Professor Julio B. CLEMPNER, PhD
E-mail: julio@clempner.name
Instituto Politécnico Nacional, Mexico
Professor Alexander S. POZNYAK, PhD
E-mail: apoznyak@ctrl.cinvestav.mx
Center for Research and Advanced Studies

ANALYZING AN OPTIMISTIC ATTITUDE FOR THE LEADER FIRM IN DUOPOLY MODELS: A STRONG STACKELBERG EQUILIBRIUM BASED ON A LYAPUNOV GAME THEORY APPROACH

***Abstract**. This paper presents a novel game theory approach for representing the Stackelberg dynamic duopoly models. The problem is fitted into a class of ergodic controllable finite Markov chains game. It is consider the case where a strong Stackelberg equilibrium point is convenient for both firms. We first analyze the best-reply dynamics of the Stackelberg duopoly model such that each firm's best-reply strategies set agree with the set of maximizers of the opponents' strategies. Then, the classical Stackelberg duopoly game is transformed into a potential game in terms of the Lyapunov theory. As a result, a duopoly model has the benefit that the best-reply dynamics results in a natural implementation of the behavior of a Lyapunov-like function. In addition, the strong equilibrium point properties of Stackelberg and Lyapunov meet in potential games. We validate the proposed method theoretically by computing the complexity and by a numerical experiment related to the duopoly model.*

***Keywords**: Dynamic duopoly model, complexity, strong Stackelberg equilibrium, Lyapunov equilibrium, Stackelberg games, Lyapunov games, Markov decision process*

JEL Classification: C72, C73, D43, E37, C32

1. Introduction

1.1 Brief review

Game theory is an increasingly important paradigm for reasoning about complex duopoly/oligopolies problems. For pioneering and innovative works on duopoly/oligopolies, see (Bös, 1991; Harris and Wiens, 1980; Ghadimi et al., 2013) and surveys can be found in (Breitmoser, 2012; De Fraja and Delbono, 1990). Cournot's duopoly model (Cournot, 1938) of quantity competition was modified by von Stackelberg (von Stackelberg, 1934), who represented the commitment that the leaders firm make to a leader policy, and the capability of the followers firms to learn about the policy during the planning phase of the game. This is, for example, the case of a mixed oligopoly market (Kalashnikova et al., 2010) there is at least one private and one public firm. In this context, private firms are assumed to be profit maximizers and in some cases they have perfect knowledge and an optimal follower behavior. While the most prominent assumption is that a public firm maximizes social welfare (consumer plus producer surplus). In the paper by Matsumura (Matsumura, 2003), the author assumes a mixed duopoly and analyses a desirable role of leader or follower for a public firm. Matsumura found that the role of the public firm should be that of the leader.

A standard way to interpret Stackelberg equilibrium is to see it as a Subgame Perfect Equilibrium (SPE) (Selten, 1965). The concept of Stackelberg strategy can be extended to allow for a non-unique "rational" response of the follower. Then, the choice of the best-reply strategy from the leader's point of view can have two different approaches: to assume a weak Stackelberg equilibrium which leads to a pessimistic attitude for the leader) or, the strong Stackelberg equilibrium (SSE) which leads to an optimistic approach (Morgan and Patrone, 2006). We are considering the case where the SPE is a SSE. Many works analyzed Stackelberg models focusing primarily on computing Strong Stackelberg Equilibrium, which forces the follower to break ties optimally for the leader selecting from its best-reply strategy set the option that maximizes the utility of the leader.

The notion of potential games were introduced by Monderer and Shapley (Monderer and Shapley, 1996) whereby the information about Nash equilibria is nested into a single real-valued function (the potential function) over the strategy space. Several definitions of potential games have been presented in the literature. For instance, Voorneveld (Voorneveld, 2000) introduced the best-reply potential games having the distinctive feature that it allows infinite improvement paths, by imposing restrictions only on paths in which players that can improve actually deviate to a best-reply. Clempner and Poznyak (Clempner and Poznyak, 2011) proved that the Lyapunov equilibrium point coincides with the Nash equilibrium point under certain conditions. Clempner and Poznyak (Clempner and Poznyak, 2013) showed that in the ergodic class of finite controllable Markov chains dynamic games the best reply actions lead obligatory to one of Nash equilibrium points. This conclusion is done by the Lyapunov Games concept which is based on the designing of an individual Lyapunov function (related with an individual cost function) which monotonically decreases (non-increases) during the game.

A result related to Cournot and potential games was presented by Clempner (Clempner, 2015) proving that the stability conditions and the equilibrium point properties of Cournot and Lyapunov meet in potential games. Dragone et al. (Dragone et al., 2012) proved that the Cournot oligopoly game with non-linear market demand can be reformulated as a best-reply potential game where the best-reply potential function is linear-quadratic in the special case where marginal cost is normalized to zero. Also, Dragone et al. (Dragone et al., 2008) identified the conditions for the existence of a best-response potential function and characterize its construction, describing then the key properties of the equilibrium presenting applications to oligopoly. Amir and Grilo (Amir and Grilo, 1996) provided different sets of minimal conditions, directly on the demand and cost functions, yielding respectively the simultaneous and the two sequential modes of play.

1.2 Motivating example

In this scenario, we will suppose that the market for a certain product is dominated by two firms, one plays a dominant role. An example might be Volkswagen, at times big enough. Followers are: Fiat, Peugeot, etc.Segments are established based on comparison to well-known brand models. For instance, we will focus on cars classified in four segments: 1) Large – cars are greater speed, capacity and occupant protection are safer designed, 2) Medium - cars are drawn with a sedan shape designed to seat four to six passengers comfortably 3) Small - cars that refer to the hatchbacks and shortest saloons marketed at low price, and 4) Mini – cars is limited to approx. 3,700 millimeters.

The companies can choose to produce a certain quantity of the product depending on the segment: Large, Medium, Small and Mini. Quantities will be denoted by (L, M, S, I) for producer 1 and (l, m, s,i) for producer 2. Actions taken by the producer depend on the segment and are determined by: high, medium, low. The market price of the product decreases with increasing total quantity produced by both producers. If the companies decide to produce a high quantity of the product, the price collapses so that profits drop to zero. Both companies know how increased production lowers product price and their profits. The utilities are as follows

Co.I\ Co.II	**l**arge	**m**edium	**s**mall	**(i)m**ini
Large	20,13	23,15	**36,17**	62,10
Medium	15,23	32,32	40,30	64,20
Small	18,36	30,40	36,36	54,20
(**I**)Mini	19,52	14,34	16,48	5,10

The leader-follower company dynamic version of the game is as follows: Company I plays first and optimal commit to a level of production given by a row.

It follows that by backward induction strategy of Company II will respond to **L** by **s**, to **M** by **m**, to **L** also by **m**, and to **I** by **l**. By looking ahead these anticipated best-reply by Company II, Company I does optimal play **L**, thehighest level of production. Then, the result is that Company I makes a profit 36 (in contrast to the 32 in the simultaneous-choice game). When Company II must play the role of the follower, its best-reply profits fall from 32 (in the simultaneous-choice game) to 17. As a result, the Stackelberg equilibrium of the game is (**L**,**s**).

1.3 Main results

This paper presents a new game theory approach for representing the Stackelberg duopoly model.

- We consider the case where a SSE point is convenient for both firms.
- The problem is fitted into a class of discrete-time, ergodic, controllable and finite Markov chains games.
- We restrict the attention of the game to Markov pure and stationary fixed-local-optimal strategies (Clempner and Poznyak, 2014) for the leader and the follower that agree with the equilibrium point.
- In order to represent the game with a strong Stackelberg equilibrium point, we propose a non-converging state-value function that fluctuates (increases and decreases) between states of the stochastic game.
- We prove that it is possible to represent that function in a recursive format using a one-step-ahead fixed-optimal policy.
- The method looks for an optimal strategy of the leader firm that coincides with the best-reply strategy, finding the highest reward strategies and exploiting the advantage of being the leader firm.
- Then, we present a method for constructing a Lyapunov-like function that explains the behavior of players in a repeated stochastic Markov chain game.
- The Lyapunov-like function replaces the recursive mechanism with the elements of the ergodic system that model how players are likely to behave in one-shot games.
- Then, the classical Stackelberg duopoly game is transformed into a potential game in terms of the Lyapunov theory.
- As a result, a duopoly model has the benefit that the best-reply dynamics results in a natural implementation of the behavior of a Lyapunov-like function.
- In addition, the strong equilibrium point properties of Stackelberg and Lyapunov meet in potential games.

1.4 Organization of the paper

The paper is organized as follows. The next section contains the formal definition of the oligopoly model considered and the mathematical background needed to understand the rest of the paper. In Section 3, we present a game theory approach for representing the Stackelberg duopoly model. Section 4 contains the theoretical results showing that the best-reply dynamics results in a natural implementation of the behavior of a Lyapunov-like function and that the Lyapunov equilibrium point coincides with the SSE point, which we consider to be the main contribution of this paper. Section 5 shows the experimental results related to the Stackelberg duopoly model. Finally, in Section 6 some concluding remarks are outlined.

2. Background and related work

The aim of this section is to introduce the duopoly model and all the structural assumptions related with the Markov model (Poznyak et al., 2000).

2.1 Basics

Firms are looking for maximum benefits. The benefits are derived from both maximum sales volume (a larger share of the market) and higher prices (higher profitability). The problem is originated by the fact that increasing profitability through higher prices can reduce the revenue by losing market share. What Cournot's approach (Cournot, 1938) does is to maximize both market share and profitability by defining optimum prices. Stackelberg games (von Stackelberg, 1934) draw attention to the fact of having truthful market information when defining a strategy, and the interdependence of each firm's strategies, when there is a market leader firm.

The dynamic of the game is described as follows. We consider $\mathcal{N} = \{1,2\}$ firms in the industry. Firms make the same homogeneous and undifferentiated product and choose a quantity to produce independently and sequentially. The single homogeneous product is sold in the market with a constant marginal cost. The firms do not collude, and they seek to maximize their profit based on their competitors' decisions. Each firm's output decision is assumed to affect the product price.

The game has a finite set S, called the *state space,* consisting of all positive integers $N \in \mathbb{N}$ of states $\{s(1),..., s(N)\}$ and it begins at the initial state $s(1)$ which (as well as the states further realized by the process). All the states are organized in a stationary *Markov chain*, which is a sequence of S-valued random variables s_n, $n \in \mathbb{N}$, satisfying the *Markov condition*:

$$p\left(s_{n+1} = s(j) \mid s_n = s(i)\right) = \pi(ij) \qquad (1)$$

The *Markov chain* can be represented by a complete graph whose nodes are the states, where each edge $(s(i), s(j)) \in S^2$ is labeled by the transition probability (1). The matrix $\Pi = (\pi(ij))_{(s(i),s(j))\in S} \in 0,1]^{S\times S}$ determines the evolution of the chain: for each $k \in \mathbb{N}$, the power Π^k has in each entry $(s(i), s(j))$ the probability of going from state $s(i)$ to state $s(j)$ in exactly k steps.

A *Markov Decision Process* is a 5-tuple

$$MDP = \{S, A, \mathrm{K}, \Pi, U\} \tag{2}$$

where: 1) S is a finite set of states, $S \subset \mathbb{N}$; 2) A is the set of actions. For each $s \in S$, $A(s) \subset A$ is the non-empty set of admissible actions at state $s \in S$. Without loss of generality we may take $A = \cup_{s \in S} A(s)$; 3) $\mathrm{K} = \{(s,a) \mid s \in S, a \in A(s)\}$ is the set of admissible state-action pairs, which is a finite subset of $S \times A$; 4) $\Pi(k) = [\pi(ij \mid k)]$ is a stationary transition controlled matrix, where

$$\pi(ij \mid k) \equiv p(s_{n+1} = s(j) \mid s_n = s(i), a_n = a(k)) \tag{3}$$

representing the probability associated with the transition from state $s(i)$ to state $s(j)$ under an action $a(k) \in A(s(i))$, $k = 1, \ldots, M$ and; 5) $U : S \times \mathrm{K} \to \mathbb{R}$ is a utility function, associating to each state a real value.

The *Markov property* of the decision process in (2) is said to be fulfilled if

$$p(s_{n+1} \mid (s_1, s_2, ..., s_{n-1}), s_n = s(i), a_n = a(k)) = p(s_{n+1} \mid s_n = s(i), a_n = a(k))$$

Each firm ι is allowed to randomize, with distribution $d_n^{(\iota)}(k_\iota \mid i_\iota)$, over the pure action choices $a^\iota(k) \in A^\iota(s^\iota(i))$, $i = \overline{1,N}$ and $k = \overline{1,M}$. Formally, the strategy (policy) $d_n^{(\iota)}(k_\iota \mid i_\iota) \equiv p^{(\iota)}(a^{(\iota)}(k) \mid s_n^{(\iota)} = s(i_\iota))$ which represents the probability measure associated with the occurrence of an action $a_n^{(\iota)}$ from state $s_n^{(\iota)} = s(i_\iota)$.

For a strategy tuple $d = (d^{(1)}, \ldots, d^{(|\mathcal{N}|)}) \in \Delta$ we denote the complement strategy $d^{(-\iota)} = (d^{(1)}, \ldots, d^{(\iota-1)}, d^{(\iota+1)}, \ldots, d^{(|\mathcal{N}|)})$ and, with an abuse of notation, $d = (d^{(\iota)}, d^{(-\iota)})$. The state $d = (d^{(1)}, \ldots, d^{(|\mathcal{N}|)})$ represents the distribution vector of strategy frequencies and can only move on Δ.

$$p(s_{n+1}^{(\iota)} = s(j_\iota) \mid s_n^{(\iota)} = s(i_\iota)) = \sum_{k_\iota = 1}^{M_\iota} p(s_{n+1}^{(\iota)} = s(j_\iota) \mid s_n^{(\iota)} = s(i_\iota), a_k^{(\iota)} = a(k_\iota)) \cdot d_n^{(\iota)}(k_\iota \mid i_\iota)$$

Let us denote the collection $\{d_n(k \mid i)\}$ by Δ_n as follows: $\Delta_n = \{d_n^{(\iota)}(k_\iota \mid i_\iota)\}_{k=\overline{1,M},\ i=\overline{1,N}}$. A *policy* $\{\Delta_n^{loc}\}_{n \geq 0}$ is said to be *local optimal* if for each $n \geq 0$ it minimizes the conditional mathematical expectation of the utility-function $U^{(\iota)}(s_{n+1})$ under the condition that the prehistory of the process

$$F_n := \{\Delta_0, P\{s_0 = s(j)\}_{j=\overline{1,N}}; \ldots; \Delta_{n-1}, P\{s_n = s(j)\}_{j=\overline{1,N}}$$

is fixed and cannot be changed hereafter, i.e., it realizes the " one-step ahead" conditional optimization rule $\Delta_n^{loc} := \arg\max_{d_n \in \Delta_n} \mathrm{E}\{U^{(\iota)}(s_{n+1}) \mid F_n\}$ where $U^{(\iota)}(s_{n+1})$ is the utility function of the player ι at the state s_{n+1}. Locally optimal policies are known as a "myopic" policies in games literature.

A *non-cooperative game* is a tuple $\mathcal{G} = \left\langle \mathcal{N}, S, \left(\Delta^{\iota}\right)_{\iota\in\mathcal{N}}, \left(\Upsilon^{\iota}\right)_{\iota\in\mathcal{N}}, \Pi, \left(U^{\iota}\right)_{\iota\in\mathcal{N}} \right\rangle$. Applying the conditional transition probability matrix given in (3) changes the conditional probability distribution at time n from $p_n^{(\iota)}(a^{(\iota)}(k) \mid s_n^{(\iota)} = s(i))$ to $p_{n+1}^{(\iota)}(a^{(\iota)}(k) \mid s_{n+1}^{(\iota)} = s(i))$. The probability of the player $\iota \in \mathcal{N}$ in the game G to find itself in the next state is as follows:

$$p_n^{(\iota)}(s_{n+1}^{(\iota)} = s(j_\iota) \mid s_n^{(\iota)} = s(i_\iota)) = \sum_{k_\iota=1}^{M_\iota} p_n^{(\iota)}(s_{n+1}^{(\iota)} = s(j_\iota) \mid s_n^{(\iota)} = s(i_\iota), a^{(\iota)}(k_\iota)) d_n^{(\iota)}(k_\iota \mid i_\iota) =$$

$$\sum_{k_\iota=1}^{M_\iota} \pi_n^{(\iota)}(i_\iota j_\iota \mid k_\iota) d_n^{(\iota)}(k_\iota \mid i_\iota)$$

In the ergodic case when the Markov chain is ergodic for any stationary strategy $d_n^{(\iota)}(k_\iota \mid i_\iota)$ the distributions $p_n^{(\iota)}(s_{n+1}^{(\iota)} = s(j_\iota))$ exponentially fast converge to their limits $p^{(\iota)}(s^{(\iota)} = s(i_\iota))$ satisfying

$$p^{(\iota)}(s^{(\iota)} = s(j_\iota)) = \sum_{i_\iota=1}^{N_\iota} \sum_{k_\iota=1}^{M_\iota} \pi_n^{(\iota)}(i_\iota j_\iota \mid k_\iota) d_n^{(\iota)}(k_\iota \mid i_\iota) p^{(\iota)}(s^{(\iota)} = s(i_\iota))$$

The market determines the price at which an output is sold by $\mathcal{P}^{(\iota)} : S \times \mathrm{K} \to \mathrm{R}$, where $\mathcal{P}^{(\iota)}(i_\iota, j_\iota \mid k_\iota)$ is commonly-known non-increasing function of total output and denotes the inverse demand function. Cost to firm ι of producing units is given by $C^{(\iota)} : S \times \mathrm{K} \to \mathrm{R}$, where $C^{(\iota)}(i_\iota, j_\iota \mid k_\iota)$ is nonnegative and nondecreasing in S (the cost functions are assumed to be the same for all firms). Profit of firm ι is given by

$$U^{(\iota)}(i_\iota, j_\iota, k_\iota) = \mathcal{P}^{(\iota)}(i_\iota, j_\iota \mid k_\iota) \mathcal{Q}^{(\iota)}(i_\iota, j_\iota \mid k_\iota) - C^{(\iota)}(i_\iota, j_\iota \mid k_\iota)$$

and if

$$\mathcal{R}^{(\iota)}(i_\iota, j_\iota \mid k_\iota) = \mathcal{P}^{(\iota)}(i_\iota, j_\iota, k_\iota) \mathcal{Q}^{(\iota)}(i_\iota, j_\iota \mid k_\iota)$$

we have that

$$U^{(\iota)}(i_\iota, j_\iota, k_\iota) = \mathcal{R}^{(\iota)}(i_\iota, j_\iota, \mid k_\iota) - C^{(\iota)}(i_\iota, j_\iota \mid k_\iota)$$

The *utility function* $\mathbf{U}^{(\iota)}$ of any fixed policy $d^{(\iota)}$ is defined over all possible combinations of states and actions, and indicates the expected value when taking action $a^{(\iota)}$ in state $s^{(\iota)}$ and following policy $d^{(\iota)}$ thereafter. In this sense, the average utility function at the state s for the firm ι, namely, $\mathbf{U}^{(\iota)}(d^{(\iota)}) := \mathrm{E}(U^{(\iota)}(s\ ) \mid d^{(\iota)})$ where $U^{(\iota)}(s\ )$ is the utility function of the ι-firm at the state s and, $\mathrm{E}(\cdot \mid d^{(\iota)})$ is the operator of the conditional mathematical expectation subject to the constraint when the mixed strategy $d^{(\iota)}$ has been applied.

2.2 Parameters of the Model

In a Stackelberg game the leader considers the best-reply of the follower, i.e. how it will respond once it has observed the "optimal" strategy of the follower. We assume that the number of agents in the model is two ($\mathcal{N} = \{1,2\}$). Player ι should be understood to refer to players in a general expression of the game**.** For the remainder of this paper, the leader firm is designated as player (1) and the follower firm as player (2). Let introduce the notations for the model.

Parameters of the firms ι : N_ι - The number of states of firm ι ; M_ι - The number of actions of firm ι ; $\pi^{(\iota)}(i_\iota j_\iota \mid k_\iota)$ - The elements of the transition matrix of firm ι ; $\mathcal{P}^{(\iota)}(i_\iota, j_\iota \mid k_\iota)$ - The elements of the price matrix of firm ι ; $\mathcal{Q}^{(\iota)}(i_\iota, j_\iota \mid k_\iota)$ - The quantity of elements of firm ι ; $C^{(\iota)}(i_\iota, j_\iota \mid k_\iota)$ - The elements of the cost matrix of firm ι ; $d^{(\iota)}(k_\iota \mid i_\iota)$ - The probability for a firm ι to apply action k_ι at state i_ι .

Then, the U -values of the leader firm can be expressed by

$$\mathbf{U}^{(1)}(c^{(1)}_{i_1,k_1}, c^{(2)}_{i_2,k_2}) = \sum_{i_1=1}^{N_1}\sum_{k_1=1}^{M_1}\sum_{i_2=1}^{N_2}\sum_{k_2=1}^{M_2} W^{(1)}(i_1,k_1,i_2,k_2)c^{(1)}(i_1 \mid k_1)c^{(2)}(i_2 \mid k_2) \tag{4}$$

where

$$W^{(1)}(i_1,k_1,i_2,k_2) = \sum_{j_1=1}^{N_1}\sum_{j_2=1}^{N_2} \mathcal{R}^{(1)}(i_1, j_1, k_1, i_2, j_2, k_2) - C^{(1)}(i_1, j_1 \mid k_1)$$

and

$$c^{(1)}(i_1 \mid k_1) = d^{(1)}(k_1 \mid i_1) p^{(1)}(i_1) \ c^{(2)}(i_2 \mid k_2) = d^{(2)}(k_2 \mid i_2) p^{(2)}(i_2)$$

as well as for the follower firm we have

$$\mathbf{U}^{(2)}(c^{(1)}_{i_1,k_1}, c^{(2)}_{i_2,k_2}) = \sum_{i_1=1}^{N_1}\sum_{k_1=1}^{M_1}\sum_{i_2=1}^{N_2}\sum_{k_2=1}^{M_2} W^{(2)}(i_1,k_1,i_2,k_2)c^{(1)}(i_1,k_1)c^{(2)}(i_2,k_2)$$

and the variable $c^{(\iota)}(i_\iota \mid k_\iota)$ is restricted by the following constraints:

$$c^{(\iota)} \in C_{adm}^{(\iota)} := \begin{cases} \sum_{i_\iota=1}^{N_\iota}\sum_{k_\iota=1}^{M_\iota} c^\iota(i_\iota \mid k_\iota) = 1, \left[c^\iota(i_\iota \mid k_\iota)\right]: c_n^\iota(i_\iota \mid k_\iota) \geq 0, \\ \sum_{k_\iota=1}^{M_\iota} c^\iota(j_\iota \mid k_\iota) = \sum_{i_\iota=1}^{N_\iota}\sum_{k_\iota=1}^{M_\iota} \pi^\iota(i_\iota, j_\iota \mid k_\iota) c_n^\iota(i_\iota \mid k_\iota) \end{cases}$$

Loosely speaking, a necessary and sufficient condition for $c^\iota(i_\iota \mid k_\iota)$ to be a Stakelberg/Cournot equilibrium point is that the solution of the following problem

$$\mathbf{U}^\iota(c^{(1)}, c^{(2)}) \to \max_{c^{(\iota)} \in C_{adm}^{(\iota)}}$$

The solution we apply to this game is that of Stackelberg equilibrium. To think about the leader-follower equilibrium point, we first consider the nature of the firms' local best-reply strategy.

3. Stackelberg Duopoly Model

The dynamics of the Stackelberg game is as follows. The leader Company plays first considering the best-reply of the follower $d^{(2)*}$. Then, the leader commits to a mixed strategy $d^{(1)*}$ (a probability distribution over deterministic schedules) that maximizes the utility, anticipating the predicted best-reply of the follower. Then, taking into the account the adversary's mixed strategy selection, the follower in equilibrium selects the expected best-reply that maximizes the utility $d^{(2)*}$.

3.1 Local Best-reply Strategy Definition for the Follower Firm

In this situation, the leader firm looks one-step-ahead to the best-reply of the follower firm and the effect that it will have on the duopoly equilibrium.

Definition 1.*The strategy*

$$d^{(2)*}(k_2 \mid i_2) := \arg \max_{d^{(2)}(k_2 \mid i_2)} \Big(\sum_{i_1=1}^{N_1}\sum_{k_1=1}^{M_1}\sum_{j_1=1}^{N_1}\sum_{j_2=1}^{N_2} \left[(\mathcal{P}^{(2)}(i_2, j_2 \mid k_2)\mathcal{Q}^{(2)}(i_2, j_2 \mid k_2) - C^{(2)}(i_2, j_2 \mid k_2) \right]$$

$$\pi^{(1)}(i_1, j_1 \mid k_1)\pi^{(2)}(i_2 j_2 \mid k_2) d^{(2)}(k_2 \mid i_2)) d^{(1)*}(k_1 \mid i_1)$$

such that

$$\sum_{k_2=1}^{M_2} d^{(2)}(k_2 \mid i_2) = 1,\ d^{(2)}(k_2 \mid i_2) \geq 0$$

is called ***the local best-reply strategy*** *for the average follower firm.*

Lemma 1 (Clempner and Poznyak, 2013)*The local best-reply strategy* $d^{(2)*}(k_2 \mid i_2)$ *for the follower firm is pure and stationary, that is,*

$$d^{(2)*}(k_2 \mid i_2) = \begin{cases} 1 & \text{if} \quad k_2 = k_2^*(i_2) \\ 0 & \text{if} \quad k_2 \neq k_2^*(i_2) \end{cases} \tag{5}$$

$$k_2^*(i_2) = \arg\max_{k_2} \Big(\sum_{i_1=1}^{N_1} \sum_{k_1=1}^{M_1} \sum_{j_1=1}^{N_1} \sum_{i_2=1}^{N_2} \sum_{j_2=1}^{N_2} \left[\mathcal{P}^{(2)}(i_2, j_2 \mid k_2) \mathcal{Q}^{(2)}(i_2, j_2 \mid k_2) - C^{(2)}(i_2, j_2 \mid k_2) \right] \cdot \\ \pi^{(1)*}(i_1 j_1 \mid k^*{}_1) \pi^{(2)}(i_2 j_2 \mid k_2) \Big)$$

3.2 Markov Chains Dynamics for the Average Follower Firm

For the local best-reply strategy $d^{(2)*}(k_2 \mid i_2)$ we have the following Markov chain dynamics for the average follower firm:

$$p^{(2)*}(j_2) = \sum_{i_1=1}^{N_1} \sum_{i_2=1}^{N_2} \sum_{k_2=1}^{M_2} \pi^{(2)}\left(i_2 j_2 \mid k_2^*(i_2)\right) d^{(2)*}\left(k_2^*(i_2) \mid i_2\right) p^{(1)*}(i_1) p^{(2)*}(i_2)$$

such that

$$\sum_{i_2=1}^{N_2} p^{(2)*}(i_2) = 1, \quad p^{(2)*}(i_2) \geq 0, j_2 = 1, \ldots, N_2$$

3.3 Market Leader's Firm Utilities Optimization

The leader then selects a strategy that maximizes its utility, anticipating the predicted response of the follower. Then, for the given strategies $d^{(1)}(k_1 \mid i_1)$ and $d^{(2)*}(k_2 \mid i_2)$ the average utility of the leader is given by

$$\mathbf{U}^{(1)}(d^{(1)}, d^{(2)*}) = \sum_{i_1=1}^{N_1} \sum_{i_2=1}^{N_2} \sum_{k_1=1}^{M_1} \sum_{k_2=1}^{M_2} W^{(1)}(i_1 \mid k_1) c_n^{(1)}(i_1 \mid k_1) c^{(2)*}(i_2 \mid k_2)$$

where

$$c^{(2)*}(i_2 \mid k_2) := p^{(2)*}(s_n = s(i_2)) d^{(2)*}(k_2 \mid i_2), \quad c^{(1)}(i_1 \mid k_1) := p^{(1)}(s_n^{(1)} = s(i_1)) d^{(1)}(k_1 \mid i_1)$$

$$W^{(1)}(i_1 \mid k_1) := \sum_{j_1=1}^{N_1} \sum_{j_2=1}^{N_2} \left[\mathcal{P}^{(1)}(i_1, j_1 \mid k_1) \mathcal{Q}^{(1)}(i_1, j_1 \mid k_1) - C^{(1)}(i_1, j_1 \mid k_1) \right] \cdot \\ \pi^{(1)}(i_1 j_1 \mid k_1) \pi^{(2)}\left(i_2 j_2 \mid k_2(i_2)\right)$$

$U^{(1)}(i_1 j_1 \mid k_1)$ is the highest utility of leader firm (1) at state $s^{(1)}(i_1)$ when the action $a^{(1)}(k_1)$ is applied while the follower firm (2) select the worse response. In other words, the leader maximizes his payoff under the " pessimistic" supposition that the followers act to his disadvantage. This pessimistic supposition is used to the define a " Stackelberg payoff" to the leader in dynamic games (Basar and Olsder, 1982).

Notice that by Lemma 1 it follows

$$\mathbf{U}_n^{(1)}(d^{(1)},d^{(2)*})\underset{n\to\infty}{\to}U(c^{(1)},c^{(2)*})=\sum_{i_1=1}^{N_1}\sum_{k_1=1}^{M_1}\sum_{i_2=1}^{N_2}\sum_{k_2=1}^{M_2}W^{(1)*}(i_1\mid k_1)c^{(1)}(i_1\mid k_1)c^{(2)*}(i_2\mid k_2^*(i_2))$$

3.4 Local Optimal Best-reply Stackelberg Strategy of the Leader Firm

The optimal Stackelber leader's firm strategy is the solution of the following optimization problem

$$\mathbf{U}^{(1)}(c^{(1)},c^{(2)*})=\sum_{i_1=1}^{N_1}\sum_{k_1=1}^{M_1}\sum_{i_2=1}^{N_2}\sum_{k_2=1}^{M_2}W^{(1)}(i_1\mid k_1)c^{(1)}(i_1\mid k_1)c^{(2)*}(i_2\mid k_2^*(i_2))\to\max_{c^{(1)}} \tag{6}$$

subject the constraints

$$\sum_{i_1=1}^{N_1}\sum_{k_1=1}^{M_1}c^{(1)}(i_1\mid k_1)=1,\ c^{(1)}(i_1\mid k_1)\geq 0$$
$$j_1:\sum_{k_1=1}^{M_1}c^{(1)}(j_1\mid k_1)=\sum_{i_1=1}^{N_1}\sum_{k_1=1}^{M_1}\pi^{(1)}(i_1,j_1\mid k_1)c^{(1)}(i_1\mid k_1) \tag{7}$$

Denote the solution of the optimization problem (6)-(7) by $c^{(1)*}(i_1\mid k_1)$.

Lemma 2 (Clempner and Poznyak, 2013)*The solution* $c^{(1)*}(i_1\mid k_1)$ *of the optimization problem (6)-(7) coincides with the best-reply strategy*

$$d^{(1)*}(k_1^*(i_1)\mid i_1)=\arg\max_{d^{(1)}(k_1(i_1)\mid i_1)}W^{(1)*}(i_1\mid k_1)=\arg\max_{d^{(1)}(k_1(i_1)\mid i_1)}\Big(\sum_{j_1=1}^{N_1}\sum_{i_2=1}^{N_2}\sum_{j_2=1}^{N_2}\sum_{k_2=1}^{M_2}$$
$$\left[\mathcal{P}^{(1)}(i_1,j_1\mid k_1)\mathcal{Q}^{(1)}(i_1,j_1\mid k_1)-C^{(1)}(i_1,j_1\mid k_1)\right]\pi^{(1)}(i_1j_1\mid k_1)\pi^{(2)}(i_2j_2\mid k_2^*(i_2))$$
$$c^{(1)}(i_1\mid k_1)c^{(2)*}(i_2\mid k_2^*(i_2))\Big)$$

Remark 1 *For any Stackelberg game if the follower follows a Markov pure stationary strategy, then there exists a Markov pure stationary strategy that is the optimal strategy for the leader.*

3.5 Markov Chains Dynamics for the Leader Firm

For the local best-reply strategy $d^{(1)*}(k_1\mid i_1)$ we have the following Markov chain dynamics for the leader firm:

$$p^{(1)*}(j_1)=\sum_{i_2=1}^{N_2}\sum_{i_1=1}^{N_1}\sum_{k_1=1}^{M_1}\pi^{(1)}(i_1j_1\mid k_1^*(i_1))d^{(1)*}(k_1^*(i_1)\mid i_1)p^{(1)*}(i_1)p^{(2)*}(i_2),\ j_1=1,\dots,N_1, \tag{8}$$

such that $\sum_{i_1=1}^{N_1}p^{(1)*}(i_1)=1,\ p^{(1)*}(i_1)\geq 0, j_1=1,\dots,N_1$

3.6 Strong Stackelberg Equilibrium and the Average Utilities

For any leader firm policy, the follower firm plays the best-reply policy defined as follows:

$$d^{(2)*}(k_2^*(i_2) \mid i_2) = \arg\max_{k_2} \; W^{(2)*}(i_2 \mid k_2)$$

The leader's optimal policy is then

$$d^{(1)*}(k_1^*(i_1) \mid i_1) = \arg\max_{k_1} \; W^{(1)*}(i_1 \mid k_1)$$

In our Stackelberg model, the leader firm chooses a strategy first, and the follower firm chooses a strategy after observing the leader's choice. Then, the pair $d^* = (d^{(1)*}(k_1^*(i_1) \mid i_1), d^{(2)*}(k_2^*(i_2) \mid i_2))$ forms a Stackelberg equilibrium. The same point d^* conforms a Strong Stackelberg Equilibrium (SSE) if both the leader firm and follower firm choose an optimal strategy and, in addition, the follower firm breaks ties optimally for the leader: $\mathbf{U}^{(1)}(d^{(1)*}, d^{(2)*}) > \mathbf{U}^{(2)}(d^{(1)*}, d^{(2)*})$ where $d^{(2)*}$ is the follower firm best-reply to $d^{(1)*}$.

The average leader utility under the strategies $c^{(1)*}$, $c^{(2)*}$ is given by

$$\mathbf{U}^{(1)}(c^{(1)*}, c^{(2)*}) = \sum_{i_1=1}^{N_1}\sum_{i_2=1}^{N_2}\sum_{j_1=1}^{N_1}\sum_{j_2=1}^{N_2}\sum_{k_1=1}^{M_1}\sum_{k_2=1}^{M_2} \left[\mathcal{P}^{(1)}(i_1, j_1 \mid k_1)\mathcal{Q}^{(1)}(i_1, j_1 \mid k_1) - C^{(1)}(i_1, j_1 \mid k_1)\right] \cdot$$
$$\pi^{(1)}\left(i_1 j_1 \mid k_1^*(i_1)\right)\pi^{(2)}\left(i_2 j_2 \mid k_2^*(i_2)\right) c^{(1)*}\left(i_1 \mid k_1^*(i_1)\right) c^{(2)*}(i_2 \mid k_2^*(i_2)) =$$
$$\sum_{i_1=1}^{N_1}\sum_{k_1=1}^{M_1}\sum_{i_2=1}^{N_2}\sum_{k_2=1}^{M_2} W^{(1)*}(i_1 \mid k_1) c^{(1)*}(i_1 \mid k_1) c^{(2)*}(i_2 \mid k_2^*(i_2))$$

At the end, the follower observes the strategy that maximizes the leader utility and in equilibrium selects the expected strategy as a response. Then, the average follower utility under the strategies $c^{(1)*}, c^{(2)*}$ is given by

$$\mathbf{U}^{(2)}(c^{(1)*}, c^{(2)*}) = \sum_{i_1=1}^{N_1}\sum_{i_2=1}^{N_2}\sum_{j_1=1}^{N_1}\sum_{j_2=1}^{N_2}\sum_{k_1=1}^{M_1}\sum_{k_2=1}^{M_2} \left[\mathcal{P}^{(2)}(i_2, j_2 \mid k_2)\mathcal{Q}^{(2)}(i_2, j_2 \mid k_2) - C^{(2)}(i_2, j_2 \mid k_2)\right] \cdot$$
$$\pi^{(1)}\left(i_1 j_1 \mid k_1^*(i_1)\right)\pi^{(2)}\left(i_2 j_2 \mid k_2^*(i_2)\right) \cdot \; c^{(1)*}\left(i_1 \mid k_1^*(i_1)\right) c^{(2)*}(i_2 \mid k_2^*(i_2)) =$$
$$\sum_{i_1=1}^{N_1}\sum_{k_1=1}^{M_1}\sum_{i_2=1}^{N_2}\sum_{k_2=1}^{M_2} W^{(2)*}(i_2 \mid k_2) c^{(1)*}\left(i_1 \mid k_1^*(i_1)\right) c^{(2)*}(i_2 \mid k_2)$$

4. Computing the Lyapunov Equilibrium Point

The aim of this section is to associate to any utility function $\mathbf{U}_n^{(1)}$ a Lyapunov-like function which monotonically increases (non-decreases) on the trajectories of the given system (Clempner and Poznyak, 2013).

4.1 Utility Function Value Iteration

In vector format, the utility can be expressed as

$$\mathbf{U}^{(1)}(c^{(1)*}, c^{(2)*}) = \left\langle \mathbf{w}_n^{(1)}, \mathbf{p}_n^{(1)} \right\rangle$$

where

$$\left(\mathbf{w}_n^{(1)}\right)_{i_1} := \sum_{k_1=1}^{M_1}\sum_{k_2=1}^{M_2}[\sum_{j_1=1}^{N_1}\sum_{j_2=1}^{N_2}\sum_{i_2=1}^{N_2}\left[\mathcal{P}^{(1)}(i_1, j_1 \mid k_1)\mathcal{Q}^{(1)}(i_1, j_1 \mid k_1)C^{(1)}(i_1, j_1 \mid k_1)\right]\cdot$$
$$\pi^{(1)}(i_1 j_1 \mid k_1)\pi^{(2)}(i_2 j_2 \mid k_2)]d_n^{(1)}(k_1 \mid i_1)d_n^{(2)}(k_2 \mid i_2)$$

and given Eq. (8) we have that

$$\mathbf{p}_{n+1}^{(1)} = \left(\mathbf{\Pi}_n^{(1)}\right)^{\mathrm{T}}\mathbf{p}_n^{(1)},\ \left(\mathbf{\Pi}_n^{(1)}\right)_{i_1} := \sum_{j_1=1}^{N_1}\sum_{k_1=1}^{M_1}\pi^{(1)}(i_1 j_1 \mid k_1)d_n^{(1)}(k_1 \mid i_1)$$

Let us introduce the following general definition of Lyapunov-like function.

Definition 2 *Let* $\mathsf{V}: S \to \mathsf{R}_+$ *be a continuous map. Then,* V *is said to be a* ***Lyapunov-like function*** *iff it satisfies the following properties :*

1. $\exists s^*$, *called below a* ***Lyapunov equilibrium point****, such that* $\mathsf{V}^\iota(s^*) = 0$,
2. $\mathsf{V}^\iota(s) > 0$ *for all* $s \neq s^*$ *and all* $\iota \in \mathsf{N}$,
3. $\mathsf{V}^\iota(s^i) \to \infty$ *if there exists a sequence* $\left\{s^i\right\}_{i=1}^{\infty}$ *with* $s^i \to \infty$ *as* $i \to \infty$ *for all* $\iota \in \mathsf{N}$,
4. $\Delta\mathsf{V}^\iota(s', s) = \mathsf{V}^\iota(s') - \mathsf{V}^\iota(s) > 0$ *for all* $s' \leq_{\mathsf{V}^\iota} s : s, s' \neq s^*$.

Given fixed history of the process $\left(\mathbf{p}_0^{(1)}, d_0^{(1)}, d_1^{(1)}, \ldots, d_{n-1}^{(1)}\right)$

$$\max_{d^{(1)}} \mathbf{U}_{n+1}^{(1)} = \sum_{i_1=1}^{N_1}\sum_{i_2=1}^{N_2}\left(\mathbf{p}_n^{(1)}\right)_{i_1}\sum_{j_1=1}^{N_1}\sum_{j_2=1}^{N_2}\sum_{k_1=1}^{M_1}\sum_{k_2=1}^{M_2}\left[\mathcal{P}^{(1)}(i_1, j_1 \mid k_1)\mathcal{Q}^{(1)}(i_1, j_1 \mid k_1) - C^{(1)}(i_1, j_1 \mid k_1)\right]\cdot$$
$$\pi^{(1)}(i_1 j_1 \mid k_1)d^{(1)}(k_1 \mid i_1)\pi^{(2)}(i_2 j_2 \mid k_2)d^{(2)}(k_2 \mid i_2)$$

and considering point (4) of Definition 2, for the local policy $d^{(1)*}(k_1 \mid i_1)$ we have

$$\sum_{j_1=1}^{N_1}\sum_{i_2=1}^{N_2}\sum_{j_2=1}^{N_2}\sum_{k_2=1}^{M_2}U^{(1)}(i_1 j_1 \mid k_1)\pi^{(1)}(i_1 j_1 \mid k_1)d^{(1)*}(k_1 \mid i_1)\pi^{(2)}(i_2 j_2 \mid k_2)d^{(2)}(k_2 \mid i_2) \geq$$

$$\sum_{j_1=1}^{N_1}\sum_{i_2=1}^{N_2}\sum_{j_2=1}^{N_2}\sum_{k_2=1}^{M_2}U^{(1)}(i_1 j_1 \mid k_1)\pi^{(1)}(i_1 j_1 \mid k_1)d^{(1)}(k_1 \mid i_1)\pi^{(2)}(i_2 j_2 \mid k_2)d^{(2)}(k_2 \mid i_2) := W^{(1)}(i_1 \mid k_1)$$

$\forall k_1 = 1,\ldots, M_1$, where $U^{(1)}(i_1 j_1 \mid k_1) = \mathcal{P}^{(1)}(i_1, j_1, k_1)\mathcal{Q}^{(1)}(i_1, j_1, k_1) - C^{(1)}(i_1, j_1 \mid k_1)$.

As a result we can state the following lemma.

Lemma 3 *Given a fixed-local policy, the U -values iteration for all state-action pairs from (4) become* $\mathbf{U}_{n+1}^{(1)} = \left\langle \mathbf{w}^{(1)*}, \mathbf{p}_n^{(1)} \right\rangle$ *where* $\mathbf{w}^{(1)*} := \left(\left(\mathbf{w}^{(1)*} \right)_1, \ldots, \left(\mathbf{w}^{(1)*} \right)_{N_1} \right)$ *and*

$$\left(\mathbf{w}^{(1)*} \right)_{i_1} := \sum_{k_1=1}^{M_1} \sum_{k_2=1}^{M_2} \sum_{j_1=1}^{N_1} \sum_{j_2=1}^{N_2} \sum_{i_2=1}^{N_2} U^{(1)}(i_1 j_1 \mid k_1) \pi^{(1)}(i_1 j_1 \mid k_1) d^{(1)*}(k_1 \mid i_1) \pi^{(2)}(i_2 j_2 \mid k_2) d^{(2)}(k_2 \mid i_2)$$

$$= \max_{d^{(1)}} W^{(1)}(i_1 \mid k_1)$$

Under the strategy $d^{(1)*}(k_1 \mid i_1)$ the state-vector $\mathbf{p}_n^{(1)}$ satisfies that

$$\mathbf{p}_{n+1}^{(1)} = \left(\mathbf{\Pi}^{(1)*} \right)^{\mathrm{T}} \mathbf{p}_n^{(1)} = \left(\left(\mathbf{\Pi}^{(1)*} \right)^{\mathrm{T}} \right)^{n+1} \mathbf{p}_0^{(1)}$$

where

$$\left(\mathbf{\Pi}^{(1)*} \right)_{i_1} := \sum_{j_1=1}^{N_1} \sum_{k_1=1}^{M_1} \pi^{(1)}(i_1 j_1 \mid k_1) d^{(1)*}(k_1 \mid i_1)$$

4.2 Recurrent Form for the Utility Function

Let us represent $\mathbf{U}_{n+1}^{(1)}$ as follows

$$\mathbf{U}_{n+1}^{(1)} = \left\langle \mathbf{w}^{(1)*}, \mathbf{p}_n^{(1)} \right\rangle = \left[\left(1 + \frac{\left\langle \left[\mathbf{\Pi}^{(1)*} - I \right] \mathbf{w}^{(1)*}, \mathbf{p}_{n-1}^{(1)} \right\rangle}{\mathbf{U}_n^{(1)}} \right) \mathbf{U}_n^{(1)} \right]$$

and denoting

$$\alpha_n^{(1)} = \frac{\left\langle \left[\mathbf{\Pi}^{(1)*} - I \right] \mathbf{w}^{(1)*}, \mathbf{p}_{n-1} \right\rangle}{\mathbf{U}_n^{(1)}} = \frac{\left\langle \left[\mathbf{\Pi}^{(1)*} - I \right] \mathbf{w}^{(1)*}, \mathbf{p}_{n-1}^{(1)} \right\rangle}{\left\langle \mathbf{w}^{(1)*}, \mathbf{p}_{n-1}^{(1)} \right\rangle}$$

we obtain

$$\mathbf{U}_{n+1}^{(1)} = (1 + \alpha_n^{(1)}) \mathbf{U}_n^{(1)} \tag{9}$$

Now we are ready to formulate the main result of this paper.

4.3 Lyapunov and Strong Stackelberg Equilibrium

Defining $\tilde{\alpha}_n^{(1)}$ as

$$\tilde{\alpha}_n^{(1)} = \begin{cases} \alpha_n^{(1)} & \text{if} \quad \alpha_n^{(1)} \geq 0 \\ 0 & \text{if} \quad \alpha_n^{(1)} < 0 \end{cases}$$

we get

$$\mathbf{U}_{n+1}^{(1)} = (1+\alpha_n^{(1)})\mathbf{U}_n^{(1)} \geq (1+\tilde{\alpha}_n^{(1)})\mathbf{U}_n^{(1)}$$

which leads to the following statement.

Theorem 1 *Let* $G = \langle \mathcal{N}, S, (\Delta^\iota)_{\iota\in\mathcal{N}}, (\Upsilon^\iota)_{\iota\in\mathcal{N}}, \Pi, (U^\iota)_{\iota\in\mathcal{N}} \rangle$ *be a duopoly Stackelberg game and let* $\mathbf{U}_{n+1}^{(1)}$ *be represented by the recursive matrix format of Eq. (9). Then, a possible Lyapunov-like function* $\mathbf{U}_n^{(1),mon}$ *(which is monotonically non-decreasing) for the leader* (1) *in* G *has the form*

$$\mathbf{U}_n^{(1),mon} = \mathbf{U}_n^{(1)} \prod_{t=1}^{n-1} (1+\alpha_t^{(1)})^{-1} = \frac{1+\tilde{\alpha}_{n-1}^{(1)}}{1+\alpha_{n-1}^{(1)}} \mathbf{U}_{n-1}^{(1),mon}, \ \mathbf{U}_0^{(1),mon} = \mathbf{U}_0^{(1)}$$

Proof. Let us consider the recursion $x_{n+1} \geq (1+\gamma_n)x_n + \eta_n$ with $\gamma_n, x_n, \eta_n \geq 0$. Defining

$$\tilde{x}_n := x_n \prod_{t=1}^{n-1} (1+\gamma_n)^{-1}, \ \eta_n := \eta_n \prod_{t=1}^{n} (1+\gamma_n)^{-1}, \ y_n = \tilde{x}_n - \sum_{t=1}^{n-1} \eta_t$$

we obtain $y_{n+1} \geq y_n$. Indeed,

$$\tilde{x}_{n+1} = x_n \prod_{t=1}^{n} (1+\gamma_n)^{-1} \geq x_n \left[\prod_{t=1}^{n} (1+\gamma_n)^{-1} \right] (1+\gamma_n) + \eta_n \prod_{t=1}^{n} (1+\gamma_n)^{-1} = \tilde{x}_n + \tilde{\eta}_n$$

which implies

$$y_{n+1} = \tilde{x}_{n+1} - \sum_{t=1}^{n} \eta_t \geq \tilde{x}_n + \eta_n - \sum_{t=1}^{n} \eta_t = \tilde{x}_n + \sum_{t=1}^{n-1} \eta_t = y_n$$

and therefore $y_{n+1} \geq y_n$. In view of this we have

$$\mathbf{U}_{n+1}^{(1),mon} = \mathbf{U}_{n+1}^{(1)} \prod_{t=1}^{n} (1+\alpha_t^{(1)})^{-1} \geq (1+\alpha_n^{(1)})\mathbf{U}_n^{(1)} \prod_{t=1}^{n} (1+\alpha_t^{(1)})^{-1} = \mathbf{U}_n^{(1)} \prod_{t=1}^{n-1} (1+\alpha_t^{(1)})^{-1} = \mathbf{U}_n^{(1),mon}$$

that proves the result.

Remark 2 Following a similar development for the follower firm we have

$$\mathbf{U}_{n+1}^{(2),mon} = \mathbf{U}_{n+1}^{(2)} \prod_{t=1}^{n} (1+\alpha_t^{(2)})^{-1} \geq (1+\alpha_n^{(2)})\mathbf{U}_n^{(2)} \prod_{t=1}^{n} (1+\alpha_t^{(2)})^{-1} = \mathbf{U}_n^{(2)} \prod_{t=1}^{n-1} (1+\alpha_t^{(2)})^{-1} = \mathbf{U}_n^{(2),mon}$$

Definition 3. *A Lyapunov game is a tuple* $G = \langle \mathcal{N}, S, (\Delta^\iota)_{\iota\in\mathcal{N}}, (\Upsilon^\iota)_{\iota\in\mathcal{N}}, \Pi, (\mathbf{U}^\iota)_{\iota\in\mathcal{N}} \rangle$ *where* $\mathbf{U}^\iota$ *is a Lyapunov-like function (monotonically increasing in time).*

Theorem 2 *If* $d^* = (d^{(1)*}(k_1^*(i_1)|\, i_1), d^{(2)*}(k_2^*(i_2)|\, i_2))$ *is a strong Stackelberg equilibrium point then the maximum is asymptotically approached (or the maximum is attained) by the Lyapunov-like function* $\mathbf{U}$*, i.e.* $\mathbf{U}(d^*) = 0$ *or* $\mathbf{U}(d^*) = Const$*, where Const is a constant.*

Proof. Suppose that d^* is a equilibrium point. We want to show that $\mathbf{U}$ has an asymptotically approached infimum (or reaches a minimum). Since, d^* is an equilibrium point by Definition 2 cannot be modified. Then it follows that the strategy attached to the action(s), following d^*, is 0. Let us consider the value of $\mathbf{U}$ cannot be modified. Since the Lyapunov-like function is a increasing function of the strategies d (by Definition 2) an infimum or a minimum is attained in d^*.

Theorem 3 The fixed-local-optimal strategy $d^* = (d^{(1)*}(k_1^*(i_1)|\, i_1),\, d^{(2)*}(k_2^*(i_2)|\, i_2))$ is a strong Stackelberg equilibrium point.

Proof. Let suppose that there exist a strategy $d^{(2)'}(k_2'(i_2)|\, i_2) \neq d^{(2)*}(k_2^*(i_2)|\, i_2)$ such that $(d^{(1)*}(k_1^*(i_1)|\, i_1),\, d^{(2)'}(k_2^*(i_2)|\, i_2))$ is an equilibrium point. Then, $\mathbf{U}(d^{(1)*}(k_1^*(i_1)|\, i_1), d^{(2)'}(k_2'(i_2)|\, i_2)) > \mathbf{U}(d^{(1)*}(k_1^*(i_1)|\, i_1),\, d^{(2)*}(k_2^*(i_2)|\, i_2))$ by Definition 3, but it is a contradiction to Theorem 2 and the fact that $d^{(2)*}(k_2^*(i_2)|\, i_2)$ is fixed-local-optimal strategy. As well as, let suppose that $d^{(2)'}(k_2'(i_2)|\, i_2)$ is a best-reply strategy to $d^{(1)'}(k_1^*(i_1)|\, i_1) \neq d^{(1)*}(k_1^*(i_1)|\, i_1)$ and $d' = (d^{(1)'}(k_1'(i_1)|\, i_1), d^{(2)'}(k_2'(i_2)|\, i_2))$ is an equilibrium point. Then, $\mathbf{U}(d') > \mathbf{U}(d^*)$, but it is a contradiction to Theorem 2 and the fact that $d^{(1)*}(k_1^*(i_1)|\, i_1)$ is fixed-local-optimal strategy.

Corollary 1 The Lyapunov-like function converges to a unique strong Stackelberg equilibrium $d^* = (d^{(1)*}(k_1^*(i_1)|\, i_1), d^{(2)*}(k_2^*(i_2)|\, i_2))$.

Proof. Let us suppose that d^* is not an equilibrium point. Therefore, it is possible to apply an output strategy to d^*. Then, it is possible to modify the utility over d^*. As a result, it is possible to obtain a higher utility, i.e. $\mathbf{U}(d^*) > 0$ or $\mathbf{U}(d^*) > C$. It is a contradiction to Theorem 4.3 and the fact that d^* is fixed-local-optimal strategy.

Theorem 4 The Lyapunov equilibrium point coincides with the strong Stackelberg equilibrium point.

Proof. Straightforward from Theorem 1, 2 and 3.

5. Numerical Example

In this Section we implement the example presented in Subsection 1.2. The segmentation of the car market is represented in Figure 1.Let $N_1 = 4$ and $N_2 = 4$ be the number of states and $M_1 = 3$ and $M_2 = 2$ be the number of actions of the Leader and the Follower respectively. The utility for the companies $\iota = 1,2$ are as follows

$$U^{(1)}(i,j,k)=\begin{bmatrix}20 & 23 & 36 & 62\\ 15 & 32 & 40 & 64\\ 18 & 30 & 36 & 54\\ 19 & 14 & 16 & 5\end{bmatrix} U^{(2)}(i,j,k)=\begin{bmatrix}13 & 15 & 17 & 10\\ 23 & 32 & 30 & 20\\ 36 & 40 & 36 & 20\\ 52 & 34 & 48 & 10\end{bmatrix}$$

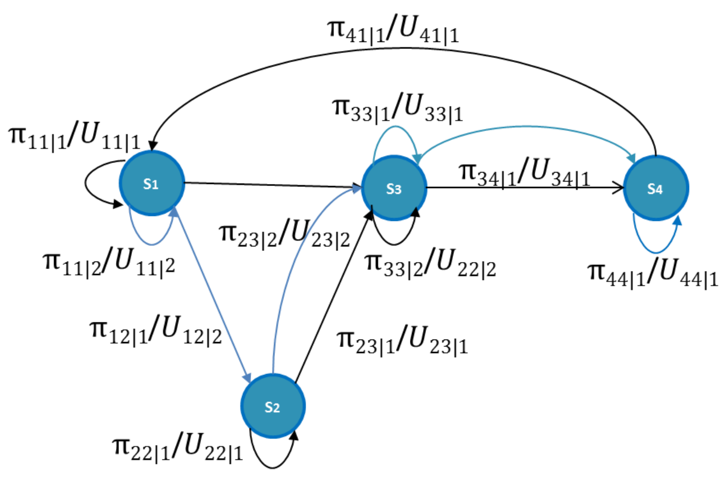
and let the transition matrix for Company 1 identified by $k_1 = 1,2,3$ be as follows

$$\pi^{(1)}(i,j,1)= \qquad \pi^{(1)}(i,j,2)=$$

$$\begin{bmatrix}0.0121 & 0.0013 & 0.8969 & 0.0897\\ 0.0129 & 0.0012 & 0.9859 & 0.0000\\ 0.4324 & 0.5405 & 0.0000 & 0.0270\\ 0.2154 & 0.0066 & 0.0607 & 0.7172\end{bmatrix}\begin{bmatrix}0.0127 & 0.0013 & 0.9658 & 0.0202\\ 0.0149 & 0.0000 & 0.9850 & 0.0001\\ 0.8117 & 0.0901 & 0.0089 & 0.0893\\ 0.8273 & 0.0202 & 0.1314 & 0.0212\end{bmatrix}$$

$$\pi^{(1)}(i,j,3)=\begin{bmatrix}0.0704 & 0.0257 & 0.9039 & 0.0001\\ 0.1440 & 0.0000 & 0.8559 & 0.0001\\ 0.8346 & 0.0000 & 0.0002 & 0.1653\\ 0.9023 & 0.0827 & 0.0146 & 0.0004\end{bmatrix}$$

Figure 1.Markov chain car market segmentation



and let the transition matrix for Company 2 given by medium and low actions are identified by $k_2 = 1,2$ be defined as follows

$$\pi^{(2)}(i,j,1)=\begin{bmatrix}0.0738 & 0.1256 & 0.8006 & 0.0001\\ 0.0423 & 0.0000 & 0.9576 & 0.0001\\ 0.3193 & 0.0001 & 0.0575 & 0.6231\\ 0.9527 & 0.0067 & 0.0406 & 0.0000\end{bmatrix}\quad \pi^{(2)}(i,j,2)=\begin{bmatrix}0.0370 & 0.0246 & 0.9376 & 0.0008\\ 0.1106 & 0.0005 & 0.8778 & 0.0111\\ 0.9710 & 0.0094 & 0.0048 & 0.0149\\ 0.8179 & 0.0082 & 0.0853 & 0.0886\end{bmatrix}$$

For the best-reply strategies $d^{(1)*}$ and $d^{(2)*}$ we have the following results (see, Clempner and Poznyak, 2014):

$$\pi^{(1)*}=\begin{bmatrix}0.0121 & 0.0013 & 0.8969 & 0.0897\\ 0.0129 & 0.0012 & 0.9859 & 0.0000\\ 0.4324 & 0.5405 & 0.0000 & 0.0270\\ 0.9023 & 0.0827 & 0.0146 & 0.0004\end{bmatrix}\quad \pi^{(2)*}=\begin{bmatrix}0.0370 & 0.0246 & 0.9376 & 0.0008\\ 0.0423 & 0.0000 & 0.9576 & 0.0001\\ 0.9710 & 0.0094 & 0.0048 & 0.0149\\ 0.9527 & 0.0067 & 0.0407 & 0.0000\end{bmatrix}$$

$$k_1^*=[1,1,1,3],\ k_2^*=[2,1,2,1],\ n_0^{(1)}=2,\ \chi_{erg}^{(1)}=0.9224\ ,\ n_0^{(2)}=2,\ \chi_{erg}^{(2)}=0.9199$$

$$w^{(1)*}=[38.1216, 39.6683, 25.4596, 18.5366],\ w^{(2)*}=[16.7971, 29.7028, 35.7990, 51.7171]$$

We show in Figure 2 and Figure 3 the state-value function behavior for both firms leader and follower and, in Figure 4 and Figure 5 the corresponding Lyapunov-like functions. The numerical results clearly show that under the same fixed-local-optimal strategy the original utility functions converge non-monotonically to the value, for the leader company, 31.8387 and the corresponding Lyapunov-like functions converge monotonically to the value 31.8387. As well as, for the follower company, 26.2819 and the corresponding Lyapunov-like functions converge monotonically to the value 26.2895 respectively which, obviously, are very close. We also concluded from the numerical example that the Lyapunov equilibrium point coincides with the SSE point.

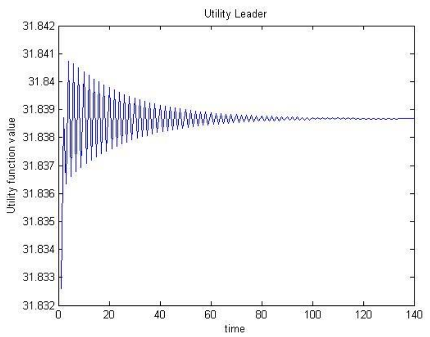
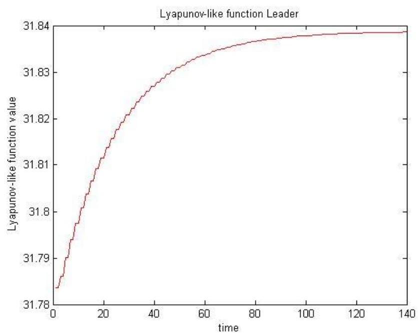
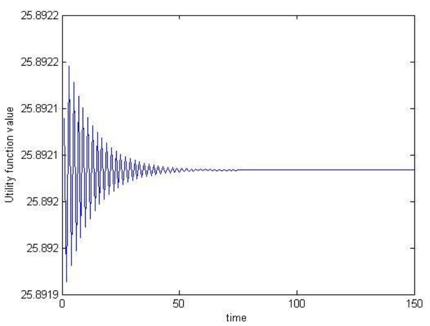
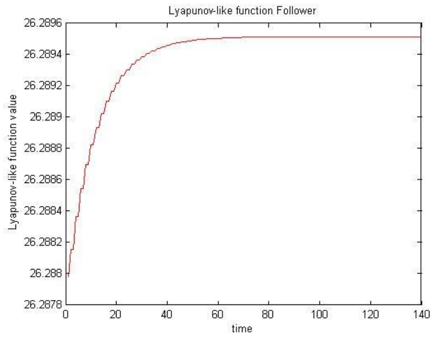
Figure 2 Non-monotonic behavior of the Leader Firm.	**Figure 3.Lyapunov-like function for the Leader Firm.**
	

Figure 4. Non-monotonic behavior of the Follower Firm.	**Figure 5.Lyapunov-like function for the Follower Firm.**
	

6. Conclusion

This paper suggested an approach where the classical Stackelberg duopoly game is transformed into a potential game in terms of the Lyapunov theory.It focused on a general class of ergodic controllable finite Markov chains game for representing Stackelberg duopoly game. The paper confronts a fundamental question of equivalent convergence of different equilibrium points. In this context, this paper provides four key contributions. First, we present a method to represent a non-converging state-value function that fluctuates (increases and decreases) between states of the stochastic game in a recursive format. Second, we propose a Lyapunov-like function that replaces the recursive mechanism with the elements of the ergodic system. Third, our main result (Theorem 4) states that the Lyapunov equilibrium point coincides with the SSE point (which is the Cournot equilibrium point). Fourth, our experimental results emphasize positive the coincidence of the equilibrium points.

REFERENCES

[1] **Amir, R. and Grilo, I. (1996),*Stackelberg versus Cournot Equilibrium*.** *Games and Economic Behavior*, 26: 1-21;

[2] **Basar, T. and Olsder, G. J. (1982),*Dynamic Noncooperative Game Theory***. *Academic Press*;

[3] **Breitmoser, Y. (2012),*On the Endogeneity of Cournot, Bertrand and Stackelberg Competition in Oligopolies***. *International Journal of Industrial Organization*, 30 (1): 16-29;

[4] **Bös, D. (1991),*Privatization: A Theoretical Treatment***. Clarendon Press;

[5] **Clempner, J. B. (2015),*Setting Cournot vs. Lyapunov Games Stability Conditions and Equilibrium Point Properties***. *International Game Theory Review*. DOI: 10.1142/S0219198915500115;

[6] **Clempner, J. B. and Poznyak, A. S. (2011),*Convergence Properties and Computational Complexity Analysis for Lyapunov Games*.** *International Journal of Applied Mathematics and Computer Science*, 21(2): 49-361;

[7]**Clempner, J. B. and Poznyak, A. S. (2014),*Simple Computing of the Customer Lifetime Value: A Fixed Local-Optimal Policy Approach***. *Journal of Systems Science and Systems Engineering*, 23 (4): 439–459;

[8]**Clempner, J. B. and Poznyak, A. S. (2013), *In Parsini T., Tempo R. (eds.), Analysis of Best-Reply Strategies in Repeated Finite Markov Chains Games***. 52nd IEEE Conference on Decision and Control, pp. 934-939, Florence, Italy, December 10-13, 2013, IEEE;

[9] **Cournot, A. A. (1938),*Recherches sur les Principes Mathematiques de la Theorie des Richesses***. Paris: Hachette;

[10] **De Fraja, G. andDelbono, F. (1990),*Game Theoretic Models of Mixed Oligopoly*.** *Journal Economic Surveys*, 4: 1-17;

[11] **Dragone, D., Lambertini, L. and Palestini, A. (2008), *A Class of Best-Response Potential Games***; working paper;

[12] **Dragone, D., Lambertini, L. and Palestini, A. (2012), *Static and Dynamic Best-Response Potential Functions for the Non-Linear Cournot Game*.** *Optimization*, 61: 1283-1293;

[13**] Ghadimi, S., Szidarovszky F.,Farahani, R. Z. and Khiabani, A. Y. (2013),*Coordination of Advertising in Supply Chain Management with Cooperating Manufacturer and Retailers***. *IMA Journal of Management Mathematics*, 24 (1): 1-19;

[14] **Harris, R. G. and Wiens, E. G. (1980),*Government Enterprise: An Instrument for the Internal Regulation of Industry***. *Canadian Journal Economics*, 13: 125-132;

[15] **Matsumura, T. (2003). *Stackelberg Mixed Duopoly with a Foreign Competitor***. *Bulletin of Economics Reserch*, 55: 275-287;

[16] **Monderer, D. and Shapley, L. S. (1996),*Potential Games***. *Games and Economic Behavior*, 14: 124-143;

[17] **Morgan, J. and Patrone, F. (2006), *Stackelberg Problems: Subgame Perfect Equilibria via Tikhonov Regularization; Vol. 8 of Advances in Dynamic Games***. *Annals of the International Society of Dynamic Games*, Birkhuser Boston;

[18] **Poznyak, A. S., Najim, K. and Gomez-Ramirez, E. (2000),*Selflearning Control of Finite Markov Chains***, Marcel Dekker, Inc., New York.

[19] **Selten, R. (1965), *Spieltheoretische Behandlungeines Oligopolmodellsmit Nachfragetragheit***. *Zeitschriftfür die gesamteStaatrwissenschaft*, 121: 301-324, 667-689;

[20] **Von Stackelberg, H. (1934), *Marktform und Gleichgewicht***, *Springer*, Vienna;

[21]**Voorneveld, M. (2000),*Best-response Potential Games***. *Economics letters*, 66: 289-295.