ISSN 0005-1179 (print), ISSN 1608-3032 (online), Automation and Remote Control, 2025, Vol. 86, No. 6, pp. 577–588.

CONTROL IN SOCIAL ECONOMIC SYSTEMS

Informational Control of Strategies in an n-Player Oligopoly Game with Reflexive Behavior

M. I. Geraskin

Samara University, Samara, Russia
e-mail: innovation@ssau.ru

Received November 29, 2024
Revised March 21, 2025
Accepted March 25, 2025

Abstract—This paper is devoted to an n-player oligopoly game with quantity competition under general demand and cost functions. Players are assumed to be reflexive: each player conjectures about the strategies of all other players. As a result, the subsets of players with different Stackelberg leadership levels are formed in this game (a game with multilevel leadership). The reflexion of players is formalized by conjectural variations, i.e., players' expectations regarding the impact of their actions on the counterparty's action. The problem of controlling the strategy of one player (the controlled player) by the other $(n-1)$ players (the Principal) is investigated, and an optimal Nash equilibrium is established in terms of the Principal's utility criterion. A hierarchical game model of players' interactions is proposed, and the dependence of the maximum of the Principal's utility function on the vector of the sums of conjectural variations (SCV) of all players is found within this model. The dependence is used to calculate the controlled player's SCV value optimizing the Principal's utility function. An informational control method is developed, enabling the Principal to induce the controlled player to choose the reaction function optimal from the former's standpoint.

Keywords: oligopoly, conjectural variation, Stackelberg leadership, hierarchical game

DOI: 10.31857/S0005117925060065

1. INTRODUCTION

The oligopoly game is an aggregative game [1], i.e., one where each player's payoff depends on the sum (aggregate) of the actions of all players. The solution of this game is the Cournot–Nash equilibrium [2, 3]. Based on a conjectural variation, H. Stackelberg was the first to define the leader's strategy in the game as opposed to the follower's strategy [4].

The prerequisite for the concept of a conjectural variation in an aggregative game was the comprehension that, when choosing their optimal actions, players will inevitably expect the optimal behavior of their rivals (i.e., they will perform reflexion). Consequently, the conjectural variation is a mathematical formalization of the mental process of reflexion [5], in this case being interpreted as a thought operation executed by some player to calculate the optimal reaction (best response) of another player to the former's action. As a rule, a quantity conjectural variation is considered, which characterizes the player's expected reciprocal change in the counterparty's action (supply quantity), optimizing the latter's utility function under the action chosen by the former. In modern research, the conjectural variation is widely used to analyze Stackelberg leadership in two directions as follows. First, an increase in the number of reflexive players leads to the emergence of multiple Stackelberg leaders in the game [6]. Second, deeper reflexion causes the emergence of higher-level leaders (multilevel leadership) [7]. The second aspect is expressed in the hierarchy of players' conjectures, bringing to the following hierarchy of their mental types: 1) a follower, who makes no conjectures regarding the strategies of the environment (its conjectural variation is therefore zero); 2) a (first-level) Stackelberg leader, who expects followers in the environment; 3) a second-level (or higher-level) Stackelberg leader, who expects first-level (or other lower-level) Stackelberg leaders in the environment. The hierarchy of mental types determines the reflexion rank of player r as the number in the described sequence of mental types. Note that the above hierarchy is constructed only in the players' beliefs (in this case, we have a game with phantom players); in fact, however, there is a nonhierarchical game with equal players, investigated in most studies of the oligopoly problem. As an exception, a hierarchical aggregative game with Principal's control was considered in [8] as an incentive problem with players (universities) institutionally dependent on the Principal (the government).

Given the above stratification of leaders, depending on the awareness of each player, players of different mental types can coexist in an oligopoly game, and a player of a definite mental type will choose a predictable action according to its conjectural variation. Therefore, it becomes possible to change, in a purposeful way, some player's action by forming a definite information field for him/her. This possibility leads to the well-known concept of informational control. The idea of informational control [9–11] is based on the formation of a purposeful sequence of opinions in a social group depending on the opinions of the so-called influence agents. Formally speaking, informational control is intended to induce purposefully the desired way of thinking, set by a control authority, for one or several players.

In the context of oligopoly games, the concept of informational control is constructed as follows. Consider a group consisting of $n-1$ players, further denoted by the symbol j. Let this group strive to achieve a favorable action of a non-group player i. To do so, the group performs actions from which player i concludes on some reflexion rank r of the group. Therefore, for player i, the optimal reflexion rank is $r+1$, which corresponds to definite values of its conjectural variations; in turn, they predetermine the desired action of this player for the group. For a particular realization of such a control process, it is necessary to find the sum of conjectural variations (SCV) of player i that are optimal (consistent) from the group's standpoint and then determine the dependences of the equilibrium actions of all players on the parameters of their mental type.

In this paper, we consider a procedure for calculating the optimal SCV value of a certain player in terms of the environment's utility criterion, a method for estimating the player's mental type corresponding to this SCV value or its reaction function, and an algorithm for calculating the group's actions inducing the required player's response.

2. THE BASIC OLIGOPOLY GAME MODEL

The game-theoretic model describes the interactions of n players representing firms in an oligopoly market. By a traditional assumption [6], these firms offer an identical product to the market with a common decreasing inverse demand function; in the case of quantity competition, they choose actions in the form of supply quantities. Players are rational, i.e., maximize individual action-concave utility functions $\pi_i(Q, Q_i) = P(Q)Q_i - C_i(Q_i)$; in addition, they are informed about the utility functions of the environment and choose their actions simultaneously, once, and independently. Then the basic model of player's action choice is described by

$$\max_{Q_i \geqslant 0} \pi_i(Q, Q_i) = \max_{Q_i \geqslant 0} \left[P(Q)Q_i - C_i(Q_i) \right], \quad i \in N = \{1, \ldots, n\}, \tag{1}$$

$$Q = \sum_{i \in N} Q_i, \tag{2}$$

where Q_i and π_i denote the action and utility function of player i; Q is the aggregate of actions (the total action of all players); N stands for the set of players; n is the number of players; $P(Q)$ is the inverse demand function, $P'_Q < 0$; finally, $C_i(Q_i)$ is the cost function of player i, $C'_{Q_i} > 0$.

For a known vector of conjectural variations, the Nash equilibrium in the game $\Gamma = \langle N, \{Q_i, i \in N\}, \{\pi_i, i \in N\}\rangle$ is determined by solving the following system of reaction equations:

$$\frac{\partial \pi_i(Q_i^*, \rho_{ij})}{\partial Q_i} = 0, \quad i, j \in N, \tag{3}$$

where $\rho_{ij} = Q'_{jQ_i}$ is the quantity conjectural variation of player i, i.e., its expectation regarding the supply quantity change of player j in response to the unit increase in the supply quantity of player i; Q_i^* is the equilibrium value.

The optimal conjectural variation (also called consistent in the literature) is calculated from equation (3) of player j, i.e., this variation corresponds to its best response. For the utility function (1), system (3) takes the form

$$P(Q) + (1 + S_i^r) Q_i P'_Q - C'_{iQ_i} = 0, \quad i \in N, \quad S_i^r = \sum_{j \in N \backslash i} \rho_{ij}^r, \tag{4}$$

where S_i^r is the SCV value of player i at a reflexion rank r.

In the case of action-independent conjectural variations ($\rho'_{ijQ_i} = 0$), SCV values at an arbitrary reflexion rank are given by the recurrent formula [7]

$$S_i^r = \left(\frac{1}{\sum_{j \in N \backslash i} \dfrac{1}{u_j - S_j^{r-1} + 1}} - 1 \right)^{-1}, \tag{5}$$

where $u_i = -1 + \frac{P'_{Q_i} + (1 + S_i^{r-1}) Q_i P''_{QQ_i} - C''_{iQ_iQ_i}}{|P'_Q|}$ is a nonlinearity coefficient expressing the influence of the nonlinearity of the demand and cost functions on the unimodality of the utility function of player i.

Due to (4), the Nash equilibrium vector $\mathbf{Q}^* = \{Q_i^*, i \in N\}$ in the n-player oligopoly game depends on the SCV vector $\mathbf{S^r} = \{S_i^r, i \in N\}$ (the conjectural variations of all players). Therefore, an inverse dependence also exists: given a known action vector $\mathbf{Q}^* = \{Q_i^*, i \in N\}$, it is possible to establish a vector $\mathbf{S^r} = \{S_i^r, i \in N\}$ inducing these actions of the players. On this basis, let us consider some tools for controlling (manipulating) the player's behavior by the environment.

3. AN OPTIMAL CONTROL MODEL FOR PLAYER'S BEHAVIOR

Consider the following modification of the basic oligopoly game model in the form of a hierarchical game. Player i is the controlled object, and its environment (i.e., the other players) acts as the control subject, also called the Principal. Thus, we study a hierarchical game of the Principal–agent type (Fig. 1). For the sake of simplicity, the environment of player i will be assigned number j (i.e., $j = \{N \backslash i\}$). The environment has a common goal: induce player i to choose an optimal action $\overline{Q}_i$ in terms of the former's utility functions. Therefore, let us define the Principal's goal function as the vector of the utility functions of the environment players. Since the latter are supposed to be identical, the goal function can be represented as a single function of the form

$$\pi^{(i)} = \pi_j, \quad j \in \{1, \ldots, i-1, i+1, \ldots, n\},$$

and briefly written as

$$\pi = \pi^{(i)}.$$

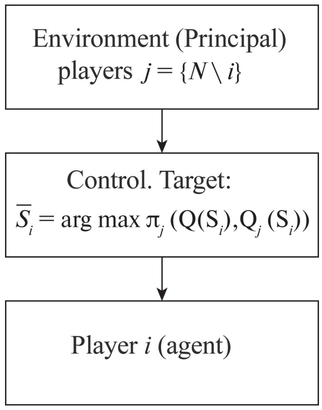

Fig. 1. The diagram of a hierarchical game.

Now we describe the main assumptions adopted in the hierarchical game analysis.

1) The utility functions of all environment players are identical, i.e., these players have the same cost functions with the same coefficient:

$$C_j(Q_j) = C_k(Q_k), \quad \pi_j(Q_j) = \pi_k(Q_k)\forall j, \quad k = \{N\backslash i\}.$$

2) The awareness of the players is described by the following awareness sets I_i and I_j:

— The awareness set of the controlled player includes the set of players and the actions and utility functions of all players:

$$I_i = \{N, Q_k, \pi_k, k \in N\}.$$

— The awareness set of the environment encompasses the set of players, the actions and utility functions of all players, the SCV values of the environment players, and the Principal's goal function:

$$I_j = \{N, Q_k, \pi_k, S_j, \pi, j \in N\backslash i, k \in N\},$$

where $\pi = \pi_j$, $j \in N\backslash i$, is the goal function of the environment (Principal), identical to the utility functions of the environment players.

3) All players calculate their conjectural variations based on the sets I_i and I_j. These variations may be optimal (i.e., consistent with the utility functions of the players) or be determined by the players using the actions of the other players. If players' actions are inconsistent with their utility functions, the players are action-oriented. At each time instant of the game, either the SCV of the controlled player or the SCV of the environment has a constant value:

$$S_i = \text{const} \vee S_j = \text{const},$$

since to change the SCV, players estimate the actions of the other players at the previous time instant.

The environment controls the behavior of player i through its reflexion in the following procedure:

— The environment calculates the target SCV value $\overline{S}_i$ of player i that is optimal in terms of the former's utility functions:

$$\overline{S}_i = \arg\max_{S_i} \pi_j(Q(S_i), Q_j(S_i)).$$

— The environment performs actions $\overline{Q}_j$ that will induce player i to choose the SCV value $\overline{S}_i$ and, consequently, the optimal action $\overline{Q}_i$ from the environment's standpoint.

The second stage of the above behavior control procedure explains the meaning of Assumption 3 in the context of the players' dual approach to estimating the conjectural variations. The environment controls player i by performing the actions $\overline{Q}_j$ determined not from the maximum of its utility function but from the condition of inducing this player to choose $\overline{S}_i$; therefore, having predicted the SCV values from the environment's utility functions, player i will arrive at a contradiction. Consequently, the player in this contradiction will favor the method for estimating the SCV values by the environment's actions. Therefore, two alternatives are considered when estimating the SCV values: if the players' actions are consistent with their utility functions, the other players will estimate the optimal SCV values; if the SCV values found by the players' actions do not coincide with the SCV values based on the utility functions, the first estimation method as more realistic will be given priority.

4. METHODS FOR CALCULATING OPTIMAL CONTROL

The behavioral control of player i is based on the dependence of each player's utility function on the SCV values of all players, see system (4). Therefore, we first derive a formula for the maximum of the environment's utility function depending on the SCV vector of all players.

Proposition 1. *The maximum value of the environment's utility function is given by*

$$\pi_j^*(\mathbf{S}^r) = P\left[Q^*(\mathbf{S}^r)\right] Q_j^*(\mathbf{S}^r) - \int\limits_0^{Q_j^*(\mathbf{S}^r)} \left[P(Q^*) + (1+S_j^r)Q_j^* P_Q'\right] dQ_j + C_j(0), \tag{6}$$

where $\mathbf{S}^r = \{S_k^r, k \in N\}$ *denotes the SCV vector.*

Proof of Proposition 1. Let us express the marginal cost from equation (4), written for the environment:

$$C'_{jQ_j} = P(Q^*) + (1+S_j^r)Q_j^* P_Q'.$$

Integration over Q_j yields

$$C_j(Q_j^*) = \int\limits_0^{Q_j^*} C'_{jQ_j} dQ_j + C_j(0) = \int\limits_0^{Q_j^*} \left[P(Q^*) + (1+S_j^r)Q_j^* P_Q'\right] dQ_j + C_j(0);$$

after substitution into the environment's utility function, we obtain

$$\pi_j^* = P(Q^*)Q_j^* - C_j(Q_j^*) = P(Q^*)Q_j^* - \int\limits_0^{Q_j^*} \left[P(Q^*) + (1+S_j^r)Q_j^* P_Q'\right] dQ_j + C_j(0),$$

where $C_j(0)$ is fixed costs. Note that in this formula, the equilibrium action of player i, Q_j^*, the equilibrium price $P(Q^*)$, and the equilibrium aggregate action Q^* are all functions of the SCV $\mathbf{S}^r = \{S_i^r, i \in N\}$. Thus, the maximum utility of the environment also depends on this vector, which finally implies (6). ∎

Let us express the SCV of player i that is optimal in terms of the environment's utility criterion:

$$\overline{S}_i = \arg\max_{S_i} \pi_j^*(\mathbf{S}^r).$$

Proposition 2. *The optimal SCV value $\overline{S}_i$ of the controlled player, in terms of the environment's utility function, is calculated from the equation*

$$2(1+S_j^r)Q_j^* Q_{jS_i}^{*'} P_Q' + \left((1+S_j^r)P''_{QS_i} + P_Q' \frac{\partial S_j^r}{\partial S_i}\right) Q_j^{*2} = 0 \tag{7}$$

under the condition

$$2\frac{\partial S_j^r}{\partial S_i}Q^{*'}_{jS_i} + (1+S_j^r)Q^{*'}_{jS_i} < 0, \tag{7a}$$

in the case of a weak impact of the SCV change on the equilibrium shift and a relatively small value of the second derivative of the environment's SCV with respect to the player's SCV compared to the first derivative.

Proof of Proposition 2.

With the Leibniz integral rule (differentiation under the integral sign) applied to the second term in (6), where the integration limits are functions of the parameter S_i, we write the necessary first-order maximum condition for the function (6):

$$\pi^{*'}_{jS_i} = P'_{S_i}Q_j^* + PQ^{*'}_{jS_i} - \left\{\left[P+(1+S_j^r)Q_j^*P'_Q\right]Q^{*'}_{jS_i} + \int\limits_0^{Q_j^*}\left[P(Q^*)+(1+S_j^r)Q_j^*P'_Q\right]'_{S_i}dQ_j\right\} = 0. \tag{7b}$$

The integral in this equation can be transformed as follows:

$$I = \int\limits_0^{Q_j^*}\left[P+(1+S_j^r)Q_j^*P'_Q\right]'_{S_i}dQ_j = \int\limits_0^{Q_j^*}\left[P'_{S_i} + (1+S_j^r)(Q^{*'}_{jS_i}P'_Q + Q_j^*P''_{QS_i}) + Q_j^*P'_Q\frac{\partial S_j^r}{\partial S_i}\right]dQ_j.$$

In the integrand, the parameters $Q_j^*, Q^*, Q^{*'}_{jS_i}, P'_{S_i}(Q^*)$, and $P'_Q(Q^*)$ characterize the equilibrium of all players, therefore being independent of the action Q_j of player j; the parameter S_j^r and hence $\frac{\partial S_j^r}{\partial S_i}$ weakly depend on the action Q_j (see the proof in [7]). Therefore, the following variables are considered to be independent of Q_j:

$$Q_j^*, Q^*, P'_{S_i}, P'_Q, \frac{\partial S_j^r}{\partial S_i}, Q^{*'}_{jS_i}, S_j^r.$$

In this case, the integral becomes

$$I = (P'_{S_i} + (1+S_j^r)Q^{*'}_{jS_i}P'_Q)Q_j^* + \left((1+S_j^r)P''_{QS_i} + P'_Q\frac{\partial S_j^r}{\partial S_i}\right)Q_j^{*2}.$$

Substituting it into (7a) gives the expression

$$\pi^{*'}_{jS_i} = P'_{S_i}Q_j^* + PQ^{*'}_{jS_i} - \left[P+(1+S_j^r)Q_j^*P'_Q\right]Q^{*'}_{jS_i} - (P'_{S_i} + (1+S_j^r)Q^{*'}_{jS_i}P'_Q)Q_j^*$$
$$-\left((1+S_j^r)P''_{QS_i} + P'_Q\frac{\partial S_j^r}{\partial S_i}\right)Q_j^{*2} = -2(1+S_j^r)Q_j^*Q^{*'}_{jS_i}P'_Q - \left((1+S_j^r)P''_{QS_i} + P'_Q\frac{\partial S_j^r}{\partial S_i}\right)Q_j^{*2}.$$

Then equation (7b), used to calculate the optimal SCV value of player i in terms of the environment's utility criterion, takes the form (7).

The second-order maximum condition for the function (6) is given by

$$\pi^{*''}_{jS_iS_i} = -\left\{2\frac{\partial S_j^r}{\partial S_i}Q_j^*Q^{*'}_{jS_i}P'_Q + 2(1+S_j^r)\left[Q_j^*Q^{*''}_{jS_iS_i}P'_Q + Q_j^*(Q^{*''}_{jS_iS_i}P'_Q + Q^{*'}_{jS_i}P''_{QS_i})\right]\right.$$
$$+\frac{\partial S_j^r}{\partial S_i}P''_{QS_i}Q_j^{*2} + (1+S_j^r)(P'''_{QS_iS_i}Q_j^{*2} + 2Q_j^*Q^{*'}_{jS_i}P''_{QS_i})$$
$$\left.+P''_{QS_i}\frac{\partial S_j^r}{\partial S_i}Q_j^{*2} + P'_Q\left(\frac{\partial^2 S_j^r}{\partial S_i^2}Q_j^{*2} + 2Q_j^*Q^{*'}_{jS_i}\frac{\partial S_j^r}{\partial S_i}\right)\right\} < 0.$$

After straightforward transformations, we obtain

$$\frac{\partial S_j^r}{\partial S_i} Q_j^* (2P''_{QS_i} Q_j^* + 4Q^{*'}_{jS_i} P'_Q)$$

$$+ (1 + S_j^r) \left(2(Q^{*'}_{jS_i})^2 P'_Q + 2Q_j^* Q^{*''}_{jS_iS_i} P'_Q + P'''_{QS_iS_i} Q_j^{*2} + 4Q_j^* Q^{*'}_{jS_i} P''_{QS_i} \right) + P'_Q \frac{\partial^2 S_j^r}{\partial S_i^2} Q_j^{*2} > 0.$$

Due to the assumption of a weak influence of the SCV change on the equilibrium shift,

$$P''_{QS_i} = 0, \quad P'''_{QS_iS_i} = 0, \quad Q^{*''}_{jS_iS_i} = 0.$$

Due to the assumption of a small value of the second derivative of the environment's SCV with respect to the player's SCV compared to the first derivative,

$$\left| \frac{\partial^2 S_j^r}{\partial S_i^2} \right| \ll \left| \frac{\partial S_j^r}{\partial S_i} \right| \Rightarrow \frac{\partial^2 S_j^r}{\partial S_i^2} \approx 0.$$

Under these assumptions, the above condition becomes

$$4 \frac{\partial S_j^r}{\partial S_i} Q_j^* Q^{*'}_{jS_i} P'_Q + 2(1 + S_j^r)(Q^{*'}_{jS_i})^2 P'_Q > 0.$$

Since $P'_Q < 0$ by the inverse demand function property and $Q^{*'}_{jS_i} > 0$ by the Stackelberg leadership property, we finally arrive at a sufficient maximum condition for the solution (7) in the form (7a). ■

Let us present a methodology for calculating the derivatives $Q^{*'}_{jS_i}$ in equation (7).

Proposition 3. *The derivatives $Q^{*'}_{jS_i}$ are the roots of the following system of linear equations:*

$$\sum_{k \in N} a_{jk} Q^{*'}_{kS_i} = b_j, \quad j \in N, \tag{8}$$

where $b_j = -\left(\frac{\partial S_j^r}{\partial S_i} Q_j^* P'_Q + (1 + S_j^r) Q_j^* P''_{QS_i} C''_{jQ_jS_i} \right),$

$$a_{jk} = \begin{cases} \gamma_{jk} + P'_Q & \text{for } j \neq k \\ \gamma_{jk} + P'_Q + (1 + S_j^r) P'_Q & \text{for } j = k, \end{cases}$$

$$\gamma_{jk} = P'_{Q_k}(Q^*) + (1 + S_j^r) \left\{ Q_j^* P''_{QQ_k} + P'_Q Q'_{jQ_jQ_k} \right\} - C''_{jQ_jQ_k}.$$

Proof of Proposition 3.

Assuming that the optimal actions of all environment players (system (4)) depend on S_i, we consider the n implicit functions

$$F_j(Q^*, S_i) = P(Q^*) + (1 + S_j^r) Q_j^* P'_Q - C'_{jQ_j} = 0, \quad j \in N.$$

In this case, the derivatives $Q^{*'}_{jS_i}$ of the implicit functions with several independent variables are calculated from the following system [12]:

$$\sum_{k \in N} \frac{\partial F_j}{\partial Q_k} \frac{\partial Q_k}{\partial S_i} + \frac{\partial F_j}{\partial S_i} = 0, \quad j \in N, \tag{8a}$$

where $\dfrac{\partial F_j}{\partial Q_k} = P'_Q(Q^*) + (1 + S^r_j)\left\{Q^*_j P''_{QQ_k} + P'_Q Q^{*'}_{jQ_k}\right\} - C''_{jQ_jQ_k} = \gamma_{jk}$,

$$\frac{\partial F_j}{\partial S_i} = P'_Q(Q^*)Q^{*'}_{S_i} + \frac{\partial S^r_j}{\partial S_i} Q^*_j P'_Q$$
$$+ (1 + S^r_j)\left\{Q^*_j P''_{QS_i} + P'_Q Q^{*'}_{jS_i}\right\} - C_{jQ_jS_i}, \quad Q^{*'}_{S_i} = \sum_{k\in N} Q^{*'}_{S_i}.$$

Here, γ_{jk} stands for the component without an explicit dependence of the desired parameters $Q^{*'}_{jS_i}$, further denoted by $x_j = Q^{*'}_{jS_i}$. In this case, system (8a) has the form

$$\sum_{k\in N} \gamma_{jk} x_k + P'_Q \sum_{k\in N} x_k + (1 + S^r_j)P'_Q x_j + \frac{\partial S^r_j}{\partial S_i} Q^*_j P'_Q + (1 + S^r_j)Q^*_j P''_{QS_i} - C''_{jQ_jS_i} = 0.$$

With $b_j = -\left(\dfrac{\partial S^r_j}{\partial S_i} Q^*_j P'_Q + (1 + S^r_j)Q^*_j P''_{QS_i} - C''_{jQ_jS_i}\right)$ and

$$a_{jk} = \begin{cases} \gamma_{jk} + P'_Q & \text{for } j \neq k \\ \gamma_{jk} + P'_Q + (1 + S^r_j)P'_Q & \text{for } j = k, \end{cases}$$

we write the following system of linear algebraic equations for the unknowns x_k: $\sum_{k\in N} a_{jk}x_k = b_j$, $j \in N$, which matches (9). ■

System (9) allows determining the derivatives $Q^{*'}_{jS_i}$ as functions of the SCV values S^r_j of all players, including the desired value $\overline{S}_i$. Thus, we have provided a method for calculating the target SCV value of the controlled player: solve equation (7) considering the derivatives $Q^{*'}_{jS_i}$ expressed through S_i from the solution of system (8).

5. A MECHANISM FOR CALCULATING OPTIMAL CONTROL

Consider a possible method for the environment to induce the controlled player to choose the target SCV value $\overline{S}_i$. As an illustration, we will interpret the considerations by the example of duopoly. Let us start with the description of the classical principle of Stackelberg leader emergence in the game of initially equal participants (Fig. 2), i.e., from the situation of Cournot responses. (Here, the equilibrium and Cournot responses are indicated by the symbol K.) As is known [13], the optimal reaction functions of the players in the linear Cournot duopoly model have the form

$$Q_1 = \frac{\alpha_1 - Q_2}{2}, \quad Q_2 = \frac{\alpha_2 - Q_1}{2},$$

where $\alpha_1 = \frac{a - B_i}{b}$, with a and b representing the maximum price and the rate of price reduction in the inverse demand function, respectively, and B_i specifying the marginal cost of player i. However, if the reaction functions were unknown to the players, they could reconstruct these functions from observations of each other's actions. When treated as a potential leader, the first player observes in the game dynamics the response of the second player: the second player takes the action Q^t_2 in response to the action M_1 and the action $Q^{(t+1)}_2$ in response to the first player's reciprocal action Q^t_1. Based on these observations, the first player (with the reaction function R^K_1) determines the second player's reaction function R^K_2 and calculates from it the conjectural variation (equal to the SCV in the duopoly) as follows:

$$S_1 = Q'_{2Q_1} = -\frac{1}{2}.$$

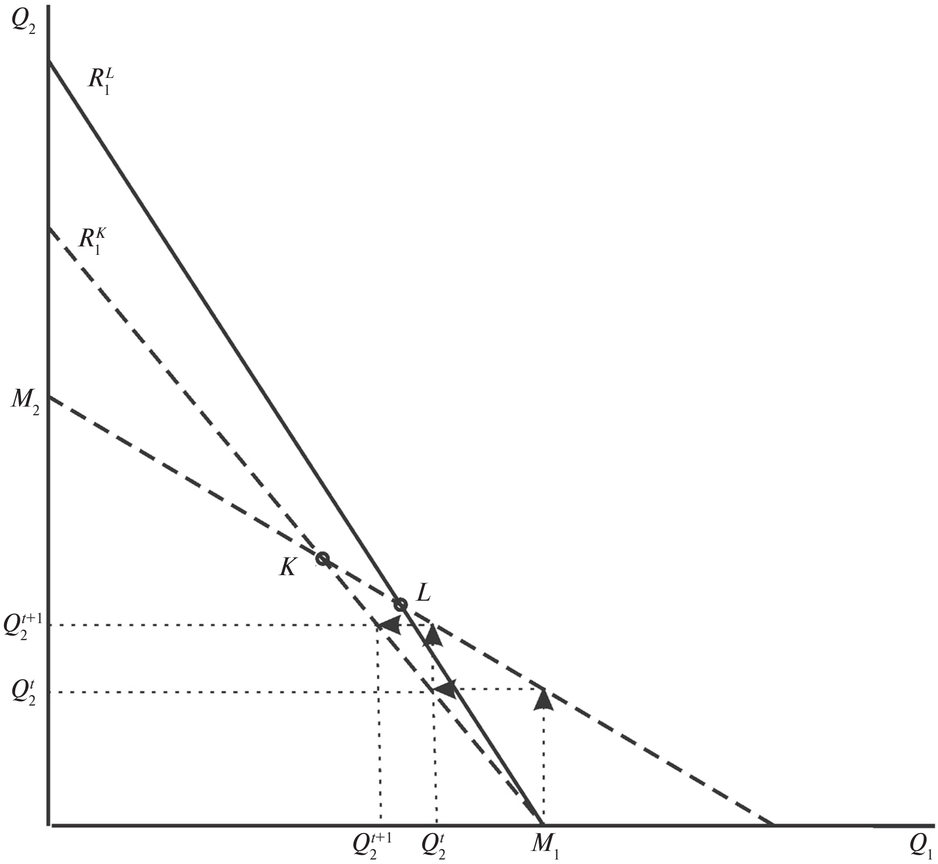

Fig. 2. The emergence of a Stackelberg leader: an illustration.

As a result, the reaction function of the first player is transformed to

$$Q_1 = \frac{\alpha_1 - Q_2}{2 + S_2} = \frac{\alpha_1 - Q_2}{2 - \frac{1}{2}};$$

and this player becomes a Stackelberg leader (in Fig. 2, its response and the corresponding equilibrium are indicated by the symbol L). In other words, observing the response $Q_2 = \frac{\alpha_2 - Q_1}{2+0}$ of the second player, the first player has revised its SCV: the SCV $S_2 = 0$ of the second player has induced the first player to set the SCV value $S_1 = -\frac{1}{2}$.

Extending this procedure to multilevel leadership, we can formulate the following law: for a certain player to change its conjectural variation to some given value corresponding to a definite-level Stackelberg leader, this player must observe another player's action corresponding to the response of a previous-level Stackelberg leader. Consequently, the other player must create the so-called phantom agent acting not according to its true reaction function, so this response will be called phantom and denoted by the symbol f. Formally, this means that to induce player i to set the SCV value $\overline{S}_i$, the environment must act according to the phantom reaction function

$$Q_j^f = \frac{\alpha_j - Q_i}{2 + S_j^f}$$

under the condition

$$S_i = Q_{jQ_i}^{f'} = -\frac{1}{2 + S_j^f} = \overline{S}_i.$$

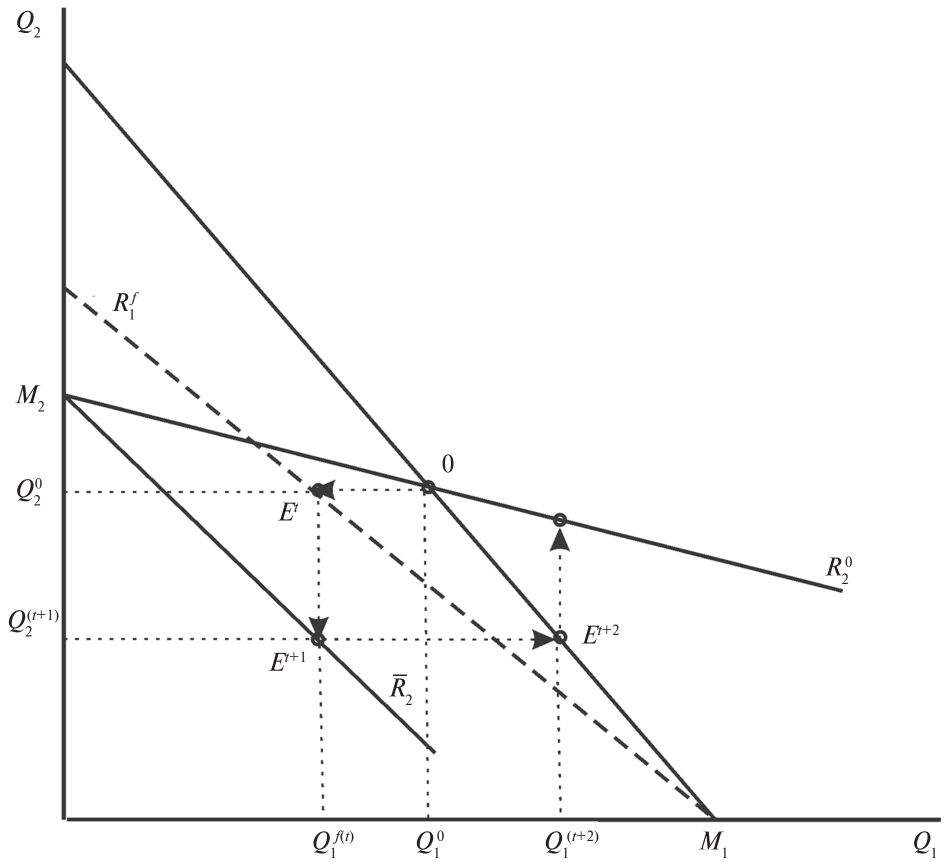

Fig. 3. The informational control process: an illustration.

Hence, the SCV of the environment for this action is given by

$$S_j^f = -\frac{1}{\overline{S}_i} - 2.$$

This general principle was proved earlier [7] as formula (5). Based on the latter, we provide a method for calculating the phantom response in the general case of nonlinear cost functions, when the reaction functions cannot be expressed explicitly. Let us summarize the considerations as follows.

Proposition 4. *The environment's phantom reaction function inducing the controlled player to set the target SCV value $\overline{S}_i$ corresponds to the environment's SCV value S_j^f calculated by solving the equation*

$$\overline{S}_i = \left(\frac{1}{\sum_{j \in N \setminus i} \dfrac{1}{u_j - S_j^f + 1}} - 1 \right)^{-1}. \tag{9}$$

Based on this principle, we describe the informational control process in the above duopoly example, assuming that the target SCV value of the controlled (second) player is $\overline{S}_2 = -\frac{3}{4}$. In other words, the first player strives to make the second player's response match a third-level Stackelberg leader. (Recall that in the linear duopoly, the leaders of the first, second, and third levels have the SCV values $-\frac{1}{2}, -\frac{2}{3}$, and $-\frac{3}{4}$, respectively.) We assign the number "0" to the initial equilibrium state, i.e., the equilibrium actions are Q_1^0 and Q_2^0, the SCV values of the players are S_1^0 and S_2^0, and the reaction functions are R_1^0 and R_2^0. The control process is illustrated in Fig. 3.

At the time instant t, the first player calculates its action using the phantom reaction function R_1^f of the second-level leader: $Q_1^{f(t)} = \frac{\alpha_1 - Q_2^0}{2-\frac{2}{3}}$. Therefore, the game state at this instant (the point E^t) is described by the action vector $(Q_1^{f(t)}, Q_2^0)$.

At the time instant $t+1$, the second player calculates $\overline{S}_2 = -\frac{3}{4}$ using this action and passes to the target reaction function $\overline{R}_2$, given by the equation $Q_2^{t+1} = \frac{\alpha_2 - Q_1^{f(t)}}{2-\frac{3}{4}}$. With this reaction function, it responds to the action $Q_1^{f(t)}$ by the action $Q_2^{t+1} = \frac{\alpha_2 - Q_1^{f(t)}}{2-\frac{3}{4}}$. At this time instant, the game state is denoted by the point E^{t+1}.

At the time instant $t+2$, the first player performs an action according to its true reaction function $Q_1^{t+2} = \frac{\alpha_1 - Q_2^{t+1}}{2+S_2^0}$ (maximizes its utility for the initial equilibrium). For the combination (Q_1^{t+2}, Q_2^{t+1}), the game state is denoted by the point E^{t+2}.

At the subsequent time instants of the game, the initial equilibrium is restored according to the above procedure (see Fig. 2). With this procedure, the first player gains an additional utility at the time instants $t+1$ and $t+2$ since the second player performs actions according to the SCV target value $\overline{S}_2$.

The game state returns to the initial equilibrium in an infinite number of steps. Therefore, the Principal's control efficiency can be assessed by the following condition:

$$\sum_{\tau=t}^{\infty} \pi^{*(\tau)} e^{-\rho\tau} - \pi^{*(0)} \sum_{\tau=t}^{\infty} e^{-\rho\tau} \geqslant 0,$$

where $\pi^{*(0)}$ is the Principal's maximum utility at the initial equilibrium; $\pi^{*(\tau)}$ is the Principal's maximum utility at a time instant τ; finally, ρ is the discount rate.

6. CONCLUSIONS

This paper has developed an informational control method for the actions of a given player in an oligopoly game model: other players perform a control action inducing this player to make an optimal response from the environment's standpoint. The foundations of this informational control are, first, the dependence of players' actions on their conjectures regarding the expected actions of counterparties and, second, the a priori unawareness of players regarding each other's conjectures due to the dual nature of their conjectural variations. On the one hand, the variations are based on the analysis of the utility functions of the environment; on the other, a player cannot ignore the nature of the responses of its environment. Therefore, the following hypothesis has been adopted above: in the case of contradiction between these two approaches, the players estimate the conjectural variations by each other's actions, which are more reliable information. Under this hypothesis, without changing its conjectural variations, the environment can perform an action as if on behalf of a phantom player that induces the controlled player to respond in a way favorable to the environment, and the latter interprets this action as a signal of a change in the environment's true reaction and performs the desired action.

The main results of this study can be summarized as follows. A hierarchical game model of players' interactions in an oligopoly has been presented, where the environment is treated as the Principal and some player as a controlled object. An explicit expression has been derived for the maximum of the environment's utility function depending on the SCV vector of all players; this expression allows finding the controlled player's SCV value optimizing the environment's utility function. A methodology for calculating the target SCV value of the controlled player from the environment's standpoint has been defined. An iterative procedure has been developed to induce the controlled player to choose a reaction function corresponding to the target SCV value; as a result, the environment maximizes its utility.

The optimal control problem has been formulated for an n-player game with general utility functions. Therefore, it is impossible to obtain explicit solutions to analyze the game results in the developed techniques and procedures. Hence, the next stage of research is to apply these general tools to particular utility functions and carry out numerical experiments.

REFERENCES

1. Anderson, S.P., Erkal, N., and Piccinin, D., Aggregative Games and Oligopoly Theory: Short-Run and Long-Run Analysis, *RAND J. Econom.*, 2020, vol. 51, no. 2, pp. 470–495.
2. Cournot, A.A., *Researches into the Mathematical Principles of the Theory of Wealth*, London: Hafner, 1960. (Original 1838.)
3. Nash, J.F.Jr., Non-cooperative Games, *Annal. Mathem.*, 1951, vol. 54, no. 2, pp. 286–295.
4. Stackelberg, H., *Market Structure and Equilibrium*, 1st ed. Translation into English, Bazin, Urch & Hill, Springer, 2011. (Original 1934.)
5. Novikov, D.A. and Chkhartishvili, A.G., *Reflexion and Control: Mathematical Models*, Leiden: CRC Press, 2014.
6. Julien, L.A., On Noncooperative Oligopoly Equilibrium in the Multiple Leader–Follower Game, *Eur. J. Oper. Res.*, 2017, vol. 256, no. 2, pp. 650–662.
7. Geraskin, M.I., The Properties of Conjectural Variations in the Nonlinear Stackelberg Oligopoly Model, *Autom. Remote Control*, 2020, vol. 81, no. 6, pp. 1051–1072.
8. Malsagov, M., Ougolnitsky, G., and Usov, A., A Differential Stackelberg Game Theoretic Model of the Promotion of Innovations in Universities, *Advances Syst. Sci. Appl.*, 2020, vol. 20, no. 3, pp. 166–177.
9. Gubanov, D.A., Novikov, D.A., and Chkhartishvili, A.G., Informational Influence and Information Control Models in Social Networks, *Control Sciences*, 2009, no. 5, pp. 28–35.
10. Gubanov, D.A., Novikov, D.A., and Chkhartishvili, A.G., *Social Networks: Models of Information Influence, Control and Confrontation*, Cham: Springer, 2019.
11. Gubanov, D.A., Petrov, I.V., and Chkhartishvili, A.G., Multidimensional Model of Opinion Dynamics in Social Networks: Polarization Indices, *Autom. Remote Control*, 2021, vol. 82, no. 10, pp. 1802–1811.
12. Korn, G.A. and Korn, T.M., *Mathematical Handbook for Scientists and Engineers. Definitions, Theorems, and Formulas for Reference and Review*, New York: Dover Publications, 2000.
13. Algazin, G.I. and Algazina, D.G., Modeling the Dynamics of Collective Behavior in a Reflexive Game with an Arbitrary Number of Leaders, *Informatics and Automation*, 2022, vol. 21, no. 2, pp. 339–375.

This paper was recommended for publication by A.G. Chkhartishvili, a member of the Editorial Board