Volume Title: Annals of Economic and Social Measurement, Volume 4, number 2

Volume Author/Editor: NBER

Volume Publisher: NBER

Volume URL: http://www.nber.org/books/aesm75-2

Publication Date: April 1975

Chapter Title: Survey of Nash and Stackelberg Equilibrim Strategies in Dynamic Games

Chapter Author: J.B. Cruz, Jr.

Chapter URL: http://www.nber.org/chapters/c10402

Chapter pages in book: (p. 339 - 344)

Annals of Economic and Social Measurement, 4/2, 1975

SURVEY OF NASH AND STACKELBERG EQUILIBRIUM STRATEGIES IN DYNAMIC GAMES*

BY J. B. CRUZ, JR.

The characteristics of Nash and Stackelberg equilibrium strategies for dynamic games are reviewed. Both strategies are appropriate when cooperation is not possible or when cooperation cannot be guaranteed. Open-loop, feedback and sampled-data strategies are distinguished by differences in the information sets available to the players. These strategies are secure against attempts by a single player to deviate from the equilibrium strategy during the time-horizon of the game.

I. INTRODUCTION

A dynamic game is a system with the following attributes:

(a) It has N persons, players, or decision-makers.

(b) Player i chooses a control variable $u^{(i)}$ from a set of admissible controls $U^{(i)}$.

(c) It has a time horizon which is defined by the interval $[t_0, t_f]$ where t_0 is known and fixed, and t_f may be fixed or free and it may be finite or infinite.

(d) It has a state $x(t)$ at time t, $t \in [t_0, t_f]$ which is an element of a finite dimensional vector space X. The evolution of the state is such that $x(t)$ is uniquely determined by the values of $u^{(i)}$ on $[t_1, t]$, $i = 1, \ldots, N$ and $x(t_1)$ for any t_1 satisfying $t_0 \leq t_1 < t$. We only consider state evolutions describable by differential equations or difference equations.

(e) Each player i has a real scalar cost function $J^{(i)}$ which is a mapping from X and $U^{(i)}, i = 1, \ldots, N$ to the set of real numbers.

(*f*) Each player i has knowledge of an information set $I^{(i)}$ which may include the differential equations for state evolution, the state x, its own cost function mapping as well as those of the other players, and control strategies of the other players. The set $\{I^{(i)}\}$ is called the information structure of the game.

(g) Each player i has a control law or strategy $\gamma^{(i)} \in \Gamma^{(i)}$ which is a mapping from the information set $I^{(i)}$ to the control space $U^{(i)}$.

A dynamic game whose state evolution is given by a differential equation

$$\dot{x} = f(x, u^{(1)}, \ldots, u^{(N)}, t) \tag{1}$$

is called a differential game. We assume that f is continuously differentiable in all its arguments and $t \in [t_0, t_f]$. We only consider the case when t_f is fixed.

It is assumed that $I^{(i)}$ for each i includes knowledge of f. Clearly, $\gamma^{(i)}$ is always part of $I^{(i)}$. Finally, we assume that $I^{(i)}$ includes $J^{(i)}$ and $x(t_0)$. Possible additional information to be included in $I^{(i)}$ will be considered later.

Two types of strategies are reviewed in this paper. One strategy called the Nash or Cournot equilibrium strategy will be considered first. The second strategy we consider is known as the Stackelberg equilibrium strategy.

* This work was supported in part by the National Science Foundation under Grant GK 36276, in part by the U.S. Air Force under Grant AFOSR-73-2570, and in part by the Joint Services Electronics Program under Contract DAAB-07-72-C-0259, with the Coordinated Science Laboratory, University of Illinois, Urbana, Illinois.

II. Nash Equilibrium Strategies

When all the cost function mappings are included in each $I^{(i)}$ in addition to the data enumerated in the previous section, and if there is a set of strategies $\gamma^{*(1)}, \ldots, \gamma^{*(N)}$ where $\gamma^{*(i)} \in \Gamma^{(i)}$ and

$$(2) \qquad J^{(i)}(\gamma^{*(1)}, \ldots, \gamma^{*(i-1)}, \gamma^{*(i)}, \gamma^{*(i+1)}, \ldots, \gamma^{*(N)}, \\ \leq J^{(i)}(\gamma^{*(1)}, \ldots, \gamma^{*(i-1)}, \gamma^{(i)}, \gamma^{*(i+1)}, \ldots, \gamma^{*(N)},$$

for all $\gamma^{(i)} \in \Gamma^{(i)}$ and for each $i = 1, \ldots, N$ then $\gamma^{*(1)}, \ldots, \gamma^{*(N)}$ is defined as a set of Nash equilibrium strategies [1, 2].

We define three types of Nash strategies. When $I^{(i)}$ for $i = 1, \ldots, N$ contain no other information, the Nash strategy is called an open-loop strategy. When $I^{(i)}$ for $i = 1, 2, \ldots, N$ include $X(t)$ at the present time t for all values of present time, the Nash strategy is called a closed loop or feedback Nash strategy. When $x(t_j)$ is included in $I^{(i)}$ for each i and each j where $\{t_j\}$ is a finite set of instants (countable set if t_f is infinite) where $t_j \leq t$ (present time) the Nash equilibrium strategy is called a sampled-data Nash strategy. Open-loop and closed-loop Nash strategies are described in [1–8, 10] and sampled-data Nash strategies are reported in [9].

The Nash strategy has the property that if all but one player use their Nash strategies, the deviating player could not decrease the value of his or her own cost function. Thus the Nash strategy safeguards against a single player deviating from the equilibrium strategy. However, two or more players could form a coalition and possibly the coalition could gain by deviating from the Nash strategy. The Nash strategy is reasonable when cooperation or coalition cannot be guaranteed and when the information structure is as stated above. Generally, the open-loop, sampled-data, and closed-loop Nash strategies yield different values for cost functions since the information structures are different.

In open-loop strategies, the players commit their control functions on $[t_0, t_f]$ before the start of the game. In sampled-data and closed-loop Nash strategies, the mappings from the space of $x(t_i)$ or $x(t)$ to $U^{(i)}$ are announced at the beginning of the game. However, the sampled-data and closed-loop Nash strategies satisfy the principle of optimality [9, 19]. This implies that if at any time t_j during the game, the ith player recomputes his or her sampled-data Nash strategy for a game starting at t_j, the desired sampled-data Nash strategy is identical to the previously computed strategy for $t \geq t_j$. This property of the Nash strategy means that the equilibrium condition is secure against any single player who may consider changing strategy *during* the game. The closed loop Nash strategy also satisfies the principle of optimality so that it is secure against any initiative of a single player to deviate from the Nash equilibrium strategy during the game. The open-loop Nash strategy is of course a special case of sampled-data Nash strategy where the only sampling time is t_0.

Necessary conditions for obtaining Nash equilibrium strategies for dynamic games have been obtained using either the variational approach or the optimal cost function approach via dynamic programming [1–4]. In general numerical algorithms based on these necessary conditions are complex and not easy to carry out [11]. When the differential equations or difference equations are linear and when the cost functions are quadratic in state and controls, the solutions may be expressed in terms of coupled Riccati-type matrix quadratic differential or difference equations [2, 3, 6–10]. For the infinite horizon case, there is no general result concerning the stability of the dynamic system. Nevertheless, in specific classes of linear-quadratic dynamic games, sufficient conditions for stability are available [12, 14].

We mention two situations in political science and economics where Nash strategies have been considered. In [12] arms race between two nations is modeled by a pair of first order differential equations where the control variables are expenditures in arms. These expenditures are chosen by the nations as Nash feedback strategies in a two-person dynamic game model. The cost functions are quadratic in arms level; expenditures, and consumption. As expected, the strategies are linear in arms level. The resulting closed-loop model gives a way of explaining an earlier model by Richardson which contains no controls [12]. A similar formulation for a discrete-time model has been considered in [13].

In [14], necessary conditions for Nash strategies have been obtained for a first order dynamic duopoly game. The concept of a dynamic demand function is introduced whereby the rate of change of price is modeled as a function of price and total quantity of goods in the market which is shared between two firms. The performance functions are the negative values of the profit. Specific results have been obtained when the dynamic demand is linear and when the cost is quadratic.

The Nash strategy for dynamic games, particularly the sampled-data Nash strategy provides an attractive conceptual tool in dynamic economic problems where Cournot equilibria are already accepted concepts for static models when the horizon is short or static models for steady state equilibria. The dynamic game models are more appropriate for intermediate length horizons or for infinite horizons where the adjustment period for reaching steady state is not to be ignored.

III. Stackelberg Equilibrium Strategies

In this section we only consider two-person games, where one player is called the leader and the other player is called the follower. The leader knows the cost function mapping of the follower but the follower may not know the cost function mapping of the leader. However, the follower knows the control strategy of the leader, and the follower always takes this into account in computing his or her strategy. If player 1 is the follower, $\gamma^{(1)}$ is restricted to those strategies which minimize $J^{(1)}$ for a given $\gamma^{(2)}$. The collection of pairs of such strategies is called the reaction set of player 1. If there exists a pair $\gamma^{*(1)}, \gamma^{*(2)}$ on the reaction set of player 1 such that

$$J^{(2)}(\gamma^{*(1)}, \gamma^{*(2)}) \leq J^{(2)}(\gamma^{(1)}, \gamma^{(2)}) \tag{3}$$

for any pair $(\gamma^{(1)}, \gamma^{(2)})$ on the reaction set of player 1, the pair $\gamma^{*(1)}, \gamma^{*(2)}$ is defined as a Stackelberg strategy with player 2 as leader. Observe that $\gamma^{*(1)}$ is the minimizing strategy of player 1 (the follower) corresponding to a strategy $\gamma^{*(2)}$ of player 2 [17–19, 22].

If no other data are included in the information sets except those described above and in Section 1, the Stackelberg strategies are called open-loop strategies. If $x(t)$ is added to the information sets of both players, the Stackelberg strategies are called closed-loop strategies. The Stackelberg strategy should be considered whenever a player has an option to declare his or her strategy in advance.

So far as the follower is concerned the Stackelberg strategy is obtained from an ordinary minimization problem and so long as the leader sticks to his or her strategy the principle of optimality applies to the follower. Thus there is no incentive for the follower to attempt to change strategy during the game, unless the leader changes strategy also. However, for the leader, the principle of optimality does not apply in general [19]. If the leader is allowed to change strategy during the game, say during sampling times of a sampled-data game, the leader will do so if the new game starting at the new time will give a lower value for the remaining cost function. Thus the strategies for all future times computed at a sampling time may not be implemented at all except for the first interval following the computation, and the strategies may fluctuate.

A modification of the Stackelberg strategy which satisfies the principle of optimality is now defined. This was previously called feedback Stackelberg strategy [19] (to distinguish it from closed loop Stackelberg strategy). However, because the terms closed-loop and feedback are interchangeably used in other contexts in many other areas, it is proposed that the modified Stackelberg strategy be called *Stackelberg equilibrium strategy*. The modified strategy is secure against potential changes by the leader during the game. Thus it is appropriate to call it an equilibrium strategy. The Stackelberg *equilibrium* strategy is simpler to compute than the Stackelberg strategy but it is still more difficult to compute compared to the Nash equilibrium strategy because of the requirement that the leader choose a strategy on the reaction curve of the follower. This imposes a complicated constraint for the optimization of the leader.

Formally, we define the Stackelberg equilibrium strategy for a discrete-time system,

$$(4) \qquad x(l+1) = f(x(l), l, u_1(l), u_2(l)), \qquad x(0) = x_0, \qquad l = 0, \ldots, N-1$$

where the state $x(l)$ and the decision variables $u_1(l)$ and $u_2(l)$ are n-dimensional, m_1-dimensional, and m_2-dimensional vectors of real numbers respectively. Let the cost functionals defined over the stages $k, \ldots, N$ be

$$(5) \qquad J^{(i)}(x(k), k, u_1, u_2) = K_i(x(N)) + \sum_{l=k}^{N-1} L_i(x(l), l, u_1(l), u_2(l)).$$

Let $u_1(l) = \gamma_{1s2}(l)$ and $u_2(l) = \gamma_{2s2}(l)$ be the Stackelberg equilibrium strategies with player 2 as leader starting at stage l in the sense to be defined here. The strategies $\gamma_{is2}(l)$ are mappings from the integer set $[l, \ldots, N-1]$ and $x(l)$ in the case of open-loop strategies, but in the case of closed loop strategies, these are mappings from the integer set $[l, \ldots, N-1]$ and the state set $\{x(k): k = l, \ldots, N-1\}$, where $x(k)$ is generated from (4), and the decision variables u_1 and u_2 belong to specified admissible sets. Denote the cost corresponding to the Stackelberg equilibrium strategies starting at stage $k+1$ by $V^{(i)}(x(k+1), k+1)$,

$$(6) \qquad V^{(i)}(x(k+1), k+1) = J^{(i)}(x(k+1), k+1, \gamma_{1s2}(k+1), \gamma_{2s2}(k+1)).$$

The admissible set for decision variables in defining the Stackelberg equilibrium strategies at stage k is restricted to the set of decision variables which are also Stackelberg equilibrium strategies at stage $k + 1$. Thus

$$J^{(i)}(x(k), k, u_1, u_2) = V^{(i)}(x(k + 1), k + 1) + L_i(x(k), u_1(k), u_2(k)). \tag{7}$$

With player 1 as follower and player 2 as leader, the Stackelberg equilibrium strategy is defined in the same way as the Stackelberg strategy in (3) except that instead of using $J^{(i)}$ in (5), we use $J^{(i)}$ in (7), with the boundary condition

$$V^{(i)}(x(N), N) = K_i(x(N)). \tag{8}$$

This definition is identical to the feedback Stackelberg strategy in [19] when the decision variables are closed-loop functions. However, in the present definition, the decision variables are not necessarily closed loop functions.

The Stackelberg equilibrium strategy is attractive when one player has enough information to be a leader. However, the potential leader may find that it is preferable to be a follower, in which case, he or she would divulge enough information for the other player to be a potential leader instead. Such a move would not necessarily convince the latter player to play as leader. However, once a player decides to lead, the other player would have no better choice than to play Stackelberg equilibrium strategy as follower assuming of course that the leader announces his strategy first.

The Stackelberg equilibrium strategy appears to be appropriate for the optimal stabilization problem considered in [15] where one player is the government policy maker and the other player is the competitive private sector. The private sector takes the government policy as given, and maximizes a consumer surplus objective function. The government policy maker must take into consideration the effect of their policy rules on the private sector's decision rules. The government policy should be chosen to maximize some welfare function, assuming that the private sector is reacting optimally. Although the Nash equilibrium strategy could be justified for this problem, the Stackelberg equilibrium strategy appears to be a more suitable concept in this case.

IV. Concluding Remarks

The properties of Nash and Stackelberg equilibrium strategies for dynamic games have been reviewed. These strategies are appropriate when cooperation is not possible or when cooperation cannot be enforced. In two-player games, these strategies are secure against attempts by a single player to deviate from its dynamic equilibrium strategy during the time horizon of the game. In economic situations where Cournot and Stackelberg equilibria are already useful concepts for static models or short-horizon models, the dynamic game models reviewed here should prove to be more useful when the horizon is not too short such that the transient adjustment period for reaching steady state is not to be ignored.

University of Illinois

References

[1] J. H. Case, "Toward a Theory of Many Player Differential Games," *SIAM J. on Control*, 7, (1969), pp. 179–97.

[2] A. W. Starr and Y. C. Ho, "Nonzero-Sum Differential Games," *Journal of Optimization Theory and Applications* (JOTA), 3, (1969), pp. 184–206.

[3] A. W. Starr and Y. C. Ho, "Further Properties of Nonzero-Sum Differential Fames," *JOTA*, 3, (1969), pp. 207–19.

[4] I. G. Sarma, R. K. Ragade, and U. R. Prasad, "Necessary Conditions for Optimal Strategies in a Class of Noncooperative N-person Differential Games," *SIAM J. on Control*, 7, (1960), pp. 637–44.

[5] Y. C. Ho, "Differential Games, Dynamic Optimization, and Generalized Control Theory," *JOTA*, 6, (1970), pp. 179–209.

[6] D. L. Lukes and D. L. Russel, "A Global Theory for Linear Quadratic Differential Games," *Journal of Mathematical Analysis and Applications*, 33, (1971), pp. 96–123.

[7] D. L. Lukes, "Equilibrium Feedback Control in Linear Games with Quadratic Costs," *SIAM J. on Control*, 9, (1971), pp. 234–52.

[8] M. H. Foley and W. E. Schmitendorf, "On a Class of Nonzero-Sum Linear Quadratic Games," *JOTA*, 7, (1971), pp. 357–77.

[9] M. Simaan and J. B. Cruz, Jr., "Sampled Data Nash Controls in Nonzero-Sum Differential Games," *Int. J. of Control*, 17, (1973), pp. 1201–09.

[10] M. Simaan and J. B. Cruz, Jr., "On the Solution of the Open-Loop Nash Riccati Equations in Linear Quadratic Differential Games," *Int. J. of Control*, 18, (1973), pp. 57–63.

[11] L. F. Pau, "Differential Game Among Sectors in a Macro-economy," IFAC/IFORS Int. Conf. on Dynamic Modeling and Control of National Economies, pp. 254–81.

[12] M. Simaan and J. B. Cruz, Jr., "A Differential Game Example of Armament Reduction," *Proc. of the 7th Annual Princeton Conf. on Inf. Sciences and Systems*, (1973), pp. 327–31. Revised version to appear in *The Review of Economic Studies*, "Formulation of Richardson's Model of Arms Race from a Differential Game Viewpoint," 1975.

[13] M. Simaan and J. B. Cruz, Jr., "Nash Equilibrium Strategies for the Problem of Armament Race and Control," to appear in *Management Science*, 1975.

[14] M. Simaan and T. Takayama, "Dynamic Duopoly Game: Differential Game Theoretic Approach," Faculty Working Paper #155, College of Commerce and Business Administration, University of Illinois, Urbana-Champaign, February 6, 1974.

[15] F. Kydland and E. C. Prescott, "Optimal Stabilization: A New Approach," presented at NBER-NSF Stochastic Control and Economic Systems Conference, Chicago, June 7–9, 1973.

[16] H. von Stackelberg, *The Theory of the Market Economy*, (1952), Oxford University Press, Oxford, England.

[17] C. I. Chen and J. B. Cruz, Jr., "Stackelberg Solution for Two-Person Games with Biased Information Patterns," *IEEE Trans. Aut. Cont.*, *AC*-17, (1972), pp. 791–98.

[18] M. Simaan and J. B. Cruz, Jr., "On the Stackelberg Strategy in Nonzero-Sum Games," *JOTA*, 11, (1973), pp. 533–55.

[19] M. Simaan and J. B. Cruz, Jr., "Additional Aspects of the Stackelberg Strategy in Nonzero-Sum Games," *JOTA*, 11, (1973), pp. 613–26.

[20] M. Simaan and J. B. Cruz, Jr., "A Stackelberg Strategy for Games with Many Players," *IEEE Trans. Aut. Cont.*, *AC*-18, (1973), pp. 322–24.

[21] T. Basar, "On the Relative Leadership Property of Stackelberg Strategies," *JOTA*, 11, (1973), pp. 655–61.

[22] M. Simaan, "Stackelberg vs. Nash Strategies in Nonzero-Sum Differential Games," *Proc. Eighth Princeton Conference on Systems and Information Sciences*, (1974).