Numerical solution of the dynamic incentive problem in discrete time taking into account the learning curve effect

Oleg Pavlov
Institute of economics and management
Samara National Research University
Samara, Russia
pavlov@ssau.ru

***Abstract*—The paper considers the dynamic incentive problem in discrete time, taking into account the learning effect. The task is formulated as a dynamic game between the leader and the performers. To solve the problem, the principle of cost recovery is applied, which reduces the original task to the optimal control problem in discrete time. Numerical solutions of the problem for various models of learning curves are obtained using the Bellman dynamic programming method. Also, the study is conducted of the discount rate's impact on the solution of the incentive problem**

Keywords—dynamic incentive problem, inverse Stackelberg game, learning curve effect, Bellman dynamic programming

I. Introduction

The article discusses the game dynamic task of the executors performing the production task in the context of new product development. The development of new products at industrial enterprises is characterized by the learning curve effect, which is that the time spent by employees (laboriousness) on performing multiple repetitive production operations is reduced.

The task of executors stimulation is one of the most important in the management theory. The management (the center) should choose such an incentive system based on the forecast of the agent's actions in order to ensure the fulfillment of their economic interests. The executor (the agent) chooses an action (volume of work) based on his economic interests.

Dynamic problems of interaction of unequal players are considered in the active systems theory [1], in the information theory of hierarchical systems [2–4] and in the dynamic games theory developed by international authors [5–11]. It should be noted that the stimulation problem in different theories has received various names. In the active systems theory it is the incentive task, in publications of foreign authors on game theory it is the inverse Stackelberg game, in the information theory of hierarchical systems it is the Germeyer game.

The active systems theory [1] offers the approach called the principle of agent's cost compensation. The center pays material remuneration to the agent, compensating his costs, in the case of choosing the optimal planned trajectory of the center and does not pay material compensation otherwise. The initial problem is divided into two tasks: the choice of the incentive system and the solution of the optimal control problem. In [12], results are presented that generalize the theorems from the monograph [1].

The hierarchical systems theory [2–4] suggests the approach that uses the center's choice of the program of joint actions with the agent and punishment for deviation from this program. When making a decision, the center proceeds from the principle of maximum guaranteed result. As a result, the initial problem is transformed into the optimal control problem.

In this article basing on the approach [1,12], the dynamic incentive problem of agents taking into account the learning curve effect is formulated and numerically solved using the Bellman dynamic programming method.

II. Statement and Algorithm for Solving the Dynamic Incentive Problem of Agents

A two-level dynamic manufacturing system consisting of a center and n independent agents is considered. Agents produce parts from which the finished product is then assembled. Labor costs and financial incentives for agents depend only on their own actions. This article applies the principle of game decomposition [1], which allows to consider the management of the i-th agent independently and not to take into account the interaction of agents with each other. The state of a dynamic production system depends on the actions of agents, and the center affects the managed system only through the payment of material remuneration to agents.

The dynamics of part production by the i-th agent is described by a discrete equation:

$$x_t = x_{t-1} + u_t, \quad t = \overline{1,T},$$

where x_t is the cumulative production volume of the part in the time period t, t is the number of the time period, u_t is the production volume of the part in the period t, T is the quantity of time periods considered.

Before the start of mass production, we know the number of manufactured parts, it is as follows:

$$x_0 = X_0 .$$

In the final time period, the cumulative volume of parts must be equal to the specified as follows:

$$x_T = X_0 + R,$$

where R is the specified number of parts.

Restrictions are imposed on the production volume of the part:

$$0 \le u_t \le X_0 + R - x_{t-1}, \quad t = \overline{1,T} .$$

The target function of the center is to maximize the discounted total difference between the income from the manufactured parts and the costs of the agent's material compensation:

$$J_p = \sum_{t=1}^{T} \frac{1}{(1+r)^t} [\, p u_t - \sigma(x_t)] \rightarrow max, \qquad (1)$$

where p is the part price, $\sigma(x_t)$ is the center incentive function, r is the center discount rate.

The incentive function of the center is a rule in accordance with which a material remuneration is assigned to the agent for the amount of work performed. The center manages the production process through the mechanism of material incentives $\sigma(x_t)$, economically encouraging agents to fulfill the planned production volumes.

The discount rate helps to take into account the time preferences of the center (agent) for the cost of cash flows. The more distant in time the cash flow, the cheaper it is for the center (agent).

The target function of the agent is to maximize the discounted total difference between material remuneration and labor costs, expressed in monetary form:

$$J_a = \sum_{t=1}^{T} \frac{1}{(1+r)^t} [\sigma(x_t) - C_t(u_t, x_{t-1})] \to max, \quad (2)$$

where r is the agent discount rate, $C_t(u_t, x_{t-1})$ is the agent labor costs.

Agent labor costs are determined by the following equation:

$$C_t(u_t, x_{t-1}) = sc_t u_t \quad (3)$$

where s is the cost of one hour per agent, c_t is the laboriousness of manufacturing the part.

The dependence of the part laboriousness on the cumulative production volume is described by various models of the learning curve given in [13]-[15].

In accordance with his economic interests, the agent selects parts production volumes that maximize his target function (2). The center's task is to choose the optimal incentive system in which the agent will produce such parts production volumes that maximize the center target function (1).

To solve the formulated control problem, the principle of cost compensation is applied [1, 12]. The solution algorithm consists in dividing the initial problem into two tasks: choosing a compensatory incentive system and solving the optimal control problem with the objective function equal to the difference between the center's income and the agent's labor costs.

1. The choice of a compensatory incentive system.

The center selects a compensatory incentive system, which consists in compensating the agent costs in the case of choosing the optimal planned production volume of the center x_t^{opt} and the absence of material payments otherwise:

$$\sigma(x_t) = \begin{cases} C_t(u_t, x_{t-1}), \ \text{если} \ x_t = x_t^{opt}, \ \text{для} \ \forall \ t = 1,T, \\ 0, \quad \text{если} \ x_t \neq x_t^{opt}, \ \text{для} \ \forall \ t = 1,T. \end{cases}$$

2. The solution of the optimal control problem with the target function equal to the difference between the center income and the agent labor costs.

To encourage the agent to choose the planned production volume, the center pays a material remuneration equal to the agent costs:

$$\sigma(x_t) = C_t(u_t, x_{t-1}). \quad (4)$$

We substitute the formula (4) into the target function of the center, taking into account (3):

$$J_p = \sum_{t=1}^{T} \frac{1}{(1+r)^t} [p - sc_t] u_t \to max.$$

Since the part price p is constant, the center can increase his profit only by minimizing the total cost of paying the agent's material remuneration. The target function of the center will take the following form:

$$J_p = \sum_{t=1}^{T} \frac{1}{(1+r)^t} sc_t u_t \to min.$$

Thus, the initial dynamic incentive task is reduced to the optimal control problem:

$$J_p = \sum_{t=1}^{T} \frac{1}{(1+r)^t} sc_t u_t \to min. \quad (5)$$

$$x_t = x_{t-1} + u_t, \quad t = 1,T, \quad (6)$$

$$x_0 = X_0, \quad (7)$$

$$x_T = X_0 + R, \quad (8)$$

$$0 \le u_t \le X_0 + R - x_{t-1}, \ t = 1,T. \quad (9)$$

The center's task is to select the optimal production volumes of parts u_t^{opt}, taking into account restrictions (9), under which the production process (6) will switch from the initial state (7) to the final state (8) and the minimum of the center's target function (5) will be achieved.

The formulated optimal control problem (5)-(9) was solved using the Bellman dynamic programming method [16], implemented in the pascal programming language.

III. The Results of the Numerical Solution of the Dynamic Incentive Problem of Agents

The numerical solution of the optimal control problem is carried out on the example of the production of parts of the enterprise Salut JSC. According to the enterprise data, regression models of the of laboriousness manufacturing parts are constructed: power, exponential and logistic.

Power-based labour input model:

$$c_t = 42{,}64 x_{t-1}^{-0{,}3}.$$

Exponential labor input model:

$$c_t = 9{,}17 + 6{,}16 e^{-0{,}03 x_{t-1}}.$$

Logistic labor input model:

$$c_t = 55{,}10 + 36{,}61 \left[\frac{1}{1 + 0{,}017 e^{0{,}05 x_{t-1}}} \right].$$

To solve the problem, the following data was used: the number of time periods T=12 months, the production volume of parts R=240 pcs., production experience before serial production $x_0 = 1$ pcs. The discrete step of changing the parts production volume when implementing the dynamic programming method is 1 pcs.

Numerical solutions of the optimal control problem for power-based, exponential and logistic models of labor input are presented in Fig. 1-3. The figures show the optimal trajectories of cumulative production volumes for various discount rates.

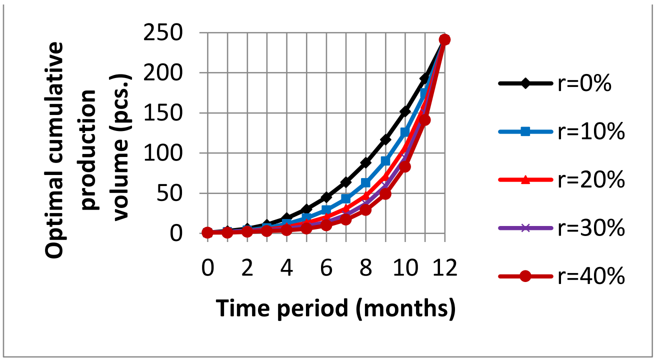

Fig. 1. The dependence of the optimal cumulative production volume on the discount rate for the power-based laboriousness model.

From an analysis of Fig. 1-2, it follows that for a power-based and exponential model of labour input, a convex curve is the optimal trajectory of the cumulative production volume. The optimal strategy of the center is the redistribution of large production volumes of parts for the last time periods in which the production laboriousness of parts is less than in the initial ones.

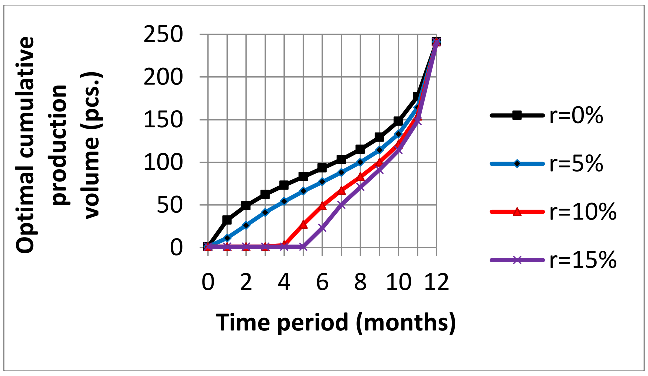

Fig. 2. The dependence of the optimal cumulative production volume on the discount rate for the logistic laboriousness model.

With an increase in the discount rate, the center's strategy to redistribute large production volumes of parts for the last time periods intensifies. This is due to the "cheaper" cost of the money that the center pays to the agent as a material reward in remote time periods. With large discount rates, the effect of deferring the production of parts from the initial time periods to later ones occurs. It is economically advantageous for the center to postpone the production of parts to late time periods, since in this case its total discounted costs will be minimal.

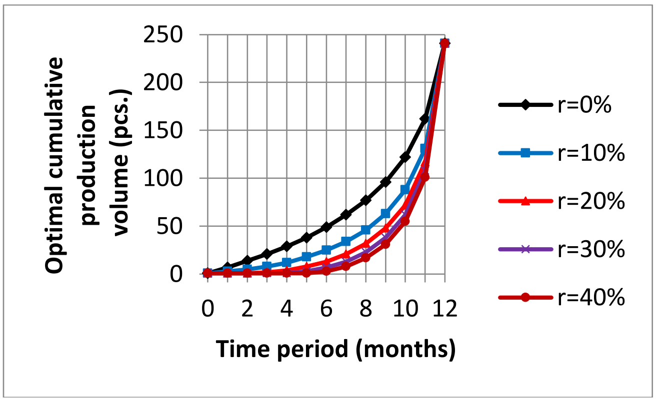

Fig. 3. The dependence of the optimal cumulative production volume on the discount rate for the exponential laboriousness model.

Analyzing Fig. 3, we conclude that for the logistic laboriousness model in the absence of discounting (*r*=0%), the optimal trajectory of cumulative production volume is the logistic curve. The optimal trajectory of the cumulative production volume consists of two sections: concave and convex.

Fig. 4 shows the optimal trajectories of production volumes for various discount rates *r*. The optimal strategy of the center in the absence of discounting (r = 0%) is: reduction of production volumes for the concave section of the optimal trajectory of the cumulative production volume and increase in production volumes for the convex section of the trajectory. The minimum of production volume corresponds to the inflection point of the optimal trajectory of the cumulative production volume.

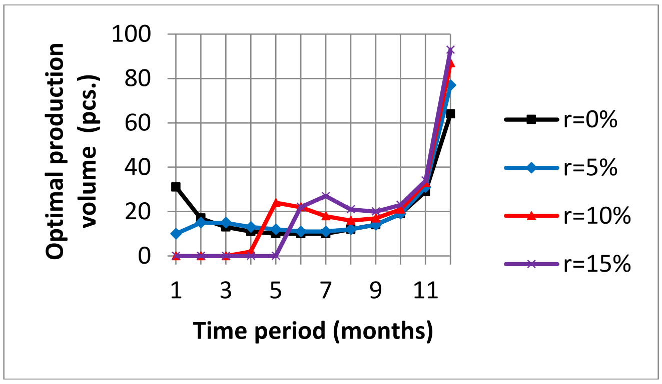

Fig. 4. The dependence of the optimal production volume on the discount rate for the logistic laboriousness model.

When discounting for the logistic model of laboriousness is taken into account, the effect of postponing the parts production from the initial time periods to later ones is also observed. Discounting leads to the appearance of the cumulative production volume of an additional convex section in the initial time periods on the optimal trajectory.

The optimal trajectory of the cumulative production volume is transformed into a curve of three sections: convex, concave and convex. The optimal strategy of the center is: on convex sections of the trajectory to increase production volumes, on concave sections - to decrease. Inflection points correspond to extreme values of production volumes.

IV. Conclusion

The paper considers the dynamic executors incentive task in discrete time, taking into account the learning curve effect. To solve the problem, the principle of cost compensation has been applied, which consists in dividing the original problem into two tasks: choosing a compensatory incentive system and solving the optimal control problem with the objective function equal to the difference between the income of the center and the labor costs of the agent.

Using the Bellman dynamic programming method, numerical solutions of the optimal control problem are obtained for various laboriousness models. The study of the impact of the discount rate on the solution of the incentive problem was conducted.

Based on a numerical study, the following conclusions are formulated:

1. The optimal strategy of the center for the power - based and exponential learning curves models is to

redistribute large production volumes of parts to the last time periods in which the production laboriousness of parts is less than in the initial ones.

2. The consideration of discounting for the power - based and exponential learning curves models leads to an even greater redistribution of the production volumes of parts over the last time periods.

3. Taking into account the discounting for all the considered learning curves models leads to the effect of postponing production from initial periods to later ones.

4. The optimal trajectory of the cumulative production volume in the case of the logistic learning curve model is a curve consisting of several convex and concave sections. The optimal strategy of the center is to increase production volumes on convex sections of the trajectory, and to decrease production volumes on concave sections. Inflection points correspond to extreme values of production volumes.

5. Taking into account the discounting for the logistic learning curve model leads to a redistribution of production volumes of parts in the middle and recent time periods.

Acknowledgment

The reported study was funded by RFBR and Samara region according to the research project № 17-46-630606.

References

[1] D.A. Novikov, M.I. Smirnov and T.E. Shokhina, "Mechanisms of Dynamic Active Systems Control," Moscow: IPU RAN, 2002, 124 p.

[2] V.A. Gorelik, M.A. Gorelov and A.F. Kononenko, "Analysis of Conflict Situations in Control Systems," Moscow: Radio i svyaz, 1991, 228 p.

[3] V.A. Gorelik and A.F. Kononenko, "Game-theoretic Models of Decision Making in Ecological–economic Systems," Moscow: Radio i svyaz, 1982, 144 p.

[4] M.A. Gorelov and A.F. Kononenko, "Dynamic conflict models III. Hierarchical games," Automation and Remote Control, vol. 2, pp. 89-106, 2015.

[5] T. Basar and G.J. Olsder, "Dynamic Noncooperative Game Theory," Philadelphia: SIAM, 1999, 519 p.

[6] E. Dockner, S. Jorgensen, N.V. Long and G. Sorger, "Differential Games in Economics and Management Science," Cambridge: Cambridge University Press, 2000, 382 p.

[7] Y.-C. Ho, P. Luh and R. Muralidharan, "Information Structure, Stackelberg Games, and Incentive Controllability," IEEE Trans. Automat. Control., vol. 26, no. 2, pp. 454-460, 1981.

[8] G.J. Olsder, "Phenomena in inverse Stackelberg games. Part 1: static problems," J. Optim. Theory Appl., vol. 143, no. 3, pp. 589-600, 2009.

[9] G.J. Olsder, "Phenomena in inverse Stackelberg games. Part 2: dynamic problems," J. Optim. Theory Appl., vol. 143, no. 3, pp. 601-618, 2009.

[10] N. Groot, B. De Schutter and H. Hellendoorn, "Reverse Stackelberg games. Part I: basic framework," Proc. of the IEEE Int. Conf. on Control Applications, pp. 421-426, 2012.

[11] N. Groot, B. De Schutter and H. Hellendoorn, "Reverse Stackelberg games. Part II: results and open issues," Proc. of the IEEE Int. Conf. on Control Applications, pp. 427-432, 2012.

[12] D.B. Rokhlin and G.A. Ougolnitsky, "Stackelberg equilibrium in a dynamic stimulation_model with complete information," Automation and Remote Control, vol. 79, no. 4, pp. 701-712, 2018.

[13] A. Badiru, "Computational survey of univariate and multivariate learning curve models," IEEE Transactions on Engineering Management, vol. 39, no. 2, pp. 176-188, 1992.

[14] L.E. Yelle, "The learning curve: historical review and comprehensive survey," Decision Sciences, vol. 10, no. 2, pp. 302-328, 1979.

[15] Y. Jaber, "Learning Curves: Theory, Models, and Applications," Boca Raton: CRC Press, 2011, 476 p.

[16] R. Bellman, "Dynamic Programming," Moscow: Izdatelstvo inostrannoy literatury, 1960.