ISSN 0005-1179 (print), ISSN 1608-3032 (online), Automation and Remote Control, 2025, Vol. 86, No. 7, pp. 644–654.

A Control Algorithm for Player's Actions in a Triopoly Game under Linear Demand and Cost Functions

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 considers a game of three oligopolistic firms (a triopoly) with the linear demand and cost functions of players. The reflexive behavior of players is investigated through a formalization with conjectural variations, i.e., players' expectations regarding the impact of their actions on the counterparty's action. A method is developed to calculate the optimal (consistent) sum of conjectural variations of a given player in terms of the utility functions of the other players (environment). An algorithm is proposed to control the player's actions by the environment: implementing this algorithm forms a given mental type of the player and predetermines the latter's goal-oriented behavior. Computer simulations are carried out for a hierarchical game in which the environment (the Principal) applies informational control to the player's actions, using the Russian telecommunications market as an example. According to the simulation results, the control algorithm is effective and increases the environment's utility.

Keywords: oligopoly, conjectural variation, hierarchical game, optimal control, telecommunications market

DOI: 10.31857/S0005117925070043

1. INTRODUCTION

The game-theoretic model of oligopoly as an economic sphere in which several firms sell an identical product formalizes the interaction of these firms as players. As a rule, players are treated as price takers: they face the equilibrium market price on the aggregate demand curve and perform game actions (choose their strategies) in the form of supply quantities of the product. The key feature of this game is that the equilibrium price depends on the actions of all players, so when choosing a strategy, players must predict the actions of their counterparties (the environment). The change in the environment's action expected by a player in response to a unit change in the latter's action is called a conjectural variation. The player's a priori unawareness regarding the conjectural variations of the environment predetermines the fundamental complexity of the oligopoly game. As is known from the classical models of A. Cournot [1] and H. Stackelberg [2], the optimal strategies of players depend on two factors: the inverse demand function for the product and the cost functions of the players. Therefore, the simpler form these functions have, the more complete and informative the analysis of oligopoly game outcomes will be.

The oligopoly game with the linear demand function for a product and linear cost functions of players is the most typical case considered by researchers [3–17], since this model is the most convenient tool for studying game strategies. In particular, dynamic oligopoly games based on finite-difference equations of players' responses [3–6] and dynamic games with differential equations describing the change process of the utility functions of players [7–9] were studied in such a formulation. A dynamic game based on fractal differential equations of changes in the players' actions was investigated as well [10]. Also, static oligopoly games were introduced and Cournot and Stackelberg equilibria were compared accordingly [11].

Applied aspects of the oligopoly game with linear demand and cost functions were considered in the context of choosing a firm's tax policy [12], assessing the consequences of merging firms [13], entering a bilateral oligopoly game by a new firm [14], selling product sets [15, 16], and adopting new technology [17].

The above studies involved a non-hierarchical game-theoretic model of oligopoly with initially equal players, even in the case of Stackelberg leadership: being not conditioned by an original hierarchy of players, leadership arises from their asymmetric awareness in game dynamics. The awareness levels were formalized using reflexion ranks r, and reflexion was understood as the process of making conjectures by a player regarding the strategies of other players (the environment) [18]. As a result of this process, each player creates a set of phantom environment players in the mind, which can be followers at reflexion rank 1, first-level Stackelberg leaders at reflexion rank 2, and $(r-1)$th-level Stackelberg leaders at reflexion rank r. In other words, at each reflexion rank, the player thinks that the environment consists of players of a lower reflexion rank. Relating these theoretical considerations to business practice, we emphasize that the decision-maker (DM) representing a firm may have no idea of Stackelberg leadership as a scientific category. Nevertheless, when choosing an appropriate action in the market, the DM must analyze the possible responses of the counterparties (i.e., must perform reflexion). Therefore, if a player correctly predicts the environment's beliefs (i.e., the phantoms coincide with the real players), he/she will dominate in equilibrium in accordance with the leadership level [19]. But in a non-hierarchical model of the game, the leader does not actually control the environment players.

However, in reality, it is easy to imagine a game situation where a certain group of players can choose a common strategy aimed at increasing their payoffs at the expense of one player. Therefore, a topical issue is to study the following hierarchical system of oligopoly players: a group of players (the Principal) controls the strategy of a certain player by choosing a group strategy, inducing this player to choose an optimal strategy in terms of the group's utility functions. In this case, the control process is also based on the above reflexive behavior of players. Below, this problem is investigated for the three-player game model with linear demand and cost functions, which makes the analysis of equilibria explicit.

2. THE NON-HIERARCHICAL GAME WITH LINEAR DEMAND AND COST FUNCTIONS (THE LINEAR MODEL)

The non-hierarchical model of an oligopoly market describes the behavior of n players (firms) supplying an identical product to the market. The consumer demand for this product is characterized by an inverse demand function $P(Q)$, decreasing with the total sales quantity $Q(P'_Q < 0)$. The players are equal and choose actions (strategies) in the form of their supply quantities Q_i, in accordance with their increasing cost functions $C_i(Q_i)$ $(C'_{Qi} > 0)$, with the goal of maximizing the utility function $\pi_i(Q, Q_i) = P(Q)Q_i - C_i(Q_i)$.

Consider the oligopoly game in the case of the linear demand and cost functions of players:

$$P(Q) = a - bQ, \quad A > 0, \quad b > 0, \quad a \gg b,$$
$$C_i(Q_i) = B_{0i} + B_iQ_i, \quad B_{0i} \geqslant 0, \quad B_i > 0,$$

where Q_i denotes the action of player i; a, b, B_{0i}, and B_i are constant coefficients expressing the parameters of the demand function (the maximum price a and the price decline rate b) and the parameters of the cost function (the fixed costs B_{0i} and the marginal costs B_i). The total action of the players is $Q = \sum_{i \in N} Q_i$.

In this case, the players seek to maximize their utility functions π_i as follows:

$$\max_{Q_i \geqslant 0} \pi_i(Q, Q_i) = \max_{Q_i \geqslant 0} [(a - bQ)Q_i - B_{0i} - B_iQ_i], \quad i \in N = \{1, \ldots, n\}, \tag{1}$$

where N stands for the set of players and n is the number of players.

The Nash equilibrium Q_i^* in the non-hierarchical game $\Gamma = \langle N, \{Q_i, i \in N\}, \{\pi_i, i \in N\}\rangle$ is calculated based on the necessary optimality conditions $\frac{\partial \pi_i(Q_i^*, \rho_{ij})}{\partial Q_i} = 0$, $i, j \in N$, which involve the conjectural variation of player i. For the system of the objective functions (1), the necessary optimality conditions have the form

$$a - bQ - b(1 + S_i^r)Q_i - B_i = 0, \quad i \in N, \quad S_i^r = \sum_{j \in N \setminus i} \rho_{ij}^r, \tag{2}$$

where S_i^r is the sum of conjectural variations of player i at reflexion rank r. Due to equations (2), *the sum of conjectural variations (SCV)* of each player is crucial for calculating the game equilibrium: the other parameters in (2) can be considered common knowledge, and the player's SCV value may be unknown to the environment.

With the player's type parameter denoted by $\alpha_i = \frac{a - B_i}{b}$, system (2) can be conveniently written as

$$\alpha_i - (2 + S_i^r)Q_i - Q_{j\sum} = 0, \quad i \in N, \tag{2a}$$

where $Q_{j\sum}$ is the total action of the environment of player i; the environment is designated as the generalized player j.

A comparative analysis of (1) and (2a) allows deriving the maximum profit of player i depending on the latter's equilibrium action, i.e., the function $\pi_i^* = \pi_i(Q_i^*)$, in the case of reflexive behavior of players.

Proposition 1. *Under the reflexive behavior of all players, the dependence of the maximum of the player's utility function on his/her equilibrium action has the form*

$$\pi_i^* = b(1 + S_i^r)(Q_i^*)^2 - B_{0i}. \tag{3}$$

Proof. From (1) it follows that

$$\pi_i = \left(a - b\left(Q_i + Q_{j\sum}\right)\right) - B_iQ_i - B_{0i} = b\left(\alpha_i - \left(Q_i + Q_{j\sum}\right)\right)Q_i - B_{0i};$$

on the other hand, by formula (2a),

$$\alpha_i - Q_i^* - Q_{j\sum}^* = (1 + S_i^r)Q_i^*.$$

Substituting this result into the former expression finally gives (3).

According to equations (2), the solution of this system (the vector of equilibrium actions $\mathbf{Q}^* = \{Q_i^*, i \in N\}$) depends on the SCV vector $\mathbf{S} = \{S_i, i \in N\}$. Therefore, if the SCV value of some player i is such that the utility functions of the other players (the environment) are maximized in equilibrium, then we can say that the environment purposefully controls this player (i.e., a hierarchical game arises).

3. THE HIERARCHICAL GAME IN THE LINEAR TRIOPOLY MODEL

Consider an oligopoly game with three players (i.e., the triopoly case). In this model, one player i opposes all other players (the environment), which have a common goal of reaching a beneficial action of the former player from their standpoint. Therefore, player i will be called the controlled player. Let the number j be associated with all environment players taken together.

We describe the awareness system in the hierarchical game with the reflexive behavior of players by the following assumptions.

1) All players are completely aware of each other's utility functions and actions and choose their actions simultaneously; the controlled player and the environment are independent of each other.

2) Environment players j are informed about each other's conjectural variations and are not informed about the SCV values of the controlled player; they choose the same strategies (actions denoted by q) and have the same SCV value (denoted by s). In addition, suppose that the environment players are identical in the type parameter α_j, $j \in N\backslash i$, with α being the average value of this parameter. Accordingly, the values of the environment's objective functions will be the same, denoted by π. The environment players can change the SCV value in a coordinated manner depending on the SCV value of the controlled player. Moreover, this dependence is inverse: increasing the SCV value of the controlled player (lowering his/her leadership level) reduces the SCV value of the environment (i.e., raises its leadership level). We formalize this assumption as follows:

$$\begin{gathered} q = Q_j, \quad \alpha = \frac{1}{n-1} \sum_{j \in N\backslash i} \alpha_j, \quad \pi = \pi_j, \\ s = S_j, \quad s' = \frac{\partial S_j}{\partial S_i} < 0, \quad s' = \text{const } \forall j \in N\backslash i. \end{gathered} \tag{4}$$

3) At each time instant (step) of the game, the controlled player i chooses an action independently of the environment's actions by considering the latter's previous actions. The controlled player is unaware of the SCV values of the environment but can indirectly estimate them by restoring the optimal reactions (best responses) from the observed actions of the environment and his/her actions. The controlled player determines the SCV value as a Stackelberg leader (i.e., by differentiating the optimal reaction function of the environment).

4) Both counterparties (the controlled player and the environment) are unaware of each other's true leadership levels: they do not know the opponent's SCV value in the game.

The last assumption is crucial for organizing an informational control process by manipulating the controlled player's actions. With a goal-oriented action, the environment can create a controlled player's belief that the former's SCV value has changed appreciably, although the true action may be unit, leaving the environment's reaction curve and SCV value actually unchanged.

Based on these assumptions, we formulate a hierarchical model of the game: if the controlled player chooses an optimal action considering the available data on the environment's SCV values, then the environment maximizes its utility functions, inducing the controlled player to choose the best SCV value from the former's standpoint. In other words, the environment solves the following problem:

$$\max_{S_i} \pi^*(q(s, S_i), \quad Q_i(s, S_i)), \quad s = s^0 = \text{const}, \tag{5}$$

provided that the controlled player chooses the strategy from the condition

$$\max_{Q_i} \pi_i(Q(s, S_i), q(s, S_i)), \tag{6}$$

where s^0 is the environment's SCV value at the initial step of the game.

Consider the environment's control process for the behavior of player i. For this purpose, we determine the optimal SCV value $\overline{S}_i$ of this player in terms of the environment's utility function (5): $\overline{S}_i = \arg\max_{S_i} \pi(q(s, S_i), Q_i(s, S_i))$.

Proposition 2. *The controlled player's SCV value maximizing the utility function* (5) *of the environment is calculated by solving the equation*

$$2(1+s)q^{*'}_{S_i} + q^* s' = 0 \tag{7}$$

under the condition

$$s' < -\frac{(1+s)q^*}{2q^{*'}_{S_i}}. \tag{7a}$$

Proof. By the first-order optimality condition applied to the environment's objective function π^* (3), we have

$$\pi^{*'}_{S_i} = 2b(1+s)q^{*'}_{s_i}q^* + b(q^*)^2 s' = 0,$$

which immediately gives (7). Let a change in the SCV value weakly impact the shift of the equilibrium $q^{*''}_{S_iS_i}$. In view of $s'' = 0$ (see (4)), we transform the second-order optimality condition

$$\pi^{*''}_{S_iS_i} = b\{2s'q^{*'}_{s_i}q^* + 2(1+s)(q^{*''}_{S_iS_i}q^* + q^{*2}) + 2q^{*'}_{S_i}q^*s' + q^{*2}s''\} < 0 \tag{7b}$$

to

$$2s'q^{*'}_{s_i} + (1+s)q^* < 0. \tag{7c}$$

The analysis of the Stackelberg leadership levels dynamics shows that $q^{*'}_{s_i} > 0$; in addition, $1+s > 0$ in the linear oligopoly model. Therefore, according to (7b), the maximum is achieved at $s' < 0$ (see Assumption (4)), particularly under condition (7a).

Note that equation (7) corresponds to the previously known formula [20] $2(1+S^r_i)Q^*_jQ^{*'}_{jS_i}P'_Q + \Big((1+S^r_i)P''_{QS_i} + P'_Q\frac{\partial S^r_j}{\partial S_i}\Big)Q^{*2}_j = 0$ for the general oligopoly problem, since $P'_Q = -b$ and $P''_{QS_i} = 0$ in the model with linear demand and cost functions under consideration.

4. OPTIMAL CONTROL METHODS

Equation (7) has been written for each environment player. The desired value $\overline{S}_i$ in (7) figures in the expressions for $q^{*'}_{s_i}, q^*$, which are necessary to calculate this unknown. Therefore, we consider a special case of a triopoly, where, without loss of generality, the second player will be the controlled one ($i = 2$) whereas the first and third players the environment ($j = 1, 3$).

Let us find the equilibrium in this game and the solution of the control problem in the hierarchical system (5).

Proposition 3. *In the triopoly game with players $j = 1, 3$ as the environment and player $i = 2$ as the controlled one, we have:*

i) The optimal reaction functions of the controlled player and the environment are given by

$$q = \frac{\alpha - Q_2}{3+s}, \quad Q_2 = \frac{\alpha_2 - 2q}{2+S_2}. \tag{8a}$$

ii) The equilibrium actions have the form

$$q^* = \frac{\alpha_2 - \alpha(2+S_2)}{2-(2+S_2)(3+s)}, \quad Q^*_2 = \frac{2\alpha - \alpha_2(3+s)}{2-(2+S_2)(3+s)}. \tag{8b}$$

iii) The controlled player's SCV value maximizing the utility function of the environment is calculated by solving the equation

$$2(1+s)\frac{zq^*-\alpha}{y} + q^*s' = 0 \tag{8c}$$

under the condition

$$\frac{\pi^{*''}_{S_iS_i}}{b} = \frac{4q^*}{y}s'\left(\frac{zq^*}{2} + (1+s)q^* - \alpha\right) + 2(1+s)q^{*2} < 0, \tag{8d}$$

where $y = 2 - (2+S_2)(3+s)$ and $z = s'(2+S_2) + 3 + s$.

Proof. In the case under consideration, system (2a) becomes

$$\begin{cases} \alpha_1 - (2+S_1)Q_1 - Q_3 - Q_2 = 0 \\ \alpha_2 - (2+S_2)Q_2 - Q_1 - Q_3 = 0 \\ \alpha_3 - (2+S_3)Q_3 - Q_1 - Q_2 = 0, \end{cases} \Rightarrow \begin{cases} 2\alpha - 2(2+s)q - 2Q_2 - 2q = 0 \\ \alpha_2 - (2+S_2)Q_2 - 2q = 0. \end{cases}$$

Consequently, the optimal reaction functions of the second player and the environment take the form (8a). Resolving them yields the equilibrium actions (8b). For the sake of simplification, let us introduce $y = 2 - (2+S_2)(3+s)$ and $z = s'(2+S_2) + 3 + s$. In this case, the derivative $q^{*'}_{S_2}$ is calculated as

$$q^{*'}_{S_2} = \frac{\alpha_2 z - \alpha[y + z(2+S_2)]}{y^2} = \frac{z[\alpha_2 - \alpha(2+S_2)]}{y^2} - \frac{\alpha y}{y^2} = \frac{zq^* - \alpha}{y}.$$

Therefore, equation (7) turns into (8c).

Now we derive the second-order optimality condition (7a) for the environment's objective function. In view of $s'' = 0$, the value $q^{*''}_{S_2S_2}$ can be expressed in the form

$$q^{*''}_{S_2S_2} = \frac{(z'q^* + zq^{*'}_{S_2})y - y'(zq^* - \alpha)}{y^2} = \frac{z'q^* + 2zq^{*'}_{S_2}}{y}$$

(more precisely than (7a)).

Substituting this expression into (7b), after appropriate transformations with $z' = 2s'$, we obtain

$$\begin{aligned} \frac{\pi^{*''}_{S_iS_i}}{b} &= q^* \left(4s' \frac{zq^* - \alpha}{y} + 2(1+s) \frac{2s'q^* + 2zq^{*'}_{S_2}}{y} \right) + 2(1+s)q^{*2} \\ &= \frac{4q^*}{y} \left(z[s'q^* + (1+s)q^{*'}_{S_2}] + s'[(1+s)q^* - \alpha] \right) + 2(1+s)q^{*2}. \end{aligned}$$

According to (7), the change $(1+s)q^{*'}_{S_2} = -\frac{s'q^*}{2}$ in the last equality finally leads to the sufficient maximum condition (8d).

Proposition 3 can be used to find $\overline{S}_2$.

Now we develop an algorithm for reaching the target value $\overline{S}_2$ from the environment's standpoint. Let an equilibrium (q^0, Q_2^0) be established at some initial step 0 of the game under some SCV values S_2^0 (the controlled player) and s^0 (the environment), which are unknown to the counterparties. Let us determine the environment's phantom reaction under which the controlled player will choose the SCV value $\overline{S}_2$. We denote by s_f the SCV value that the environment must have for the controlled player to determine his/her SCV value as $\overline{S}_2$. It will be called *the phantom SCV value of the environment.*

Proposition 4. *In the triopoly game, the controlled player $i = 2$ establishes the response according to $\overline{S}_2$ under the phantom SCV of the environment $(j = 1, 3)$ if and only if*

$$s_f = -\frac{1}{\overline{S}_2} - 3. \tag{9}$$

Proof. If the controlled player performs the action Q_2 and the environment responds with the action q, the former will calculate the SCV value by differentiating the environment's reaction function. The environment's phantom reaction $q_f(Q_2)$ is calculated by substituting the unknown value s_f into (8a): $q_f = \frac{\alpha - Q_2}{3+s_f}$. Then $S_2 = q'_{Q_2} = \left(\frac{\alpha - Q_2}{3+s_f}\right)'_{Q_2} = -\frac{1}{3+s_f}$. Therefore, if the environment's goal is to induce the controlled player's choice $\overline{S}_2$, then from $\overline{S}_2 = -\frac{1}{3+s_f}$ we obtain the environment's SCV value (9) required.

Based on Proposition 4, we organize the following iterative control process with time instants (steps) $0, t, t+1, t+2, \ldots$, indicated by the superscripts of the players' actions ($t > 0$).

1. At step t, the environment performs an action calculated according to its phantom reaction function:

$$q^t = \frac{\alpha - Q_2^0}{3 + s_f} = -\overline{S}_2(\alpha - Q_2^0).$$

2. At step $t+1$, the controlled player calculates the SCV value (9) given this action, which is equal to $\overline{S}_2$, and then responds according to the new reaction function $Q_2^{t+1} = \frac{\alpha_2 - 2q}{2+\overline{S}_2}$:

$$Q_2^{t+1} = \frac{\alpha_2 - 2q^t}{2 + \overline{S}_2}.$$

3. At step $t+2$, the environment performs an action according to its true reaction function:

$$q^{t+2} = \frac{\alpha_2 - Q_2^{t+1}}{3 + s^0}.$$

4. At step $t+3$, the controlled player calculates the SCV value (9) given this action, which is equal to S_2^0, and then returns to the original reaction function $Q_2^{t+3} = \frac{\alpha_2 - 2q}{2+S_2^0}$.

5. At steps $t+4, t+5, \ldots$, the environment and the controlled player repeat Stages 3 and 4, and the game equilibrium returns to the initial state (q^0, Q_2^0).

At steps $t+1$ and $t+2$ of this process, the environment achieves the desired result and receives the maximum profit. As is known [3–6], the above iterative process converges to the equilibrium (8b) fast enough; therefore, suppose that the game state has already stabilized starting from step $t+3$. Consequently, the environment receives an additional payoff at steps $t+1$ and $t+2$ but may incur losses at step t. The payoffs at different steps must be brought to a comparable level based on discounting $\pi_d(t) = \pi^t e^{-\rho t}$, commonly used in dynamic oligopoly models [9]. To assess the effectiveness of the iterative control process, we compare the maximum values of the environment's utility function at steps 0, t, $t+1$, and $t+2$: control is effective if

$$\pi^{*t} e^{-\rho t} + \pi^{*t+1} e^{-\rho(t+1)} + \pi^{*t+2} e^{-\rho(t+2)} \geqslant \pi^{*0}(e^{-\rho t} + e^{-\rho(t+1)} + e^{-\rho(t+2)}),$$

where ρ denotes a discounting factor.

5. EXPERIMENTAL ANALYSIS OF OPTIMAL CONTROL

Consider a numerical experiment for the optimal control process. Computer simulations were carried out on the 2016–2021 demand and operator cost data for the Russian telecommunications market [22]. The linear demand function for the voice traffic of mobile operators was constructed in the form

$$P(Q) = a - bQ, \quad a = 1.6, \quad b = 0.000001;$$

the cost functions of mobile operators were taken without fixed costs due to the latter's negligible impact on the equilibrium:

$$C_i(Q_i) = B_i Q_i, \quad B_1 = 0.0005, \quad B_2 = 0.0018, \quad B_3 = 0.0004,$$

where $i = 1$ corresponds to MTS, $i = 2$ to MegaFon, and $i = 3$ to VimpelCom.

As in Section 4, the second player is the controlled one, with the action denoted by Q_2, and the first and third players represent the environment ($j = 1, 3$), with the same actions denoted by q. Figure 1 shows the optimal reaction lines of the players (indicated by R) in the cases of Cournot equilibrium (the SCV values are $s = S_2 = 0$) and three-party Stackelberg leadership (the SCV values are $s = S_2 = -0.5$); the latter case is treated as the initial state of the game (point 0).

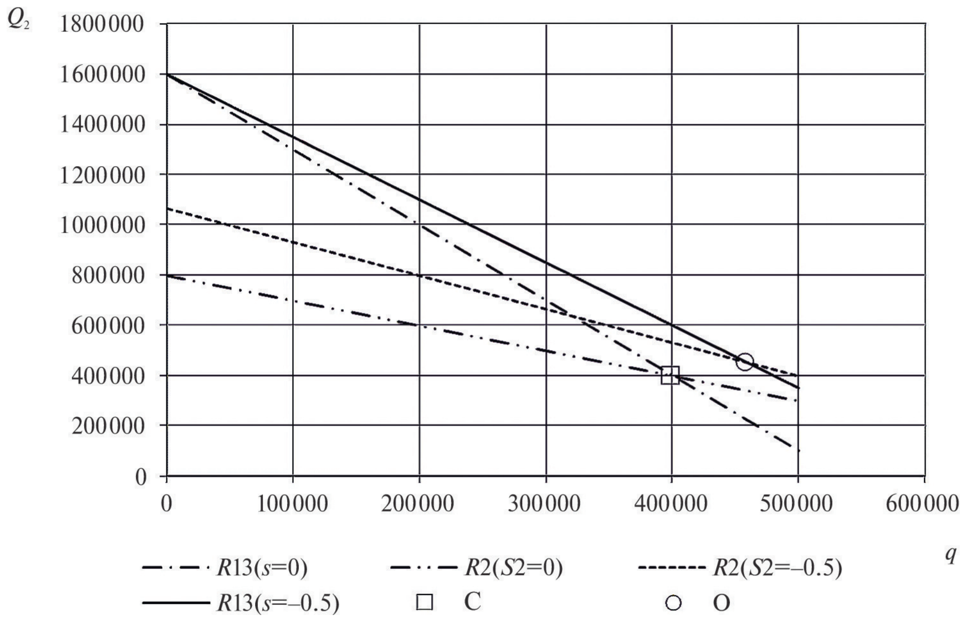

Fig. 1. Cournot equilibrium (C) and three-party Stackelberg leadership (0).

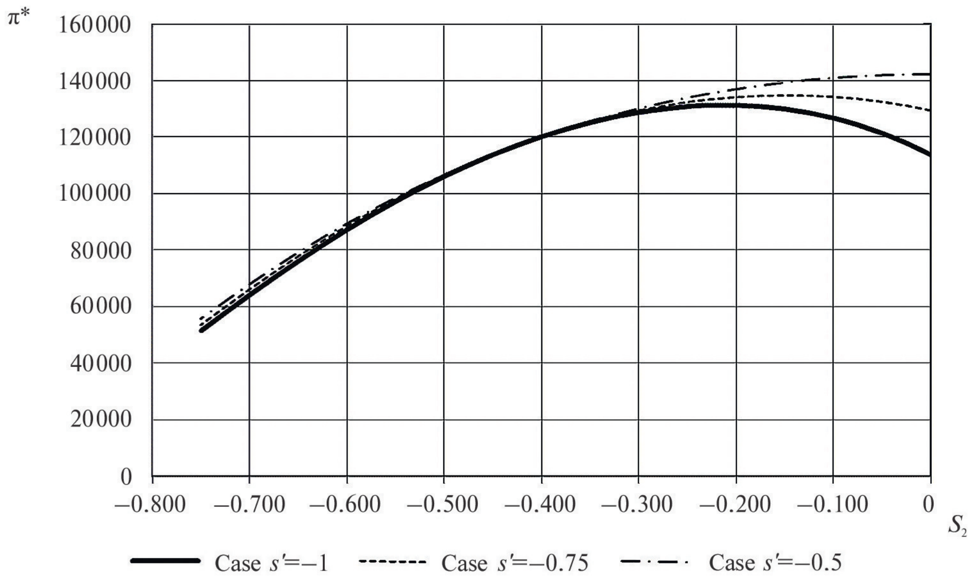

Fig. 2. The environment's utility functions depending on the controlled player's SCV value for different s'.

Table 1 and Fig. 2 show the impact of changes in the environment's response on the corresponding changes in the controlled player's response, expressed by s'. The following model dependence was used:

$$s = -1 + s' S_2,$$

in which $s \in (-1, 0]$ for $S_2 \in (-1, 0]$ and $s' \in (-1, 0]$. According to the data in Table 1, reducing the rate of change of the environment's SCV value (the absolute value of s') shifts the target SCV value $\overline{S}_2$ of the controlled player closer to zero and increases the maximum utility π^* of the environment. The graphs in Fig. 2 demonstrate the existence of a maximum of the environment's utility function for certain SCV values of the controlled player. Note that the second-order optimality condition (8d) was also verified.

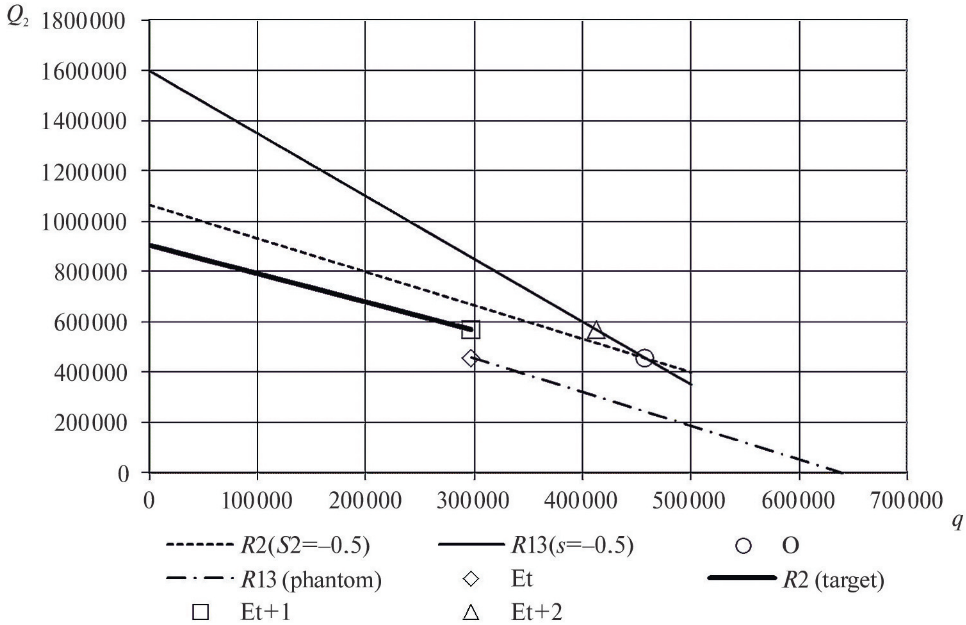

Fig. 3. The informational control process from Stackelberg leadership (0).

Figure 3 demonstrates the informational control process from the state of three-party Stackelberg leadership (point 0) for the case $s' = -1$. At step 1, the environment's action according to its phantom reaction function R13(phantom) transfers the game to the state E^t; and the controlled player performs an action corresponding to the initial position 0 in the game. This change in the environment's action allows the controlled player to establish that the former's behavior corresponds to the SCV value $\overline{S}_2$; therefore, at step 2, this player changes his/her reaction function to the one desired for the environment, $R2$(target). Then the game evolves to the state $E^{(t+1)}$, and the environment's action remains at the level E^t. Responding at step 3, the environment returns to its true reaction function $R13(s = -0.5)$, since the environment calculated its optimal target SCV value of the controlled player precisely for its initial SCV value. As a result, the game passes into the state $E^{(t+2)}$. The further process of restoring the initial equilibrium at point 0 is not shown here: at step $t + 2$, the players' actions are already close to this point.

The economic effect of informational control is presented in Table 2 at a discount rate of 10%. The total discounted utility of the environment for three steps of the control process exceeds the total discounted value of its utility in the initial state of the game, which confirms the validity of

Table 1. The impact of changes in the environment's SCV value on the environment's optimum when changing the controlled player's SCV value

Case	s'	$\overline{S}_2$	q^*	π^*
1	-1	-0.2338	5 079 027	128 969
2	-0.75	-0.1478	518 714	134 506
3	-0.5	-0.002	533 454	142 260

Table 2. The economic effect of informational control, $\rho = 0.1$

Step	$e^{-\rho t}$	Nominal values		Discounted values	
		π^{*0}	π^{*t}	π^{*0}	π^{*t}
t	0.905	104 761	163 474	94 821	147 963
$t+1$	0.819	104 761	129 758	85 823	106 301
$t+2$	0.741	104 761	85 005	77 680	63 031
Sum				258 324	317 295

such control from the environment's standpoint. Note an increased utility of the environment at step t, which is not always the case; in the situation under consideration, the first action of the environment was less than in the initial equilibrium, resulting in an equilibrium price increase; if the first action of the environment were the opposite, its utility would decrease.

6. CONCLUSIONS

This paper has demonstrated the fundamental possibility of implementing informational control of player's actions in a triopoly game by other players (the environment). The linear models of demand and cost functions have been considered; in this case, explicit expressions for the optimal reactions and equilibrium actions of players have been derived, an informational control algorithm has been developed, and the effect of informational control has been illustrated. The following key results have been obtained.

The dependence of the maximum of the player's utility function on his/her equilibrium action under the reflexive behavior of all players has been established, showing that the player's payoff depends both on his/her SCV value and the SCV value of the environment through the former's equilibrium action. Hence, by varying the SCV value of the environment, this player can increase his/her maximum payoff, which is the basis to control the environment through its SCV value as the control parameter. However, one player cannot manipulate the other players, while a group of players, acting in concert, can change the beliefs of one player.

Therefore, a hierarchical game has been investigated in which the environment controls a certain player by finding the optimal SCV value for him/her. An equation has been obtained for calculating the optimal SCV value of this player in terms of the environment's utility function. To motivate the controlled player to choose the required SCV value, the method of optimal reactions has been used, with explicit formulas derived in the triopoly case. In the three-player game (one controlled player and two opponents with a common goal), equilibrium actions have been determined, and an equation has been compiled for calculating the optimal SCV value of the controlled player.

In the triopoly model under consideration, a phantom reaction-based method has been developed to determine the environment's action inducing the controlled player to use the optimal SCV value from the environment's standpoint. This method has been developed into an algorithm to model the action sequence of the environment and the controlled player in the informational control process; the effectiveness of this process has been assessed.

Computer simulations have been carried out to demonstrate the effectiveness and economic impact of the control algorithm. Although the controlled player receives higher profit only at two steps of the game, this outcome is quite significant from a practical point of view. In real business, the game moment often corresponds to the publication period of the company's financial statements (annually). Therefore, one or two years in which environment players receive predominant payoffs can lead to large financial losses for the controlled player and even force the latter out from the market. Consequently, the results of this study have practical value.

REFERENCES

1. Cournot, A.A., *Researches into the Mathematical Principles of the Theory of Wealth*, London: Hafner, 1960. (Original 1838.)
2. Stackelberg, H., *Market Structure and Equilibrium*, 1st ed. Translation into English, Bazin, Urch & Hill, Springer, 2011. (Original 1934.)
3. Algazin, G.I. and Algazina, D.G., Reflexion Processes and Equilibrium in an Oligopoly Model with a Leader, *Autom. Remote Control*, 2020, vol. 81(7), pp. 1258–1270.

4. Algazin, G.I. and Algazina, Y.G., To the Analytical Investigation of the Convergence Conditions of the Processes of Reflexive Collective Behavior in Oligopoly Models, *Autom. Remote Control*, 2022, vol. 83, no. 3, pp. 367–388.

5. 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.

6. Algazin, G.I. and Algazina, D.G., Aggregative Oligopoly Games with Linear Demand and Cost Functions of Agents, in *MAK: Matematiki – Altaiskomu krayu* (MAK: Mathematicians to Altaiskii Krai), Barnaul, 2024, no. 6, pp. 99–101.

7. Xiao, Y., Zhang, S., and Peng, Y., Dynamic Investigations in a Stackelberg Model with Differentiated Products and Bounded Rationality, *J. Comput. Appl. Math.*, 2022, vol. 414, art. no. 114409.

8. Ougolnitsky, G. and Gorbaneva, O., Sustainability of Intertwined Supply Networks: A Game-Theoretic Approach, *Games*, 2022, vol. 13, no. 3, art. no. 35.

9. Ougolnitsky, G.A. and Usov, A.B., The Interaction of Economic Agents in Cournot Duopoly Models under Ecological Conditions: A Comparison of Organizational Modes, *Autom. Remote Control*, 2023, vol. 84, no. 2, pp. 175–189.

10. Al-Khedhairi, A., Dynamical Study of Competition Cournot-like Duopoly Games Incorporating Fractional Order Derivatives and Seasonal Influences, *Int. J. Nonlinear Sci. Numer. Simulation*, 2020, vol. 21, no. 3–4, pp. 339–359.

11. Zouhar, J. and Zouharova, M., Stackelberg versus Cournot Duopoly with Asymmetric Costs: Primary Markups, Entry Deterrence, and a Comparison of Social Welfare and Industry Profits, *Econom. Theory Bulletin.*, 2020, vol. 8, pp. 89–96.

12. Collie, D.R., Taxation under Oligopoly in a General Equilibrium Setting, *J. Public Economic Theory*, 2019, vol. 21, no. 4, pp. 738–753.

13. Ferrarese, W., When Multiple Merged Entities Lead in Stackelberg Oligopolies, *Rev. Industr. Organization*, 2020, vol. 56, no. 1, pp. 131–142.

14. Mukherjee, A. and Zeng, C., Social Desirability of Entry in a Bilateral Oligopoly – The Implications of (Non) Sunk Costs, *Mathem. Social Sci.*, 2022, vol. 118, pp. 12–19.

15. Zhou, J., Mixed Bundling in Oligopoly Markets, *J. Econom. Theory*, 2021, vol. 194, art. no. 105257.

16. Shuai, J., Yang, H., and Zhang, L., Dominant Firm and Competitive Bundling in Oligopoly Markets, *Games Econom. Behav.*, 2022, vol. 132, pp. 421–447.

17. Zhang, Y., When Should Firms Choose a Risky New Technology? An Oligopolistic Analysis, *Econom. Modelling*, 2020, vol. 91, pp. 687–693.

18. Novikov, D.A. and Chkhartishvili, A.G., *Reflexion and Control: Mathematical Models*, Leiden: CRC Press, 2014.

19. 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.

20. Geraskin, M.I., Informational Control of Strategies in an n-Player Oligopoly Game with Reflexive Behavior, *Autom. Remote Control*, 2025, vol. 86, no. 6, pp. 577–588.

21. Geraskin, M.I., Quantity Conjectural Variations in Oligopoly Games under Different Demand and Cost Functions and Multilevel Leadership, *Autom. Remote Control*, 2024, vol. 85, no. 7, pp. 711–724.

22. Geraskin, M.I. and Moiseeva, K.S., Analysis of Demand and Cost Trends of Telecommunications Market Operators in Russia, *Applied Mathematics and Control Sciences*, 2023, no. 3, pp. 157–174.

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