CHEATING IN ADAPTIVE GAMES MOTIVATED BY ELECTRICITY MARKETS

Ioannis Kordonis, Ph.D. Candidate[1], *George P. Papavassilopoulos, Professor*[1]

[1]Electrical and Computer Engineering Department, National Technical University of Athens, Greece.

ABSTRACT

We study Dynamic Game situations with incomplete structural information, motivated by problems arising in electricity market modeling. Some Adaptive 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, in order to assess those strategies, two criteria are stated: Firstly, how far the cost of each player is from the cost of her best response in the sense of the Nash equilibrium. Secondly, we consider the case where the first player follows the adaptive strategy and the second player implements the best response to the first player. Then, the criterion depends on the difference of the cost of the first player comparing with the cost in case where both players follow their adaptive control laws. This difference may be positive or negative. We then examine a smaller class of strategies, called the pretender 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 is pretending, she can achieve the same cost as if she were Stackelberg leader. The situation where all the players are pretending is then considered. The effects of adaptation and cheating, when the number of players in the game becomes large, is examined in a simple example.

Index Terms— Adaptive Games, Cheating, Bounded Rationality, Electricity Markets - Power Grid.

1. INTRODUCTION

In a number of real game situations there is 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 producer firms are competing repeatedly over time and each firm knows its own costs but not the costs of the other firms ex. [1], [2]. Such strategic interactions over time can be described by dynamic games with incomplete structural information.

Two difficult problems arise in the problem of finding a Nash equilibrium in the case of dynamic games with incomplete structural information. The first is due to the "Witsenhausen effect" [3], i.e. the current action of each player affects the future state estimation of the other players. The second is due to the "Dual Control effect" [4], i.e. that the current action of a player affects the quality of his own future parameter estimation. Due to the later difficulty, the Optimization problem has not been solved analytically even for the single player (control) case, except only of a few special cases [5].

In this context, theory of Adaptive Games was introduced [8] - [10]. Each player solves an optimal control problem using an estimated value for the gains or the types of the other players and then updates the estimated values based on the measurements. The adaptive strategies that try to approximately solve the optimization problems are an expression of the Bounded rationality [6], [7] of the players and they are not in equilibrium. These works prove finiteness of the costs and under certain conditions, the asymptotic convergence of the adaptive strategies to the full information Nash strategies.

Are these adaptive 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 on 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 "cheat", i.e. use the knowledge of the adaptation law of the first player to manipulate her. The topic of this work is to study phenomena like "cheating" and the implications that may have to the costs of the participants of the game.

2. ADAPTIVE STRATEGIES AND ASSESSMENT

In this section, we define two criteria in order to assess a set of adaptive strategies. At first, a general form of strategies is considered. For simplicity reasons, the criteria are stated for two player games.

Let us first define formally the game. There are two players of types θ_1, θ_2. Each player knows her own type and θ_1, θ_2 are part of a random vector $\bar{\theta} = [\theta_1 \ \ \theta_2 \ \ \theta]^T \in \bar{\Theta} = \Theta_1 \times \Theta_2 \times \Theta$ having a commonly known distribution.

Authors e-mails: I. Kordonis: jkordonis1920@yahoo.com, G. P. Papavassilopoulos: yorgos@netmode.ntua.gr.

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.

The dynamics have the form:

$$x_{k+1} = f(x_k, \bar{\theta}, u_k^1, u_k^2, w_k), \quad (1)$$

where u_k^1, u_k^2 are the action variables of the players and w_k a random disturbance.

The cost functions have the form:

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

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

$$J_i = \lim_{T\to\infty} E\left[\frac{1}{T}\sum_{k=0}^{T} L_i(x_k, \theta_i, u_k^1, u_k^2)\right]. \quad (3)$$

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:

Information Structure 1: $\bar{I}_k^{i,new} = (x_k, u_{k-1}^i)$, or

Information Structure 2: $\bar{I}_k^{i,new} = (x_k, u_{k-1}^i, u_{k-1}^{-i})$.

The information that each player possesses at time step k has the form $I_k^i = (\bar{I}_0^{i,new}, \ldots, \bar{I}_k^{i,new})$.

Let us now describe a general form of strategies of the players: $s^i = (\gamma_1^i, \gamma_2^i, \ldots)$, $i = 1, 2$, where γ_k^i is a function having the form:

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

We shall focus on "state feedback" strategies, where all the previous information is used only for "adaptation". Specifically,

$$u_k^i = \gamma_k^i(x_k, \theta_i, \hat{\theta}_k^i), \quad (4)$$

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

$$\hat{\theta}_{k+1}^i = \phi^i(\hat{\theta}_k^i, \bar{I}_{k+1}^{i,new}, \theta_i). \quad (5)$$

For infinite horizon games, the following property is quite interesting see for example [11]:

Property 1: The adapted values $\hat{\theta}_k^1, \hat{\theta}_k^2$ converge to some limits $\hat{\theta}_\infty^1, \hat{\theta}_\infty^2$, such that the feedback (no memory) strategies $\gamma_k^i(x_k, \theta_i, \hat{\theta}_\infty^i)$, $i = 1, 2$ constitute a Perfect Nash equilibrium for the complete information game.

The criteria depend on the best response of each player to the opponent's strategy. Thus, in order to state the criteria, the best response $\bar{s}^i$, $i = 1, 2$ of each player given the strategy of the other player s^{-i}, is considered.

The first criterion states that each player does not have a lot to profit from moving to her best response. We call it the opportunity criterion.

Opportunity Criterion *with parameter* a: For some $a > 0$ it holds:

$$J^i(s^i, s^{-i}) \le J^i(\bar{s}^i, s^{-i}) + a, \quad (6)$$

for any θ_i, $i = 1, 2$.

The second criterion states that if the other player moves to her best response, the first player does not have a lot to lose. Let us call it the conservative criterion.

Conservative Criterion *with parameter* b: For some $b > 0$, it holds:

$$J^i(s^i, \bar{s}^{-i}) \le J^i(s^i, s^{-i}) + b, \quad (7)$$

for any θ_i, $i = 1, 2$.

Definition 1 *A pair of strategies s^1, s^2 is a, b - not sensitive to cheating if either the Opportunity Criterion with parameter a or the Conservative Criterion with parameter b holds.*

Remark 1 *Definition 1 borrows some ideas form satisficing based decision making [6], where the values of a, b have roles related to the satisfactory levels.*

Remark 2 *Definition 1 corresponds to conservative players. Particularly, it states that each player believes either that the opponent will not have enough motivation to change her strategy or that if she has, this change would not increase the cost of the former player a lot. For less conservative players, the "either, or" of the definition should be replaced by "and". An alternative definition would involve any "better response" $\tilde{s}^i$ instead of the best response $\bar{s}^i$.*

Remark 3 *A class of games which satisfy Definition 1 is Team Games. More generally, if $max|J_1(s_1, s_2) - J_2(s_2, s_2)| \le max\{a, b\}$, then the game is a, b not sensitive to cheating.*

The verification of Definition 1 is not easy due to the fact that the optimal control problems involved are quite difficult. This definition can be numerically checked in some examples, however the analysis is quite lengthy and thus, it is not included in this paper.

3. PRETENDERS STRATEGIES

In this section we concentrate on a special class of cheating strategies. Namely, the pretenders' strategies. Particularly, the cheating player pretends to have a false type. For simplicity, we assume that the θ is commonly known, i.e. $\Theta = \{\theta\}$.

The general form of a pretender's strategy that corresponds to (4) is:

$$u_k^i = \gamma_k^i(x_k, \theta_k^{i,pr}, \hat{\theta}_k^i), \quad (8)$$

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

$$z_{k+1}^i = \phi^{i,pr}(z_k^i, \bar{I}_{k+1}^{i,new}, \theta_i), \quad (9)$$

$$\theta_k^{i,pr} = \psi^i(\theta_i, z_k^i). \quad (10)$$

Equations (8)-(10) represent a cheating 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.

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. In the spirit of Property 1, we analyze the following situation. Player 1 has revealed all the useful information for θ_2 and the pretended type of player 1 has converged to $\theta_\infty^{1,pr}$. Player 2 reacts to a player of type $\theta_\infty^{1,pr}$. Furthermore, the pair of strategies $\gamma^1(x_k, \theta_\infty^{1,pr}, \hat{\theta}_\infty^1), \gamma^2(x_k, \theta_2, \hat{\theta}_\infty^2)$ constitute a Perfect Nash equilibrium for the game with full information and types $\theta_\infty^{1,pr}, \theta_2$.

In order to define the Optimal Stationary Pretending, the following assumption is made:

Assumption 1: For any pair of types θ_1, θ_2, there exist a unique Perfect Nash equilibrium of the full information game. Let us denote by $\gamma^{i,N}_{\theta_1,\theta_2}(x_k),\ i = 1, 2$ the pair of strategies constituting the Nash equilibrium.

Assumption 1 is not unusual in static or dynamic games.

The optimal pretending for the player 1 is given by:

$$\theta_\infty^{1,pr} = arg \min_{\tilde{\theta}_1 \in \Theta_1} J_1(\gamma^{1,N}_{\tilde{\theta}_1,\theta_2}, \gamma^{2,N}_{\tilde{\theta}_1,\theta_2}) \tag{11}$$

It is interesting to compare the cost that the cheating player 1 attains with the cost of the full information game having 1 as Stackelberg leader. Let us denote by $\gamma^{i,S}_{\theta_1,\theta_2},\ i = 1, 2$, a pair of strategies constituting a feedback Stackelberg equilibrium with 1 as leader.

Proposition 1 *It holds:*

$$J_1(\gamma^{1,S}_{\theta_1,\theta_2}, \gamma^{1,S}_{\theta_1,\theta_2}) \leq J_1(\gamma^{1,N}_{\theta_\infty^{1,pr},\theta_2}, \gamma^{1,N}_{\theta_\infty^{1,pr},\theta_2}). \tag{12}$$

An equality is attained is there exist a $\tilde{\theta}_1 \in \Theta_1$ *such that* $\gamma^{1,S}_{\theta_1,\theta_2} = \gamma^{1,N}_{\tilde{\theta}_1,\theta_2}$.

The proof of Proposition 1 is easy and thus it is omitted.

Remark 4 *Proposition 1 roughly says that if there is enough uncertainty and only one player pretends, then:*
"The pretender becomes the leader".

3.2. Pretenders' Equilibrium

The case where both players are pretending is then analyzed. Particularly, it is defined the notion of pretenders equilibrium, under the Assumption 1.

Definition 2 *A pair of pretending types* $(\theta_\infty^{1,pr}, \theta_\infty^{2,pr})$ *and a pair of strategies* $(\gamma^{1,p}_{\theta_1,\theta_2}, \gamma^{2,p}_{\theta_1,\theta_2})$ *constitute a pretenders' equilibrium if* $(\gamma^{1,p}_{\theta_1,\theta_2}, \gamma^{2,p}_{\theta_1,\theta_2})$ *is a Nash equilibrium for the full information game with types* $(\theta_\infty^{1,pr}, \theta_\infty^{2,pr})$ *and it holds:*

$$\theta_\infty^{i,pr} = arg \min_{\tilde{\theta}_i \in \Theta_i} J_i(\gamma^{i,N}_{\tilde{\theta}_i,\theta_{-i}}, \gamma^{-i,N}_{\tilde{\theta}_i,\theta_{-i}}), \tag{13}$$

for $i = 1, 2$.

Remark 5 *The definitions and the reasoning of Section 3 can be extended to the many players case.*

Fig. 1. The blue line corresponds to best response with no pretending players, the red with player 1 pretending, the purple with player 2 and the green with both players pretending.

4. SPECIAL CASES

4.1. Two Player Quadratic Games

A two players, static, quadratic game is studied. The cost function is given by (3) and the instantaneous costs by:

$$L_1 = (u^1 - \theta_1)^2 + (u^1 - u^2)^2, \tag{14}$$
$$L_2 = (u^2 - \theta_2)^2 + (u^2 - u^1)^2, \tag{15}$$

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

$$u^{i,N} = \frac{2}{3}\theta_i + \frac{1}{3}\theta_{-i},\ \ i = 1, 2 \tag{16}$$

Several adaptive (iterative) techniques for the incomplete information game were studied in [12]. Probably, the simplest one is the best response map:

$$u_k^i = (\theta_i + \hat{\theta}_k^i)/2, \tag{17}$$
$$\hat{\theta}_k^i = u_{k-1}^{-i}. \tag{18}$$

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 analyze the situation of player 1 cheating against player 2 and player 2 following (17), (18). Due to the fact that the map $\theta_1 \mapsto u^{1,N}$ in (16) 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:

$$\theta_\infty^{1,pr} = \frac{6}{5}\theta_1 - \frac{1}{5}\theta_2. \tag{19}$$

Player 1 can use several ways to learn θ_2 in order to implement her cheating policy. One way is to use only the last iteration. Particularly:

$$z_{k+1}^{1,1} = z_k^{1,2}, \tag{20}$$
$$z_{k+1}^{1,2} = u_k^1, \tag{21}$$
$$\theta_k^{1,pr} = \frac{6}{5}\theta_1 - \frac{1}{5}(2u_{k-1}^2 - z_k^{1,1}). \tag{22}$$

An alternative way is to use Recursive Least Squares (RLS).

Figure 1 shows the action trajectories when no player, one player and both players are pretending. The parameters are $\theta_1 = 1$ and $\theta_2 = -1.3$.

4.2. A Quadratic Game With Many Players

In this subsection, we study the pretenders' equilibria of an N players generalization of the game studied in the previous subsection. The instantaneous costs have the form:

$$L_i = (u^i - \theta_i)^2 + (u^i - \sum_{j=1}^{N} a_j u^j)^2, \quad (23)$$

where $\sum_{j=1}^{N} a_j = 1$ and a_j represents the "strength" of the player j.

The unique feedback Nash equilibrium of the full information game has the form:

$$u^i = (\theta_i + \sum_{j=1}^{N} a_j \theta_j)/2 \quad (24)$$

Simple but lengthy calculations show that if all the players are pretending, the conditions for optimal pretending are given by:

$$\theta_\infty^{i,pr} = \frac{1+a_i}{1+a_i^2/2}\theta_i - \frac{a_i}{1+a_i^2/2}\sum_{j\neq i} a_j \theta_\infty^{j,pr}, \quad (25)$$

for $i = 1, \ldots, N$. Equation (25) shows that weak players tend to be truthful. It is not difficult to show that:

Proposition 2 *Denoting by* $a = \max_{i=1,\ldots,N}\{a_i\}$ *it holds:*

$$|\theta_\infty^{i,pr} - \theta_i| \leq \frac{2a}{1-a} \max_{i=1,\ldots,N} |\theta_i| \quad (26)$$

Proposition 2 is proved using contraction mapping ideas and shows that if there exist a large population of weak players (ex. $a_i \sim 1/N$) then the benefit from pretending is not big.

5. FUTURE WORK

Future work involves the numerical study of Definition 1 in simple examples and the derivation of sufficient conditions. The study of more examples, such as cheating in Linear Quadratic games with unknown cost parameters and the study of games with local coupling is also of interest. Another possible point is the extension of the pretenders' equilibrium definition for cases with multiple equilibria and the study of the possibility of equilibrium selection through cheating. Finally, the convergence analysis of the schemes proposed is of certain interest.

6. REFERENCES

[1] S.A. Gabriel, A.J. Conejo, B.F. Hobbs, D. Fuller, and C. Ruiz *Complementarity Modeling in Energy Markets*, Springer, New York, NY, USA, 2012.

[2] C. Skoulidas, C. Vournas, and G. P. Papavassilopoulos "Adaptive Game Modeling of Deregulated Power Markets," *IEEE Power Engineering Review*, vol. 22, pp. 42-45, Sept. 2002

[3] H. S. Witsenhausen, "A counterexample in stochastic optimum control," in *SIAM Journal on Control*, Vol. 6, pp.131147, 1968.

[4] A.A Feldbaum, "Dual control theory," in *Automation and Remote Control*, Vol. 21, pp.847-1039, 1960.

[5] B. Wittenmark, "Adaptive dual control," in *in Control Systems, Robotics and Automation, Encyclopedia of Life Support Systems (EOLSS), Oxford*, UK: Eolss Publishers, 2002.

[6] H. A. Simon, "Theories of bounded rationality". In *C. B. McGuire & R. Radner (Eds.), Decision and organization: Volume 2* pp. 161-176. Minneapolis: University of Minnesota Press. 1986

[7] R. Selten, "What is bounded rationality?" *In Bounded rationality: The adaptive toolbox*, ed. G. Gigerenzer and R. Selten, 13-36. Cambridge, MA: MIT Press 2001

[8] Y. M. Chan and J. B. Cruz, "Decentralized stochastic adaptive Nash games", *Optimal Control Applications and Methods*, vol. 4, pp. 163-178, 1983.

[9] G. P. Papavassilopoulos, "Adaptive Games", *Stochastic Processes in Physics and Engineering*, vol 42, pp. 223-236, Springer Netherlands, 1988

[10] W.Y. Yang and G.P. Papavassilopoulos, "On a class of decentralized discrete-time adaptive control problems", *Proc. of 24th Conf. on Decision and Control*, Fort Lauderdale Dec. 1985.

[11] W.Y. Yang and G.P. Papavassilopoulos, "Decentralized Adaptive Control in a Game Situation for Discrete-Time,Linear, Time Invariant Systems", *Proc. of 1994 American Cotnrol Conference* , Baltimore, Maryland, June 29-July 1, 1994.

[12] G. P. Papavassilopoulos, "Iterative techniques for the Nash solution in quadratic games with unknown parameters", *SIAM J. Control and Optimization* vol. 34 no. 4 pp 821-834, 1986.