profits by retailers whose activities are simply selling electricity to the consumers after purchasing it from the power exchange.

⁷ Additional simulations were run considering different values for these parameters, without significant effects on the results.

Case 1: We have two retailers that compete in the retail market (Bertrand model: they compete on prices). The action of retailer 1 (and 2) is choosing the optimal markup with reference to the wholesale price, as defined in formula (8). If Δ_{F1} (Δ_{F2}) = 0, it means that the retailer is selling at a price which is approximately equal to the average cost of purchasing in the Power Exchange.⁸

Table 3. Q -tables of retailers and No. of consumers selection in equilibrium under case 1 when only retailers are agents

<i>Q-table of retailer 1</i>						<i>Q-table of retailer 2</i>					
<i>States</i>	Δ_{F1}					<i>States</i>	Δ_{F2}				
	-0.2	0	0.1	0.3	0.5		-0.2	0	0.1	0.3	0.5
-0.2	-10102	395.37	3596.4	1694.4	660.24	-0.2	-9016	256.14	3278.6	1787.7	824.19
0	-8692	241.87	3313.5	1226.2	293.67	0	-8710	343.51	3189.5	1059.2	324.49
0.1	-11751	344.94	3367.4	453.1	338.85	0.1	-11759	336.95	3354.4	345.66	331.64
0.3	-11474	550.1	3170.6	2390.6	670.39	0.3	-10934	548.05	3305.9	2443.9	735.73
0.5	-8312	493.37	3485.6	1236.8	1301.8	0.5	-10240	359.32	3275	1177.6	896.71
<i>No. of consumers selecting retailer 1 in equilibrium:</i> 479						<i>No. of consumers selecting retailer 2 in equilibrium:</i> 521					

Theory predicts that in Bertrand model both retailers make zero profits. Actually, in our model we restrict the choice of Δ_{F1} (Δ_{F2}) to a discrete action set, so the best strategies of both retailers should be the lowest possible delta above 0. The simulation confirms the expected result, and for both retailers the action [Δ_{F1} (Δ_{F2}) = 0.1] results to be the best action in every state. This has a nice interpretation in terms of game theory, because the action [Δ_{F1} (Δ_{F2}) = 0.1] can be seen as a dominant strategy for both retailers; the couple [Δ_{F1} = 0.1; Δ_{F2} = 0.1] is therefore a Nash Equilibrium in dominant strategies. That bridges an interesting parallel between game theory equilibrium and multi-agents equilibrium (convergence).

Case 2: Now, we consider different pricing options for the two retailers. In particular, let us assume that retailer 1 offers a TOU rate, with different night and day prices, while retailer 2 offers flat rate. The action of retailer 1 is now choosing the degree of price differentiation between day and night. In this case, we impose the following restriction: $\Delta_{N1} = -2 * \Delta_{D1}$ in order to reduce the number of states and actions and therefore the computational requirements. We still get a Nash equilibrium in dominant strategies [Δ_{D1} = 0.2; Δ_{F2} = 0.1]. In this case, the heterogeneity in consumer load profiles becomes relevant. The first retailer, choosing his optimal combination between day and night prices, can get more profits than the second retailer (who offers flat rate), though he is actually serving

⁸ We investigate both the case when the pattern of discrete possible action is symmetric or asymmetric among retailers (or firms), and we got the same results.

a lower number of consumers (393 vs. 607). Comparing case 2 with case 1 (comparing the Q -values in these 2 cases), we see that however retailer 1 makes more or less the same profits (similar Q -values) as before, while retailer 2 makes less profits. Summarizing, it seems that competing on different pricing schemes even reduce the profits for retailer instead of producing benefits.

Table 4. Q -tables of retailers and No. of consumers selection in equilibrium under case 2 when only retailers are agents

<i>Q-table of retailer 1</i>						<i>Q-table of retailer 2</i>					
<i>States</i>	Δ_{D1}					<i>States</i>	Δ_{F2}				
	<i>-0.1</i>	<i>0</i>	<i>0.1</i>	<i>0.2</i>	<i>0.4</i>		<i>-0.2</i>	<i>0</i>	<i>0.1</i>	<i>0.3</i>	<i>0.5</i>
<i>-0.2</i>	-4079	248.13	3353	1882.4	299.33	<i>-0.1</i>	-8793	197.53	3379.8	671.41	242.69
<i>0</i>	-4792	223.12	3636.6	1443.8	1314.9	<i>0</i>	-10194	216.4	3048.4	524.92	256.63
<i>0.1</i>	-5689	336.69	3408.2	349.49	532.93	<i>0.1</i>	-11760	332.18	3560.4	465.74	338.52
<i>0.3</i>	-3806	320.13	3257.3	1227.5	373.79	<i>0.2</i>	-11055	350.93	3329.2	830.3	346.65
<i>0.5</i>	-4804	282.66	3119.8	1125.1	2987.5	<i>0.4</i>	-10153	273.97	3478.7	1596.4	292.67
<i>No. of consumers selecting retailer 1 in equilibrium:</i> 393						<i>No. of consumers selecting retailer 2 in equilibrium:</i> 607					

Case 3: As in case 3, we consider different pricing options offered by the two retailers. However, now retailer 1 can choose separately 2 different deltas, one for the day and one for the night price. His action is thus the combination of two $\Delta_{D1}(\Delta_{N1})$, resulting in many more possibilities.

Again, we get a Nash Equilibrium in dominant strategies [$\Delta_{D1}=0$, $\Delta_{N1}=1.2$; $\Delta_{F2}=0.1$]. Profit for retailer 2 is very similar to case 2, while retailer 1 can make a little higher profit (Q -value=3760 vs. 3000). This is due to the higher flexibility in choosing prices; however, this does not change the previous insight that competition on different price schedules does not seem to help to make profits. Again, retailer 1 makes more profits though retailer 2 has more consumers.

It is interesting to note that the profits for retailer 1 comes from the high delta on night prices ($\Delta_{N1}=1.2$ means that night prices are, in average, 2.2 times the cost of purchasing electricity from the EPX during the night). Retailer 1 can get these profits and still capture the consumers which have a more night oriented (and thus cheaper) load profile. Retailer 2 will instead serve the more daily oriented (and more expensive) consumers.

Table 5. Q -tables of retailers and No. of consumers selection in equilibrium under case 3 when only retailers are agents

<i>Q-table of retailer 1</i>								
<i>States</i>	Δ_{D1}/Δ_{N1}							
	<i>-0.2</i>	<i>-0.2</i>	<i>-0.2</i>	<i>-0.2</i>	<i>0.0</i>	<i>0.0</i>	<i>0.0</i>	<i>0.0</i>
	<i>-0.6</i>	<i>0.0</i>	<i>0.6</i>	<i>1.2</i>	<i>-0.6</i>	<i>0.0</i>	<i>0.6</i>	<i>1.2</i>
<i>-0.2</i>	-7717	-8730	-4396	-2066	-1216	393	2084	3983
<i>0.0</i>	-8600	-8583	-3359	-1890	-1180	419	1774	3668
<i>0.1</i>	-13661	-10407	-6976	-3455	-2871	409	3067	3761
<i>0.3</i>	-13741	-10313	-6831	-3266	-2881	565	2938	4047
<i>0.5</i>	-8494	-6816	-3517	-1725	-1502	373	835	3996
	<i>0.2</i>	<i>0.2</i>	<i>0.2</i>	<i>0.2</i>	<i>0.4</i>	<i>0.4</i>	<i>0.4</i>	<i>0.4</i>
	<i>-0.6</i>	<i>0.0</i>	<i>0.6</i>	<i>1.2</i>	<i>-0.6</i>	<i>0.0</i>	<i>0.6</i>	<i>1.2</i>
<i>-0.2</i>	3490	3344	2925	2617	2583	804	1960	1113
<i>0.0</i>	1298	3287	3493	3055	2838	2363	2483	446
<i>0.1</i>	2867	3016	2254	454	1980	541	496	444
<i>0.3</i>	2817	2886	2822	2451	3064	2791	3039	643
<i>0.5</i>	3260	3300	3186	2675	2952	1696	2511	451

<i>Q-table of retailer 2</i>							
<i>States</i>		Δ_{F2}					
		<i>-0.2</i>	<i>0.0</i>	<i>0.1</i>	<i>0.3</i>	<i>0.5</i>	
<i>-0.2</i>	<i>-0.6</i>	-9921	300	1877	1276	639	
<i>-0.2</i>	<i>0.0</i>	-7178	164	2424	1376	499	
<i>-0.2</i>	<i>0.6</i>	-7497	5	1700	1315	356	
<i>-0.2</i>	<i>1.2</i>	-5508	191	1722	1090	236	
<i>0.0</i>	<i>-0.6</i>	-7378	147	1493	1426	587	
<i>0.0</i>	<i>0.0</i>	-8550	278	1799	1501	1335	
<i>0.0</i>	<i>0.6</i>	-7686	10	1806	1162	216	
<i>0.0</i>	<i>1.2</i>	-11792	161	1564	342	167	
<i>0.2</i>	<i>-0.6</i>	-10994	78	2027	1310	352	
<i>0.2</i>	<i>0.0</i>	-11713	281	1947	1546	389	
<i>0.2</i>	<i>0.6</i>	-11631	337	1921	1515	815	
<i>0.2</i>	<i>1.2</i>	-11702	351	1656	1269	564	
<i>0.4</i>	<i>-0.6</i>	-10918	-191	2222	2122	853	
<i>0.4</i>	<i>0.0</i>	-11263	104	1663	1463	1489	
<i>0.4</i>	<i>0.6</i>	-9817	199	1890	1599	550	
<i>0.4</i>	<i>1.2</i>	-5758	141	1902	1209	661	

<i>No. of consumers selecting retailer 1 in equilibrium:</i>	<i>418</i>
<i>No. of consumers selecting retailer 2 in equilibrium:</i>	<i>582</i>

4.3 Both consumers and retailers are agents.

This is the natural union between situation A and B, with the difference that states for consumers are now represented by the combination of retailers strategies, and the action of consumers is choosing among the two retailers.

The analysis in this situation repeats the one seen above with the only difference that now also consumers are included in the multi-agent framework. The main goal is to see if in this context retailers can have the possibility to make more profits.

Case 1: Both retailers offer flat rate, and the price schemes are the same.

When consumers are agents, we get a fundamentally different result in the choice of retailers and consequently in their profits. In fact, the Nash Equilibrium now is $[\Delta_{F1}= 0.3; \Delta_{F2}= 0.3]$, and retailers can make higher profits than Case 1 when consumers are not agents. This can be interpreted in the sense that retailers can exploit the difficulty by the consumers in learning their best actions. When both retailers are offering $\Delta_{F1}(\Delta_{F2}) = 0.3$, no one will have incentive to lower the price, because even if a lower price would be more convenient for all consumers, only part of them will be aware of that, or however they will need time to learn it.

Table 6. *Q*-tables of retailers and No. of consumers selection in equilibrium under case 1 when both retailers and consumers are agents

<i>Q</i> -table of retailer 1						<i>Q</i> -table of retailer 2					
States	Δ_{F1}					States	Δ_{F2}				
	-0.2	0	0.1	0.3	0.5		-0.2	0	0.1	0.3	0.5
-0.2	-	-	-	-	-	-0.2	-	-	-	-	-
	7971.4	985.36	5193.3	9662.9	6531.5		9115.6	978.15	5397.6	10077	6052.1
0	-	-	-	-	-	0	-	-	-	-	-
	7302.7	1033.8	5748.8	9163.1	5279.5		-9100	1066	5971	9645.1	4648.9
0.1	-	-	-	-	-	0.1	-	-	-	-	-
	8187.7	1098.5	5734.7	10127	4914.8		9066.1	991.18	5650.6	10763	4675.9
0.3	-	-	-	-	-	0.3	-	-	-	-	-
	9825.8	1009.1	6378.5	10188	4577.7		9863.4	997.53	6156	9631.5	4329.3
0.5	-	-	-	-	-	0.5	-	-	-	-	-
	9112.6	1189	6304.1	9330.9	5559		9353.8	1100.2	5811.7	9605.3	5143.6
No. of consumers selecting retailer 1 in equilibrium: 507						No. of consumers selecting retailer 2 in equilibrium: 493					

Case 2: Retailer 1 offers time of use (TOU) rates (with $\Delta_{NI} = 2*\Delta_{DI}$) and retailer 2 offers flat rate.

We get a Nash Equilibrium at $[\Delta_{DI}= 0.4; \Delta_{F2}= 0.3]$, with substantial profits for both retailers, but lower than in case 1. As in the case without consumers as agents, here retailer 1 (choosing TOU pricing) can get higher profits than Retailer

2 (flat rate); in this case the number of consumer is more or less equally split between the retailers.

Table 7. Q -tables of retailers and No. of consumers selection in equilibrium under case 2 when both retailers and consumers are agents

<i>Q-table of retailer 1</i>						<i>Q-table of retailer 2</i>							
<i>States</i>	Δ_{D1}					<i>States</i>	Δ_{F2}						
	<i>-0.1</i>	<i>0</i>	<i>0.1</i>	<i>0.2</i>	<i>0.4</i>		<i>-0.2</i>	<i>0</i>	<i>0.1</i>	<i>0.3</i>	<i>0.5</i>		
<i>-0.2</i>	-6865.9	820.14	4628.3	7766.9	7617.3	<i>-0.1</i>	-10133	555.46	3613.6	7270.6	4531.4		
<i>0</i>	-6890.3	830.27	4508.7	7271.7	8441.4	<i>0</i>	-10163	437.14	4071.2	7471.9	4657.6		
<i>0.1</i>	-6780.6	844.07	4640	7359.9	8938.4	<i>0.1</i>	-10218	508.33	4226.3	7256.1	4520.3		
<i>0.3</i>	-6847.4	860.03	4497.2	7398.1	8694.2	<i>0.2</i>	-10358	87.381	3652.2	4115.9	4618.8		
<i>0.5</i>	-6513.3	1006.2	4442.4	7509	9007.3	<i>0.4</i>	-10202	357.89	4434.1	7071	4812.9		
<i>No. of consumers selecting retailer 1 in equilibrium:</i>						533	<i>No. of consumers selecting retailer 2 in equilibrium:</i>						466

Note that, with respect to the case without consumers as agents, we have retailer 1 and retailer 2 choosing a higher delta (maximum price differentiation between day and night for retailer 1; $\Delta_{D1} = 0.3$ with respect to 0.1 for retailer 2). This allows gaining more profits: again, consumers need time to learn their best actions and thus give more opportunities of profits to the retailers.

Case 3: Retailer 1 offers time of use (TOU) rate (with 2 deltas), and retailer 2 offers flat rate.

With respect to case 2, the higher flexibility in the pricing choice of Retailer 1 allows him to increase his profit. Also Retailer 2 appears to benefit from the change in strategies by Retailer 1. Note that however, comparing the profits with the Case 1, they appear to be lower: this confirms the impressions that retailers can not gain from competing on different price schedules.

Note that in this case, even if we still have a Nash Equilibrium [$\Delta_{D1} = 0.2$, $\Delta_{N1} = 1.2$; $\Delta_{F2} = 0.3$], now the equilibrium is not in dominant strategies. In particular, the best strategy of Retailer 1 depends on the State (and therefore, it depends on the action chosen by retailer 2). This means that the results we get produce a Nash Equilibrium, but the concept is "weaker" than the case of Nash Equilibrium in dominant strategies, because the equilibrium will be achieved conditional to the hypothesis of rational expectation (Retailer 1 must believe that Retailer 2 will act in a rational way, choosing $\Delta_{F2} = 0.3$).

Table 8. Q -tables of retailers and No. of consumers selection in equilibrium under case 3 when both retailers and consumers are agents

<i>Q-table of retailer 1</i>									
<i>States</i>		Δ_{D1}/Δ_{N1}							
		<i>-0.2</i>	<i>-0.2</i>	<i>-0.2</i>	<i>-0.2</i>	<i>0.0</i>	<i>0.0</i>	<i>0.0</i>	<i>0.0</i>
		<i>-0.6</i>	<i>0.0</i>	<i>0.6</i>	<i>1.2</i>	<i>-0.6</i>	<i>0.0</i>	<i>0.6</i>	<i>1.2</i>
<i>-0.2</i>		-7767	-5159	-2944	-1330	-1124	755	1604	4011
<i>0.0</i>		-5475	-4925	-2170	-1230	-1412	879	2124	3692
<i>0.1</i>		-7761	-5035	-3337	-1169	-1343	958	3002	5103
<i>0.3</i>		-11580	-8607	-5238	-2395	-2162	923	4063	7073
<i>0.5</i>		-9126	-7156	-4319	-1665	-1561	1218	3560	6662
		<i>0.2</i>	<i>0.2</i>	<i>0.2</i>	<i>0.2</i>	<i>0.4</i>	<i>0.4</i>	<i>0.4</i>	<i>0.4</i>
		<i>-0.6</i>	<i>0.0</i>	<i>0.6</i>	<i>1.2</i>	<i>-0.6</i>	<i>0.0</i>	<i>0.6</i>	<i>1.2</i>
<i>-0.2</i>		5584	5062	4972	7167	7161	8888	7083	6474
<i>0.0</i>		4188	5977	6742	10553	6641	6546	7675	5603
<i>0.1</i>		5645	5826	7061	6351	8460	6901	6252	6039
<i>0.3</i>		6504	7799	7612	10027	7035	7559	7139	3943
<i>0.5</i>		6553	8016	8208	9407	8045	8179	8061	7134
<i>Q-table of retailer 2</i>									
<i>States</i>		Δ_{F2}							
		<i>-0.2</i>	<i>0.0</i>	<i>0.1</i>	<i>0.3</i>	<i>0.5</i>			
<i>-0.2</i>	<i>-0.6</i>	-6680	613	3744	7438	5633			
<i>-0.2</i>	<i>0.0</i>	-5936	433	3558	7042	5519			
<i>-0.2</i>	<i>0.6</i>	-5267	667	2870	7571	6376			
<i>-0.2</i>	<i>1.2</i>	-5108	585	3182	7483	4754			
<i>0.0</i>	<i>-0.6</i>	-5615	728	3843	7061	3867			
<i>0.0</i>	<i>0.0</i>	-5615	534	3858	7148	4134			
<i>0.0</i>	<i>0.6</i>	-5419	654	3272	7295	4950			
<i>0.0</i>	<i>1.2</i>	-5943	745	3822	7497	5171			
<i>0.2</i>	<i>-0.6</i>	-6155	901	4381	6874	5688			
<i>0.2</i>	<i>0.0</i>	-8742	423	5610	4223	3914			
<i>0.2</i>	<i>0.6</i>	-10252	617	4964	4735	3902			
<i>0.2</i>	<i>1.2</i>	-9706	725	6078	7939	4255			
<i>0.4</i>	<i>-0.6</i>	-9988	624	5050	7904	4523			
<i>0.4</i>	<i>0.0</i>	-9582	678	5250	8320	5276			
<i>0.4</i>	<i>0.6</i>	-8627	713	4889	7645	6000			
<i>0.4</i>	<i>1.2</i>	-7915	1083	5296	7111	5363			
<i>No. of consumers selecting retailer 1 in equilibrium:</i>		<i>508</i>							
<i>No. of consumers selecting retailer 2 in equilibrium:</i>		<i>488</i>							

5. CONCLUSIONS

This paper proposes a multi-agent framework to model a liberalized retail electricity market. We run simulations on both consumers and retailers behavior to study the impact of having consumers with heterogeneous load profiles on their choices among alternative pricing structures and on retailers' strategies. The results highlight some possible departures from the theoretical predictions.

When we assume perfect competition among retailers and consumers only learn their best pricing option from the experience on previous bills, flat price will still be a rational choice for consumers with more peaky load profiles. In theory, flat rate should disappear in a world without transaction costs and without risk aversion; however, this would not be the case in our multi-agent system. This is because our framework allows to capture the idea that, even if an agent behaves rationally, he must consider the probability that the others are not doing so.

When it comes to the retailer behavior, the multi-agent system converges to Q -values which can be interpreted as a game theoretical structure reflecting the basic Bertrand model assumptions. When customers behavior is ignored, the Nash equilibrium is consistent with the Bertrand predictions. Furthermore, the differentiation about pricing schemes among retailers does not lead to an increase in profits; stated better, competitive screening on a load profile basis does not help in making more profits. However, when also consumer behavior is included in our multi-agent system, the results highlight a departure from the Bertrand prediction of zero profits. In a world where consumers can make short term errors (and can learn from them), retailers are able to charge higher prices and make substantial profits. It is interesting to note that, even in our simplified model with assumptions coherent with the perfect competition model, the retail market does not lead to the perfect competition outcome because of the unique hypothesis that consumers may not be able to select the best retailer at any time, even if he is actually rational.

Overall, the simulations in our paper suggest that in a liberalized retail market retailers may be able to charge substantial markups to consumers. The retailers may not have the interest in introducing time-varying tariff structure, unless other competitors do. Thus, some form of regulation may be needed to enhance competition and the possibility of choice for consumers, especially at the first stage of the liberalization.

The issue of retail competition and consumer heterogeneity remains a promising area for further research, which should include for example the joint modeling of different customer dimensions and different market segments (such as residential, industrial and commercial users). This can spread light also on the relevance of consumers' aggregation policies.

6. APPENDIX: NOTATIONS OF FORMULAS

δ_W : Weekly load random factor. When δ_W is 0, it means that the weekly load profile is the same every week.

e_W : Weekly load random term. e_W obeys a uniform distribution $U(-\delta_W \cdot L_{DW}, \delta_W \cdot L_{DW})$.

δ_D : Daily load random factor. When δ_D is 0, it means the everyday load profile is the same during one week.

e_D : Daily load random term. e_D obeys a uniform distribution $U(-\delta_D \cdot L_{iDW}, \delta_D \cdot L_{iDW})$.

δ_{PD} : Day price random factor. When is δ_{PD} 0, it means that the day price is strictly proportional to daily load and no random factor is introduced.

e_{PD} : Day price random term. e_{PD} obeys a uniform distribution $U(-\delta_{PD} \cdot P_{DW}, \delta_{PD} \cdot P_{DW})$.

δ_{PN} : Night price random factor. When is δ_{PN} 0, it means that the night price is strictly proportional to night load and no random factor is introduced.

e_{PN} : Night price random term. e_{PN} obeys a uniform distribution $U(-\delta_{PN} \cdot P_{NW}, \delta_{PN} \cdot P_{NW})$.

δ_{Per} : This parameter is for state identification. It means the times of average flat rate. When δ_{Per} is 1.0, it means that the consumer state is divided into 1) his flat rate is lower than the average system flat rate, and 2) his flat rate is equal or higher than the average system flat rate. When δ_{Per} is 1.9, it means that the consumer state is divided into 1) his flat rate is lower than 1.9 times of the average system flat rate, and 2) his flat rate is equal or higher than 1.9 times of the average system flat rate.

d_{WM} : The maximum percentage of the daily load the customer may be willing to shift in order to save money.

L_D : System everyday load during the day.

L_N : System everyday load during the night.

L_{iD} : Individual load of consumer i^{th} during the day.

L_{iN} : Individual load of consumer i^{th} during the night.

L_{DW} : A 1000·1 vector denoting the 1000 consumers' individual average weekly daily load profile.

P_D : System everyday price during the day.

P_N : System everyday price during the night.
 P_F : Every day weighted average between day and night price (every day flat rate when all consumers are under flat rate and there is perfect competition in the retail market).
 P_{F-ret1} : Flat price offered by retailer 1.
 P_{F-ret2} : Flat price offered by retailer 2.
 P_{D-ret1} : Day price offered by retailer 1 (TOU pricing).
 P_{N-ret1} : Night price offered by retailer 1 (TOU pricing).
 \mathbf{R}_{con} : A 1000·1 vector denoting the 1000 consumers' reward.
 R_{ret1} : The reward of retailer1.
 R_{ret2} : The reward of retailer2.
 Δ_{F1} : % make up of retailer 1 over the average weekly cost of buying electricity from the EPX.
 Δ_{F2} : % make up of retailer 2 over the average weekly cost of buying electricity from the EPX.
 Δ_{D1} : % make up of retailer 1 over the average weekly cost of buying electricity from the EPX during the day.
 Δ_{N1} : % make up of retailer 1 over the average weekly cost of buying electricity from the EPX during the night.
 $*_W$: Weekly average of everyday values taken by *.
 For example:
 P_{DW} : System weekly average price during the day.
 P_{NW} : System weekly average price during the night
 P_{FW} : System weekly average flat rate.

7. REFERENCE

- [1]. Abrate G., Pricing Electricity in a Liberalized Market: a Demand Side Approach, PhD Dissertation, University of Pavia, 2005.
- [2]. Borenstein S., "The Long-Run Effects of Real-Time Electricity Pricing", 2005, Energy Journal 26(3).
- [3]. Holland S.P. and E.T. Mansur, "The Distributional and Environmental Effects of Time-Varying Prices in Competitive Electricity Markets" (May 17, 2005). Center for the Study of Energy Markets. Paper CSEMWP-143. <http://repositories.cdlib.org/ucei/csem/>.
- [4]. Sutton R.S. and A. G. Barto, Reinforcement Learning: An Introduction, MIT Press, Cambridge, MA, 1998.
- [5]. Xiong G., O. Shigeru and F. Hideki, "Multi-agent Based Experiments on Uniform Price and Pas-as Bid Electricity Auction Markets", Electric Utility Deregulation, Restructuring and Power Technologies, 2004. (DRPT 2004). Proceedings of the 2004 IEEE International Conference on Volume. 1, 5-8 April 2004 Page(s) : 72 - 76 Vol.1.
- [6]. Faruqui A. and S. George, "Quantifying Customer Response to Dynamic Pricing", 2005, Electricity Journal 18(4), 53-63.
- [7]. Shy O., "Dynamic Peak-Load Pricing", available at: <http://econ.haifa.ac.il/~ozshy/ozshy.html>.
- [8]. Severin Borenstein and Stephen P. Holland, "On the Efficiency of Competitive Electricity Markets With Time-Invariant Retail Prices", Center for the Study of Energy Markets (CSEM), CSEM WP 116.
- [9]. Brennan T.J., "Demand-Side Management Programs Under Retail Electricity Competition", Discussion Paper 99-02, Resources for the Future, website: <http://www.rff.org>
- [10]. Blumsack S.A., J. Apt and L.B. Lave, "Comments on Wholesale and Retail Electricity Competition", Electric Energy Market Competition Interagency Task Force and the Federal Energy Regulatory Commission, Docket No. AD05-17-000.
- [11]. Joskow, P. and J. Tirole (2004) "Retail Electricity Competition," mimeo, MIT and IDEI.
- [12]. Hunt S. (2002), Making Competition Work in Electricity, Wiley Publishers.
- [13]. Watanabe I., K. Okada; K. Tokoro and S. Matsui, "Adaptive multiagent model of electric power market with congestion management", Evolutionary Computation, 2002. CEC '02. Proceedings of the 2002 Congress on Volume 1, 12-17 May 2002 Page(s):523 - 528 Digital Object Identifier 10.1109/CEC.2002.1006289

- [14]. Bagnall, A.J. and G.D. Smith, "A multiagent model of the UK market in electricity generation, Evolutionary Computation", IEEE Transactions on Volume 9, Issue 5, Oct. 2005 Page(s):522 - 536, Digital Object Identifier 10.1109/TEVC.2005.850264
- [15]. Bunn, D.W. and F.S. Oliveira, "Agent-based simulation-an application to the new electricity trading arrangements of England and Wales", Evolutionary Computation, IEEE Transactions on Volume 5, Issue 5, Oct. 2001 Page(s):493 - 503 , Digital Object Identifier 10.1109/4235.956713
- [16]. Fujii, Y., T. Okamura, K. Inagaki and K. Yamaji, "Basic analysis of the pricing processes in modeled electricity markets with multi-agent simulation", Electric Utility Deregulation, Restructuring and Power Technologies, 2004. (DRPT 2004). Proceedings of the 2004 IEEE International Conference on Volume 1, 5-8 April 2004 Page(s):257 - 261 Vol.1
- [17]. Praca, I., C. Ramos, Z. Vale and M. Cordeiro, "MASCEM: a multiagent system that simulates competitive electricity markets", Intelligent Systems, IEEE [see also IEEE Intelligent Systems and Their Applications], Volume 18, Issue 6, Nov-Dec 2003 Page(s):54 - 60, Digital Object Identifier 10.1109/MIS.2003.1249170