

# Pretending in Dynamic Games, Alternative Outcomes and Application to Electricity Markets

Ioannis Kordonis<sup>1</sup>  · Alexandros C. Charalampidis<sup>1</sup> · George P. Papavassilopoulos<sup>2</sup>

© Springer Science+Business Media, LLC 2017

**Abstract** This work studies dynamic game situations with incomplete structural information, motivated by problems arising in electricity market modeling. Some adaptive/learning strategies are considered as an expression of the bounded rationality of the participants of the game. The adaptive strategies are typically not in Nash equilibrium. Thus, the possibility of manipulation appears. That is, a player may use the dynamic rule of the opponent in order to manipulate her. We focus on a smaller class of manipulating strategies, called pretending strategies, where each player acts as if she had different, not real, preferences. It turns out that under certain technical conditions, if only one player pretends, she can achieve the same cost as if she were the Stackelberg leader. The situation where all the players are pretending is then considered, and an auxiliary game, called pretenders' game, is introduced. A class of quadratic games is then studied, and several relations among pretending and Stackelberg leadership are derived. A linear quadratic environmental game is also studied. We then study some competitive electricity market models. Particularly, a supply function model and the market mechanism described in Rasouli and Teneketzis (electricity pooling markets with strategic producers possessing asymmetric information ii: inelastic demand, [arXiv: 1404.5539](https://arxiv.org/abs/1404.5539), 2014)

---

This research has been cofinanced by the European Union (European Social Fund—ESF) and Greek national funds through the Operational Program “Education and Lifelong Learning” of the National Strategic Reference Framework (NSRF)—Research Funding Program: THALES. Investing in knowledge society through the European Social Fund and the program ARISTEIA, project name HEPHAISTOS.

---

✉ Ioannis Kordonis  
jkordonis1920@yahoo.com

Alexandros C. Charalampidis  
alexandros.charalampidis@centralesupelec.fr

George P. Papavassilopoulos  
yorgos@netmode.ntua.gr

<sup>1</sup> CentraleSupélec, Avenue de la Boulaie, 35576 Cesson-Sévigné, France

<sup>2</sup> School of Electrical and Computer Engineering, National Technical University of Athens, 9 Iroon Polytechniou str., Postal Code 157 80 Athens, Greece

are considered. It turns out that pretending may increase competition or cooperation and in some cases pretending may cause behaviors making the system not working at all.

**Keywords** Dynamic Games with Incomplete Information · Learning/Adaptation in Games · Manipulation · Pretending · Electricity Markets

## 1 Introduction

In a number of actual interactive decision situations (game situations) there are a number of decision makers that interact strategically over time, but each one of them has only a partial knowledge of the intentions of the others. A particular example is an electricity market where several production firms are competing repeatedly over time and each firm knows its own costs but not the costs of the other firms [17,42]. Such strategic interactions over time can be described by dynamic games with incomplete structural information.

Quite often it is very difficult to find a Nash equilibrium for dynamic or repeated games with incomplete information. Two difficult problems may arise; the first is due to the “Witsenhausen effect” [44], i.e., the current action of each player affects the future state or parameter estimation of the other players. The second is due to the “dual control effect” [13], i.e., that the current action of a player affects the quality of his own future parameter estimation. The latter difficulty makes the optimization problems involved very difficult. Even in the degenerate game case (one player game) the optimal control problem with unknown parameters has not been solved analytically, except only for a few special cases [45].

In this context, it has been proposed that the players (instead of trying to find the equilibrium strategies) use some simple deterministic or stochastic dynamic rules which determine the future actions of the players based on their past actions. The use of such rules, usually, reflects the bounded rationality [38,41] of the participants of the game, i.e., their inability to solve very difficult problems. These rules can be models of learning, adaptation, evolution or imitation [16,40,47] ch. 7, [10,27,31]. Most of the research on dynamic rules focuses on convergence issues. However, a set of such strategies, typically, is not in Nash equilibrium.

A set of dynamic rules often converges to an equilibrium or a correlated equilibrium of a simplified game. For example, in repeated finite games no regret learning rules converge to a correlated equilibrium of the static full information game ([21]) and in the same class of games a set of uncoupled stochastic rules converges probabilistically to a Nash equilibrium of the corresponding static game [14]. Such results are often used to justify (or as an alibi for) the use of the equilibrium notions of the simplified games (for example, static full information) to the repeated or dynamic incomplete information game situation.

Are these adaptive/learning strategies a reasonable prediction for the evolution of the game? Particularly, if the players knew that they are going to implement those strategies, would they stick to them? When a player is implementing an adaptive control law, she may be viewed by another player as a system under control. That is, the other player may try to use the knowledge of the adaptation law of the first player to manipulate her. The topic of this work is to study phenomena occurring when the players are trying to manipulate one another and the implications that manipulation may have to the costs of the participants of the game.

### 1.1 Contribution

At first, a general form of adaptive/learning strategies is presented and an optimal control problem for strategic manipulation is stated. Then a smaller class of manipulating strategies

is considered, called the “pretender’s” strategies. The manipulating player implements her learning/adaptation rule as if she had a different, not real type (preferences). Outcomes alternative to the Nash equilibrium are proposed for the cases in which one player pretends and all the players pretend. Under some technical conditions, if only one player pretends, she may achieve the same outcome as if she were a Stackelberg leader. We then study the case in which all the players are pretending, and some possible limit points are identified, using an auxiliary game called the “pretenders’ game.”

In order to gain some intuition, the notions described are applied to a very simple class of repeated quadratic opinion games, where analytical results are possible. At first, games with two players are analyzed and relationships of pretending and Stackelberg leadership are derived. It turns out that pretending enhances competition among the players. In the case of a repeated Stackelberg interaction, it turns out that if the Stackelberg follower is pretending, the follower can achieve approximately the same cost as if she were acting on the Stackelberg leader’s behalf.

The cases of repeated quadratic opinion games with many players having symmetric and asymmetric interactions are then analyzed. In the symmetric case, it turns out that when the number of players increases, the motivation to pretend decreases, the value of pretending decreases, and the outcome approaches the Nash equilibrium.

In the asymmetric influence case, the motivation of each player to pretend depends on a measure of the centrality of her position in the network of interactions. In the case where there is one influential player and many players each one of which has a rather small influence, only the influential player has a non-negligible motivation to pretend and the outcome approaches the Stackelberg equilibrium with the influential player as the leader. Thus, a “natural” Stackelberg leadership appears.

A linear quadratic game with two players is then analyzed. The game describes the interaction of two countries emitting pollutants. In this example, each player has a motivation to pretend to have a lower discount factor, i.e., to pretend that she cares less about the future. The outcome of the game when the players are pretending is to emit more and have a more polluted environment. In all the examples of quadratic and linear quadratic games analyzed, pretending is not a cooperative action.

Two competitive electricity market models are then studied. At first, a linear supply function model is considered. In this game situation, the players have a motivation to pretend to have a larger production cost in order to increase the price. Pretending in this model is a cooperative action, i.e., the pretending of the one player is beneficial to the other. The outcome when all the players pretend, is that the players produce less, the price is higher, and their profit increases, compared to the Nash case.

In the second electricity market model a mechanism from the literature is considered [34]. The mechanism is designed in order to work optimally in the Nash equilibrium. However, each one of the participants has a motivation to pretend. Furthermore, if all the participants pretend the prices diverge and the system is not working at all.

## 1.2 Related Literature

A related topic is the literature on reputation effects [15, 26, 28]. These models consider a long-run, patient player and many short-lived or not patient opponents. Assuming that there is a possibility (even with small probability) that the patient player has a “committed type” (or tough type), the patient player may gain in equilibrium a strategic advantage, similar to the Stackelberg leadership. A model of two-sided reputation was studied in [3], resembling a war of attrition situation. The basic difference of the current work with the theory of

reputation games, is that in reputation games the roles of the players are asymmetric and a priori specified.

A very interesting related stream of research is the theory of indirect evolution ([12, 18, 19, 22, 23, 25]). These models assume rational players the preferences of which are affected by evolutionary forces. The evolution of preferences may explain phenomena such as altruism, spite, overconfidence, fairness, and reciprocity ([23]). In [1] it is shown that if both the preferences (or the perceptions) and the actions are evolving, the players would behave in an effectively rational way. The basic differences of the current work from the theory of indirect evolution are that the current work studies dynamic games, as well as repeated, and does not assume a full rationality for the participants. Another related research area is the theory of coevolutionary games [33, 43]. In coevolutionary games in addition to the players' strategies, other properties such as the structure of the network describing the interactions or the players payoffs evolve, leading to different outcomes from the evolutionary games. An interesting class of telecommunication games is studied in [2, 4, 5]. In these examples, the players are better off, if instead of optimizing against their own preferences, they have an amount of altruism, i.e., optimize against a weighted sum of the payoffs of all the players. This effect is called "Paradox in Cooperation."

A closely related topic is the theory of strategic delegation [39]. This theory studies situations where the acting party delegates its decision to another party having different interests.

Another related topic is the control of dynamic rules and strategic teaching. In [35], the optimal control of the best reply is studied and in [36] it is shown that in all the uncoupled learning rules the players have an incentive to deviate. Furthermore, in [37] it is shown that in a mixed population of myopic best response and mimic players, the imitators are better off. The difference of the current work with the theory of strategic teaching is that first it does not assign different roles to the players a priori and second that it suggests some alternative outcomes for the game.

In this work, we study the relationship among pretending and Stackelberg leadership. Endogenous models for Stackelberg leadership were studied in [7, 20].

### 1.3 Organization

The rest of this paper is organized as follows: In Sect. 2, the game and learning models are described. The pretenders' strategies are studied in Sect. 3. The opinion quadratic games are studied in Sect. 4. Section 5 studies the LQ environmental game. Section 6 studies some electricity market games. Particularly, Sect. 6.1 studies a supply function model and Sect. 6.2 a mechanism presented in [34]. Finally, Sect. 7 summarizes the main contributions and proposes some future work.

## 2 Game Description and Adaptive Strategies

Let us start describing an  $N$ -player game. The players have types  $\theta_1, \dots, \theta_N$ . Each player knows her own type  $\theta_i$ , and  $\theta_1, \dots, \theta_N$  are part of a random vector  $\bar{\theta} = [\theta_1 \dots \theta_N \ \theta]^T \in \bar{\Theta} = \Theta_1 \times \dots \times \Theta_N \times \Theta$  which follows a commonly known distribution.

The dynamics has the form:

$$x_{k+1} = f(x_k, \bar{\theta}, u_k^1, \dots, u_k^N, w_k), \quad (1)$$

where  $u_k^1, \dots, u_k^N$  are the action variables of the players and  $w_k$  is a random disturbance with a commonly known distribution.

The cost functions have the form:

$$J_i = E \left[ \sum_{k=0}^T \rho_{\theta_i}^k L_i \left( x_k, \theta_i, u_k^1, \dots, u_k^N \right) \right], \quad (2)$$

where  $\rho_{\theta_i} \in (0, 1]$  is a discount parameter and  $T$  can have a finite or an infinite value. An alternative formula for the costs is:

$$J_i = \lim_{T \rightarrow \infty} E \left[ \frac{1}{T} \sum_{k=0}^T L_i \left( x_k, \theta_i, u_k^1, \dots, u_k^N \right) \right]. \quad (3)$$

Equations (1), (2) or (1), (3) can describe dynamic games as well as repeated static games.

Each player receives at each time step an information vector according to one of the following information structures:

**Information Structure 1:**

$$\bar{I}_k^{i, new} = \left( x_k, u_{k-1}^i \right) \text{ or} \quad (4)$$

**Information Structure 2:**

$$\bar{I}_k^{i, new} = \left( x_k, u_{k-1}^i, u_{k-1}^{-i} \right), \quad (5)$$

where the  $u^{-i}$  stands for  $(u^1, \dots, u^{i-1}, u^{i+1}, \dots, u^N)$ . The information that each player possesses at time step  $k$  has the form  $I_k^i = \left( \bar{I}_0^{i, new}, \dots, \bar{I}_k^{i, new} \right)$ .

The strategies of the players are given by  $s^i = (\gamma_1^i, \gamma_2^i, \dots)$ ,  $i = 1, \dots, N$ , where  $\gamma_k^i$  is a function having the form:

$$\gamma_k^i : (x_0, x_1, \dots, x_k, \theta_i) \mapsto u_k^i \in U^i.$$

We will focus on “state feedback” strategies, where all the previous information is used only for “adaptation.” Specifically,

$$u_k^i = \gamma_k^i \left( x_k, \theta_i, \hat{\theta}_k^i \right), \quad (6)$$

where  $\hat{\theta}_k^i$  is the adapted parameter of Player  $i$ . We assume that the adapted parameter evolves according to a dynamic equation:

$$\hat{\theta}_{k+1}^i = \phi^i \left( \hat{\theta}_k^i, \bar{I}_{k+1}^{i, new}, \theta_i \right). \quad (7)$$

For infinite horizon games, the following property, which holds true for several learning rules, is quite interesting (see for example [46]):

**Property 1:** The adapted parameters  $\hat{\theta}_k^1, \dots, \hat{\theta}_k^N$  converge to some limits  $\hat{\theta}_\infty^1, \dots, \hat{\theta}_\infty^N$  such that the feedback (no memory) strategies  $\gamma_k^i(x_k, \theta_i, \hat{\theta}_\infty^i)$ ,  $i = 1, 2$  constitute a strongly time consistent Nash equilibrium ([8]) for the corresponding complete information game.

Most of the research in adaptive/learning rules focuses on convergence issues, and several converging learning rules have been proposed. Consider such a set of adaptive/learning rules in form (6), (7), and focus on Player  $i$ 's perspective. Player  $i$  faces a stochastic control problem with dynamics having state variable  $(x_k, \hat{\theta}_k^{-i})$  and incomplete state observation  $y_k = x_k$ . Furthermore, she has only a probabilistic description of the unknown parameters  $\theta, \theta_{-i}$ . Thus, Player  $i$  can use the information available in order to try to manipulate the

overall system. However, the control problem described is, in general, very difficult. In the following section, we restrict ourselves to a simpler class of manipulating strategies.

### 3 Pretending Strategies

In this section we focus on a special class of manipulation strategies called the pretender's strategies. Particularly, the manipulating player pretends to have a false type, probably in a response to the other players' actions. For simplicity, we assume that the value of  $\theta$  is commonly known, i.e.,  $\Theta = \{\theta\}$ . In what follows, only games with infinite horizon will be considered.

The general form of a pretender's strategy that corresponds to (6), (7) is:

$$\hat{\theta}_{k+1}^i = \phi^i \left( \hat{\theta}_k^i, \bar{I}_{k+1}^{i,new}, \theta_k^{i,pr} \right), \quad (8)$$

$$u_k^i = \gamma_k^i \left( x_k, \theta_k^{i,pr}, \hat{\theta}_k^i \right), \quad (9)$$

where the pretended type  $\theta_k^{i,pr}$  is given as an output of a system:

$$z_{k+1}^i = \phi^{i,pr} \left( z_k^i, \bar{I}_{k+1}^{i,new}, \theta_i \right), \quad (10)$$

$$\theta_k^{i,pr} = \psi^i \left( \theta_i, z_k^i \right). \quad (11)$$

Equations (8)–(11) represent a manipulating player who pretends to have a type that depends on her real type and a new, probably augmented, adapted parameter  $z_k^i$ . That is, in order to pretend adequately, it is probably useful to accumulate more information.

In what follows, we consider the stationary cases where one, some or all the players are pretending. These stationary outcomes are possible limit points of learning/adaptation rules, when the players are pretending.

#### 3.1 Optimal Stationary Pretending

We consider the possible limit points of the pretending strategies, assuming games with infinite horizon and long-run average cost. We assume that only Player  $i_0$  is pretending. In the spirit of Property 1, we analyze the following situation. Player  $i_0$  has discovered all the useful information for  $\theta_{-i_0}$ , and the pretended type of Player  $i_0$  has converged to  $\theta_{\infty}^{i_0,pr}$ . The rest of the players ( $-i_0$ ) react to a player of type  $\theta_{\infty}^{i_0,pr}$ . Furthermore, the set of strategies  $\gamma^{i_0}(x_k, \theta_{\infty}^{i_0,pr}, \hat{\theta}_{\infty}^{i_0})$ ,  $\gamma^j(x_k, \theta_j, \hat{\theta}_{\infty}^j)$ ,  $j \neq i_0$  constitute a strongly time consistent Nash equilibrium for the corresponding game with full information and types  $\theta_{\infty}^{i_0,pr}$ ,  $\theta_{-i_0}$ .

In order to define the optimal stationary pretending, the following assumption is made:

**Assumption 1** For any set of types  $\theta_1, \dots, \theta_N$ , there exists a unique a strongly time consistent Nash equilibrium of the corresponding full information game. Let us denote by  $\gamma_{\theta_1, \dots, \theta_N}^{i,N}(x_k)$ ,  $i = 1, \dots, N$  the set of strategies constituting the Nash equilibrium. Furthermore, fixing any subset of strategies,  $\gamma^i$  for  $i \in I_0 \subset \{1, \dots, N\}$  there is a unique set of strategies  $(\gamma^j)_{j \notin I_0}$  such that each  $\gamma^j$  is a best response (B.R.) to  $\gamma^1, \dots, \gamma^{j-1}, \gamma^{j+1}, \dots, \gamma^N$ .  $\square$

The optimal stationary pretending for Player  $i_0$  is given by:

$$\theta_{\infty}^{i_0, pr} = \operatorname{argmin}_{\tilde{\theta}_{i_0} \in \Theta_{i_0}} J_{i_0} \left( \gamma_{\tilde{\theta}_{i_0}, \theta_{-i_0}}^{i_0, N}, \gamma_{\tilde{\theta}_{i_0}, \theta_{-i_0}}^{-i_0, N} \right). \quad (12)$$

It is interesting to compare the cost that the pretending player  $i_0$  attains with the cost of the full information game having Player  $i_0$  as Stackelberg leader. Let us first recall the Stackelberg equilibrium notion under Assumption 1 (see for example [8]).

**Definition 1** Consider a set of strategies  $(\gamma_{\theta_{i_0}, \theta_{-i_0}}^{i_0, S}, \gamma_{\theta_{i_0}, \theta_{-i_0}}^{-i_0, S})$  within the feedback (no memory) class of strategies. This set constitutes a Stackelberg equilibrium with Player  $i_0$  as the leader if:

$$\gamma_{\theta_1, \dots, \theta_N}^{i_0, S} \in \operatorname{argmin}_{\gamma^{i_0}} \left\{ J_{i_0}(\gamma^{i_0}, \gamma^{-i_0}) : \text{for all } j \neq i_0, \gamma^j \text{ is a B.R. to } \gamma^{-j} \right\}. \quad (13)$$

and for all  $j \neq i_0$ ,  $\gamma_{\theta_{i_0}, \theta_{-i_0}}^{j, S}$  is a best response to  $\gamma^{-j}$ .  $\square$

We will use the term Stackelberg games, for games in which one player called the Stackelberg leader announces her strategy and sticks to it, while the other player or players decide their actions taking into account leader's announcements.

**Proposition 1** If Player  $i_0$  pretends optimally, then:

$$J_{i_0} \left( \gamma_{\theta_{i_0}, \theta_{-i_0}}^{i_0, S}, \gamma_{\theta_{i_0}, \theta_{-i_0}}^{-i_0, S} \right) \leq J_1 \left( \gamma_{\theta_{i_0}^{i_0, pr}, \theta_{-i_0}}^{i_0, N}, \gamma_{\theta_{i_0}^{i_0, pr}, \theta_{-i_0}}^{i_0, N} \right). \quad (14)$$

An equality is attained if there exist a  $\tilde{\theta}_{i_0} \in \Theta_{i_0}$  such that  $\gamma_{\theta_{i_0}, \theta_{-i_0}}^{i_0, S} = \gamma_{\tilde{\theta}_{i_0}, \theta_{-i_0}}^{i_0, N}$ .

*Proof* The strategy class  $\{\gamma_{\tilde{\theta}_{i_0}, \theta_{-i_0}}^{i_0, N} : \tilde{\theta}_{i_0} \in \Theta_{i_0}\}$  is a subclass of the feedback (no memory) strategy class. Furthermore, using Assumption 1, there is a unique set of strategies  $(\gamma_{\tilde{\theta}_{i_0}, \theta_{-i_0}}^{-i_0, N})$  such that the strategy of all the players except  $i_0$  be a best response to the strategies of the others. Hence, the minimum in (13) is less than or equal to the minimum in (12).  $\square$

*Remark 1* Proposition 1 says that if there is enough uncertainty and only one player pretends, then this player attains the same cost, as if she were a Stackelberg leader.  $\square$

### 3.2 The Pretenders' Game

In this subsection, we study the case where two or more players are pretending. Under Assumption 1, an auxiliary game, called the pretenders' game, can be defined.

**Definition 2** For the game described by (1), (3), under Assumption 1, the corresponding pretenders' game involves the same players, the action of Player  $i$  is  $\theta^{i, pr} \in \Theta_i$  and the cost is given by:

$$\bar{J}_i(\theta^{i, pr}, \theta^{-i, pr}) = J_i \left( \gamma_{\theta^{i, pr}, \theta^{-i, pr}}^{i, N}, \gamma_{\theta^{i, pr}, \theta^{-i, pr}}^{-i, N}, \theta_i \right), \quad (15)$$

where  $\theta_i$  is the actual type. Let us also introduce the  $I_0$ -constrained pretenders' game as the corresponding game where the actions of a set of players are constrained to coincide with the true value of the players' types, i.e.,  $\theta^{i, pr} = \theta_i$ , for  $i \in I_0$ .

An interesting point is the Nash equilibrium of the pretenders' game. This point can serve as an alternative prediction of the outcome of the original game.

**Definition 3** For the game described by (1), (3), under Assumption 1, a Nash pretenders' outcome is the set of strategies  $(\gamma_{\theta_{\infty}^{1,pr}, \dots, \theta_{\infty}^{N,pr}}^{1,N}, \dots, \gamma_{\theta_{\infty}^{1,pr}, \dots, \theta_{\infty}^{N,pr}}^{N,N})$ , where  $(\theta_{\infty}^{1,pr}, \dots, \theta_{\infty}^{N,pr})$  is an equilibrium of the corresponding pretenders' game. The constrained Nash pretenders' outcome is defined similarly assuming that  $(\theta_{\infty}^{1,pr}, \dots, \theta_{\infty}^{N,pr})$  is an equilibrium of the constrained pretenders' game.

The Nash pretenders' outcome is interesting as a possible limit point of a learning/adaptive algorithm, where each of the players pretends dynamically to have a false type.

*Remark 2* The pretending outcomes described, i.e., the optimal stationary pretending for one player and the Nash pretenders' outcome when all the players are pretending, are at least as reasonable as the application of the original dynamic rules, due to the fact that each player may have a motivation to pretend.

Let us note that there are also other works studying deviations from best response. For example, in smooth fictitious play, for repeated finite games (ex. Ch 4 of [16]), the players are slightly deviating from their best response toward a randomized policy. This is different from pretending policies, due to the fact that in the later case the deviations from best response are used to induce a desired behavior for the other players.  $\square$

*Remark 3* Let us comment on the relationship with the work in [4]. This work identifies an interesting phenomenon in routing games with partially altruistic players: In some cases “When a user increases its degree of cooperation while other users keep their degree of cooperation unchanged, leads to performance improvement of that user.” To be more precise, if  $J_i$  is the “selfish” cost of Player  $i$ , [4] considers the case where the players are partially altruistic, that is each player tries to minimize:

$$\tilde{J}_i = (1 - a_i)J_i + a_i J_{-i}, \quad (16)$$

where  $a_i$  is the degree of cooperation. It is then shown numerically that, when the players apply strategies constituting a Nash equilibrium of the game described by (16), it is possible that  $J_i$  decreases, as  $a_i$  increases. This phenomenon is called the “Paradox in cooperation.” In this context the notion of the price of unilateral altruism was introduced [5].

Consider a situation where the degree of cooperation of Player 2 is commonly known, whereas the degree of cooperation of Player 1 is known only to Player 1 and there is a “Paradox in cooperation.” Then, Player 1 has an incentive to pretend to be more altruistic than she actually is. Furthermore, even if Player 2 knew the actual degree of cooperation of Player 1, she would have an incentive to accept that Player 1 is more altruistic; otherwise, she would induce a less altruistic behavior to Player 1 and thus hurt herself. Finally, Player 2 cannot distinguish between the case that Player 1 is altruistic and tries to reach a Nash equilibrium (of the corresponding full information game) and the case that Player 1 is not altruistic and pretends.  $\square$

## 4 Quadratic Opinion Games

In this section, a simple class of repeated, quadratic games is analyzed. In this class of games, exact results can be obtained. We first analyze a two-player simultaneous move (Nash)



repeated game and then a sequential move (Stackelberg) repeated game. We then move to symmetric and asymmetric games with many players. The results of this section help us to get some intuition on the relationships of pretending and strategic leadership. Throughout this section we assume long-time average costs.

#### 4.1 A Two-Player Simultaneous Move Game

A two-player repeated opinion game in strategic form is studied. The cost function is given by (3) and the instantaneous costs by:

$$L_1 = (u^1 - \theta_1)^2 + (u^1 - u^2)^2, \quad (17)$$

$$L_2 = (u^2 - \theta_2)^2 + (u^2 - u^1)^2, \quad (18)$$

where  $(\theta_1, \theta_2) \in \mathbb{R}^2$ . We assume an information structure given by (5). The full information (static) game has a unique Nash equilibrium:

$$u^{i,N} = \frac{2}{3}\theta_i + \frac{1}{3}\theta_{-i}, \quad i = 1, 2. \quad (19)$$

Several adaptive (iterative) techniques were studied in [32]. Probably, the simplest one is the best response map:

$$u_k^i = (\theta_i + \hat{\theta}_k^i) / 2, \quad (20)$$

$$\hat{\theta}_k^i = u_{k-1}^{-i}. \quad (21)$$

If both players follow their best response maps, their actions will converge to the Nash equilibrium of the full information static game.

Let us then study the case where Player 1 tries to manipulate Player 2 and Player 2 follows (20), (21). Due to the fact that the map  $\theta_1 \mapsto u^{1,N}$  in (19) is onto, Proposition 1 applies. Thus, the feedback Stackelberg cost for Player 1 is feasible through pretending.

The optimal pretending for Player 1 is given by:

$$\theta_\infty^{1,pr} = \frac{6}{5}\theta_1 - \frac{1}{5}\theta_2. \quad (22)$$

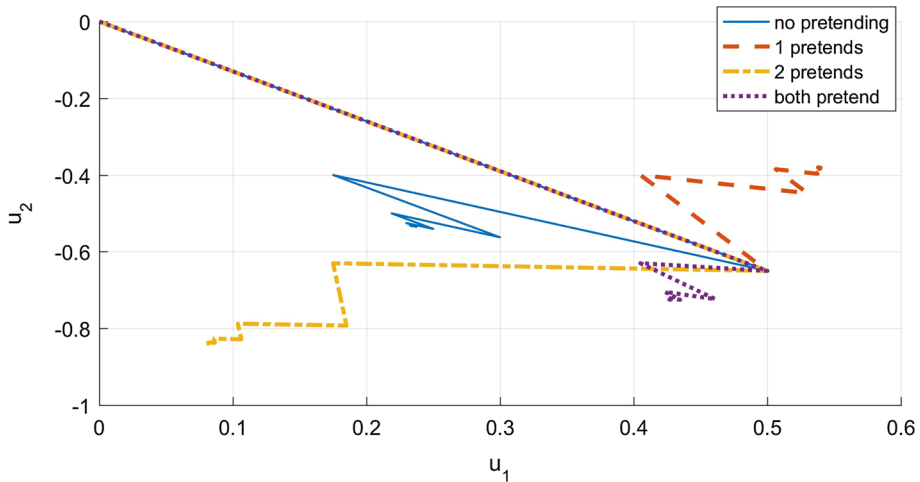
Player 1 can use several ways to learn  $\theta_2$  in order to implement her pretending policy. One way is to use only the last iteration. Particularly:

$$z_{k+1}^{1,1} = z_k^{1,2}, \quad (23)$$

$$z_{k+1}^{1,2} = u_k^1, \quad (24)$$

$$\theta_k^{1,pr} = \frac{6}{5}\theta_1 - \frac{1}{5}\left(2u_{k-1}^2 - z_k^{1,1}\right). \quad (25)$$

Figure 1 shows the action trajectories when no player, one player and both players are pretending. The parameters are  $\theta_1 = 1$  and  $\theta_2 = -1.3$ . If no player is pretending, then the dynamic rules lead to the Nash equilibrium. If a single player is pretending, the dynamic rule converges to the Stackelberg equilibrium having the pretending player as a leader. Finally, in the case where both players are pretending the dynamic rule converges to the Nash pretenders' outcome.



**Fig. 1** The *solid line* corresponds to the best response with no pretending players, the *dashed line* line to Player 1 pretending, the *dash-dot line* line to Player 2 do so and the *dotted line* to the case in which both players pretend. It turns out that the outcomes in the several cases studied are widely different

## 4.2 Two-Player Sequential Move Games

We then study a dynamic game which represents the repeated version of a static Stackelberg game (or equivalently a repeated extensive form game). Particularly, Player 1 acts at time instants 1, 3, 5, ... and Player 2 at 2, 4, 6, ... The feedback Nash equilibrium of the dynamic game corresponds to the Stackelberg equilibrium of the static game.

The instantaneous costs are given by:

$$L^1(k) = \begin{cases} (x_k - 1)^2 + \theta_1 (x_k - u_k^2)^2, & k = 2, 4, \dots \\ 0 & \text{otherwise} \end{cases} \quad (26)$$

$$L^2(k) = \begin{cases} (u_k^2 - 2)^2 + \theta_2 (x_k - u_k^2)^2, & k = 2, 4, \dots \\ 0 & \text{otherwise} \end{cases} \quad (27)$$

and the state equation by:

$$x_{k+1} = \begin{cases} u_k^1 & \text{if } k = 1, 3, 5, \dots \\ \emptyset & \text{if } k = 2, 4, 6 \end{cases} \quad (28)$$

It is not difficult to prove that the Nash equilibrium (which coincides with the Stackelberg equilibrium of the corresponding static game) is given by:

$$u_k^1 = \frac{(1+\theta_2)^2 + 2\theta_1}{(1+\theta_2)^2 + \theta_1}, \quad k = 1, 3, 5, \dots \quad (29)$$

$$u_k^2 = \frac{2+\theta_2 x_k}{1+\theta_2}, \quad k = 2, 4, 6, \dots \quad (30)$$

To prove this fact, observe that (30) represents a best response for Player 2 at time step  $k = 2\nu$ , given the action of Player 1 at time step  $2\nu - 1$ , where  $x_k = u_{k-1}^1$ . Then, substituting (30) into (26) and minimizing for  $x_k = u_{k-1}^1$ , we get (29).

A simple dynamic rule that leads to the Nash equilibrium is given by:

$$u_1^1 = 1, \quad (31)$$

$$u_k^1 = \frac{(1 + \hat{\theta}_1)^2 + 2\theta_1}{(1 + \hat{\theta}_1)^2 + \theta_1}, \quad k = 3, 5, 7 \dots \quad (32)$$

$$u_k^2 = \frac{2 + \theta_2 x_k}{1 + \theta_2}, \quad k = 2, 4, 6 \dots \quad (33)$$

where  $\hat{\theta}_1$  is the least squares estimate of  $\theta_2$  based on (30). To see that (31)–(33) converge to the Nash equilibrium in a finite number of steps, first observe that at time step  $k = 3$ , Player 1 will know the value of  $\theta_2$  and that for  $\hat{\theta}_1 = \theta_2$  (32), (33) coincide with (29), (30).

A very simple manipulation rule for Player 2 is to use (33) with  $\theta_k^{2,pr} = (\varepsilon - 1)$  in place of  $\theta_2$ , where  $\varepsilon$  is a small positive constant. Using these rules, after two time steps, the actions of the players will converge to:

$$u^1 = \frac{\varepsilon^2 + 2\theta_1}{\varepsilon^2 + \theta_1}, \quad (34)$$

$$u^2 = \frac{\varepsilon^2 + \varepsilon + 2\theta_1}{\varepsilon^2 + \theta_1}. \quad (35)$$

As  $\varepsilon$  decreases and approaches 0, actions of both of players approach the value 2. In this case, the cost of Player 2 approaches 0.

**Remark 4** The Stackelberg leader (Player 1) behaves approximately as the follower wants her to behave. The situation is quite similar to an inverse Stackelberg game with Player 2 as the leader [24]. Thus, Player 2 can achieve a very strong form of leadership and in this sense, we may conclude that:

“At least in some cases, regarding leadership, the best use of the information available is much more important than the commitment or the timing of the game.”

**Remark 5** Player 2 can achieve approximately the same cost, as if she was acting on Player 1’s behalf. An interesting question arises: “In which sense does this outcome express the ‘free will’ of Player 1”?

### 4.3 Many Player Symmetric Games

We then move to a slightly more general case, where there are  $N$  symmetrically interacting players. The instantaneous costs have the form:

$$L_i = (u^i - \theta_i)^2 + \left( u^i - \frac{1}{N-1} \sum_{j \neq i} u^j \right)^2. \quad (36)$$

The corresponding full information game has a unique feedback Nash equilibrium. It is not difficult to show that the equilibrium is given by:

$$u^i = \frac{N\theta_i + \sum_{j \neq i} \theta_j}{2N-1}. \quad (37)$$

We then consider the pretenders' game. Using  $\theta^{i,pr}$  in the place of  $\theta^i$  in (37) and substituting back in (36) we get the costs in the corresponding pretenders' game:

$$\bar{J}_i = \frac{1}{(2N-1)^2} \left[ \left( N\theta^{i,pr} + \sum_{j \neq i} \theta^{j,pr} - (2N-1)\theta^i \right)^2 + \left( (N-1)\theta^{i,pr} - \sum_{j \neq i} \theta^{j,pr} \right)^2 \right]. \quad (38)$$

Considering the partial derivative of  $\bar{J}_i$  with respect to  $\theta^{i,pr}$  and solving for  $\theta^{i,pr}$  we obtain the following characterization of the equilibrium of the pretenders' game:

$$\theta^{i,pr} = \frac{2N^2 - N}{2N^2 - 2N + 1} \theta^i - \frac{1}{2N^2 - 2N + 1} \sum_{j \neq i} \theta^{j,pr}. \quad (39)$$

Hence, the equilibrium of the pretenders' game is given by:

$$\theta^{i,pr} = \left( 1 + \frac{1}{2N} \right) \theta^i - \frac{1}{2N(N-1)} \sum_{j \neq i} \theta^j. \quad (40)$$

**Remark 6** Equation (40) shows that when the interactions are symmetric, the players tend to pretend less as the number of the players increases. Furthermore, the value of pretending decreases and the Nash pretenders' outcome approaches the Nash equilibrium, as the number of players increases.

#### 4.4 Many Player Asymmetric Games

We then consider a quadratic game with  $N$  players having asymmetric interactions. The instantaneous cost of each player is given by:

$$L_i = \left( u^i - \theta_i \right)^2 + \left( u^i - \sum_{j=1}^N a_{ij} u^j \right)^2, \quad (41)$$

where  $a_{ij}$  represents the influence of the actions of Player  $j$  to Player  $i$  or, in the opinion game setting, how much Player  $i$  is interested in Player  $j$ 's opinion. We assume that  $\sum_{j=1}^N a_{ij} = 1$  and that  $a_{ii} = 0$ , for any  $i = 1, \dots, N$ . The matrix  $A = [a_{ij}]$  will be used. We further assume that  $\theta_i \in [-M, M]$ .

The Nash equilibrium is characterized by:

$$u = \frac{1}{2} \theta + \frac{1}{2} A u, \quad (42)$$

where  $\theta = [\theta_1, \dots, \theta_N]^T$  and  $u = [u^1, \dots, u^N]^T$ . There exists a unique Nash equilibrium:

$$u = \frac{1}{2} (I - \frac{1}{2} A)^{-1} \theta = \frac{1}{2} A^+ \theta, \quad (43)$$

where  $A^+ = I + \frac{1}{2} A + \frac{1}{4} A^2 + \dots$ . The series in  $A^+$  converges due to the fact that  $A$  is a stochastic matrix.

For each Player  $i$ , let us introduce the following quantity:

$$C_i = e_i^T \left( \sum_{t=2}^{\infty} A^t / 2^t \right) e_i. \quad (44)$$

We call  $C_i$  the “centrality of Player  $i$ .” The optimal pretending of each player will be expressed in terms of  $C_i$ .

*Remark 7* The quantity  $C_i$  has some similarities with the Bonacich centrality measures [9] and the PageRank index [30]. The centrality  $C_i$  may have the following representation: Assume that the matrix  $A$  stands for the transition probabilities of a Markov chain. Consider a random walk on a graph, each node of which represents a player. At each time step, the random walk with probability  $1/2$  stops and with probability  $1/2$  jumps probabilistically to a new node according to  $A$ . Then,  $C_i$  is the expected number of times that a random walk started at the node of Player  $i$  will visit that node again.  $\square$

Let us focus on a certain Player  $i$ . Consider a profile of (real or pretended) types  $\tilde{\theta} = [\tilde{\theta}_1, \dots, \tilde{\theta}_{i-1}, \theta_i, \tilde{\theta}_{i+1}, \dots, \tilde{\theta}_N]^T$ . The pretended type of Player  $i$  can be written as  $\theta_i^{i,pr} = \theta_i + \delta^i$ . The cost of Player  $i$  in the pretenders’ game is given by:

$$\begin{aligned} \tilde{J}_i = & \left( \frac{1}{2} e_i^T A^+ \tilde{\theta} - \theta_i + \frac{1}{2} e_i^T A^+ e_i \delta^i \right)^2 \\ & + \left( \frac{1}{2} e_i^T A^+ \tilde{\theta} + \frac{1}{2} e_i^T A^+ e_i \delta^i - \frac{1}{2} e_i^T A A^+ \tilde{\theta} - \frac{1}{2} e_i^T A A^+ e_i \delta^i \right)^2. \end{aligned}$$

Observing that  $e_i^T A^+ e_i = 1 + C_i$ , that  $e_i^T A A^+ e_i = 2C_i$  and that

$$\alpha = \frac{1}{2} e_i^T A^+ \tilde{\theta} - \theta_i = - \left[ \frac{1}{2} e_i^T A^+ \tilde{\theta} - \frac{1}{2} e_i^T A A^+ \tilde{\theta} \right],$$

the cost is written as:

$$\tilde{J}_i = \left( \frac{1 + C_i}{2} \delta^i + \alpha \right)^2 + \left( \frac{1 - C_i}{2} \delta^i - \alpha \right)^2. \quad (45)$$

Hence, the optimal pretending is given by:

$$\delta^{i,*} = \frac{2C_i}{1 + C_i^2} \alpha = \frac{2C_i}{1 + C_i^2} \left[ \frac{1}{2} e_i^T A^+ \tilde{\theta} - \theta_i \right]. \quad (46)$$

**Proposition 2** *There exists an equilibrium in the pretenders’ game.*

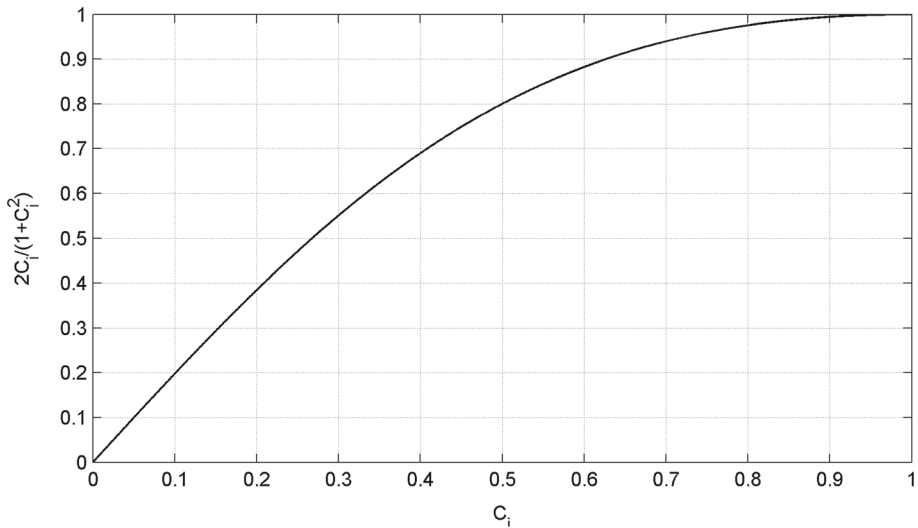
*Proof* Consider the set  $B = [-2M, 2M]^N$ . The following claim will be used:

*Claim:* If  $\tilde{\theta} \in B$ , then it holds  $\theta_i + \delta^{i,*} \in [-2M, 2M]$ , for each player  $i$ .  $\square$

To prove the claim, observe that  $C_i \in [0, 1]$  and thus  $2C_i/(1 + C_i^2) \in [0, 1]$ . Furthermore,  $\frac{1}{2} e_i^T A^+ \tilde{\theta} \in [-2M, 2M]$ . Hence,

$$\theta_i + \delta^{i,*} = \left( \frac{2C_i}{1 + C_i^2} \right) \frac{1}{2} e_i^T A^+ \tilde{\theta} + \left( 1 - \frac{2C_i}{1 + C_i^2} \right) \theta_i.$$

The fact that  $\theta_i + \delta^{i,*}$  is a convex combination of two quantities in  $[-2M, 2M]$  completes the proof of the claim.



**Fig. 2** The plot of the function  $f(C_i) = \frac{2C_i}{1+C_i^2}$

The claim shows that the best response maps  $B$  into  $B$ . Furthermore,  $B$  is compact and convex and the map is continuous. Thus, Brouwer fixed point theorem applies and the pretenders' game has an equilibrium.  $\square$

Equation (46) shows that the pretending of each player depends on its centrality  $C_i$ . The plot of the function  $f(C_i) = 2C_i / (1 + C_i^2)$  is shown in Fig. 2. Furthermore,  $\alpha$  is bounded:  $|\alpha| \leq 2M$ . Thus, a player with low centrality tends to pretend less and the value of pretending is small.

In the following example, a game involving an influential player and many less influential players is analyzed.

*Example 1* Consider a game with  $N + 1$  players. The instantaneous cost for Player 0 is given by:

$$L_0 = (u^0 - \theta_0)^2 + \left( u^0 - \frac{1}{N} \sum_{j=1}^N u^j \right)^2. \quad (47)$$

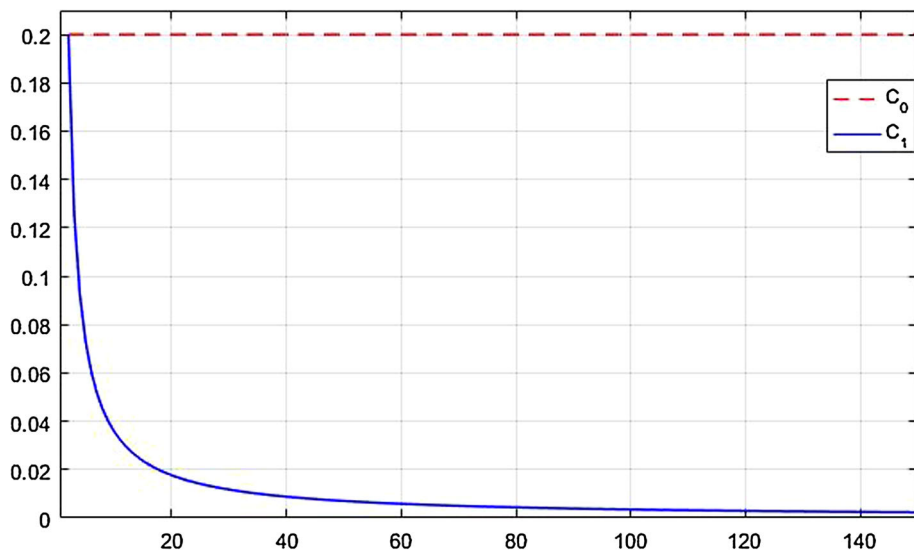
The instantaneous cost for each one of the players  $i = 1, \dots, N$  is given by:

$$L_i = (u^i - \theta_i)^2 + \left( u^i - au^0 - \frac{1-a}{N-1} \sum_{j \notin \{0,i\}} u^j \right)^2. \quad (48)$$

The matrix  $A$  has the form:

$$A = \begin{bmatrix} 0 & \frac{1}{N} \mathbf{1}^T \\ a \mathbf{1} & \frac{1-a}{N-1} (\mathbf{1} \mathbf{1}^T - I) \end{bmatrix}.$$

Figure 3 illustrates the centrality  $C_0$  of the influential player 0 and the centrality  $C_i$  of any other player  $i$ , as the number of players increases. The centrality computations were made for  $a = 0.5$ .



**Fig. 3** The centrality of the influential player  $C_0$  and the centrality  $C_i$  of another player in Example 1

As the number of players increases, only the influential player has a non-negligible motivation to pretend. Furthermore, the Nash pretenders' outcome approaches the Stackelberg equilibrium with the influential player as the leader.  $\square$

*Remark 8* In Example 1, the players do not have differences in timing, commitment or the processing of the available information. However, the Nash pretenders' outcome approaches the Stackelberg equilibrium as the number of players increases. This Stackelberg leadership appears intrinsically, without assuming different roles for the players a priori. Thus, pretending may serve as a natural Stackelberg leadership explanation.

Another way to understand, why when there is only one influential player she may acquire a leader role, is using the centrality measure  $C_i$ . A high value for  $C_i$  means that there are some players who are affected by Player  $i$ 's pretended type, while at the same time Player  $i$  is interested in the actions of those same players. A small player is less able to affect through pretending the actions of the players which in turn affect her.

## 5 An LQ Non-Cooperative Pollution Control Game

In this section, we consider a very simple linear quadratic pollution control game with two players. The full information, continuous-time analogue of this model was analyzed in [11]. Each Player  $i = 1, 2$  stands for a country. The dynamics is given by:

$$x_{k+1} = ax_k + u_k^1 + u_k^2. \quad (49)$$

The state variable  $x_k$  represents the pollution stock, and the actions of the players  $u_k^1$  and  $u_k^2$  stand for the emissions of the two countries. The pollution stock of the environment decreases naturally with a rate  $0 < a < 1$ .

The cost functions of the players are given by:

$$J_i = \sum_{k=0}^{\infty} \rho_i^k \left[ qx_k^2 + (u_k^i)^2 - u_k^i \right]. \quad (50)$$

Cost (50) comprises the effects of two factors. The economic part is expressed by the difference  $(u_k^i)^2 - u_k^i$ , whereas the term  $qx_k^2$  corresponds to the environmental cost.

We assume that both players care the same about pollution, i.e., the parameter  $q > 0$  is common among the players and that they have the same economic loss from reducing their emissions. However, they may have a different discount factor  $\rho_i \in [0.9, 0.99]$ . We assume that the players know the parameters  $q$  and  $a$  and each player knows its own discount factor. We assume, however, that each player is not aware of the discount factor of the other player.

Consider a pair of strategies  $u_k^1 = K_1^1 x_k + K_2^1$ ,  $u_k^2 = K_1^2 x_k + K_2^2$  constituting a Nash equilibrium of the corresponding full information game. Then, this pair is characterized by:

$$p^i = q + \frac{\rho_i p^i}{1 + \rho_i p^i} (\bar{a}^i)^2, \quad (51)$$

$$c_1^i = \frac{\rho_i \bar{a}^i (2p^i v^i + c_1^i)}{1 + \rho_i p^i}, \quad (52)$$

$$K_1^i = -\frac{\rho_i \bar{a}^i p^i}{1 + \rho_i p^i}, \quad (53)$$

$$K_2^i = \frac{1}{2} - \rho_i \frac{2p^i v^i + c_1^i}{2(1 + \rho_i p^i)}, \quad (54)$$

where  $\bar{a}^i = a + K_1^{-i}$  and  $v^i = K_2^{-i} + 1/2$  and  $p^1, p^2 > 0$ . The details of the optimal control problems are given in “Appendix sections A and B.”

**Proposition 3** *If  $0 < a < 1$ , there exists a solution of (51)–(54) with  $p^1, p^2 > 0$ . Thus, there exists a Nash equilibrium.*

*Proof* See “Appendix C.” □

Assuming that each player is not aware of the discount factor of the other player, a simple adaptive rule, based on an estimation  $\hat{\rho}_{-i}$  of the opponent’s discount factor, may be used. The rule for Player  $i$  is described in Algorithm 1.

### Algorithm 1:

- 1: At step 0, use  $K_1^i(0), K_2^i(0)$  which is a best response to  $u_k^{-i} \equiv 0$ .
- 2: At time step  $k \geq 1$ , find a value for  $\hat{\rho}_{-i}$  that minimizes the following criterion:

$$\hat{\rho}_{-i} \in \arg \min_{\rho \in [0.9, 0.99]} \left[ \left( u_k^{-i} - K_1^{-i, \rho, \star} x_{k-1} + K_2^{-i, \rho, \star} \right)^2 \right], \quad (55)$$

where  $(K_1^{-i, \rho, \star}, K_2^{-i, \rho, \star})$  is the best response of a player with discount factor  $\rho$  to a strategy  $(K_1^i(k-1), K_2^i(k-1))$ .

- 3: Apply a strategy  $K_1^i(k), K_2^i(k)$  which is the best response to an opponent policy given by  $K_1^{-i, \hat{\rho}_{-i}, \star}$  and  $K_2^{-i, \hat{\rho}_{-i}, \star}$ .
- 4: Move to the next step  $k+1$  and jump to Step 2.



**Remark 9** Step 2 of Algorithm 1 (also Step 2 of Algorithm 2) estimates the discount factor of the opponent based only on her last action. Of course several alternatives are available, but (55) is probably the simplest one.  $\square$

A simple pretending strategy for Player  $i$  is given in Algorithm 2. In Algorithm 2, Player  $i$  uses the estimated discount factor of her opponent in order to pretend optimally.

**Algorithm 2:**

- 1: At step 0, use  $K_1^i(0)$ ,  $K_2^i(0)$  which is a best response to  $u_k^{-i} \equiv 0$ .
- 2: At time step  $k \geq 1$ , find  $\hat{\rho}_{-i}$  according to (55) and fix an “assumed” opponent strategy  $K_1^{-i, \hat{\rho}_{-i}, \star}$  and  $K_2^{-i, \hat{\rho}_{-i}, \star}$ , as in Algorithm 1.
- 3: Find a pretended discount  $\rho^{i, pr}$  that solves:

$$\begin{aligned} & \underset{\rho}{\text{minimize}} \quad J_i(\gamma^i, F(\gamma^i, \hat{\rho}_{-i})) \\ & \text{s.t.} \quad \gamma^i = F(K_1^{-i, \hat{\rho}_{-i}, \star}, K_2^{-i, \hat{\rho}_{-i}, \star}, \rho) \\ & \quad \quad 0.9 \leq \rho \leq 0.99 \end{aligned} \tag{56}$$

where  $F(K_1, K_2, \rho)$  or  $F(\gamma, \rho)$  is the best response of a player with discount factor  $\rho$  to an opponent with strategy given by  $K_1$ ,  $K_2$  or  $\gamma$ . The computation of  $J_i(\gamma^i, \gamma^{-i})$  is given in Section 1 of Appendix.

- 4: Use the strategy  $F(K_1^{-i, \hat{\rho}_{-i}, \star}, K_2^{-i, \hat{\rho}_{-i}, \star}, \rho^{i, pr})$ .
- 5: Move to the next step  $k + 1$  and jump to Step 2.

**Example 2** In this example we consider two players with discount factors 0.95. Furthermore, we assume that the parameters have values  $q = 0.1$  and  $a = 0.8$ .

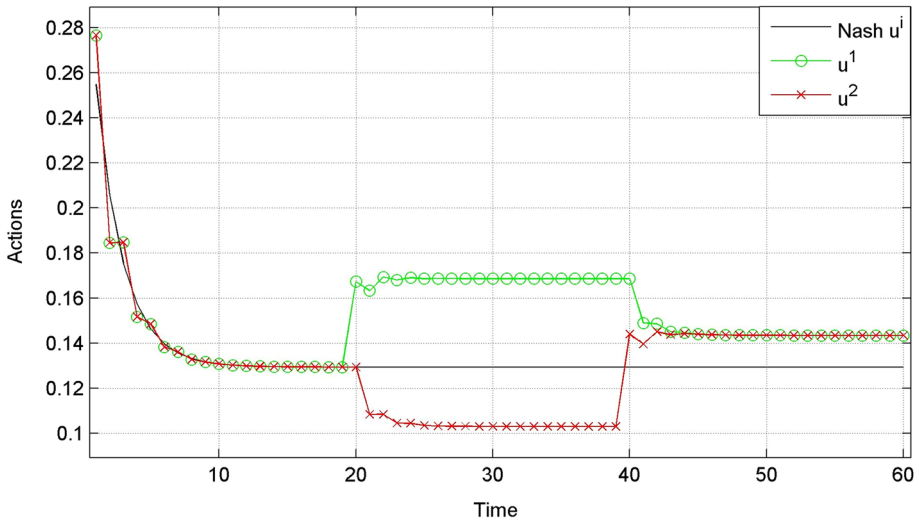
At first, no player is pretending. After time step 20, Player 1 pretends applying Algorithm 2 and after time step 40 both players apply Algorithm 2.<sup>1</sup> For comparison, the trajectories of the players’ actions and state variable for the Nash equilibrium of the full information game are also presented. In Fig. 4 the actions of the players are presented. Fig. 5 shows the evolution of the pollution stock and Fig. 6 presents the evolution of the instantaneous cost for both players.

**Remark 10** The numerical simulations show that the application of the dynamic rule described in Algorithm 1 has very similar results to the Nash equilibrium of the corresponding full information game. In the case of pretending, the manipulating player has the motivation to pretend that she has a lower discount factor (i.e., she cares less about the future) and relies to the other player to reduce its emissions. If one or both players pretend, the pollution stock increases. The non-pretending player needs to reduce its emissions, and there is also a larger pollution. Thus, pretending in this example is not a cooperative action.

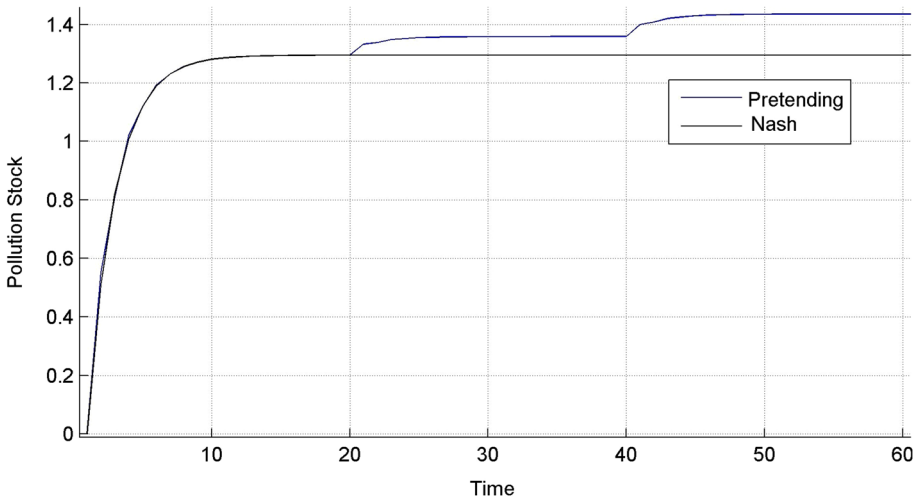
## 6 Electricity Market Models

In this section two competitive market models are analyzed. The first considers a linear supply function duopoly model. A simple pretending algorithm is proposed, and the numerical

<sup>1</sup> The reason we do not assume that none, one or both players pretend from the beginning is to use a single pair of figures to illustrate the results.

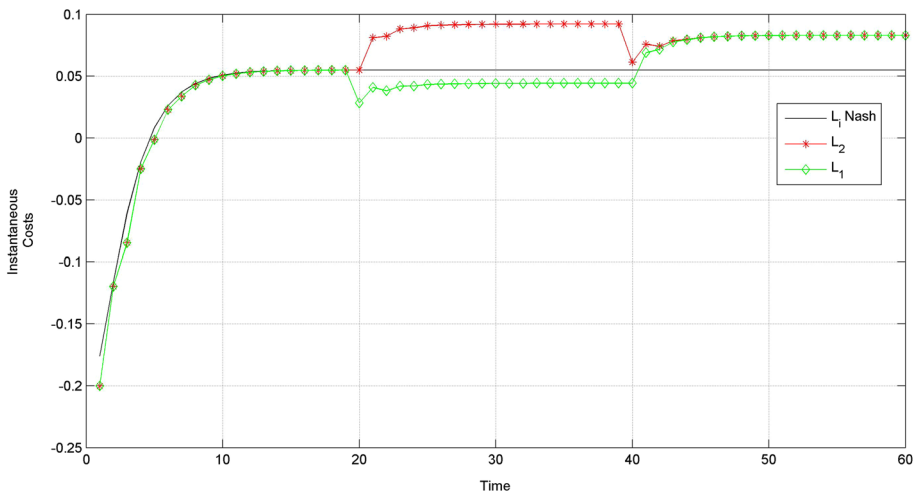


**Fig. 4** The evolution of players' actions in Example 2. In the first period of time, in which no one pretends the emissions of the players are identical and their value is very close to the full information Nash equilibrium. After time step 20, Player 1 pretends to be shortsighted and her pollution emissions increase. This fact makes Player 2 to reduce her emissions. After time step 40, both players pretend to be shortsighted. Now, Player 2 increases her emissions and makes Player 1 to decrease them. In this case, the total emissions are larger than the case where no one pretends



**Fig. 5** The evolution of the pollution stock in Example 2. The pollution stock increases after time step 20 and increases further after time step 40, where both players pretend to have a lower discount factor. In both cases pollution stock increases, despite the fact that only one player is increasing her emissions, while the other is decreasing

simulation results show that pretending is a cooperative action, i.e., it increases the firms' profits, while the price also increases. The second example considers a mechanism from the literature [34]. It is shown numerically that the attempt of the players to manipulate may make the mechanism not working at all.



**Fig. 6** The instantaneous costs of the players in Example 2. In the first part of the simulation, where both the players follow their adaptive rules with no pretending, the instantaneous costs are very close to their Nash values. After time step 20, Player 1 attains a lower cost, but Player 2 is bound to emit less and also have a more polluted environment. Thus, the instantaneous cost of Player 2 is larger than its Nash value and this difference is larger than the improvement for Player 1's cost. After time step 40, the instantaneous cost is improved for Player 2, but remains worse than the Nash cost for both the players

## 6.1 Linear Supply Function Model

In this subsection, we assume that there are two competing production firms. Each one of them proposes a linear supply function [6] that is at each price level produces an energy quantity proportional to that price. We also consider linear demand of the form  $q = A - Bp$ , where  $q$  is the total energy demand and  $p$  the price. The instantaneous cost of each player is given by:

$$L_i = c_i q_i^2 - p q_i, \quad (57)$$

where  $q_i$  is the amount of energy produced by Player  $i$ , the quantity  $c_i q_i^2$  represents the production cost and  $p q_i$  the revenues for Player  $i$ .

Each player chooses the constant  $u_i$  of the linear supply function, that is the quantity produced depends on the price according to the linear relation  $q_i = u_i p$ . We assume that each player knows her own production cost but not the production cost of her opponent.

Simple manipulations show that the price and the instantaneous cost of the players depend on  $u_1, u_2$  according to:

$$p = \frac{A}{u_1 + u_2 + B}, \quad (58)$$

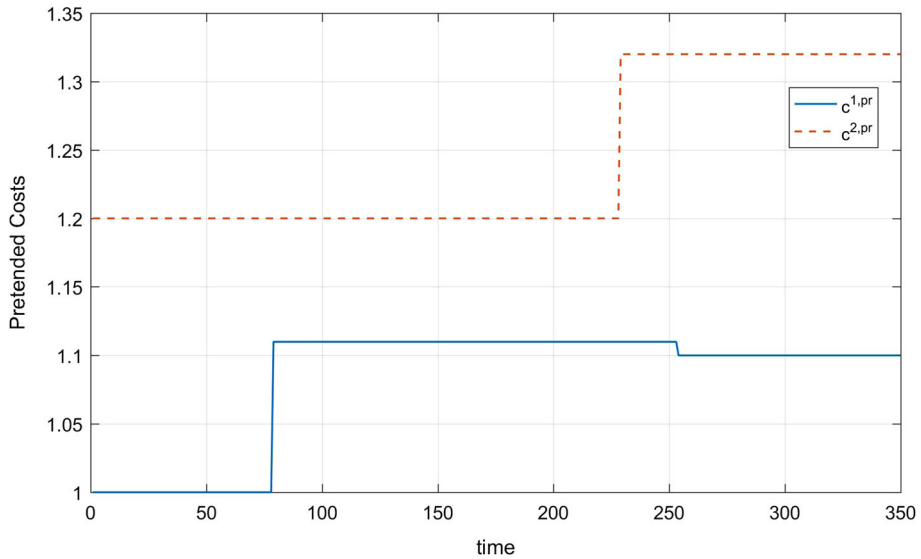
and

$$L_i = A^2 \frac{c_i u_i^2 - u_i}{(u_1 + u_2 + B)^2}. \quad (59)$$

The best response maps are given by:

$$u_1^+ = \frac{B + u_2}{2c_1 u_2 + 2c_1 B + 1}, \quad (60)$$

$$u_2^+ = \frac{B + u_1}{2c_2 u_1 + 2c_2 B + 1}. \quad (61)$$



**Fig. 7** The pretended production costs  $c^{1,pr}$ ,  $c^{2,pr}$  in Example 3. Player 1 implements Algorithm 3 from the beginning, but the steps 6 and 7 are used for the first time at  $k = 78$ . Both of the players have an interest to pretend that they have a larger production cost compared to their actual. After time 200, where Player 2 is using Algorithm 3 and moves to her pretending strategy, the motivation for Player 1 to pretend is slightly reduced and thus the next time she reviews her pretended cost, i.e., at time step  $k = 253$ , she moves slightly toward her real cost

It is easy to see that the best response map is contractive with respect to the infinity norm. Thus, if the players use repeatedly (60), (61) the system will approach the Nash equilibrium of the static full information game.

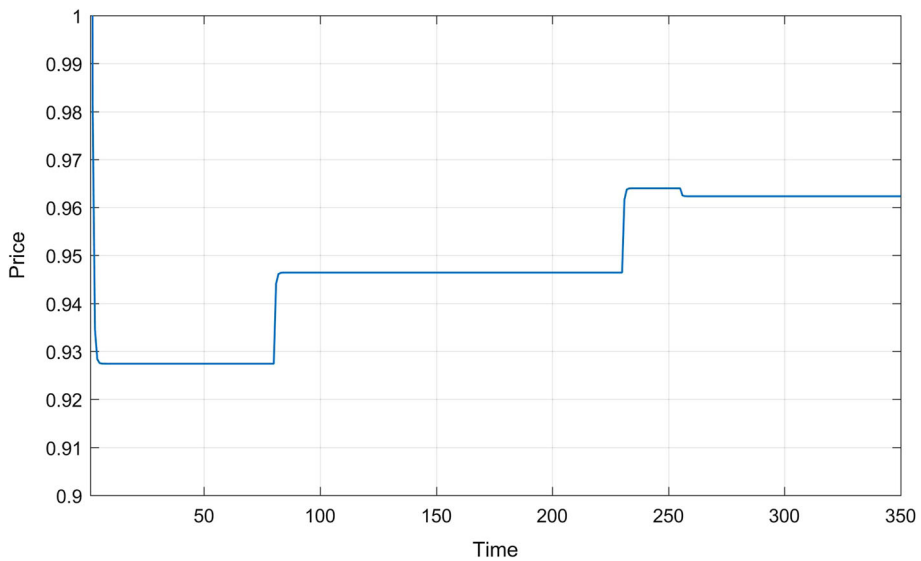
A simple pretending strategy for Player 1 is described in Algorithm 3.

### Algorithm 3:

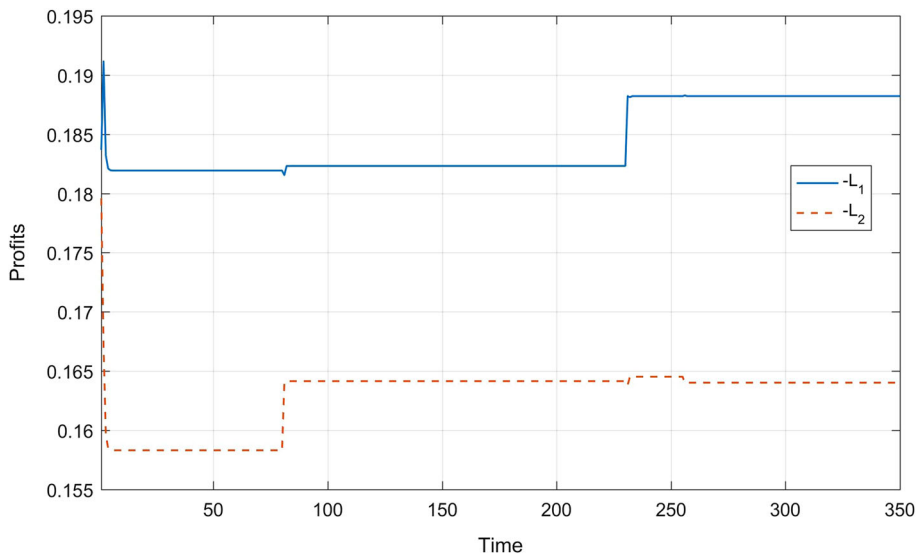
- 1: Choose  $c_0^{1,pr} = c_1$ .
- 2: At time step  $k$  choose  $u_1$  according to (60) with  $c_k^{1,pr}$  in the place of  $c_1$ .
- 3: Estimate  $c_2$  from the last  $M$  measurements according to (61) using least squares.
- 4: Set  $c_{k+1}^{1,pr} = c_k^{1,pr}$ .
- 5: For some small constant  $\varepsilon$ , with probability  $1 - \varepsilon$ , jump to step 2.
- 6: Implement repeatedly (60), (61) with the estimated value of  $c_2$  on a grid of several  $c_1$  values.
- 7: Choose  $c_{k+1}^{1,pr}$  the value of  $c_1$  which minimizes the actual cost and jump to Step 2.

A similar pretending algorithm may be used also by Player 2.

**Example 3** Assume that  $A = 1$ ,  $B = 0.5$ ,  $c_1 = 1$ ,  $c_2 = 1.2$ . Player 1 implements Algorithm 3, from the time step  $k = 1$ , and Player 2 uses the best response until time step  $k = 200$  and Algorithm 3 afterward. For both players, we assume that Algorithm 3 is implemented with  $\varepsilon = 0.01$ . Figure 7 shows the evolution of pretended production costs, Fig. 8 presents the evolution of prices, and Fig. 9 shows the evolution of the player instantaneous profits (negative total costs).



**Fig. 8** The time evolution of the price in Example 3. Price increases, compared to the Nash equilibrium price as players pretend to have a higher cost. The price is slightly reduced at time step  $k = 253$ , following Player 2's pretended cost reduction



**Fig. 9** The time evolution the profits ( $-L_i$ ) in Example 3. The profits of both players compared to the Nash equilibrium increase. At the first adjustment of pretended cost of Player 1, she increases her profits after two time steps, while the profits for Player 2 increase immediately. The profits of Player 1 have this transient response, due to the fact that steps 6 and 7 of Algorithm 3 make computations for the corresponding equilibrium and not the dynamic response. The interesting feature of this figure is that the non-pretending player benefits more from pretending than the pretender. Thus, pretending in this case is a cooperative action

*Remark 11* In this example it happens that the Nash pretenders' outcome induces lower costs (higher value) for both players, compared to the Nash equilibrium of the corresponding static full information game. Furthermore, each player has a motivation to pretend that she has

a higher production cost, in order to drive the prices higher. This in turn benefits the other player, for two different reasons. First, she is going to acquire a larger market share and second she is going to sell this larger amount of energy at a higher price. Thus, pretending is a cooperative action, i.e., each player prefers the other player to pretend than not to pretend.  $\square$

## 6.2 The Rasouli–Teneketzis Mechanism

In this subsection we analyze an example of an electricity market having many competing production firms. At first, we briefly review a market mechanism described in [34]. An algorithm to approximate the Nash equilibrium is then proposed and tested numerically. The ability of the participants to manipulate the rules of their opponents is then studied.

There are  $N$  energy producers  $P_1, \dots, P_N$ . Each one of them produces a quantity of energy  $q^i$  and proposes a price  $p^i$  that is the action of a player is given by  $u^i = [q^i \ p^i]^T$ . Let us denote the total demand by  $D_0 > 0$ . The production cost for Player  $P_i$  is given by  $C_i(q^i) = c_i(q^i)^2$ . The instantaneous cost of each player is given by:

$$L_i(q^i, t^i) = C_i(q^i) - t^i, \quad (62)$$

where  $t^i$  is the payment to the producer  $i$  made by the system operator. The payment  $t^i$  is given by:

$$t^i = p^{i+1} q^i - (p^i - p^{i+1})^2 - 2p^i \zeta^2, \quad (63)$$

where  $p^{N+1} \equiv p^1$  and

$$\zeta = \sum_{i=1}^N q^i - D_0, \quad (64)$$

is the power surplus (or deficit if it is negative).

[34] studies the non-trivial Nash equilibria of the game, i.e., the Nash equilibria such that it does not hold  $q^1 = \dots = q^N = 0$ . It is shown that in any such equilibrium the total production coincides with the demand, the proposed prices are identical, and the total energy is produced with the smallest total cost (the sum of  $C_i$  is minimum). Furthermore, a non-trivial Nash equilibrium always exists. In this sense, the non-trivial Nash equilibria of the game produce “socially optimal” outcomes.

We then propose a very simple algorithm (Algorithm 4). This algorithm is found numerically to converge asymptotically to a non-trivial Nash equilibrium of the game.

### Algorithm 4:

- 1: Initialize the proposed prices to a small positive constant  $p_0^i = \varepsilon$ .
- 2: At time step  $k$  the energy production is the best response:

$$q_k^i = \arg \min_{q'} \{C_i(q') - p_{k-1}^{i+1} q'\}. \quad (65)$$

- 3: If  $\zeta < 0$  (i.e., there is a power deficit) increase the prices by a constant step  $\delta_1$ , i.e.,

$$p_{k+1}^i = p_k^i + \delta_1, \quad (66)$$

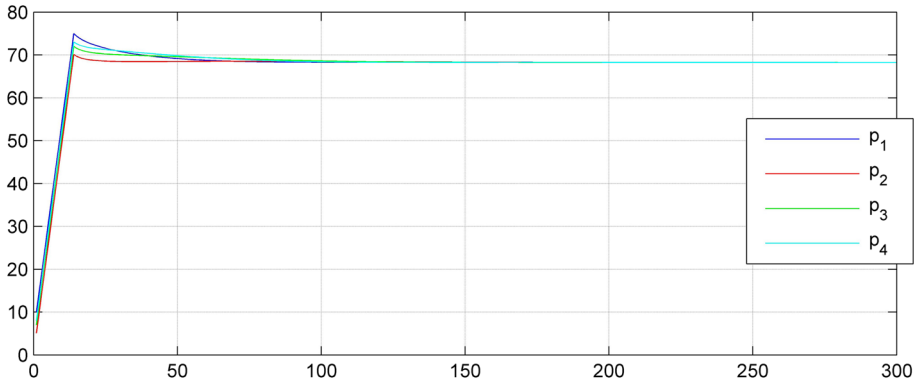
where  $\delta_1$  is a small positive scalar.

- 4: If  $\zeta \geq 0$ , the new values of the proposed prices move toward the best responses:

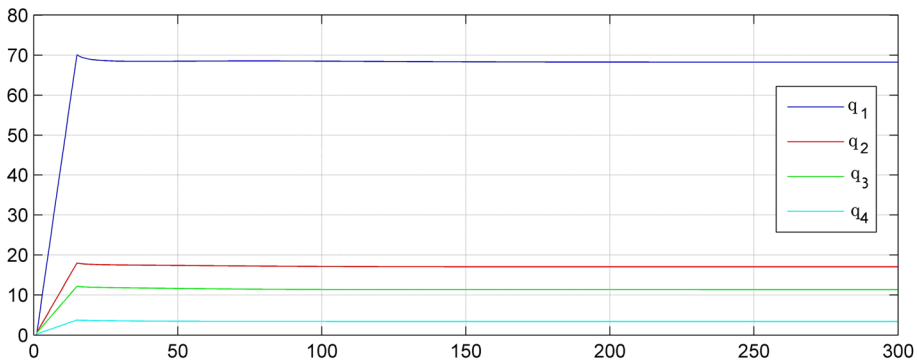
$$p_i^{k+1} = (1 - \delta_2) p_i^k + \delta_2 \arg \min_{p'} \{(p' - p_{i+1}^k)^2 + 2p'(\zeta^k)^2\}, \quad (67)$$

where  $\delta_2 \in (0, 1)$ .

- 5: Move to the time step  $k + 1$  and jump to Step 2.



**Fig. 10** The evolution of the prices in Example 4. Until approximately  $k = 20$ , the power produced does not meet the demand, and thus, the prices are increased according to (66). After  $k = 20$ , the prices are adjusted according to (67). After a transient phase, all the prices converge to the same limit

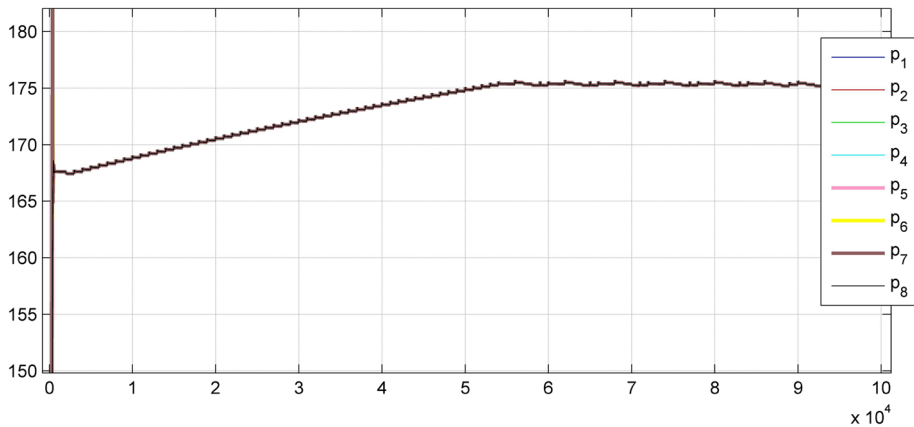


**Fig. 11** The evolution of the produced quantities in Example 4. Until approximately  $k = 20$ , due to the increasing prices the energy quantities following (65) are also increasing linearly. After  $k = 20$  the quantities converge rather quickly to their equilibrium values

**Lemma 1** *If Algorithm 4 converges, then the limit point is a non-trivial Nash equilibrium of the game.*

*Proof* First, assume that the dynamics described by Algorithm 4 converges to a fixed point,  $(p_i^\infty, q_i^\infty)_{i=1}^N$ . For this fixed point, it holds  $\zeta \geq 0$ , due to the fact that if  $\zeta < 0$ , we would have  $p_{k+1}^i = p_k^i + \delta_1$ , for each  $i$  and  $(p_i^\infty, q_i^\infty)_{i=1}^N$  would not be a fixed point. Thus, using (65), we conclude that  $q_i^\infty$  is a best response to the actions of the others. Furthermore, the prices  $p_i^\infty$  are fixed if they correspond to a best response of the actions of the other players. Thus any fixed point of (65) and (67) corresponds to a Nash equilibrium of the game. Observe that the trivial Nash equilibrium is not a fixed point. Thus, any fixed point of the dynamics is a non-trivial Nash equilibrium. It remains to show that the limit of the dynamics is a fixed point. For the proof see “Appendix E.”  $\square$

**Example 4** In this example there are four production firms with production costs  $c_1 = 0.5$ ,  $c_2 = 2$ ,  $c_3 = 3$ ,  $c_4 = 10$ . Figures 10 and 11 show the evolution of the energy production and the proposed prices assuming that the players follow Algorithm 4.  $\square$



**Fig. 12** The evolution of the prices in Example 5. Only Player 1 is pretending. Only Player 1 is pretending. The pretending of a single player is sufficient to increase the prices for all the players. Thus, the pretending of a player is beneficial to the others

A very simple manipulating strategy is then described. A manipulating Player  $i$  implements Algorithm 4 as if she had a different (not real) production cost  $c^{i,pr}$ . The pretended type is updated rarely but periodically according to the rule:

$$\begin{aligned} c^{i,pr}(k+1) &= c^{i,pr}(k), \quad \text{if } k+1 \neq tT, \\ c^{i,pr}((t+1)T) &= c^{i,pr}(tT) - \rho \operatorname{sat}_{[-M,M]} \left( \frac{L_i(tT+T-1) - L_i(tT-1)}{c^{i,pr}(tT+T-1) - c^{i,pr}(tT-1)} \right), \end{aligned} \quad (68)$$

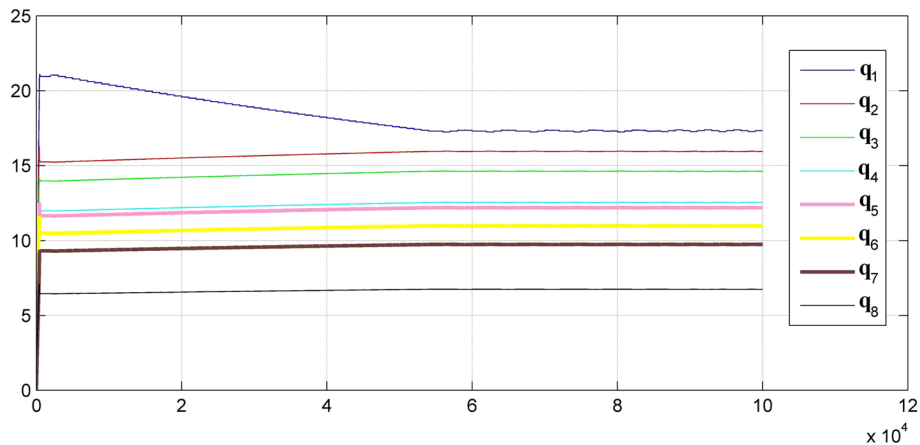
where  $\rho$  is a positive constant,  $T$  is the time interval, and  $\operatorname{sat}_{[-M,M]}(\cdot)$  is the saturation function:

$$\operatorname{sat}_{[-M,M]}(z) = \begin{cases} M & \text{if } z > M \\ z, & \text{if } |z| \leq M \\ -M & \text{if } z < -M \end{cases}$$

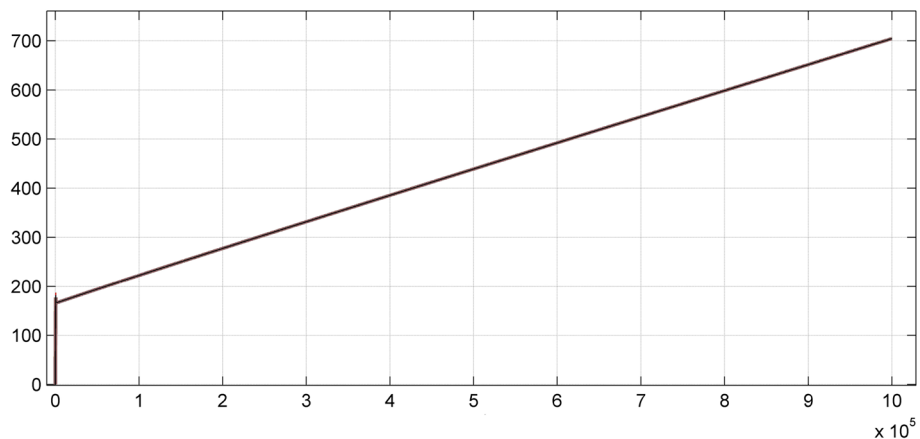
**Remark 12** Pretending rule (68) does not estimate the production cost of the other players. Instead, it uses a gradient decent like rule to minimize the cost with respect to the pretended cost. A learning scheme very much related to (68) is directional learning [29]. Particularly, in [29] a player increases the probability of contributing to a game of public goods, if a previous round switch from non-contributing to contributing increased her payoff or a switch from contributing to non-contributing reduced her payoff. Similarly, a player reduces her probability of contributing if a previous round switch from contributing to non-contributing increased her payoff or a switch from non-contributing to contributing reduced her payoff.  $\square$

**Example 5** In this example there are eight production firms. We assume that only Player 1 (having the smaller production cost) is pretending implementing (68). Figure 12 illustrates the evolution of the prices, and Fig. 13 the evolution of production quantities. It turns out that the prices proposed by the several players converge to identical values. Furthermore, the pretending of Player 1 reduces her market share but increases the prices. Thus, pretending is slightly beneficial for Player 1, whereas it is more beneficial for the other players which sell





**Fig. 13** The evolution of energy production in Example 5. Only Player 1 is pretending. We may observe that the market share of Player 1 is reduced. However, the prices are increased and the production cost is reduced. This makes the pretending player slightly better off. From the point of view of the other players, both the prices and the market share are increasing. Thus, pretending of Player 1 makes them significantly better off

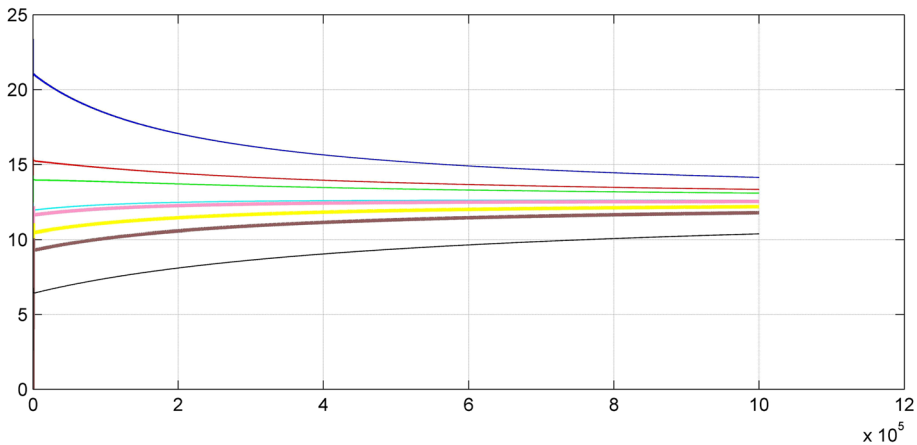


**Fig. 14** The evolution of the prices in Example 6. All the players are pretending. Each player has a motivation to pretend that she has a production cost larger than her actual. Furthermore, pretending of the other players generates the incentive for a player to pretend that she has an even larger production cost. This dynamics leads to a price divergence

energy at higher prices and also get larger market shares. Thus, in this case pretending is a cooperative behavior.  $\square$

**Example 6** This example continues Example 5 assuming that all the players are pretending. Figure 14 shows the evolution of the prices, and Fig. 15 the evolution of quantities. In this example, price grows unbounded. The reason is that the optimal pretended production cost for each player increases when the pretended costs of the others increase. This dynamics leads to price divergence. Therefore, the mechanism in this case is not working at all.  $\square$

**Remark 13** The mechanism was designed in order to have a “socially optimal” behavior in the Nash equilibrium. However, each one of the players has a motivation to pretend. If all



**Fig. 15** The evolution of the energy production in Example 6. All the players are pretending. As the prices increase, the differences in actual production cost become less important and the production quantities approach one another

the players pretend, the system is not working at all, as illustrated in the last example where the prices diverge.  $\square$

## 7 Conclusion

We studied the possibility of manipulation in game situations where the players are using a dynamic rule to determine their future actions. Some alternative outcomes were proposed, and the results were applied in opinion quadratic games, a linear quadratic environmental game and two electricity market models.

Several relationships among pretending and leadership were identified in the repeated quadratic games examples. Particularly, in the two-player simultaneous move repeated game the pretender has the same cost as if she were Stackelberg leader, in the repeated Stackelberg game if the follower pretends she may achieve the same cost as if she were the leader in the corresponding inverse Stackelberg game and in the case where there is an influential player and many players of low influence the Nash pretenders' outcome approaches the Stackelberg equilibrium as the number of players increases.

In the environmental game example pretending leads to a less cooperative behavior and more pollution. In the electricity market games, pretending may enhance the cooperation among the players, increase the prices or even make the system not working at all.

The alternative outcomes proposed are by no means the only reasonable ones, and further work is needed. For example, how would a player react knowing that her opponent tries to manipulate her? Another direction of future research is the generalization of the results about centrality to dynamic game situations. Furthermore, it would be interesting to study how to design the network of interactions such that the motivation to manipulate is not large.

## Appendix A: Scalar LQ Optimal Control With an Additive term

Consider the scalar system:

$$x_{k+1} = \bar{a}x_k + \bar{u}_k + v, \quad (69)$$

where  $v$  is a constant. Consider also the quadratic criterion:

$$J = \sum_{k=0}^{\infty} \rho^k (qx_k^2 + \bar{u}_k^2). \quad (70)$$

Simple manipulations show that the optimal cost is given by:

$$J^*(x) = px^2 + c_1x + c_2, \quad (71)$$

and the constants  $p$  and  $c_1$  satisfy:

$$p = q + \rho \bar{a}^2 p - \frac{\rho^2 \bar{a}^2 p^2}{1 + \rho p} = q + \frac{\rho p}{1 + \rho p} \bar{a}^2, \quad (72)$$

$$c_1 = \rho \bar{a} (2pv + c_1) - \frac{\rho^2 \bar{a} p (2pv + c_1)}{1 + \rho p} = \frac{\rho \bar{a} (2pv + c_1)}{1 + \rho p}. \quad (73)$$

Furthermore, the optimal controller is given by:

$$\bar{u}_k = -\frac{\rho \bar{a} p}{1 + \rho p} x_k - \rho \frac{2pv + c_1}{2(1 + \rho p)}. \quad (74)$$

## Appendix B: The Optimal Control Problems of Section 5

The optimal control problems of Sect. 5 are then reduced to a pair of problems in the form which was studied in Subsection A.

Without loss of generality let us state the optimal control problem of Player 1. Assume that Player 2 uses a control policy in the form  $u_k^2 = K_1^2 x_k + K_2^2$ . Then the cost function given by (50) is written as:

$$J_1 = \sum_{k=0}^{\infty} \rho_1^k [qx_k^2 + (u_k^1 - 1/2)^2] - \frac{1}{4(1 - \rho)}.$$

Ignoring the constant term, and setting  $\bar{a} = a + K_1^2$ ,  $\bar{u}_k = u_k^1 - 1/2$  and  $v = K_2^2 + 1/2$ , the problem reduces to the minimization of (70) subject to (69). The optimal controller is given by:

$$u_k^1 = \bar{u}_k + 1/2. \quad (75)$$

## Appendix C: Proof of Proposition 3

Let us first observe that (51), (53) do not depend on (52), (54). Thus, let us first show the existence of a solution of (51), (53). Consider the compact and convex domain  $D \subset \mathbb{R}^4$  with:

$$D = [0, a] \times [0, a] \times [0, q + 1] \times [0, q + 1]. \quad (76)$$

Consider also the mapping  $f : D \rightarrow \mathbb{R}^4$ , with:

$$f([\bar{a}^1, \bar{a}^2, p^1, p^2]^T) \mapsto [\bar{a}^{1,+}, \bar{a}^{2,+}, p^{1,+}, p^{2,+}]^T, \quad (77)$$

where  $p^{i,+}$  is the unique positive solution of:

$$p^{i,+} = q + \frac{\rho_i p^{i,+}}{1 + \rho_i p^{i,+}} (\bar{a}^i)^2, \quad (78)$$

and

$$\bar{a}^{i,+} = a - \frac{\rho_{-i} \bar{a}^{-i} p^{-i}}{1 + \rho_{-i} p^{-i}}. \quad (79)$$

It is not difficult to see that  $f$  maps  $D$  into  $D$ . For, observe that  $p^{i,+} > 0$  and  $p^{i,+} < q + 1$  and that  $\bar{a}^{i,+} \in (0, a)$ . Applying Brouwer's fixed point theorem, we conclude that there exists a solution of (51), (53).

Then, assume  $p^i, K_1^i, i = 1, 2$  is a solution of (51), (53). Equation (52) is equivalent to:

$$c_1^i = \frac{2\rho_i p^i \bar{a}^i}{1 + \rho_i p^i - \rho_i \bar{a}^i} v^i = \lambda_i v^i. \quad (80)$$

Using (54) and  $v^{-i} = K_2^i + 1/2$ , we obtain:

$$\rho_i \frac{2p^i + \lambda_i}{2(1 + \rho_i p^i)} v^i + v^{-i} = 1. \quad (81)$$

It is not difficult to see that  $\rho_i \frac{2p^i + \lambda_i}{2(1 + \rho_i p^i)} < 1$ . For, observe that this inequality is equivalent to  $\rho_i \lambda_i < 2$ , which can be written as:

$$\rho_i^2 p^i \bar{a}^i < (1 + \rho_i p^i - \rho_i \bar{a}^i),$$

which is trivially true, due to the fact that  $\bar{a}^i < 1$  and  $\rho_i < 1$ .

Hence, the pair of equations obtained by (81) substituting  $i = 1$  and  $i = 2$  has a unique solution. Substituting to (80) and (54) a solution of equations (51)–(54) is obtained.

## Appendix D: Cost Computations

Some computations, needed to implement Algorithm 2, are then presented. Particularly, the value of the cost function:

$$J = \sum_{k=0}^{\infty} \rho^k [qx_k^2 + u_k^2 - u_k],$$

is computed under the dynamics:

$$x_{k+1} = \bar{a}x_k + u_k + v, \quad (82)$$

and the control law  $u_k = K_1 x_k + K_2$ .

The equilibrium point of the dynamics is given by  $\bar{x} = \bar{v}/(1 - \bar{a})$ , where  $\bar{a} = \bar{a} + K_1$  and  $\bar{v} = v + K_2$ . The computation is simplified using the quantity  $z_k = x_k - \bar{x}$  which evolves according to  $z_{k+1} = \bar{a}z_k$ .

Simple manipulations show that  $J$  can be written as:

$$J = \sum_{k=0}^{\infty} \rho^k [s_1 z_k^2 + s_2 z_k + s_3],$$

where  $s_1 = q + K_1^2$ ,  $s_2 = q\bar{x} + K_1^2 \bar{x} + K_1 K_2 - K_1/2$  and  $s_3 = q\bar{x}^2 + (K_1 \bar{x} + K_2)^2 - K_1 \bar{x} - K_2$ . Thus, assuming that  $\bar{a} < 1$ , it holds:

$$J = \frac{s_1 z_0^2}{1 - \rho \bar{a}^2} + \frac{s_2 z_0}{1 - \rho \bar{a}} + \frac{s_3}{1 - \rho}. \quad (83)$$

To compute  $J_i(\gamma^i, \gamma^{-i})$  with  $\gamma^j(x_k) = K_1^j x_k + K_2^j$  and  $j = 1, 2$ , set  $v = K_2^{-i}$  and  $\bar{a} = a + K_1^{-i}$ . The rest of the computations are performed with  $K_1 = K_1^i$ ,  $K_2 = K_2^i$ .

## Appendix E: Proof of Lemma

It remains to show that the limit of the sequence is a fixed point. Assume that the sequence  $(p_i^k, q_i^k)_{i=1}^N$  obtained by iterations of Algorithm 4 converges to a limit  $(p_i^\infty, q_i^\infty)_{i=1}^N$ . Then, there exists a  $k_0$  such that for any  $k \geq k_0$  it holds  $\zeta_k \geq 0$ . Indeed, otherwise Step 3 of the algorithm would be used infinitely often and we would have  $\|(p_i^{k+1}, q_i^{k+1})_{i=1}^N - (p_i^k, q_i^k)_{i=1}^N\| \geq \delta_1$  for an infinite number of  $k$ . Thus, the sequence would not converge. Therefore, after  $k_0$  the terms of the sequence satisfy (65) and (67).

Observe that the function:

$$F\left(\left(q_i^k, p_i^k\right)_{i=1}^N\right) = \left(p_i^{k+1}, q_i^{k+1}\right)_{i=1}^N,$$

where  $p_i^{k+1}, q_i^{k+1}$  are given by (65) and (67) is continuous. Thus,

$$F\left(\left(p_i^\infty, q_i^\infty\right)_{i=1}^N\right) = \left(p_i^\infty, q_i^\infty\right)_{i=1}^N,$$

which completes the proof.

## References

1. Acemoglu D, Yildiz M (2001) Evolution of perceptions and play. Mimeo, MIT. Available online <https://economics.mit.edu/files/200>
2. Antoniadis P, Fdida S, Griffin C, Jin Y, Kesidis G (2013) Distributed medium access control with dynamic altruism. In: Zheng J, Mitton N, Li J, Lorenz P (eds) Ad hoc networks. Springer, Berlin, pp 29–42
3. Atakan AE, Ekmekci M (2013) A two-sided reputation result with long-run players. J Econ Theory 148(1):376–392
4. Azad AP, Altman E, El-Azouzi R (2010) Routing games: from egoism to altruism. In: Modeling and optimization in mobile, ad hoc and wireless networks (WiOpt), proceedings of the 8th international symposium on, IEEE, 2010, pp 528–537
5. Azad AP, Musacchio J (2011) Unilateral altruism in network routing games with atomic players. In: Network games, control and optimization (NetGCooP), 5th international conference on, IEEE, 2011, pp 1–8
6. Baldick R, Grant R, Kahn E (2004) Theory and application of linear supply function equilibrium in electricity markets. J Regul Econ 25(2):143–167
7. Basar T (1973) On the relative leadership property of stackelberg strategies. J Optim Theory Appl 11(6):655–661
8. Basar T, Olsder GJ (1999) Dynamic noncooperative game theory, 2nd edn. SIAM, Philadelphia, PA
9. Bonacich P (1987) Power and centrality: a family of measures. Am J Sociol 92:1170–1182
10. Chan YM, Cruz J (1983) Decentralized stochastic adaptive nash games. Optim Control Appl Methods 4(2):163–178
11. Dockner EJ, Van Long N (1993) International pollution control: cooperative versus noncooperative strategies. J Environ Econ Manag 25(1):13–29
12. Dufwenberg M, Güth W (1999) Indirect evolution vs. strategic delegation: a comparison of two approaches to explaining economic institutions. Eur J Political Econ 15(2):281–295
13. Feldbaum AA (1960) Dual control theory. Autom Remote Control 21(9):874–1039
14. Foster DP, Young HP (2006) Regret testing: learning to play nash equilibrium without knowing you have an opponent. Theor Econ 1:341–367
15. Fudenberg D, Levine D (1989) Reputation and equilibrium selection in games with a patient player. Econometrica 57(4):759–778
16. Fudenberg D, Levine DK (1998) The theory of learning in games, vol 2. MIT press, Cambridge
17. Gabriel SA, Conejo AJ, Fuller JD, Hobbs BF, Ruiz C (2012) Complementarity modeling in energy markets, vol 180. Springer, New York
18. Güth W (1995) An evolutionary approach to explaining cooperative behavior by reciprocal incentives. Int J Game Theory 24(4):323–344

19. Güth W, Peleg B (2001) When will payoff maximization survive? An indirect evolutionary analysis. *J Evol Econ* 11(5):479–499
20. Hamilton JH, Slutsky SM (1990) Endogenous timing in duopoly games: Stackelberg or cournot equilibria. *Games Econ Behav* 2(1):29–46
21. Hart S, Mas-Colell A (2000) A simple adaptive procedure leading to correlated equilibrium. *Econometrica* 68(5):1127–1150
22. Heifetz A, Shannon C, Spiegel Y (2007) The dynamic evolution of preferences. *Econ Theory* 32(2):251–286
23. Heifetz A, Shannon C, Spiegel Y (2007) What to maximize if you must. *J Econ Theory* 133(1):31–57
24. Ho YC, Luh PB, Olsder GJ (1982) A control-theoretic view on incentives. *Automatica* 18(2):167–179
25. Konigstein M, Muller W (2000) Combining rational choice and evolutionary dynamics: the indirect evolutionary approach. *Metroeconomica* 51(3):235–256
26. Kreps DM, Wilson R (1982) Reputation and imperfect information. *J Econ Theory* 27(2):253–279
27. Lasaulce S, Tembine H (2011) Game theory and learning for wireless networks: fundamentals and applications. Academic Press, Waltham
28. Mailath GJ, Samuelson L (2006) Repeated games and reputations, vol 2. Oxford University Press, Oxford
29. Nax HH, Perc M (2015) Directional learning and the provisioning of public goods. *Sci Rep* 5:8010
30. Page L, Brin S, Motwani R, Winograd T (1999) The pagerank citation ranking: bringing order to the web. Technical Report, Stanford InfoLab
31. Papavassilopoulos GP (1988) Adaptive games. In: Alberverio S, Blanchard P, Hazewinkel M, Streit W (eds) Stochastic processes in physics and engineering. Springer, Berlin, pp 223–236
32. Papavassilopoulos GP (1986) Iterative techniques for the nash solution in quadratic games with unknown parameters. *SIAM J Control Optim* 24(4):821–834
33. Perc M, Szolnoki A (2010) Coevolutionary games—a mini review. *BioSystems* 99(2):109–125
34. Rasouli M, Teneketzis D (2014) Electricity pooling markets with strategic producers possessing asymmetric information ii: inelastic demand. *arXiv arXiv: 1404.5539*
35. Schipper BC (2011) Strategic control of myopic best reply in repeated games. Available at SSRN 1804667
36. Schipper BC (2015) Strategic teaching and learning in games. *CoRR*, abs/1504.06341, [Online]. Available <http://arxiv.org/abs/1504.06341>
37. Schipper BC (2009) Imitators and optimizers in cournot oligopoly. *J Econ Dyn Control* 33(12):1981–1990
38. Selten R (2001) What is bounded rationality. In: Bounded rationality: the adaptive toolbox. MIT Press, pp 13–36
39. Sengul M, Gimeno J, Dial J (2012) Strategic delegation a review, theoretical integration, and research agenda. *J Manag* 38(1):375–414
40. Shoham Y, Leyton-Brown K (2008) Multiagent systems: algorithmic, game-theoretic, and logical foundations. Cambridge University Press, Cambridge
41. Simon HA (1972) Theories of bounded rationality. *Decis Organ* 1:161–176
42. Skoulidas C, Vournas C, Papavassilopoulos GP (2002) Adaptive game modeling of deregulated power markets. In: IEEE power engineering review, pp. 42–45
43. Szolnoki A, Perc M (2008) Coevolution of teaching activity promotes cooperation. *New J Phys* 10(4):043036
44. Witsenhausen HS (1968) A counterexample in stochastic optimum control. *SIAM J Control* 6(1):131–147
45. Wittenmark B (2002) Adaptive dual control. In: Control systems, robotics and automation, encyclopedia of life support systems (EOLSS), Developed under the auspices of the UNESCO. Eolss Publishers, Oxford
46. Yang WY, Papavassilopoulos GP (1994) Decentralised adaptive control in a game situation for discrete-time, linear, time-invariant systems. In: American control conference, vol. 3. IEEE, pp 3394–3399
47. Young HP (2004) Strategic learning and its limits. Oxford University Press, New York