Adv Syst Sci Appl 2020; **03**; 166-177
Published online at https://ijassa.ipu.ru.

A Differential Stackelberg Game Theoretic Model of the Promotion of Innovations in Universities

Mukharbek Malsagov[1*], Gennady Ougolnitsky[2], Anatoly Usov[3]

[1)] *Ingush State University, Nazran, Russia*
E-mail: mmm1956@bk.ru
[2)] *Southern Federal University, Rostov-on-Don, Russia*
E-mail: gaugolnickiy@sfedu.ru
[3)] *Southern Federal University, Rostov-on-Don, Russia*
E-mail: abusov@sfedu.ru

Abstract: A two-level dynamical game theoretic model "federal state - universities" in open-loop strategies is built and investigated. A refinement of the electronic learning courses, and their differentiation using modern methods of information technologies made by competing a la Cournot universities (agents) are treated as innovative investments. The algorithms of building Nash (in the non-coalitional game of agents) and Stackelberg (in the hierarchical game of the Principal with the grand coalition of agents) equilibria are proposed and implemented. For the solution of the respective dynamic control problems the Pontryagin maximum principle and simulation modeling are used. A comprehensive analysis of the received results is given.

Keywords: Nash equilibrium, Stackelberg equilibrium, Stackelberg differential games, innovations, method of qualitatively representative scenarios in simulation modeling

1. INTRODUCTION

A review of applications of the game theoretic models to the analysis of innovations is presented in [5]. The authors divide three levels of these applications: (1) intra-organizational level within a firm, where players are innovators, project managers, and resource administrators; (2) inter-organizational level where the players are competitive firms; (3) meta-organizational level, where players are a social planner (innovation policy maker, government, a social or government institution, e.g., a research foundation) and an aggregate innovative entrepreneur.

A traditional approach to the building and investigation of the differential game models is exposed in [1,4]. An original method of solution of the Stackelberg differential games based on building of a mutually benefit program of actions and punishment in the case of deviations, is described in [9].

In our papers [13,16] the approach of [9] is extended for the case of several followers with consideration of the requirements of sustainable development of the controlled dynamic system [12]. Ougolnitsky and Usov [14] have proposed a method of qualitatively representative scenarios in simulation modeling. This method allows for a good enough qualitative forecast of the dynamics of a controlled system by means of few scenarios of the computer simulation.

An interesting dynamic game theoretic model of the oligopolistic investment to the product differentiation is studied in [2, 3]. The paper [2] considers open-loop strategies, and the paper [3] analyzes closed-loop ones. Cellini and Lambertini [2, 3] have found that in the case of open-loop strategies the Nash stable volume of investments to the product differentiation (innovations) increases with a number of firms, i.e. the level of market

[*] Corresponding author: mmm1956@bk.ru

competition (Arrow hypothesis). In contrary, in the case of closed-loop strategies the dependence is inverse (Schumpeter hypothesis). The authors notice that innovation activity in the product differentiation can be considered as a contribution of private investors to the production of a public good. This fact embeds their paper to the context of public goods economics. Interesting models of Stackelberg oligopoly are proposed and investigated in [6-8].

This paper develops the papers [2, 3] with consideration of other mentioned sources in the following directions. First, its problem domain is a development of the electronic learning courses in universities, where a modification of the courses is treated as an innovative activity. There is no doubt in the actuality of this problem. On March 25, 2020 the IT holding TalentTech and on-line universities "Netology" and EdMarket have presented the research results about the on-line education market[†]. As the authors notice, a volume of the Russian market of on-line education in the b2c segment was 38,5 billion rubles in the end of 2019. In the end of the year 2023, according to that forecast, its value will be equal to 60 billion rubles a year. The global market of on-line education up to the year 2023 tends to the amount $282,62 billion.

An even higher estimate is given by the Interfax Academy. According to its report, the Russian market of the on-line education after 2019 equals to 45–50 billion rubles, and in 2020 will be equal to the amount 55–60 billion rubles, the annual growth is 20–25%. The global on-line education market is estimated by the value $74 billion (about 4,8 trillion rubles) after the year 2019, so the potential for Russia is great[‡]. It is evident that the growth of the distant forms of education due to COVID-19 pandemics will essentially enforce the noticed trends.

Second, we propose a hierarchical setup of the problem with the federal state as a leader (Principal), and competitive universities as followers (agents). A description of a university as active system [11] is given in [10]. The hierarchical impact of the Principal to the agents may be administrative (compulsion) or economic one (impulsion). In the former case the Principal impacts to the sets of feasible strategies of the agents, and in the latter case - to their payoff functionals [12].

Thus, the contribution of the paper has four aspects. First, we consider a combination of the aggregative non-cooperative oligopolistic game of the agents with the Stackelberg game of the type "Principal-agents". Second, the parameter of the demand function varies in time, and the character of this variation depends on the agents' actions (strategies) in the form of a differential equation. Third, the agents choose both outputs and investments, i.e. an agent's strategy includes a parameter of its cost function called the constant cost. Fourth, the following approach is used. From the point of view of the agents, their interaction is modeled as a game in normal form where the Nash equilibrium is built. From the point of view of the Principal it is supposed that the agents form the grand coalition, and respectively a Stackelberg two-person game "Principal - coalition of agents" arises. This permits to avoid a challenging question about what should be considered as a best response of several agents to the Principal's strategy.

The rest of the paper is organized as follows. In the Section 1 the setup of a dynamic problem of hierarchical control is given. In the Section 2 the Nash equilibrium for two symmetrical agents with administrative and economic impact of the Principal in open-loop strategies is built. Algorithms of building the Stackelberg equilibrium by means of the method of qualitatively representative scenarios in simulation modeling are presented. The Section 4 describes a numerical simulation in the problem of building the Stackelberg equilibrium. The Section 5 contains a comparative analysis of the received results, and the conclusions are formulated.

[†] http://neorusedu.ru/news/rossijskij-rynok-onlajn-obrazovaniya-ozhidaet-burnyj-rost

[‡] https://www.kommersant.ru/doc/4275439

2. THE PROBLEM SETUP

Let us consider the following hierarchical modification of the model proposed by Cellini and Lambertini [2, 3]. Several universities (agents) competing a la Cournot develop electronic learning courses for sale. A refinement of the courses, and its differentiation using modern methods of learning and information technologies are treated as innovative investments. The differentiation of the courses may be considered as a public good, and the respective investments - as a private production of the public good [3]. On the higher (relative to the agents) control level a Principal (federal state or its authorized bodies) is situated. The Principal tends to increase a public good (with possible additional consideration of its own interests) by means of administrative or economic control methods. In the case of administrative impact (compulsion) the Principal bounds from below the contributions of the agents to the innovative development (production of the public good) that incurs control cost. In the case of economic impact (impulsion) the Principal grants the agents on the base of an available budget. In both cases the Principal and the agents use open-loop strategies. The game is played on a finite interval of time $[0,T]$.

In the case of compulsion the model for *n* agents has the form:

- the Principal's payoff functional

$$J_0 = \int_0^T e^{-\rho t}\left[\sum_{i=1}^n(\pi_i(t)-s_i(t)) - Z\left(\sum_{i=1}^n y_i(t)\right)\right]dt + G_0(T) \to \max \tag{2.1}$$

- the Principal's control constraints

$$0 \le y_i(t) \le K_{\max}; \tag{2.2}$$

- the agents' payoff functionals

$$J_i = \int_0^T e^{-\rho t}(\pi_i(t)+s_i(t))dt + G_i(T) \to \max; \tag{2.3}$$

- the agents' control constraints

$$y_i(t) \le k_i(t) \le K_{\max};\ 0 \le q_i(t) \le Q_{\max}; \tag{2.4}$$

- the equation of system dynamics [9, 10]

$$\frac{dD}{dt} = -\frac{K(t)}{1+K(t)}D(t); \qquad K(t) = \sum_{i=1}^n k_i(t);\ D(0) = B; \tag{2.5}$$

- the current payoff function of the i-th agent

$$\pi_i(t) = [p_i(t)-c_i]q_i(t) - k_i(t) \tag{2.6}$$

In the Principal's payoff function the profit π_i reflects an approximate estimate of the positive externality from the universities' activity, namely, the GNP growth due to increasing of the educational level of the society;

- the inverse demand function [2,3];

$$p_i(t) = A - Bq_i(t) - D(t)\sum_{j\neq i} q_j(t) \tag{2.7}$$

- the payoffs of the Principal and the agents in the moment of time *T* respectively

$$G_i(T) = [A - Bq_i(T) - D(T)\sum_{j\neq i} q_j(T) - c_i]q_i(T), i = 1,2,...,n;$$

$$G_0(T) = \sum_{i=1}^n G_i(T) = \sum_{i=1}^n [A - Bq_i(T) - D(T)\sum_{j\neq i} q_j(T) - c_i]q_i(T).$$

J_0, J_i - the total payoff functionals of the Principal and the agents; $D(t) \in [0,B]$ - a symmetrical degree of substitutability between any pair of courses. If $D(t) = B$ then the courses are completely homogeneous. If $D(t) = 0$ then the courses are unique and each agent is a monopolist [2,3]; $q_i(t)$ - an output level of the *i*-th agent at constant returns to scale, then total operative cost per period are $C_i(t) = c_i q_i(t), c_i \in (0, A_i)$; $k_i(t)$ - individual investment of the i-th agent to the innovative development, $K(t)$ - the overall industry expenditure, $K_{\max} = \text{const}$ - an upper bound for any $k_i(t)$; $Q_{\max} = \text{const}$ - an upper bound for any $q_i(t)$; functions $s_i(t)$ are the Principal's grants to the i-th agent (in the case of compulsion they are given); T – the length of the game. Then, $y_i(t)$ are the Principal's lower bounds established for $k_i(t)$; $\rho \in (0,1)$ - a discount factor; $A > 0,\ B > 0$ - demand parameters; $Z(\cdot)$ - a convex increasing administrative control cost function, $Z(0) = 0$; the function $Z(\cdot)$ is supposed to be linear $Z(x) = Ex; E = \text{const}$. The equation of dynamics (5) is treated as a production function with the output $-\dot{D}/D(t)$ created by the input $K(t)$. This technology can be shown to exhibit decreasing returns to scale w.r.t. $K(t)$. Thus $D(t)$ is non-increasing function of time which tends to zero when $K(t) \to \infty$ [3].

In the case of impulsion the model has the form:
- the Principal's payoff functional

$$J_0 = \int_0^T e^{-\rho t} \sum_{i=1}^n [\pi_i(t) - s_i(t)] dt + G_0(T) \to \max \tag{2.8}$$

- the Principal's control constraints

$$s_i(t) \ge 0; \sum_{j=1}^n s_j(t) \le S; \tag{2.9}$$

- the agents' payoff functionals

$$J_i = \int_0^T e^{-\rho t} (\pi_i(t) + s_i(t)) dt + G_i(T) \to \max \tag{2.10}$$

- the agents' control constraints

$$0 \le k_i(t) \le K_{\max}; \quad 0 \le q_i(t) \le Q_{\max}. \tag{2.11}$$

Here $s_i(t)$ are the Principal's control to be determined; S is a total volume of the Principal's grants (with consideration of possible savings).

All input functions of the model are supposed to be continuous, and the controls of the agents and the Principal belong to the class of piecewise continuous functions. We investigate the model (2.1)–(2.7) for compulsion and (2.5)-(2.11) for impulsion from the point of view of different control agents. From the point of view of the agents there is a non-cooperative n person game which solution is supposed to be a Nash equilibrium. From the point of view of the Principal there is a Stackelberg game. Let us assume in this case that the agents cooperate (create the grand coalition) and have the summary payoff functional in the form

$$J_A = \sum_{i=1}^n J_i = \sum_{i=1}^n \left(\int_0^T e^{-\rho t} (\pi_i(t) + s_i(t)) dt + G_i(T) \right) \to \max. \tag{2.12}$$

Other model relations do not change. Thus, we have a Stackelberg game between the Principal and the grand coalition of agents. This game has the following information structure:

1. The Principal chooses its open-loop strategies $s_i = s_i(t)$ for impulsion or $y_i = y_i(t)$ for compulsion, $t \in [0,T], i = 1,2,...,n$.

2. Given these strategies $s_i = s_i(t)$ for impulsion or $y_i = y_i(t)$ for compulsion, the agents within the grand coalition choose their open-loop strategies $k_i = k_i(t), q_i = q_i(t),\ t \in [0,T], i = 1,2,...,n$, using (2.12). As an integrand function in (2.12) depends continuously on its arguments, and the domains of feasible controls of the agents (2.4) or (2.11) are non-empty closed sets, the problem of determination of the agents' best response to any Principal's strategy is resolvable.

3. The Principal maximizes its payoff functional J_0 (2.1) or (2.8) for the worst best response of the coalition of agents to its strategy $R(s_1,...,s_n)$ or $R(y_1,...,y_n)$.

4. The received set of strategies $(s_1^*,...,s_n^*,k_1^*,...,k_n^*,q_1^*,...,q_n^*)$ for impulsion or $(y_1^*,...,y_n^*,k_1^*,...,k_n^*,q_1^*,...,q_n^*)$ for compulsion is the Stackelberg equilibrium.

3. BUILDING THE NASH EQUILIBRIUM

Let us first investigate the model from the point of view of the agents when they use open-loop strategies. The Principal's strategies are supposed to be given. This interpretation corresponds to the case of an indifferent Principal without its own objectives. Then we receive a differential n-person game (2.3)-(2.7) in the case of compulsion and (2.5) – (2.7), (2.10), (2.11) in the case of impulsion. Its solution is assumed to be a Nash equilibrium. To build it we use the Pontryagin maximum principle [15]. The Hamilton function of the i-th agent both for compulsion and impulsion has the form:

$$H_i(k_i(t),q_i(t),\lambda_i(t),D(t),t)=\left(A-Bq_i(t)-D\sum_{\substack{j\neq i\\ j=1}}^{n}q_j(t)-c_i\right)q_i(t)-k_i+s_i(t)+\lambda_i(t)\frac{D\sum_{i-1}^{n}k_i(t),}{1+\sum_{j=1}^{n}k_j(t)),}$$

where $\lambda_i(t)$ is a conjugate variable (as function of time). From the necessary condition of extremum $\begin{cases}\dfrac{\partial H_i}{\partial k_i}=0\\ \dfrac{\partial H_i}{\partial q_i}=0\end{cases}$, $i=1,2,...,n$ in the case of symmetrical agents

($c_i\equiv c;\ H_i\equiv H;\ s_i\equiv s; k_i\equiv k;\ ,q_i\equiv q; \lambda_i\equiv\lambda; G_i=G;\ i=i=1,2,...,n$)

we receive the system of equation for determination of their control variables

$$\frac{\partial H}{\partial k}=-1-\frac{\lambda D}{(1+nk)^2}=0;\ \frac{\partial H}{\partial q}=A-2Bq(t)-D(n-1)q(t)-c=0. \tag{3.1}$$

Thus

$$k(t)=\frac{-1+\sqrt{-\lambda D}}{n};\ q(t)=\frac{A-c}{2B+D(n-1))}. \tag{3.2}$$

Besides, we have the system of differential equations

$$-\frac{\partial H}{\partial D}=\frac{d\lambda}{dt}=-\left(\frac{A-C}{2B+D}\right)^2+\lambda\left(1-\frac{1}{\sqrt{-\lambda D}}\right);\quad \lambda(T)=\frac{\partial G}{\partial D(T)}=\frac{(n-1)(c-A)}{2B+D(T)(n-1)}; \tag{3.3}$$

$$\frac{dD}{dt}=-1+\frac{1}{\sqrt{-\lambda D}};\ D(0)=B.$$

From (3.1) we receive

$$\frac{\partial^2 H}{\partial k^2}=\frac{2\lambda Dn}{(1+nk)^3};\quad \frac{\partial^2 H}{\partial q^2}=-2B-D(n-1)<0;\ \frac{\partial^2 H}{\partial k\partial q}=0.$$

Therefore, the following proposition is proved.

Proposition (3.1):

The formulas (3.2), (3.3) determine the point of maximum of the Hamilton function for some value t if the system (3.2), (3.3) has a solution and the values (3.2) belong to the domains of feasible controls (2.4) for compulsion or (2.11) for impulsion.

If these conditions are not satisfied for some t then the Hamilton function attains its maximum for this t at one of the bounds of the segments (2.4) or (2.11).

An analytical investigation of the system of equations (3.3) is impossible due to the form of the equation of dynamics (2.5) and presence of the third summand in (2.7) that is principal for the considered problem setup. Thus the system (3.3) was analyzed numerically by the method of shooting. The system of equations (3.3) has a solution not for all values of the model input parameters. The Picard theorem of existence and uniqueness of solution of the system of differential equations is not satisfied, the right hand sides are defined only for the negative values of $\lambda(t)D(t)$. Besides, in the neighborhood of zero the Lipschitz conditions on λ and D are not satisfied. Therefore, the range of input model parameters for which the system (3.3) has the unique solution was determined numerically.

A numerical identification of the model has a testing character and provides a reasonable relation of the model parameters and variables that allows for acceptable qualitative conclusions about the comparative analysis of the results of numerical modeling. Thus, the professors of Southern Federal University develop annually about 100 new learning courses, therefore the maximal possible value is taken $Q_{\max} = 100$. An average cost of the development is equal to 50-80 thousand rubles, so $K_{\max} = 3000$, $A = 900$, $c = 50$ (thousand rubles per year). The value of parameter B varied in the range $B = 1-15$ (thousand rubles per year). We divided the cases of small investment of the Principal ($S = nK_{\max}/10$), middle investment ($S = nK_{\max}/5$), and considerable investment ($S = nK_{\max}$), as well as different cases of the agents' control by the Principal: namely, the value of Z varied from 2 thousand rubles per year (soft control) till 10 thousand rubles per year (hard control). The discount factor was estimated as $\rho = 0.04$ according to the annual inflation rate. The period of modeling was equal to 3 years, i.e. $T = 1095$ days for two universities ($n = 2$). Other parameters were evaluated by the experts (Table 3.1).

Table 3.1. Test values of the model parameters

Parameter	$Q_{\max}$	$c_i \equiv c$	A	B	$K_{\max}$	ρ	$s_i \equiv s$	Z
Value	100	50	900	5	3000	0.04	600	2
Dimension	-	thousand rubles per year	thousand rubles per year	thousand rubles per year	thousand rubles per year	-	thousand rubles per year	-

For determination of the range of model parameters in which the system (3.3) has the only solution we have implemented about 80 numerical simulations with variation of the values: $c \in [10, 200]$; $B \in [1,250]$; $A \in [100,1000]$ (thousand rubles per year); n from 2 till 100.

As a result it was established that for any input data there is a moment of time $t_1 > 0$: exists a solution of (3.3): $\lambda(t) < 0;\ D(t) \ge 0$ *for any* $t \in [t_1, T)$. The moment t_1 belongs to the range from 950 till 1000 days and depends not essentially on the parameters A, B, c, n. Its value slightly increases with n or B, decreases with A, and does not depend on c. Some of the results are presented in Table 3.2.

Table 3.2. Dependence of the moment of time t_1 on model parameters

n	A	c	B	t_1
2	110	100	10	997
2	110	100	1	999
2	110	100	25	989
2	110	100	250	980
2	200	100	10	999

2	1000	100	10	1000
2	300	10	10	1000
2	300	200	10	1000
2	210	200	10	995
10	200	100	10	999
10	110	100	25	1000
100	110	100	25	953

Analysis of compulsion from the point of view of the agents is reduced to (3.2), (3.3) in the case when the functions $s_i(t), i=1,2,...,n$, are identically equal to zero.

Example 3.1.

For the values from Table 3.1 for impulsion and $n=2$: $J_i=33$ *(thousand rubles), i=1,2. For the values from Table 3.1 and* $B=2,\ 8$ *(thousand rubles per year)* $J_i=81.5$ *and* 21.2 *(thousand rubles) respectively.*

For the values from Table 3.1 and $c=10{,}100$ (thousand rubles per year) we receive $J_i=36$ and 29 (thousand rubles), and in the case $s_i=30{,}300$ (thousand rubles per year) the payoffs are $J_i=31$ and 32 thousand rubles respectively. The growth of the total operative cost (value c) or demand parameter (B) implies decreasing of the agents payoffs and value of output (q), and it is profitable for the agents to develop homogeneous non-unique courses. In the majority of investigated numerical examples the agents' investment to the development of new courses (k) is absent. Increasing of the Principal's grants tends expectably to the increase of the agents' payoffs (not very big), and their strategies practically not change.

For compulsion with the data from Table 3.1 consider the case of soft ($y_i \equiv y = K_{\max}/8$ or $K_{\max}/4$) and hard ($y_i \equiv y = 3K_{\max}/8$ or $K_{\max}/2$) control of the agents' investment by the Principal. Here the agents are obliged to invest in the development of new courses but this only diminish their payoffs. Thus, the values of data from Table 3.1 give the payoffs $J_i=31$ and 28 thousand rubles for the soft control, and 24 and 20 thousand rubles for the hard control respectively.

In the case of an indifferent Principal and absence of its control of the agents they do not invest for a broad class of model parameters. Thus, farther we consider the case of an interested Principal. Suppose that the agents form the grand coalition that have the summary payoff functional (2.12). It is difficult to find analytically the Stackelberg equilibrium in the game of the Principal with the grand coalition of agents. So, we have calculated it numerically by means of the method of qualitatively representative scenarios in simulation modeling [14].

4. BUILDING THE STACKELBERG EQUILIBRIUM IN THE CASE OF GRAND COALITION OF AGENTS

Here we consider a model (2.1), (2.2), (2.4) – (2.7), (2.12) for compulsion and (2.5)-(2.9), (2.11), (2.12) for impulsion. Its solution is the Stackelberg equilibrium given that the agents form the grand coalition. We use the QRS method in simulation modeling. Let $\Omega = V_1 \times ... \times V_n \times U_1 \times ... \times U_n$ be the set of possible outcomes of the Stackelberg game with n agents where $V_i, U_i, i=1,2,...,n$ are sets of the feasible strategies of the Principal and the agents respectively. In this case

$$V_i = \left\{ v_i = \begin{cases} s_i : s_i \ge 0, \sum_{i=1}^{n} s_i \le S, \text{ impulsion} \\ y_i : 0 \le y_i \le K_{\max}, \text{ compulsion} \end{cases} \right\};$$

$U_i = \{u_i = (k_i, q_i): y_i \le k_i \le K_{\max}; 0 \le q_i \le Q_{\max}\}$ and $V_i = V_i(V_1, V_2, ..., V_{i-1})$ $i = 2,3,...,n$.

Denote

$V = V_1 \times ... \times V_n = \{(v_1, v_2, ..., v_n): v_i \in V_i\}$; $U = U_1 \times ... \times U_n = \{(u_1, u_2, ..., u_n): u_i \in U_i\}$.

An empirical QRS method is based on the hypothesis that $\forall i = 1,2,...,n,\ V_i = V_i^{QRS};\ U_i = U_i^{QRS}$, where the sets $V_i^{QRS}; U_i^{QRS}$ contain qualitatively representative strategies of the i-th agent and of the Principal in relation to the i-th agent. It is supposed that the cardinality of the sets $V_i^{QRS}; U_i^{QRS}$ $i = \overline{1,n}$ is small: $|V_i^{QRS}| = K_1$; $|U_i^{QRS}| = K_2$, where numbers K_1, K_2 are small. Then $V_1^{QRS} \times ... \times V_n^{QRS} \times U_1^{QRS} \times ... \times U_n^{QRS} = QRS$ is the QRS set of the game, and its cardinality is equal to $m = |QRS| = \prod_{i=1}^{n} |V_i^{QRS}||U_i^{QRS}| = K_1^n K_2^n$.

Definition 4.1 [8]:

A set $QRS = \{(v,u)^{(1)}, (v,u)^{(2)}, ..., (v,u)^{(m)}\}$ is called the QRS set of a difference Stackelberg game with precision Δ if:

(a) for any two elements $(v,u)^{(i)}, (v,u)^{(j)} \in QRS$ it is true that $|J_0^{(i)} - J_0^{(j)}| > \Delta$;

(b) for any element $(v,u)^{(l)} \notin QRS$ there is an element $(v,u)^{(j)} \in QRS$ such that $|J_0^{(l)} - J_0^{(j)}| \le \Delta$.

Here $J_0^{(i)}, J_0^{(j)}, J_0^{(l)}$ are the Principal's payoffs; $J_0^{(s)} = J_0(v_1^{(s)}, ..., v_n^{(s)}, u_1^{(s)}, ..., u_1^{(s)})$, $s = i, j, l; \Delta > 0$.

The reason of the QRS set is that the strategies from it imply essential difference between the Principal's payoffs, while this difference between any other scenario and one of the scenarios from QRS set is not essential. Thus, instead of deliberately impossible complete enumeration or even enumeration of many scenarios it is sufficient to consider a few scenarios from QRS set. Evidently, the constant Δ should be small enough, say, not greater than 10% of the typical values of payoffs. For the evaluation we use an indicator $\max\limits_{\substack{(u,v) \in QRS \\ (u,v) \ne (u_0, v_0)}} \frac{|J_0(u_0, v_0) - J_0(u,v)|}{|J_0(u_0, v_0)|}$.

Because the QRS method is empirical, it is important to choose appropriate initial scenarios as candidates to the QRS set. Let us choose the following scenarios.

In the case of compulsion the Principal bounds from below the values of investment of the agents to the innovative development. Suppose that the Principal's strategies do not vary with time, and $y_i(t) \equiv y_i \in \{0, \frac{K_{\max}}{2n}, \frac{K_{\max}}{n}\}$. The case $y_i = 0$ corresponds to the absence of control; $y_i = \frac{K_{\max}}{2n}$ describes a soft control, and $y_i = \frac{K_{\max}}{n}$ describes a hard control of the agents by the Principal. In the case of impulsion the Principal determines the grants allocated to the agents, and the total amount of all grants cannot exceed the whole budget S. The Principal's strategies $s_i(t) \equiv s_i \in \{0, s_i^{(1)}, s_i^{(2)}\}$ are taken so that $\sum_{i=1}^{n} s_i^{(1)} = \frac{S}{2}$; $\sum_{i=1}^{n} s_i^{(2)} = S$. The case $s_i = 0$ corresponds to the absence of grants, the case $s_i = s_i^{(1)}$ describes a middle level of granting, and the case $s_i = s_i^{(2)}$ reflects a considerable amount of the grants. The strategies of the agents describe the outputs $q_i \in \{0, \frac{Q_{\max}}{2}, Q_{\max}\}$ and the investments

$k_i \in \{0, \frac{K_{\max}}{2n}, \frac{K_{\max}}{n}\}$. For simplicity it is assumed that the strategies do not vary with time. Then we receive $m = |QRS| = 27^n$. For small values $n < 6$ the computer simulations are quite implementable.

With consideration of the proposed strategies from QRS set the state variable is found analytically as $D(t) = D(0)e^{-\frac{Kt}{1+K}}$; $K = \sum_{i=1}^{N} k_i$.

5. NUMERICAL CALCULATION OF THE STACKELBERG EQUILIBRIUM

The computer simulations were implemented on personal computer with microprocessor A10 Intel Pentium G4620 with operative memory 4 Gb using an object oriented programming language C# according to the proposed algorithm. An average time of one simulation for the construction a QRS set was less than three seconds. The received results were analyzed by the following criteria:

(a) summary discounted payoff of the Principal J_0 calculated by formulas (2.1) or (2.8);

(b) index of system compatibility [12]: $I = J_0^* / J_{\max}$, where $J_{\max} = \max_{u \in U} \max_{v \in V} J_0(u,v)$, $J_0^* = J_0(u^*, v^*)$ is the Principal's payoff in the Stackelberg equilibrium.

This index demonstrates how necessary is Principal's presence in a control system. The closer is the value of I to one, the more the system is compatible, and the less is a need in the hierarchical control by the Principal.

For an initial QRS set the conditions (a) and (b) from the Definition 1 are checked. The value of Δ is selected so that a difference between the Principal's payoffs for two any strategies from the initial QRS set does not exceed 10%. Then given the model parameters the second condition from the Definition 1 is checked. If necessary, the initial QRS set is extended or narrowed. In the case of extension the initial QRS set is added by new strategies. They correspond to the values situated between the previous ones (the strategies are supposed to be constant in time). Then the computer simulations are implemented. We have realized about 150 numerical calculations for three agents that form the grand coalition. The model parameters varied in the following range: $5 \le c \le 200$; $2 \le B \le 25$; $100 \le A \le 1000$; $20 \le Q_{\max} \le 200$; $100 \le K_{\max} \le 3000$; $1 \le Z \le 15$ for compulsion; $50 \le s \le 900$ for impulsion.

In all cases the index of system compatibility is equal to one, and the system is completely compatible. Besides, in the case of impulsion on a short period of modeling (up to 4 years) both for the Principal and for the agents it is not profitable to invest to the development of innovative technologies (learning courses).

Below the results of computer simulations for several sets of input data are presented. Let the grand coalition includes three agents, and $T = 1460$ days (4 years).

Example 5.1:

The values of model parameters are $c_1 = 50$; $c_2 = 100$; $c_3 = 10$; $S = 6000$; $A = 900$ *(thousand rubles per year).*

We varied the values B, $Q_{\max}$, $K_{\max}$. The results for impulsion in the case $y_i = 0$ $(i = 1,2,3)$ are given in Table 5.1. In the majority of calculations $D(t) \equiv B$.

Table 5.1. Results of numerical simulations (Example 2) for impulsion

B	$Q_{\max}$	$K_{\max}$	J_0
2	100	3000	207587

5	100	3000	117313
8	100	3000	73703
12.5	100	3000	398515
5	300	3000	63138
5	50	3000	96271
5	100	1000	117317
5	100	5000	117317
5	300	1000	63309
5	50	1000	96271

For small values of the parameter *B* (less than 17 thousand rubles per year) it is profitable for the agents to develop new non-unique courses. When the value of parameter *B* decreases (it is less than 1 thousand rubles per year), the number of developed courses is growing up to the maximal value $Q_{\max}$. When the value of parameter *B* increases, the agents stop the development. For a broad class of model parameters the payoff of the grand coalition of agents coincide with the Principal's payoff. To create an incentive for development of new unique courses it is necessary to increase the considered period of time or introduce additional requirements of sustainable development of the system controlled by the Principal.

The condition of sustainable development is taken in the form $\sum_{i=1}^{n} q_i \geq Q^{\lim}$ where $Q^{\lim}$ is a given constant that determines a minimal acceptable level of investment.

In the case of compulsion the calculations considered (a) middle ($s_i = 30$ thousand rubles per year) and (b) considerable ($s_i = 600$ thousand rubles per year) Principal's grants allocated to the agents, $i = 1,2$. Notice that in this setup the interests of the Principal and the agents coincide. For all studied examples $y_i \equiv 0$, $i = 1,2,..,n$, and the results are similar to the case of impulsion. The only essential distinction is a difference between the payoffs of the players that is determined by the Principal's grants allocated to the agents. As the interests of the Principal and the agents coincide, the Principal does not constraint the agents' investment from below.

6. CONCLUSION

A two-level control system of innovation in the universities is investigated. The Nash equilibrium in the game of agents in normal form is built analytically for a specific class of input functions by means of the Pontryagin maximum principle. An algorithm of building the Stackelberg equilibrium in the game of the Principal with the grand coalition of agents is proposed and implemented numerically by means of the method of qualitatively representative scenarios in simulation modeling. The following conclusions are received.

1. A demand parameter *B* in this model characterizes a degree of substitutability between two different courses in the initial moment of time. For the considered period of time (4 years) it is not profitable for the agents to develop new courses for big values of the parameter *B*.
2. For the chosen form of payoff functionals a need in control of the agents by the Principal is absent. The interests of the Principal and the agent coincide, and any compulsion is not required.
3. For successful promotion of the innovative learning courses in the universities it is necessary to formulate a condition of sustainable development as an obligatory one for the

Principal. After that, the interests of the Principal and the agents cease to coincide, and the Principal should provide the conditions of innovations.

The setup of dynamic problem of hierarchical conflict control in open-loop strategies is quite complicated, and many research directions stay to be open for further deep investigation. First of all, it concerns a theoretical substantiation of the existence of Nash equilibrium. It is also useful to study the solution's behavior on the bound of the domain of feasible controls where the switching points are possible. The situation is complicated by the game theoretic setup due to which the mentioned peculiarities could differ for different players, and the right hand sides of their optimal control problems can be discontinuous. At last, other forms of the inverse demand function can be used.

The main research tool, especially in a Stackelberg game, was the method of qualitatively representative scenarios in simulation modeling. Its idea is to receive a qualitatively acceptable description of the dynamics of a controlled system by means of the enumeration of a few impact scenarios found from reasonable considerations. In a Stackelberg game a set of scenarios is considered to be qualitatively representative if the Principal's payoffs for these scenarios differ essentially (more than a constant Δ), and for any other scenario exists a scenario from this set such the difference of the respective payoffs is less than Δ (i.e. the difference is not essential). These conditions are based on the well known notions of internal and external stability of a set and seem to be natural. However, any objective estimates of the value of Δ are still absent. From engineering and physical analogies it seems sufficient that the value of Δ does not exceed 5-10% from the typical values of payoffs received in numerical simulations. Certainly, this hypothesis needs to be refined and substantiated.

ACKNOWLEDGEMENTS

This work was supported by the Russian Science Foundation, project #17-19-01038.

REFERENCES

1. Basar T., Olsder G.J. (1999). *Dynamic Non-Cooperative Game Theory,* SIAM, 1999. - 520 p.
2. Cellini R., Lambertini L. (2002). A differential game approach to investment in product differentiation, *Journal of Economic Dynamics and Control*, 27(1), 51-62.
3. Cellini R., Lambertini L. (2004). Private and social incentives towards investment in product differentiation, *International Game Theory Review*, 6(4), 493–508.
4. Dockner E., Jorgensen S., Long N.V., Sorger G. (2000). *Differential Games in Economics and Management Science*. Cambridge University Press, 2000. – 396 p.
5. Dubina I. (2011). Foreign Investigations in the Field of Game-Theoretic Analysis of Innovations, *Automation and Remote Control*, 72(8), 1753-1765.
6. Geras'kin M. I. (2018). Modeling reflexion in the non-linear model of the Stackelberg three-agent oligopoly for the Russian telecommunication market, *Automation and Remote Control*, 79(5), 841–859.
7. Geraskin M. I. (2020). Reflexive games in the linear Stackelberg duopoly models under incoincident reflexion ranks, *Automation and Remote Control*, 81(2), 302–319.
8. Geraskin M. I. (2020). The properties of conjectural variations in the nonlinear Stackelberg oligopoly model, *Automation and Remote Control*, 81(6), 1051–1072.
9. Gorelov M.A., Kononenko A.F. (2015). Dynamic models of conflicts. III. Hierarchical games, *Automation and Remote Control,* 76(2), 264-277.
10. Malsagov M. Kh. (2019). Sustainable Management in Universities, *Advances in Systems Science and Applications*, 9(3), 65-79.
11. Mechanism design and management: *Mathematical methods for smart organizations* (2013). Ed. by D. Novikov. New York: Nova Science Publishers. – 163 p.
12. Ougolnitsky G.A. (2011). *Sustainable Management*. New York: Nova Science Publishers, 2011. – 288 p.

13. Ougolnitsky G.A., Usov A.B. (2016). Solution algorithms for differential models of hierarchical control systems, *Automation and Remote Control*, 77(5), 872-880.
14. Ougolnitsky G.A., Usov A.B. (2018). Computer Simulations as a Solution Method for Differential Games. In M.D. Pfeffer and E. Bachmaier (Eds.). *Computer Simulations: Advances in Research and Applications* (pp. 63-106). N.Y.: Nova Science Publishers.
15. Pontryagin L. S., Boltyanskii V. G., Gamkrelidze R. V., Mishchenko E. F. (1962). *The Mathematical Theory of Optimal Processes*. New York/London. John Wiley & Sons. – 369 p.
16. Ugol'nitskii G.A., Usov A.B. (2014). Equilibria in models of hierarchically organized dynamic systems with regard to sustainable development conditions, *Automation and Remote Control*, 75(6), 1055-1068.