Studies on Horizontal Competition among Homogenous Retailers Based on Agent-based Approach

Ming Xie, J. Chen

Research Center for Contemporary Management, Tsinghua University
School of Economics and Management, Tsinghua University
Beijing 100084, China
{xiem1, chenj}@em.tsinghua.edu.cn

ABSTRACT

This paper adopts agent-based simulation to study the horizontal competition among homogenous price-setting retailers in a one-to-many supply chain (one supplier and multiple retailers). We model the supplier and retailers as agents, and design their behavioral rules respectively. The results show that although the agents learn individually based on their own experiences, the system shows an asymptotic convergence, which approaches Nash equilibrium. Based on these results, we discuss the effects of the retailers' horizontal competition on the retail price, retailers' and supplier's profits.

Keywords: supply chain, agent-based simulation, horizontal competition, Nash equilibrium

1. INTRODUCTION

The concept of supply chain was introduced several years ago and has been widely studied in academic researches and industrial applications. A supply chain is a network composed of suppliers, manufacturers, retailers, etc., which cooperate to offer a kind of goods or service. The partners in a supply chain cooperate to gain extra profits and also compete for more profits. These cooperations and competitions, along with the various uncertainties, make the research on supply chain a quite complex and difficult task. [1] shows that in the academic study of supply chain, only some relatively simple issues are discussed, and many assumptions and constraints are imposed to these models.

Besides analytical methods, simulation is also an important research methodology. Compared with analytical methods, it is more suitable to study complex issues. Simulation has been widely applied in various engineering fields, and recently some social scientists have tried it in social science research and have got valuable results. The most prominent research in this direction is experimental economics, e.g., [2], [3] and [4].

There are also researches that use simulation to study management issues. For example, discrete event simulation [5] has a relatively long history in operations research. But such simulations are focused on the sequence and/or causal relationships of events, which is not suited to describe the competition and cooperation relationships between the partners in a supply chain. On the contrary, agent-based modeling adopts a bottom-up approach and focuses on the design of agents' individual behaviors, and watches what systemic macro behaviors evolve from individuals' micro dynamics [6]. This is quite suitable for modeling complex systems, such as supply chain systems [7].

This paper adopts agent-based simulation to study horizontal competitions in supply chains. Horizontal competition is a hot spot and also a nodus in supply chain research. It is generally analyzed with non-cooperative game theory. But just as we have mentioned before, the application of such an analytical approach is limited into relatively simple models. [8] studies a supply chain model with multiple retailers, each of whom faces a newsboy decision problem. Based on this model, [9] adopts discrete choice model to describe the dynamic process of arrival and choice behavior of each customer. As in [8], the retail prices are fixed, and the retailers only determine order quantity. We know that in many cases retail price affects customer demand. So the model in [9] should be extended to a more sophisticated model with multiple competing price-setting retailers, i.e., retailers can adjust their retail prices to influence customer demands. But this model is almost intractable with analytical methods. In this model, each retailer's profit is determined not only by his own retail price and order quantity, but also on other retailers'. The competition becomes more intense, and things become more interesting: how will these price-setting retailers make their decisions? Will the retailers find appropriate retail price and order quantity? How will the retailers' profits and the supplier's profits be? In this paper, we study such a model and try to answer these questions through agent-based simulation. We model each retailer and the supplier into an agent and design learning rules for it. The simulation results show that although the agents learn individually based on their own experiences, each of them finally gets a relatively stable policy, and the whole system asymptotically converges to an absorbing state which approaches the Nash equilibrium. Based on these results, we also discuss how the retail price and supplier/retailer profits are affected by the competitions between the retailers.

The rest of this paper is organized as follows: Section 2 describes the model settings; Section 3 shows the simulation results and provides detailed analysis; and Section 4 concludes this paper.

2. MODEL SETTINGS

The basic structure of the model is shown in Figure 1.

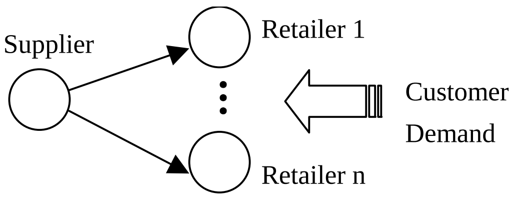

Figure 1 Model Structure

This model runs for many periods. In each period, retailers buy products from the supplier, and then sold them to customers. The supplier charges a fixed wholesale price of 1, and has infinite capacity. Every period, each retailer decides his retail price and order quantity. The products are perishable; all the unsold products are disposed with salvage value 0, and all the unsatisfied customer demands are lost. From analytical view, in each period each retailer faces a price-setting newsboy model. The only cost of each retailer is purchasing cost, occurs when receiving products from the supplier. Each retailer obtains revenue by selling products to customers. The sequence of events of this model is as follows: on each period, each retailer places an order to the supplier simultaneously; then the supplier makes enough products and sends products to each retailer with the quantity that the retailer has ordered; then customers arrive, choose retailers and buy products.

2.1 Customer Behavior

In this model, we adopt the discrete choice model [10] to represent customers' arrival and retailer-selecting behaviors. We assume that in each period, there is a population of customers, whose number is a Poisson random variable with mean 1000. The arrival of the customers is a Poisson process, and the service time of each customer is zero. Each customer desires only one product. When each customer arrives, he chooses a retailer from the series of retailers $N = \{1,2,...,N\}$ to purchase one product. Each customer associates a utility U_j with each retailer $j \in N$. In additions, there is a no selection option, denoted $j = 0$, with associated utility U_0, which means the utility of not purchase at all. A customer chooses the retailer with the highest utility among the series of retailers and buys a product from him. For example, if $MaxU_j = U_2$, the customer chooses the second retailer. But if $MaxU_j = U_0$, the customer returns home without buying anything. If a customer is not satisfied, i.e., if the retailer is out of stock, the customer makes a choice again.

The utility U_j is decomposed into two parts: one part, denoted u_j, represents the nominal (expected) utility; the other part, denoted ξ_j, is a zero-mean random variable representing the difference between U_j and u_j . Thus, $U_j = u_j + \xi_j$. Nominal utility u_j is determined by some identified factors. In this paper, we define $u_j = \beta a_j - \gamma p_j$, $u_0 = \beta a_0$, where a_j is a quality index and p_j is the price for product j. β and γ represent the sensitivities of the customer to the factors. The noise part ξ_j represents the effects of those unidentified factors, and is often modeled as a Gumbel random variable with distribution

$$P(\xi_j \le x) = \exp(-\exp(-(x/\mu) - \tau))$$

with mean zero and variance $\mu^2\pi^2/6$ (τ is the Euler's constant). Here, it is obvious that the retailer with higher retail price will have less utility for a customer, and so will be less likely to be chosen.

Different with classical utility maximization, in this model we assume U_j to be unobservable a priori, so that an individual's choice is uncertain. When a customer arrives, he makes a choice based only on knowledge of the public information of each retailer's price. We also assume all the customers are homogenous, and $\beta = \gamma = 1$, $a_j = 7.06$, $a_0 = 4$, $\mu = 0.847$ (the data comes from the model in [9]).

2.1 Retailer Behavior

We model each of the retailers as an adaptive agent. Here, we first assume that each retailer has an aspiration for more profits. No matter what policy, if it can bring more profits, the retailer will adopt it, without considering its social effects.

We assume that the retailers have minimal information on the environment: they do not know each other agent's policies, payoffs and costs. All that one retailer knows is its own experiences. Also a retailer can observe other agents' historical prices, yet we assume he does not know to utilize this information under such a complex environment. Each retailer continuously learns from his own experiences, and calibrates his policies to obtain as more profits as possible. The basic idea of each agent's learning mechanism is that: each agent has an infinite set of actions. At each stage, he chooses the action that it assesses to have the highest payoff. This idea of learning has a standard implementation – the Q-learning in Machine Learning [11]. In our model, each retailer has two decision variables: retail price and order quantity. Although these two variables are generally all continuous, we assume only discrete values can be chosen. We define the minimal difference between two possible prices to be 0.1, i.e., the prices should be 2.3, 3.4, etc. We also limit retail price into the range (1, 6) (It can be tested that when price $\geq$ 6, customers will seldom buy a product; if the price $\leq$ 1, the retailer will get non-positive profit, which is not rational.). Thus each agent has 50 possible retail prices.

Similarly, we assume the minimal batch size of the retailers' orders is 5. We limit the order quantity into the range [0, 2000]. Then each retailer has 400 order quantity choices.

We further model each of the price and quantity decision separately into a decision "black box", which uses Q-learning algorithm to adaptively find best choices. These decision black boxes are also learning automata in artificial intelligence [12]. When a learning automaton adopts Q-learning algorithm, it keeps a Q value for each action, such that $Q^i(a^i)$ represents the expected reward that the automata i believes it will obtain by playing action a^i. In this paper, we use the following formula to update the Q value of each action:

$$Q^i(a^i_{t+1}) = Q^i(a^i_t) + \alpha\left(r^i(a^1_t,\ldots,a^n_t) - Q^i(a^i_t)\right) \qquad (1)$$

Here, a^i_t means the action that the agent i has chosen at time t. For the actions that the agent did not choose, the Q value does not change. $r^i(a^1_t,\ldots,a^n_t)$ is the payoff the agent gets after time t, α is step size. Each automaton adopts softmax policy to choose actions based on their Q values. At beginning, each action's initial Q value is set randomly. The readers can refer to [11] for a detailed discussion on Q-learning algorithm.

Here we will make a further assumption on the decision behavior of each agent: each agent believes that it faces an environment with rather high uncertainty; it does not know when the game will be over, and it knows that it is hard to make predictions. So it does not take into account the possible future earnings. Thus the agents are assumed to be myopic, and repeatedly play a multi-person game in which all players choose actions simultaneously in each round.

2.3 Supplier Behavior

We model the supplier as a non-adaptive agent. In each period, she only does routine works: receives retailer's orders, produces enough products and ships them to the retailers.

2.4 Experiment Design

We carry out the simulation on the platform of Swarm developed by Santa Fe Institute. We use the number of retailers to represent the intensity of the retailers' competition. We select 5 representative retailer numbers 1, 2, 5, 10, and 20 to carry out the simulation one by one. For each retailer number, we run the simulation for 5 times. Each time the simulation is run for 8,000 periods to allow asymptotic behavior to emerge if it is present, and each time we use different initial action weights of each agent and different random seeds. Due to the size limitation of this paper, we will not present the program code.

3. SIMULATION RESULTS

We discuss the simulation results in two steps: convergence and insights.

3.1 Convergence

Figure 2, 3, 4, 5, 6 presents the time series of the retailers' moving average profits (time span is 20) when the numbers of retailers are 1, 2, 5, 10, and 20. The horizontal axis represents time, which is denoted with the number of time grains (There are only three numbers, 20000, 40000 and 60000. They are the time represented by time grain number. In this simulation, one period consists 8 time grains) and the vertical axis represents retailers' moving average profits.

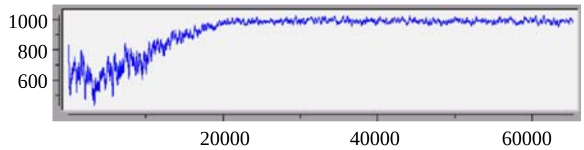

Figure 2 Time series moving average profits with 1 player

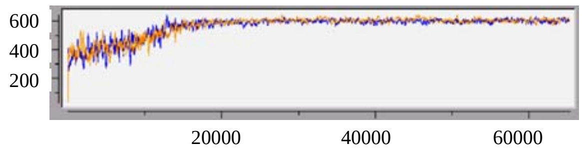

Figure 3 Time series moving average profits with 2 players

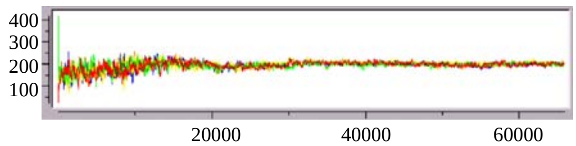

Figure 4 Time series moving average profits with 5 players

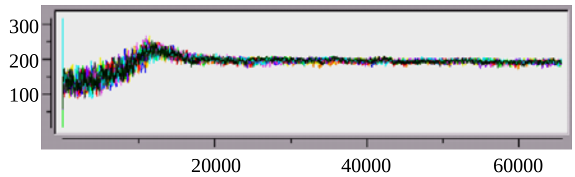

Figure 5 Time series moving average profits with 10 players

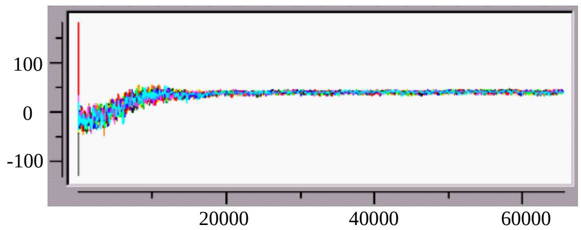

Figure 6 Time series moving average profits with 20 players

We can see from the result data that each run shows asymptotical convergence. When retailer number is the same, different run converges to the same steady state and each retailer finally finds the same policy. For example, in the 2-retailer case, each retailer's final price has mean 2.63 with standard derivation 0.05; final order quantity has mean 399 with standard derivation 18. This result seems self-evident due to the homogeneity of the retailers. But in fact, each retailer has its own individual decision boxes. It is quite possible that different retailer converges to different steady states, or even there is no convergence at all. In this model, each agent's initial state is not identical (each action's initial Q value is set randomly), and each retailer's learning process is also not identical (this is shown in Figure 7, which is an amplification of Figure 3's initial part). So the result of convergences reflects an intrinsic property of the systems.

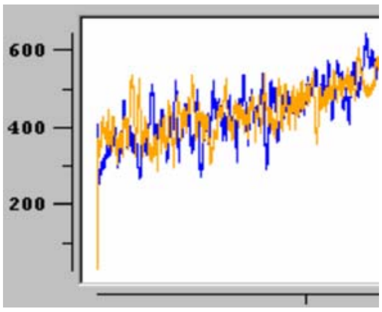

Figure 7 Initial learning dynamics in 2-retailer case

This intrinsic property is that the systems have Nash equilibria, which can be approximated by these steady states.

In this model, each agent has two learning automata, one for retail price and the other for order quantity. We implement the learning mechanism of each learning automaton using Q-learning combined with softmax selection. According to [11], a learning automaton with this learning mechanism approaches best response strategy when the temperature parameter (a parameter in softmax) approaches zero. At steady states, each learning automaton converges to stationary strategies, and the temperature parameter becomes small, so each learning automaton approaches its optimal response. According to theorems on Nash equilibria, if a game with multiple players choosing their best responses converges, it converges to Nash equilibira. Then we know that all the learning automata are in Nash equilibria at steady states.

Each agent consists of two learning automata which have identical payoffs (because these two learning automata update their Q values based on the same agent's payoff value, see (1)). Then the two learning automata of each agent also play a cooperative game with the same aim of making the agent better. So, at steady states all the learning automata approaches Nash equilibrium means two things: first, each agent sticks to stationary price and quantity policies; second, each agent finds its best combination of price policy and quantity policy given other agents' stationary policies. In other words, at steady states, no agent has incentives to change its stationary policy, i.e., the agents in the game approaches Nash equilibria.

We can conduct "turbulence tests" to see whether the equilibrium is stable. Suppose when the system arrives at a steady state, one retailer suddenly loses confidence on his stationary policies, and wants to explore more efficient ones. This is implemented by changing the temperature parameter of softmax selection. We carry out a turbulence test on a 2-retailer equilibrium, and the result shows strong stability (see Figure 8).

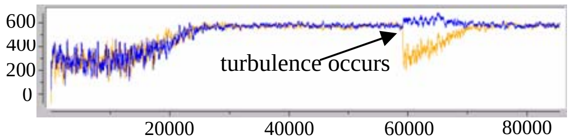

Figure 8 Turbulence test on 2-retailer equilibrium

It is notable that during the learning process, each retailer's average profit is not always increasing. See Figure 5, there is a "peak" where all retailers' profits are higher than in equilibrium. It seems strange that the agents do not learn best strategies. This phenomenon is common in multi-agent learning. Consider the Prisoners' Dilemma, each player gets highest payoff when they cooperate, but it is not Nash equilibrium: each of them has incentive to betray given the other one cooperate. The "peak" in Figure 5 stays in the similar situation: although each retailer gets rather high profits, he has incentive to change his policy to get even higher profits. All the retailers are myopic, change policies for next period's more profits, but they are finally all worse off and approach the Nash equilibrium.

The processes of how the system arrives at steady states are different in different simulation runs. Figure 9 is another run for 10-retailer simulation, which has no prominent "peaks" as in Figure 5. But although the processes of convergence are different, the final states are always the same.

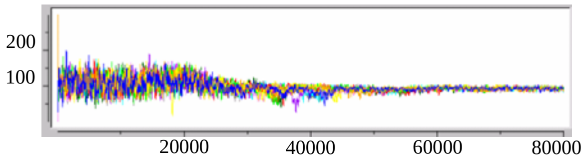

Figure 9 Another run for 10-retailer simulation

The fact that the system approaches Nash equilibrium does not mean that we are designing particular agents to find Nash equilibria. All the behavioral assumptions of the agents in this model are abstractions of real supply chain partners. We use learning algorithms to imitate the decision process of a supply chain partner who faces complex environments and with less information. The convergence means that in this model retailers will adopt stationary policies after a long period of learning and competition. This offers us a basis to investigate how the retailer price and retailer/ supplier profits are affected by the competition among the retailers: we will discuss these issues at steady states.

3.2 Insights

First we investigate how retailer competition affects retail prices. We get the retail prices at steady states of each simulation run, and list the data in table 1 and Figure 10 (These data and all the data in latter tables and figures are the average of 5 simulation runs, see section 2.4).

Table 1 Retail price under different retailer numbers

Retailer Number	1	2	5	10	20
Retail Price	2.85	2.63	2.14	2.01	1.89

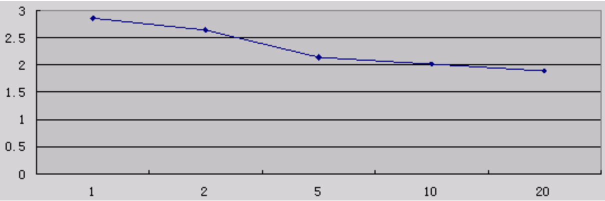

Figure 10 Retail price under different retailer numbers

It can be seen that as retailer number increases (the competition becomes more intense), the retail price decreases. So customer benefits from retailers' competition. This is correspondent with intuition. Yet we observe that although retail price decreases, it does not get too low: when there are 20 retailers, retail price is still 1.88. In other words, the price war is quite mild, and no fierce lowering-price activity appears. This is because of the assumptions on agent behaviors: aspiration on more profits and myopic. Such agents will not adopt policies that decrease short-time profits to defeat other agents.

Table 2 and Figure 11 show supplier revenue under different retailer numbers.

Table 2 Supplier revenue under different retailer numbers

Retailer Number	1	2	5	10	20
Supplier Revenue	578	800	985	1023	1044

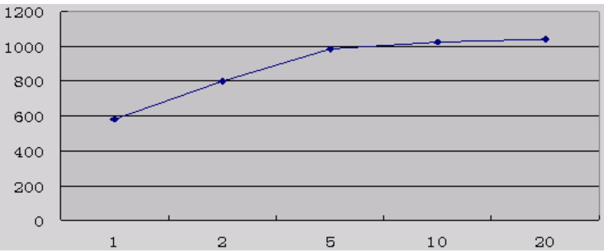

Figure 11 Supplier revenue under different retailer numbers

The supplier's revenue increases with the number of retailers, i.e., competition among the agents are beneficial to the supplier. The reason is as follows: retailer competition decreases retail price, so customer demand increases; accordingly, retailers will order more products from the supplier, and supplier's wholesale price is fixed, thus supplier gets more revenue.

We use following Table 3 and Figure 12 to study how competition affects the retailers.

Table 3 Retailer profit under different retailer numbers

Retailer Number	1	2	5	10	20
Total Profit	991	1193	1010	915	806
Single Profit	991	596	202	92	40

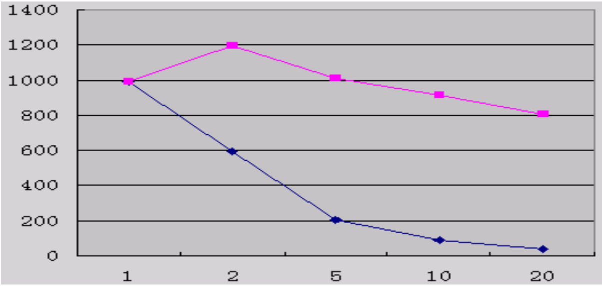

Figure 12 Retailer profit under different retailer numbers

Each retailer's profit decreases with the number of retailers in the market. In this model, the mean of total customer demand does not change and all the retailers share this market. The more the retailers, the smaller each market share.

Figure 12 also shows that retailers' total profit is maximal when there are two retailers and it decreases when there are more than two retailers. The decrease of retailers' total profit can be attributed to the competition among the retailers. Figure 10 show that retailer competition can decrease retail price. This has two effects: (1) decrease retailer profit directly, and (2) increase customer demand and thus increase retailer profit. Notice that in this model, each customer has the choice of not buying. Thus if there are more retailers, a customer is less likely to leave without buying anything, i.e., retailers' total profit increases. This can be viewed as the benefits of offering customers more selection options. We denote this effect as (3). General literature on supply chain management ignores this effect, yet following we will see that this effect can play an important role in certain cases. Whether retailers' total profit increases or decreases depends on the trade off between effect (1) (decreases retailers' total profit) and effect (2) plus (3) (increase retailers' total profit). From Figure 12 we can see when there are more than 2 retailers, the retailers' total profit decreases, which implies that effect (1) dominates effect (2) plus (3). Yet retailers' total profit is more in 2-retailer case than in 1-retailer case. This is different with general economic conclusions which believe a monopoly often obtains maximum market profits. Effect (3) plays an important role in the difference between this model and classical economic model, and makes effect (2) plus (3) dominate in this case.

4. CONCLUSIONS AND FUTURE RESEARCH DIRECTIONS

In this paper, we investigate the horizontal competition in a one-supplier-multiple-retailer supply chain. We adopt agent-based simulation, which is quite different with classical methods. Based on the assumptions of the agents' behaviors, we endow each agent with a reasonable learning mechanism. The results show that the system converges asymptotically to steady states that approach Nash equilibria. Under these steady states, we analyze how retailer price and supplier/retailer profits are affected by the competition among the price-setting retailers, and notice the important role of customer choice behavior which is generally ignored in supply chain research.

This paper also inspires us an in-depth research on the issues of horizontal cooperation among the retailers. From former analysis we know that retailers' total profit decreases with retailer number. This gives the retailers strong incentives to collaborate with each other. We will address this issue in future work.

Furthermore, this paper inspires us to study vertical competition and cooperation in the supply chain. Here we get the conclusion that the supplier will benefit from the competition between the retailers. How things would be if the supplier can adjust wholesale price? And, when the supplier knows that retailer competition is beneficial for her, she will have incentive to prevent the collaborations among the retailers. So we can investigate what measures the supplier can take to prevent such collaborations.

ACKNOWLEDGEMENTS

This work is partly supported by the National Science Foundation of China (70231010, 70321001).

REFERENCES

[1] Cachon, G.P., *Supply Chain Coordination with Contracts*, in *Handbooks in Operations Research and Management Science: Supply Chain Management*, S. Graves and T.d. Kok, Editors. 2002, North-Holland.
[2] Kirman, A.P. and N.J. Vriend, *Evolving market structure: An ACE model of price dispersion and loyalty.* Journal of Economic Dynamics & Control, 2001. **25**: p. 459-502.
[3] Chen, S.-H. and C.-H. Yeh, *Evolving Traders and the Business School with Genetic Programming: A New Architecture of the Agent-Based Artificial Stock Market.* Journal of Economic Dynamics & Control, 2001. **25**: p. 363-393.
[4] Tesfatsion, L., *Agent-Based Computational Economics: Growing Economies from the Bottom Up.* Artificial Life, 2002. **8**: p. 55-82.
[5] Banks, J., et al., *Discrete-Event System Simulation (3rd Edition)*. 2000: Prentice Hall.
[6] Parunak, H.V.D., R. Savit, and R.L. Riolo. *Agent-Based Modeling vs. Equation-Based Modeling: A Case Study and Users' Guide*. in *Proceedings of Multi-agent systems and Agent-based Simulation (MABS'98)*. 1998.
[7] Chen, J. and M. Xie. *Research on a Systematic Approach for Building Multi-Agent Supply Chain Models*. in *The 2nd International Workshop on Meta-synthesis and Complex System*. 2002. Shanghai, China.
[8] Lippman, S. and K. McCardle, *The Competitive Newsboy.* Operations Research, 1997. **45**: p. 54-65.
[9] Mahajan, S. and G.v. Ryzin, *Inventory Competition under Dynamic Consumer Choice.* Operations Research, 2001. **49**(5): p. 646-657.
[10] Train, K., *Discrete Choice Methods with Simulation*. 2003: Cambridge University Press.
[11] Sutton, R.S. and A.G. Barto, *Reinforcement Learning: An Introduction*. 1998: MIT Press, Cambridge, MA.
[12] Narendra, K.S. and M.A.L. Thathachar, *Learning Automata: An Introduction*. 1989, NJ: Prentice-Hall.
[13] Leslie, D.S. and E.J. Collins, *INDIVIDUAL Q-LEARNING IN NORMAL FORM GAMES.* Unpublished, 2004.