Distribution-Free Learning in Inventory Competition

Ashkan Zeinalzadeh
Department of Electrical Engineering, University of Hawaii at Mānoa, 2540 Dole Street, Honolulu, HI 96822, USA, ashkan@hawaii.edu

Aydın Alptekinoğlu
Edwin L. Cox School of Business, Southern Methodist University, 6212 Bishop Boulevard, Dallas, TX 75275, USA, aalp@smu.edu

Gürdal Arslan
Department of Electrical Engineering, University of Hawaii at Mānoa, 2540 Dole Street, Honolulu, HI 96822, USA, gurdal@hawaii.edu

Consider n firms locked in inventory competition. They each stock one product, and these products are substitutes. Customers who encounter a stockout in the store of their first choice, either flock to one of the other firms in fixed (deterministic) proportions or exit the market. We are interested in whether these firms can learn the equilibrium behavior over multiple periods by simply observing their own sales and without knowing anything about each other. Our main result involves proposing two simple learning rules for these firms and formally establishing that their inventory decisions, if generated by either one of these rules, would converge with probability one to the Nash equilibrium of the single-period inventory competition game.

Key words: inventory competition, learning, equilibrium

1. Introduction

Our primary concern in this paper is whether multiple firms holding inventories of substitutable goods and competing for demand on the basis of availability can iteratively learn to mimic Nash equilibrium behavior merely by observing their own sales and nothing else.

Existence and nature of equilibrium in inventory competition has been studied from a traditional game theory perspective (Parlar 1988, Lippman and McCardle 1997, Mahajan and van Ryzin 2001, Avsar and Baykal-Gursoy 2002, Netessine and Rudi 2003, Bernstein and Federgruen 2004, Netessine et al. 2006, Olsen and Parker 2008), which typically requires the parties to know a lot about each other and also their own demand not just sales. In this paper we take a learning approach with minimal informational requirements on the players. Our work extends Burnetas and Smith (2000) to multiple periods with inventory carryover, which significantly increases the practical relevance of learning in game theory to inventory control.

We use the following notation throughout the paper: LHS and RHS stand for "left hand side" and "right hand side," respectively; ODE stands for "ordinary differential equation;" iid stands for "independently and identically distributed;" $:=$ states a definition; $\equiv$ means identically equal to; $-i$ means all indices other than i, i.e., $x_{-i} = (x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n)$ for $x \in \mathbb{R}^n$; $I\{\cdot\}$ is the indicator function, which equals 1 if the condition within brackets holds, and 0 otherwise; F_X denotes the cumulative distribution function of a random variable X; $P[\cdot]$ denotes probability; $E[\cdot]$ denotes expectation; $\mathbb{R}$, $\mathbb{R}_+$, $\mathbb{R}_{++}$ denote the real, nonnegative real, and positive real numbers, respectively; for a positive integer n, $\mathbb{R}^n$ denotes the n-fold product $\mathbb{R} \times \cdots \times \mathbb{R}$; $\mathbb{R}^n_+$ and $\mathbb{R}^n_{++}$ are defined analogously; $(\cdot)^T$ denotes transpose; for $y, z \in \mathbb{R}^n$, $y \leq z$ or $z \geq y$ means $z - y \in \mathbb{R}^n_+$, and $y < z$ or $z > y$ means $z - y \in \mathbb{R}^n_{++}$; $\mathbf{0}$ and $\mathbf{1}$ denote all-zeros and all-ones vectors of appropriate dimension, respectively; $\mathbf{I}$ denotes the identity matrix of appropriate dimension; for $x \in \mathbb{R}^n$ and $p \geq 1$, $\|x\|_p := (|x_1|^p + \cdots + |x_n|^p)^{1/p}$ is the p-norm in $\mathbb{R}^n$; for a closed and convex set $H \subset \mathbb{R}^n$, and $x \in \mathbb{R}^n$, $\Pi_H[x] := \operatorname{argmin}_{\bar{x} \in H} \|x - \bar{x}\|_2$ is a projection onto H; for $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$, $(x)^+ := (\max(0, x_1), \ldots, \max(0, x_n))$; $\lceil \cdot \rceil$ denotes integer ceiling; $\operatorname{diag}(x_1, \ldots, x_n)$ denotes the diagonal matrix with entries $x_1, \ldots, x_n$; for a function $f : S \to \mathbb{R}$, where S is an open subset of $\mathbb{R}^n$, ∇f denotes the gradient vector of first-order partial derivatives, and $\nabla^2 f$ denotes the Hessian matrix of second-order partial derivatives; $A \Rightarrow B$ or $B \Leftarrow A$ means "A implies B;" $A \Leftrightarrow B$ means "A and B are equivalent statements."

2. Model

Our model of inventory competition involves multiple firms and a single product that they each carry. These products are substitutes: they may or may not be differentiated, but some customers substitute one for the other. A natural special case of our models is applicable to multiple retailers carrying exactly the same product. For brevity, our model speaks of a single product as in this special case, even though they apply to any substitutable products each sold by a separate firm.

2.1. One-Shot Inventory Competition Model

Consider a collection of firms, labeled as $1, \ldots, N$. Each firm i has zero initial inventory and places replenishment orders without observing the ordering decisions of the other firms. This brings firm i's inventory level from zero to $y_i \geq 0$ at the cost of $c_i y_i$, where $c_i \geq 0$ is the variable order cost per unit.

Then, firm i receives demand d_i from the customers for whom firm i is the first choice. We assume that $d := (d_1, \ldots, d_N)$ is an $\mathbb{R}_+^N$-valued exogenous random vector, and its distribution is independent of the firms' decisions. Note that we allow $d_1, \ldots, d_N$ to be correlated. If the first-choice demand d_i cannot be fully satisfied by firm i, i.e., $d_i > y_i$, then a fixed proportion $\alpha_{j,i} \in [0,1]$ of excess demand $(d_i - y_i)^+$ switches from firm i to firm j. Clearly, $\sum_j \alpha_{j,i} \leq 1$, where $\alpha_{i,i} := 0$. (The remainder of the first-choice demand is lost for all firms.) Therefore, the total demand faced by firm i is

$$\bar{d}_i\left(y_{-i}\right) := d_i + \sum_j \alpha_{i,j}\left(d_j - y_j\right)^+ \tag{1}$$

(We will suppress the dependence of $\bar{d}_i$ on y_{-i} when there is no possibility of confusion.) After the realization of demand, including the switching behavior, firm i collects a total revenue of $r_i \min\left(y_i, \bar{d}_i\right) + q_i\left(y_i - \bar{d}_i\right)^+$, where $r_i \geq 0$ is the revenue per unit sales, and $q_i \geq 0$ is the salvage value of unsold goods per unit. As a result, firm i's expected profit is given as

$$g_i\left(y_i, y_{-i}\right) := E\left[r_i \min\left(y_i, \bar{d}_i\right) + q_i\left(y_i - \bar{d}_i\right)^+\right] - c_i y_i$$

where the expectation is taken with respect to d. We rewrite firm i's expected profit in the more convenient form as

$$g_i\left(y_i, y_{-i}\right) = \left(r_i - c_i\right) y_i - \left(r_i - q_i\right) E\left[\left(y_i - \bar{d}_i\right)^+\right]. \tag{2}$$

We refer the noncooperative game characterized by the expected profits (2) and the strategy sets $\mathbb{R}_+, \ldots, \mathbb{R}_+$ as the inventory competition game. To ensure sensible parameter values and avoid trivial cases, we will make the following assumption throughout the paper.

Assumption 1. *$q_i, c_i, r_i, d_i, \alpha_{i,j} \in \mathbb{R}_+$, $q_i < c_i < r_i$, $\sum_\ell \alpha_{\ell,i} \leq 1$, for all i, j.*

Our modeling setup and the accompanying assumptions, such as the customer switching behavior, are standard in the inventory competition literature.

The central concept in noncooperative game theory is that of Nash equilibrium. In our context, a Nash equilibrium is a collection of firms' inventory decisions $(y_1^*, \ldots, y_N^*)$ such that

$$g_i\left(y_i^*, y_{-i}^*\right) = \max_{y_i \in \mathbb{R}_+} g_i\left(y_i, y_{-i}^*\right), \qquad \text{for all } i.$$

In other words, a Nash equilibrium is a decision profile that is optimal for each firm provided that the other firms do not deviate from their Nash equilibrium decisions. We will henceforth refer to a Nash equilibrium simply as an equilibrium.

An equilibrium can also be thought of as a fixed point of the best response correspondence $BR = (BR_1, \ldots, BR_N)$ which maps $\mathbb{R}_+^N$ into the subsets of $\mathbb{R}_+^N$ and is defined by

$$BR_i\left(y_{-i}\right) := \operatorname{argmax}_{y_i \in \mathbb{R}_+} g_i\left(y_i, y_{-i}\right), \qquad \text{for all } i \text{ and } y_{-i} \in \mathbb{R}_+^{N-1}. \tag{3}$$

In other words, $y^* \in \mathbb{R}_+^N$ is an equilibrium if and only if

$$y^* \in BR\left(y^*\right).$$

We refer the reader to Appendix A for the properties of BR.

It can be shown that the inventory competition game introduced above always possess an equilibrium; moreover, the uniqueness of equilibrium can be shown under mild conditions; see Parlar (1988) for two firms, Parlar and Wang (1994) for three firms, Karjalainen (1992), Netessine and Rudi (2003) for arbitrary number of firms, Lippman and McCardle (1997), Mahajan and van Ryzin (2001), Netessine et al. (2006) for generalizations of demand generation and substitution models. For completeness, we reproduce a proof of existence from the literature; see for example Theorem 5 in Lippman and McCardle (1997).

Proposition 1. *If Assumption 1 holds, the inventory competition game has an equilibrium.*

Proof. Using the definitions of continuity and concavity, it is straightforward to see that (a) for all i, g_i is continuous in $\mathbb{R}_+^N$, and (b) for all i and $y_{-i} \in \mathbb{R}_+^{N-1}$, $g_i\left(\cdot, y_{-i}\right)$ is concave in $\mathbb{R}_+$.

Furthermore, Appendix A shows that BR maps $\mathbb{R}_+^N$ into the subsets of a closed convex set Y defined as (13). Hence, firms' strategies can be restricted, without loss of generality, to Y. As such, any inventory competition game satisfying Assumption 1 is a "concave game" in the sense of Rosen (1965). Theorem 1 in Rosen (1965) shows that concave games possess equilibria.

It appears that, in the case of arbitrary number of heterogenous firms, there are only two papers dealing with the uniqueness of equilibrium, namely Karjalainen (1992), and Netessine and Rudi (2003). Karjalainen (1992) is an unpublished report to which we do not have access. Netessine and Rudi (2003) shows the uniqueness of equilibrium when the demand has (strictly positive) density, and

$$\max_i \sum_j \alpha_{i,j} < 1 \qquad \text{or} \qquad \max_j \sum_i \alpha_{i,j} < 1. \tag{4}$$

We prove the uniqueness under weaker conditions.

Assumption 2. *F_d is continuous and* $\mathbf{0} \le \hat{y} < \check{y} \Rightarrow F_d(\hat{y}) < F_d(\check{y})$.

Under Assumptions 1 and 2, we establish that BR is a contractive mapping from $\mathbb{R}_+^N$ into a compact set; see Appendix A for the proof. This result can be obtained under a relaxation of Assumption 2 that F_d has the required properties only in a sufficiently large but bounded subset of $\mathbb{R}_+^N$, by following along the same lines of the proof. This relaxation would accommodate, in particular, some demand distributions with bounded support. However, for ease of presentation, we prefer to work under Assumption 2.

Proposition 2. *If Assumptions 1 and 2 hold, the inventory competition game has a unique equilibrium; moreover, the best response iterations $y^{\ell+1} = BR(y^\ell)$, for $\ell \ge 1$, starting from any $y^1 \in \mathbb{R}_+^N$ converges to the equilibrium.*

Proof. Appendix A shows that, under Assumptions 1 and 2, BR is a contractive mapping from $\mathbb{R}_+^N$ into Y, where Y is the compact set defined as (13). Hence, BR must have a unique fixed point in Y, which can be obtained by the successive approximation iterations starting from an arbitrary point in $\mathbb{R}_+^N$; see Theorem 1 in Edelstein (1962).

The results of Proposition 2 can be proven without the continuity of F_d provided

$$\|\alpha\|_{\mathbb{R}^{N^2}} < 1 \tag{5}$$

where α is the matrix whose (i,j)-th entry equals $\alpha_{i,j}$ and $\|\cdot\|_{\mathbb{R}^{N^2}}$ is a matrix norm induced by a monotone1. norm in $\mathbb{R}^N$. It is easy to see that (4) is sufficient but not necessary for (5). At the end of Appendix A, we point out how to modify the proof to show that BR is a contractive mapping from $\mathbb{R}_+^N$ into Y, when F_d is strictly monotonic in $\mathbb{R}_+^N$ and (5) holds.

Having discussed the basics of equilibria, we move on to the issue of how an equilibrium can arise in an actual play of an inventory competition game involving imperfect firms. Traditionally, an equilibrium is justified as the predicted outcome of a noncooperative game if it is "common knowledge" that all players are rational and know the utility functions of their own as well as their opponents'; see Başar and Olsder (1999), Fudenberg and Tirole (1991). In contrast, an equilibrium can also be justified if it emerges as the long-term outcome of an iterative procedure whereby players with limited rationality and information grope for individual optimality by making repeated decisions based on their observations of the past play. This is the essence of what is known as the "learning in games" approach for which we refer the reader to the books Young (2004, 1998), Fudenberg and Levine (1998), Hofbauer and Sigmund (1998), Weibull (1995) and the references therein. Similarly, this paper takes a learning approach to inventory competition and deals with the question of whether or not firms can learn to maximize their expected profits in a repeated inventory competition where each firm 1) observes only its own past decisions and total demands (or instead only its own past decisions and sales), and 2) knows only its own critical ratio $\gamma_i := \frac{r_i - c_i}{r_i - q_i}$, in particular without any knowledge of any firm's demand distributions including its own.

2.2. Repeated Inventory Competition Model

We now consider an infinitely repeated inventory competition such that, in each period $t \in \{1, 2, \dots\}$, the same sequence of events described in the one-shot inventory competition model occur. Namely, in each period, firms simultaneously place inventory replenishment orders which are fulfilled instantaneously, then the first-choice demands are realized, next the demand switchings occur, subsequently the unsold goods are salvaged (hence no inventory is carried to the next period), and finally the profits are made. We assume that the first-choice demands $d_t := (d_{1,t}, \ldots, d_{N,t})$ satisfy the following assumption.

Assumption 3. *$\{d_t\}_{t \geq 1}$ is an iid sequence in $\mathbb{R}_+^N$ with the cdf F_d satisfying Assumption 2.*

Let $y_{i,t}$ and $d_{i,t}$ denote the inventory level and the first-choice demand of firm i in period t, respectively. Then, $\bar{d}_{i,t} := d_{i,t} + \sum_j \alpha_{i,j} (d_{j,t} - y_{j,t})^+$ and $s_{i,t} := \min \{y_{i,t}, \bar{d}_{i,t}\}$ denote the total demand and the sales of firm i in period t, respectively. We assume that firm i's observation history in period t before choosing $y_{i,t}$ is $(y_{i,1}, \bar{d}_{i,1}, \ldots, y_{i,t-1}, \bar{d}_{i,t-1})$ (or instead $(y_{i,1}, s_{i,1}, \ldots, y_{i,t-1}, s_{i,t-1})$). The observation history (in particular the past total demands or sales) of a firm contains implicit information about the past decisions of the other firms. However, each firm does not have any other information about the other firms. In fact, the firms need not even be aware of each other or the fact that they are involved in an inventory competition. Hence, we argue that a firm would find it too difficult to influence the (future) decisions of the other firms. In this setup, it is plausible for each firm to (incorrectly) assume that it is facing an iid (total) demand with an unknown distribution. Therefore, from any individual firm's viewpoint, learning in a repeated inventory competition seems to be a process of adjusting the inventory level in each period based on the past inventory levels and the total demands or sales to meet the future demands whose distribution is iid but completely unknown.

A firm observing its total demands or sales can make inventory decisions not only to increase its future profits but also to probe for information on the distribution generating its demands. A firm can also employ a dynamic programming approach to strike an optimal balance between making profits and probing for more information; this is a viable option primarily when the demand distribution is known to belong to a parameterized family of distributions. However, we will further assume that the firms in our setup are myopic, that is, each firm makes an inventory decision in each period to maximize its expected profit only in the current period. Myopic decision making can actually be optimal in certain cases; for example, when future rewards are sufficiently discounted. Also, firms can choose to be myopic in the interest of simplicity. Our motivation for assuming myopia in the repeated inventory competition is tractability, since otherwise we would be led to a learning problem in a dynamic game with quite complicated decision rules for the firms. In the following sections, we will introduce learning models for repeated inventory competition in view of the discussion above.

3. Learning in Repeated Inventory Competition

We will start with the well-known idea of gradient ascent, that is, the adjustment of a decision variable in a direction in which the objective function increases; see Flam (2002). Assume, for the time being, that the inventory levels of firm i's competitors are constant at some $\hat{y}_{-i} \in \mathbb{R}_+^{N-1}$, i.e., $y_{-i,t} = \hat{y}_{-i}$ for all $t \geq 1$, and that the cdf of $\bar{d}_{i,t}(\hat{y}_{-i})$ now generating firm i's demands, denoted by $F_{\bar{d}_i(\hat{y}_{-i})}$, for all t, is known. According to gradient ascent in this case, firm i would adjust its inventory level $y_{i,t}$ as

$$y_{i,t+1} = y_{i,t} + a_t \frac{\partial g_i}{\partial y_i}(y_{i,t}, \hat{y}_{-i}) \tag{6}$$

where $a_t > 0$ is the step size in period t and the gradient $\frac{\partial g_i}{\partial y_i}(y_{i,t}, \hat{y}_{-i})$ in period t equals

$$\frac{\partial g_i}{\partial y_i}(y_{i,t}, \hat{y}_{-i}) = (r_i - c_i) - (r_i - q_i) F_{\bar{d}_i(\hat{y}_{-i})}(y_{i,t}).$$

By absorbing the constant $(r_i - q_i)$ in a_t, we can rewrite (6) as

$$y_{i,t+1} = y_{i,t} + a_t \left(\gamma_i - F_{\bar{d}_i(\hat{y}_{-i})}(y_{i,t}) \right). \tag{7}$$

It is known that if a_t is chosen properly then the recursion (7) will drive $\gamma_i - F_{\bar{d}_i(\hat{y}_{-i})}(y_{i,t})$ to zero, at which the optimality is achieved.

We now consider the case where the inventory levels of firm i's competitors are constant at $\hat{y}_{-i}$ but $F_{\bar{d}_i(\hat{y}_{-i})}$ is not known. In this case, if firm i had access to the realizations of its total demands, then it would be reasonable for firm i to replace $F_{\bar{d}_i(\hat{y}_{-i})}(y_{i,t})$ in (7) with the one-sample approximation

$$I\left\{ \bar{d}_{i,t}(\hat{y}_{-i}) < y_{i,t} \right\}. \tag{8}$$

The strict inequality in (8) is justified due to the continuity of $F_{\bar{d}_i(\hat{y}_{-i})}$, which follows from Assumption 3. Note that the one-sample approximation given in (8) equals

$$I\{s_{i,t} < y_{i,t}\}$$

which requires only the sales information. Therefore, a firm i observing its past sales can adjust its inventory according to the following approximate gradient ascent process

$$y_{i,t+1} = y_{i,t} + a_t(\gamma_i - I\{s_{i,t} < y_{i,t}\}). \tag{9}$$

The process (9) belongs to the well-known class of stochastic approximation (SA) processes, and hence its convergence properties can be studied using the theory developed for the SA processes; see Kushner and Yin (2003), Benveniste et al. (1990), Benaïm (1999). For example, if $y_{-i,t}$ is constant and a_t satisfies certain conditions such as $\sum_t a_t = \infty$ and $\sum_t a_t^2 < \infty$, then the convergence of (9) can be established by the ODE method for the SA processes; see Kushner and Yin (2003). One natural choice for the step size that satisfies the conditions of the SA theory is $a_t = \frac{1}{t}$.

We should point out that the RHS of (9) needs to be projected onto $\mathbb{R}_+$ to prevent $y_{i,t} < 0$. This leads to the process

$$y_{i,t+1} = (y_{i,t} + a_t(\gamma_i - I\{s_{i,t} < y_{i,t}\}))^+, \qquad y_{i,1} \geq 0. \tag{10}$$

Alternatively, we consider a close variant of the process (10) given by

$$y_{i,t+1} = y_{i,t} + a_t y_{i,t}(\gamma_i - I\{s_{i,t} < y_{i,t}\}), \qquad y_{i,1} > 0. \tag{11}$$

The interpretation of (11) is that, if all the goods are sold in one period then the inventory level in the next period increases by a factor of $a_t\gamma_i$; otherwise it decreases by a factor of $a_t(1-\gamma_i)$. The process (11) has a superfluous rest point at 0. However, it also guarantees $y_{i,t} > 0$, for all t, whenever $y_{i,1} > 0$ and $\sup_t a_t < \min_i \frac{1}{1-\gamma_i}$. The process (11) is the same as the one considered in Burnetas and Smith (2000) for a single firm. An elaborate proof of convergence of (11) for a single firm is given in Burnetas and Smith (2000) using a generalization of the martingale convergence theory. We claim that the convergence of (11) for a single firm can be established easily by the ODE method for the SA processes; see Kushner and Yin (2003), Benaïm (1999). Our goal in this paper, however, is significantly more ambitious. Namely, we wish to prove convergence of inventory decisions of multiple firms in a repeated inventory competition when each firm i adjusts its inventory decisions according to (11).

The convergence of (11) in the case of a single firm is a very desirable property, however, it is by mo means sufficient for the convergence of the inventory decisions made by multiple firms using (11). The literature on learning in games has many examples where more than two players adjusting their decisions by a learning process which is convergent in the single player case exhibit non-convergence or even chaos; see Shapley (1964), Jordan (1993), Foster and Young (79–96), Hart and Mas-Colell (2003), Sato et al. (2002). The reason for this is that, in the case of multiple firms, the learning processes of the firms are coupled, i.e., firm i's learning process is affected by firm j's learning process and vice versa. Hence, what each firm i is learning to optimize constantly changes due to another firm j's learning process. In repeated inventory competition, this means that each firm using (11) is actually learning to optimize in a non-stationary environment. Furthermore, the source of this non-stationarity faced by a firm is not a simple one, rather, it is another firm that is also learning to optimize its own expected profit in a complicated non-stationary environment. Therefore, there is no a priori reason to expect the convergence of the firms' decisions in a repeated inventory competition when each firm adjusts its decisions using (11). It turns out, however, that the convergence of such a multi-firm learning process can still be proven using some of the recent advancements in the SA theory; see Kushner and Yin (2003), Benaïm (1999).

The following assumption is needed to prove the main result.

Assumption 4. $0 \le a_t < \min_i \frac{1}{1-\gamma_i}$, *for all* t, *and* $\sum_t a_t = \infty$, $\sum_t a_t^2 < \infty$.

Assumption 4, except the upper bound on a_t, is one of the standard assumptions on the step size used in the literature to prove the convergence of the SA processes. The upper bound on a_t is necessary to keep the inventory decisions generated by (11) positive; however, it can be removed if the inventory decisions are generated by (10).

We now state the main result of this section whose proof is provided in Appendix C.

THEOREM 1. *Consider a repeated inventory competition model where all firms generate their inventory decisions $y_t := (y_{1,t}, \ldots, y_{N,t})$ by either (10) or (11). If Assumptions 1, 3 and 4 hold, then y_t converges, with probability one, to the unique equilibrium of the corresponding one-shot inventory competition model.*

Appendix A: Best Response Function

CLAIM 1.

1. *If Assumption 1 holds, then, for all i and $y_{-i} \in \mathbb{R}_+^{N-1}$,*

$$BR_i(y_{-i}) = \left[\min\left\{\xi : F_{\bar{d}_i(y_{-i})}(\xi) \geq \gamma_i\right\}, \inf\left\{\xi : F_{\bar{d}_i(y_{-i})}(\xi) > \gamma_i\right\}\right]. \tag{12}$$

Consequently, BR maps $\mathbb{R}_+^N$ into the subsets of Y defined as

$$Y := \bigtimes_i \left[\min\left\{\xi : F_{d_i}(\xi) \geq \gamma_i\right\}, \inf\left\{\xi : F_{\bar{d}_i(\mathbf{0})}(\xi) > \gamma_i\right\}\right]. \tag{13}$$

2. *If Assumptions 1 and 2 hold, then BR has the following properties in $\mathbb{R}_+^N$.*
   (a) *BR is single valued, and hence maps $\mathbb{R}_+^N$ into Y*
   (b) *BR is nonincreasing, i.e., $\check{y} \geq \hat{y} \Rightarrow BR(\check{y}) \leq BR(\hat{y})$*
   (c) *BR is contractive, i.e., $\check{y} \neq \hat{y} \Rightarrow \|BR(\check{y}) - BR(\hat{y})\|_{\mathbb{R}^N} < \|\check{y} - \hat{y}\|_{\mathbb{R}^N}$.*

Proof.

1. In the literature, the optimization problem in (3), for any i and any $y_{-i} \in \mathbb{R}_+^{N-1}$, is referred to as a *newsboy problem*. It is known that, under Assumption 1, any $y_i \in \mathbb{R}$ that belongs to the RHS of (12) is a solution to such a newsboy problem.

2. We begin by observing that, for all i and $y_{-i} \in \mathbb{R}_+^{N-1}$, (a) the strict monotonicity of F_d in $\mathbb{R}_+^N$ (in the sense of Assumption 2) implies the strict monotonicity of $F_{\bar{d}_i(y_{-i})}$ in $\mathbb{R}_+$, and (b) the continuity of F_d implies the continuity of $F_{\bar{d}_i(y_{-i})}$.

(a) Assume that, for some i and $y_{-i} \in \mathbb{R}_+^{N-1}$, $BR_i(y_{-i})$ is not a singleton. Pick $y_i^1, y_i^2 \in BR_i(y_{-i})$ such that $y_i^1 < y_i^2$. We must have $F_{\bar{d}_i(y_{-i})}(y_i^1) = F_{\bar{d}_i(y_{-i})}(y_i^2) = \gamma_i$. However, this contradicts the strict monotonicity of $F_{\bar{d}_i(y_{-i})}$ in $\mathbb{R}_+$. Hence, BR is single valued in $\mathbb{R}_+^N$.

(b) It follows from the single-valuedness in $\mathbb{R}_+^N$ and (12).

(c) Fix $y^1, y^2 \in \mathbb{R}_+^N$ such that $y^1 \neq y^2$. If $\sum_j \alpha_{i,j} \left|y_j^1 - y_j^2\right| = 0$, then $BR_i\left(y_{-i}^1\right) = BR_i\left(y_{-i}^2\right)$. Hence, we focus on any firm i for which $\sum_j \alpha_{i,j} \left|y_j^1 - y_j^2\right| > 0$. From the continuity of $F_{\bar{d}_i\left(y_{-i}^\ell\right)}$, we have, for all ℓ,

$$P\left[\bar{d}_i\left(y_{-i}^\ell\right) \leq BR_i\left(y_{-i}^\ell\right)\right] = \gamma_i. \tag{14}$$

Also, since

$$\left|\bar{d}_i\left(y_{-i}^1\right) - \bar{d}_i\left(y_{-i}^2\right)\right| \leq \sum_j \alpha_{i,j} \left|y_j^1 - y_j^2\right| \tag{15}$$

we have

$$\bar{d}_i\left(y_{-i}^1\right) \leq BR_i\left(y_{-i}^1\right) \qquad \Rightarrow \qquad \bar{d}_i\left(y_{-i}^2\right) \leq BR_i\left(y_{-i}^1\right) + \sum_j \alpha_{i,j} \left|y_j^1 - y_j^2\right|. \tag{16}$$

Moreover, if $\hat{y} := \left(0, \ldots, 0, BR_i\left(y_{-i}^1\right), 0, \ldots, 0\right)$ and $\check{y} := \hat{y} + \frac{1}{2} \frac{\sum_j \alpha_{i,j}\left|y_j^1 - y_j^2\right|}{\sum_j \alpha_{i,j}} \mathbf{1}$, then

$$\hat{y} < d \leq \check{y} \qquad \Rightarrow \qquad \begin{pmatrix} \bar{d}_i\left(y_{-i}^1\right) \not\leq BR_i\left(y_{-i}^1\right) \\ \text{and} \\ \bar{d}_i\left(y_{-i}^2\right) \leq BR_i\left(y_{-i}^1\right) + \sum_j \alpha_{i,j} \left|y_j^1 - y_j^2\right| \end{pmatrix}.$$

Since $\mathbf{0} \leq \hat{y} < \check{y}$, we have $F_d(\hat{y}) < F_d(\check{y})$, due to Assumption 2. Therefore

$$P\left[\bar{d}_i\left(y_{-i}^2\right) \leq BR_i\left(y_{-i}^1\right) + \sum_j \alpha_{i,j} \left|y_j^1 - y_j^2\right|\right] > \gamma_i. \tag{17}$$

Now, from (14) and (17), we have $BR_i\left(y_{-i}^2\right) < BR_i\left(y_{-i}^1\right) + \sum_j \alpha_{i,j} \left|y_j^1 - y_j^2\right|$. By a symmetric argument, we obtain

$$\left|BR_i\left(y_{-i}^2\right) - BR_i\left(y_{-i}^1\right)\right| < \sum_j \alpha_{i,j} \left|y_j^1 - y_j^2\right|. \tag{18}$$

This leads us to the desired result

$$\left\|BR\left(y^2\right) - BR\left(y^1\right)\right\|_1 < \quad \max_i \sum_j \alpha_{j,i} \Bigg) \left\|y^1 - y^2\right\|_1 \leq \left\|y^1 - y^2\right\|_1.$$

Finally, we show how to modify the proof if $\|\alpha\|_{\mathbb{R}^{N^2}} < 1$ but the continuity of F_d, which implies the continuity of $F_{\bar{d}_i(y_{-i})}$, is not assumed. Since F_d is strictly monotonic in $\mathbb{R}_+^N$, $F_{\bar{d}_i(y_{-i})}$ is strictly monotonic $\mathbb{R}_+$, for all i and $y_{-i} \in \mathbb{R}_+^{N-1}$. Therefore, BR is single valued and non-increasing in $\mathbb{R}_+^N$. To show that BR is contractive, we replace the equality (14) with an inequality where LHS≥RHS. Since (15)-(16) still hold, we replace the strict inequality (17) with an inequality where LHS≥RHS. Accordingly, we replace the strict inequality (18) with an inequality where LHS≤RHS. This leads us to

$$\left\|BR\left(y^2\right) - BR\left(y^1\right)\right\|_{\mathbb{R}^N} \leq \|\alpha\|_{\mathbb{R}^{N^2}} \left\|y^1 - y^2\right\|_{\mathbb{R}^N}$$

where $\|\cdot\|_{\mathbb{R}^N}$ is some monotone norm in $\mathbb{R}^N$ and $\|\cdot\|_{\mathbb{R}^{N^2}}$ is the matrix norm induced by $\|\cdot\|_{\mathbb{R}^N}$ such that $\|\alpha\|_{\mathbb{R}^{N^2}} < 1$. Therefore, BR is not only a contractive mapping but also a contraction mapping from $\mathbb{R}_+^N$ into Y.

Appendix B: Proof of Theorem 1

We provide a detailed convergence proof only for the process (11) as the convergence of the process (10) can be proven in a completely analogous manner. The only modification worth mentioning is that the process (10) can hit the boundary of $\mathbb{R}_+^N$, and $\mathbb{R}_+^N$ is positively invariant and contained in the domain of attraction for the equilibrium under the mean ODE corresponding to (10).

B.1. Proof of convergence of the process (11)

We first consider the truncated process

$$y_{t+1} = \Pi_H\left[y_t + a_t Y_t\right], \qquad y_1 \in \mathbb{R}_{++}^N \tag{19}$$

where $H := \{y \in \mathbb{R}^N : \mathbf{0} \leq y \leq k_H \mathbf{1}\}$ for some constant $k_H \in \mathbb{R}_+$ and $Y_t = (Y_{1,t}, \ldots, Y_{N,t})$ is such that, for all i,

$$Y_{i,t} := y_{i,t}\left(\gamma_i - I\left\{s_{i,t} < y_{i,t}\right\}\right).$$

Let $h\left(y\right) = \left(h_1(y), \ldots, h_N\left(y\right)\right)$ be such that, for all i,

$$h_i\left(y\right) := y_i\left(\gamma_i - F_{\bar{d}_i\left(y_{-i}\right)}\left(y_i\right)\right).$$

With this, we have

$$E\left[Y_t \mid \mathcal{F}_t\right] = h\left(y_t\right)$$

where $\mathcal{F}_t$ is the sigma-algebra generated by $(y_1, Y_1, \ldots, Y_{t-1})$. Hence, the "mean" behavior of (19) is described by the ODE2.

$$\dot{y} = h(y). \tag{20}$$

We assume that k_H is sufficiently large so that H is invariant under (20) and the unique equilibrium y^* of the one-shot inventory competition model (see Proposition 2) is in H.

A key result in the literature on the convergence of the SA processes, namely Theorem 2.1 in Chapter 5 of Kushner and Yin (2003), states that the conditions below are sufficient for the convergence of the process (19) to y^* with probability one.

1. $\sup_t E\left[\|Y_t\|_2^2\right] < \infty$
2. $E[Y_t \mid \mathcal{F}_t] = h(y_t)$ for a continuous function h
3. $a_t \geq 0$, for all t, $\sum_t a_t = \infty$, $\sum_t a_t^2 = \infty$
4. $y_t \in H \cap \mathbb{R}^N_{++}$, for all $t \geq 2$
5. $y(\cdot) \equiv y^*$ is an asymptotically stable solution of the mean ODE (20) with the domain of attraction $\mathbb{R}^N_{++}$.

It is clear that conditions (1)-(4) are satisfied. Furthermore, Claim 2 in subsection B.2 shows that condition (5) is also satisfied. Therefore, if H is sufficiently large, then the process (19) converges, with probability one, to y^*.

Finally, Claim 3 in subsection B.3 implies that, for any $y_1 \in \mathbb{R}^N_{++}$,

$$\lim_{k_H \uparrow \infty} P\left[\mathbf{0} < y_t < k_H \mathbf{1}, \text{ for all } t \mid y_1\right] = 1.$$

Therefore, the process (11) also converges, with probability one, to y^*.

B.2. Asymptotic stability of the mean ODE (20)

Claim 2. *Consider the mean ODE (20), $\dot{y} = h(y)$.*

1. *For any initial value $y(0) \in \mathbb{R}^N$, a solution $y(\cdot) : [0, \infty) \to \mathbb{R}^N$ exists. Moreover, any solution $y(\cdot)$ satisfies*

$$\left(y(t_0) \in \mathbb{R}^N_{++}, \text{ for any } t_0 \geq 0\right) \quad \Rightarrow \quad \left(y(t) \in \mathbb{R}^N_{++}, \text{ for all } t \geq t_0\right). \tag{21}$$

2. *$y(\cdot) \equiv y^*$, is an asymptotically stable solution (in the sense of Liapunov) with the domain of attraction R^N_{++}.*

Proof.

1. The first statement follows from the continuity of h and the fact that h grows at most linearly. The second statement follows from the fact that there exists a $y^0 \in \mathbb{R}^N_{++}$, such that, for all $y \in \mathbb{R}^N$,

$$y_i \leq y_i^0 \quad \Rightarrow \quad h_i(y) \geq y_i \gamma_i / 2. \tag{22}$$

2. Since $h(y^*) = \mathbf{0}$, $y(\cdot) \equiv y^*$ is a solution.

Fix $\epsilon \in (0,1)$. Consider the following recursions, for $\ell \geq 1$,

$$\begin{aligned} y^{\ell+1} &= BR\left(y^{\ell}\right) \\ y^{\ell+1,\epsilon} &= BR\left(y^{\ell,\epsilon}\right) - (-\epsilon)^{\ell}\, \mathbf{1} \end{aligned}$$

where $y^1 = y^{1,\epsilon} = \mathbf{0}$. Assume that $\epsilon > 0$ is sufficiently small so that $y^{\ell,\epsilon} \in \mathbb{R}^N_{++}$, for all $\ell \geq 2$.

Then, from the monotonicity and the contractiveness of BR, we have

$$\mathbf{0} = y^1 < y^3 \leq y^5 \leq \cdots \leq y^* \leq \cdots \leq y^6 \leq y^4 \leq y^2 \tag{23}$$

$$\mathbf{0} = y^{1,\epsilon} < y^{3,\epsilon} \leq y^{5,\epsilon} \leq \cdots \leq y^* \leq \cdots \leq y^{6,\epsilon} \leq y^{4,\epsilon} \leq y^{2,\epsilon} \tag{24}$$

$$y^{\ell,\epsilon} \leq y^{\ell} - \epsilon^{\ell-1}\mathbf{1}, \qquad \text{for all odd } \ell \geq 3 \tag{25}$$

$$y^{\ell} + \epsilon^{\ell-1}\mathbf{1} \leq y^{\ell,\epsilon}, \qquad \text{for all even } \ell \geq 2 \tag{26}$$

$$\lim_{\ell} y^{\ell} = y^* \tag{27}$$

$$\left\| y^{\ell} - y^{\ell,\epsilon} \right\|_{\mathbb{R}^N} \leq \tfrac{\epsilon}{1-\epsilon} \left\| \mathbf{1} \right\|_{\mathbb{R}^N}, \qquad \text{for all } \ell \geq 1 \tag{28}$$

Next, there exist scalars $\rho_1, \rho_2, \dots > 0$ such that, for all $y \in \mathbb{R}^N$ and $\ell \geq 1$,

$$\text{if } \ell \text{ is odd}, \quad \left(y^{\ell,\epsilon} \leq y \text{ and } y_i^{\ell+1,\epsilon} \leq y_i\right) \quad \Rightarrow \quad h_i(y) \leq -\rho_\ell y_i \tag{29}$$

$$\text{if } \ell \text{ is even}, \quad \left(y \leq y^{\ell,\epsilon} \text{ and } y_i \leq y_i^{\ell+1,\epsilon}\right) \quad \Rightarrow \quad h_i(y) \geq \rho_\ell y_i. \tag{30}$$

To see this, let $\ell \geq 1$ be odd and $\left(y^{\ell,\epsilon} \leq y \text{ and } y_i^{\ell+1,\epsilon} \leq y_i\right)$. Then

$$\gamma_i - F_{\bar{d}_i\left(y_{-i}\right)}(y_i) \leq \gamma_i - F_{\bar{d}_i\left(y_{-i}^{\ell,\epsilon}\right)}\left(y_i^{\ell+1,\epsilon}\right) = \gamma_i - F_{\bar{d}_i\left(y_{-i}^{\ell,\epsilon}\right)}\left(BR_i\left(y_{-i}^{\ell,\epsilon}\right) + \epsilon^{\ell}\right). \tag{31}$$

Since $BR_i\left(y_{-i}^{\ell,\epsilon}\right)$ is the unique solution to $\gamma_i - F_{\bar{d}_i\left(y_{-i}^{\ell,\epsilon}\right)}(\cdot) = 0$, the RHS of (31) is upper bounded by $-\rho_\ell$ for some $\rho_\ell > 0$. The case of even ℓ follows similarly.

Now, let

$$B_0^\epsilon := \mathbb{R}_{++}^N$$

$$B_\ell^\epsilon := \left\{y \in \mathbb{R}_{++}^N : y^{\ell,\epsilon} \le y \le y^{\ell+1,\epsilon}\right\}, \qquad \text{for all odd } \ell \ge 1$$

$$B_\ell^\epsilon := \left\{y \in \mathbb{R}_{++}^N : y^{\ell+1,\epsilon} \le y \le y^{\ell,\epsilon}\right\}, \qquad \text{for all even } \ell \ge 2.$$

The sets B_ℓ^ϵ, for all $\ell \ge 0$, are positively invariant under (20), i.e., any solution $y(\cdot)$ of (20) satisfies

$$\left(y(t_0) \in B_\ell^\epsilon, \text{ for any } t_0 \ge 0\right) \quad \Rightarrow \quad \left(y(t) \in B_\ell^\epsilon, \text{ for all } t \ge t_0\right).$$

The invariance of B_0^ϵ follows from (21). To show the invariance of B_ℓ^ϵ for $\ell \ge 1$, assume that ℓ is odd and $y(t_0) \in B_\ell^\epsilon$. Let $t_1 \ge t_0$ be the first time at which $y(\cdot)$ hits the boundary of B_ℓ^ϵ. If $y_i(t_1) = y_i^{\ell+1,\epsilon}$, then $\dot{y}_i(t_1) < 0$ from (29). If $y_i(t_1) = y_i^{\ell,\epsilon}$ (which can happen only for $\ell > 1$), then $\dot{y}_i(t_1) > 0$ from (30). The case of even ℓ is handled similarly.

From (27)-(28), given $\epsilon_1 > 0$, if $\epsilon > 0$ and $\ell \ge 1$ are sufficiently small and large, respectively, then $\sup_{y \in B_\ell^\epsilon} \|y - y^*\|_{\mathbb{R}^N} < \epsilon_1$. Also, from (23)-(26), for any $\ell \ge 1$, B_ℓ^ϵ contains a neighborhood of y^*, i.e., $\{y \in \mathbb{R}^N : \|y - y^*\|_{\mathbb{R}^N} \le \delta_1\} \subset B_\ell^\epsilon$, for some $\delta_1 < 0$. Therefore, given any solution $y(\cdot)$ of (20) and any $\epsilon_1 > 0$, there exists a $\delta_1 > 0$ such that

$$\|y(0) - y^*\|_{\mathbb{R}^N} \le \delta_1 \qquad \Rightarrow \quad \sup_{t \ge 0} \|y(t) - y^*\|_{\mathbb{R}^N} \le \epsilon_1.$$

This shows the stability in the sense of Liapunov.

Finally, (22)-(30) also imply that, for any solution $y(\cdot)$ of (20) with $y(0) \in \mathbb{R}_{++}^N$, there exist finite scalars $T_0^\epsilon(y(0)), T_1^\epsilon, T_2^\epsilon, \dots \ge 0$ such that

$$\left(y(t_0) \in B_\ell^\epsilon, \text{ for any } t_0 \ge 0, \ \ell \ge 0\right) \quad \Rightarrow \quad \left(y(t) \in B_{\ell+1}^\epsilon, \text{ for all } t \ge t_0 + T_\ell^\epsilon\right).$$

Therefore,

$$y(0) \in \mathbb{R}_{++}^N \quad \Rightarrow \quad \lim_t y(t) = y^*.$$

B.3. Properties of the truncated process (19)

CLAIM 3. *Consider the process (19) as $k_H \uparrow \infty$. For any $y_1 \in \mathbb{R}^N_{++}$,*

$$\lim_{k_H \uparrow \infty} P\left[\sup_{i,t} y_{i,t} = k_H \;\middle|\; y_1\right] = 0.$$

Proof. Fix $\epsilon \in (0,1)$ and let $y^{2,\epsilon}$ be as in the proof of Claim 2. Let

$$U(y) := \sum_i y_i \phi\left(y_i - y_i^{2,\epsilon}\right)$$

where $\phi : \mathbb{R} \to \mathbb{R}_+$ is a twice continuously differentiable function satisfying

$$\phi(x) = \begin{cases} 0 & x \le 0 \\ \text{increasing} & 0 \le x \le 1 \\ 1 & x \ge 1 \end{cases}$$

For example, $\phi(x)$ can be taken as $6x^5 - 15x^4 + 10x^3$ for $0 \le x \le 1$. Note that $U(y)$ is twice continuously differentiable and approximates $\sum_i y_i I\left\{y_i \ge y_i^{2,\epsilon} + 1\right\}$. From Taylor's theorem, we have

$$\begin{aligned} E\left[U(y_{t+1}) - U(y_t) \mid \mathcal{F}_t\right] &\le E\left[U(y_t + a_t Y_t) - U(y_t) \mid \mathcal{F}_t\right] \\ &= a_t \left(\nabla U(y_t)\right)^T h(y_t) + \frac{1}{2} a_t^2 E\left[(Y_t)^T \nabla^2 U(\bar{y}) Y_t \mid \mathcal{F}_t\right] \end{aligned}$$

where $\bar{y}$ is on the line segment joining y_t and $y_t + a_t Y_t$. Since $\nabla^2 U(y)$ vanishes unless $y^{2,\epsilon} \le y \le y^{2,\epsilon} + \mathbf{1}$, there exists a constant $k_1 \in \mathbb{R}_+$, independent of k_H, such that

$$E\left[U(y_{t+1}) - U(y_t) \mid \mathcal{F}_t\right] \le a_t \left(\nabla U(y_t)\right)^T h(y_t) + k_1 a_t^2. \tag{32}$$

Now, let $V_t(y) := (1 + U(y)) \prod_{s=t}^{\infty} (1 + k_1 a_s^2)$. It follows from (32) that

$$E\left[V_{t+1}(y_{t+1}) - V_t(y_t) \mid \mathcal{F}_t\right] \le a_t \left(\nabla U(y_t)\right)^T h(y_t).$$

Since $\left(\nabla U(y_t)\right)^T h(y_t) \le 0$ and $V_1(y_1) < \infty$, $(V_t(y_t), \mathcal{F}_t)$ is a supermartingale; see Theorem 2.2 in Chapter II of Nevel'son and Has'minskiĭ (1973). A well-known inequality for nonnegative supermartingales (see (1.6) on page 98 of Kushner and Yin (2003)) implies

$$P\left[\sup_t V_t(y_t) \ge k_H \;\middle|\; y_1\right] \le \frac{V_1(y_1)}{k_H}$$

which leads us to the desired result.

Acknowledgments

This work was supported by NSF Grant #ECCS-0547692.

References

Avsar, Z. M., M. Baykal-Gursoy. 2002. Inventory control under substitutable demand: A stochastic game application. *Naval Research Logistics* **49**(4) 359–375.

Başar, T., G. J. Olsder. 1999. *Dynamic Noncooperative Game Theory*. SIAM, Philadelphia, PA.

Benaïm, M. 1999. Dynamics of stochastic approximation algorithms. Jacques Azéma, Michel Émery, Michel Ledoux, Marc Yor, eds., *Séminaire de Probabilités XXXIII*, *Lecture Notes in Mathematics*, vol. 1709. Springer Berlin / Heidelberg, 1–68. URL `http://dx.doi.org/10.1007/BFb0096509`.

Benveniste, A., M. Métivier, P. Priouret. 1990. *Adaptive algorithms and stochastic approximations*. Springer-Verlag, New York, NY.

Bernstein, F., A. Federgruen. 2004. Dynamic inventory and pricing models for competing retailers. *Naval Research Logistics* **51**(2) 258–274.

Burnetas, A. N., C. E. Smith. 2000. Adaptive ordering and pricing for perishable products. *Operations Research* **48**(3) 436–443.

Edelstein, M. 1962. On fixed points and periodic points under contractive mappings. *Journal of the London Mathematical Society* **37**(1) 74–79.

Flam, S.D. 2002. Equilibrium, evolutionary stability and gradient dynamics. *International Game Theory Review* **4**(4) 357–370.

Foster, D.P., H.P. Young. 79–96. On the nonconvergence of fictitious play in coordination games. *Games and Economic Behavior* **25**.

Fudenberg, D., D.K. Levine. 1998. *The Theory of Learning in Games*. MIT Press, Cambridge, MA.

Fudenberg, D., J. Tirole. 1991. *Game Theory*. MIT Press, Cambridge, MA.

Hart, S., A. Mas-Colell. 2003. Uncoupled dynamics do not lead to Nash equilibrium. *American Economic Review* **93**(5) 1830–1836.

Hofbauer, J., K. Sigmund. 1998. *Evolutionary Games and Population Dynamics*. Cambridge University Press, Cambridge, UK.

Jordan, J. 1993. Three problems in game theory. *Games and Economic Behavior* **5** 368–386.

Karjalainen, R. 1992. The newsboy game. Tech. rep., Wharton School, University of Pennsylvania.

Kushner, H. J., G. G. Yin. 2003. *Stochastic approximation and recursive algorithms and applications*. 2nd ed. Springer-Verlag, New York, NY.

Lippman, S. A., K. F. McCardle. 1997. The competitive newsboy. *Operations Research* **45**(1) 54–65.

Mahajan, S., G. van Ryzin. 2001. Inventory competition under dynamic consumer choice. *Operations Research* **49**(5) 646–657.

Netessine, S., N. Rudi. 2003. Centralized and competitive inventory models with demand substitution. *Operations Research* **51**(2) 329–335.

Netessine, S., N. Rudi, Y. Wang. 2006. Inventory competition and incentives to back-order. *IIE Transactions* **38**(11) 883–902.

Nevel'son, M. B., R. Z. Has'minskiĭ. 1973. *Stochastic approximation and recursive estimation*. Translations of mathematical monographs, Volume 47, American Mathematical Society, Providence, Rhode Island.

Olsen, T. L., R. P. Parker. 2008. Inventory management under market size dynamics. *Management Science* **54**(10) 1805–1821.

Parlar, M. 1988. Game theoretic analysis of the substitutable product inventory problem with random demands. *Naval Research Logistics* **35**(3) 397–409.

Parlar, M., Q. Wang. 1994. A three-person game theory model arising in stochastic inventory control theory. *European Journal of Operational Research* **76**(1) 83–97.

Rosen, J. B. 1965. Existence and uniqueness of equilibrium points for concave $n-$person games. *Econometrica* **33**(3) 520–534.

Sato, S., E. Akiyama, J.D. Farmer. 2002. Chaos in learning a simple two person game. *Proceedings of the National Academy of Sciences*, vol. 99(7). 4748–4751.

Shapley, L.S. 1964. Some topics in two-person games. L.S. Shapley M. Dresher, A.W. Tucker, eds., *Advances in Game Theory*. University Press, Princeton, NJ, 1–29.

Weibull, J.W. 1995. *Evolutionary Game Theory*. MIT Press, Cambridge, MA.

Young, H. P. 1998. *Individual Strategy and Social Structure*. Princeton University Press, Princeton, NJ.

Young, H. P. 2004. *Strategic Learning and Its Limits*. SIAM, Oxford University Press Inc.